

# ENDPOINT ESTIMATES AND SPARSE DOMINATION IN NONHOMOGENEOUS TREES

JOSÉ M. CONDE ALONSO, FILIPPO DE MARI, MATTEO MONTI, ELENA RIZZO, MARIA  
VALLARINO

ABSTRACT. We prove endpoint and sparse-like bounds for Bergman projectors on nonhomogeneous, radial trees  $\mathcal{X}$  that model manifolds with possibly unbounded geometry. The natural Bergman measures on  $\mathcal{X}$  may fail to be doubling, and even locally doubling, with respect to the right metric in our setting. Weighted consequences of our sparse domination results are also considered, and are in line with the known results in the disk. Our endpoint results are partly a consequence of a new Calderón-Zygmund theory for discrete, non-locally doubling metric spaces.

## INTRODUCTION

Discrete models of continuous objects are omnipresent in harmonic analysis. Many operators of interest can often be difficult to study directly due to issues related to convergence in their definition, subtle singular behavior, or other measure theoretic/functional analytic complications. Discrete models can remedy all these by providing objects which are easier to handle and that are defined and act on a measure space that can be identified with  $\mathbb{N}$ . This allows one to isolate the questions that one is interested in answering from the technicalities that surround them. Examples of successful discrete models include Haar shifts and other dyadic objects like square functions or martingale transforms (see [24] for a general account or [17] for an outstanding result), wavelet transforms [14], or weighted graphs that model weighted manifolds [18].

We are interested in discrete models of the hyperbolic disk and other noncompact symmetric manifolds. The natural discrete structure to use in that case is that of infinite trees. Trees are connected, loop-free graphs which may have a distinguished vertex 0. They are called homogeneous when all vertices have the same number of neighbors. It is known that homogeneous trees can be thought of as a discrete analogue of the hyperbolic disk, as

---

Conde-Alonso was partially supported by grants CNS2022-135431 and RYC2019-027910-I (Ministerio de Ciencia, Spain). De Mari, Monti, Rizzo and Vallarino were partially supported by the Project “Harmonic analysis on continuous and discrete structures” (bando Trapezio Compagnia di San Paolo CUP E13C21000270007) and are members of the Gruppo Nazionale per l’Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM). Monti is supported by the PRIN 2022 project “TIGRECO—Time-varying signals on Graphs: REal and COmplex methods” funded by the European Union—Next Generation EU (grant 20227TRY8H, CUP F53D23002630001).

Mathematics Subject Classification: 05C05; 46E22; 43B35.

made precise in, for example, [8]. In the same spirit, more general trees can be thought of as discrete counterparts of more general Riemannian manifolds. In this work, we consider trees  $\mathcal{X}$  that are radial. We say that  $\mathcal{X}$  is radial when each vertex  $x$  has exactly  $q(x) + 1$  neighbors in  $\mathcal{X}$ , and there exists  $\tilde{q} : \mathbb{Z}_+ \rightarrow \mathbb{Z}_+$  such that

$$q(x) = \tilde{q}(|x|) \quad \text{for all } x \in \mathcal{X},$$

where  $|x|$  denotes the distance from  $x$  to 0. Therefore, all vertices  $x$  at the same distance from 0 have the same number of neighbors. Radial trees can be seen as analogues of noncompact rotationally symmetric Riemannian manifolds (see [25, Section 2.3]), and so our results might shed light on their analysis at infinity.

Radial trees  $\mathcal{X}$  can be equipped with a measure from a natural class. Fix  $\alpha > 1$ . We consider a measure on  $\mathcal{X}$  given by

$$\mu(\{x\}) = \prod_{\ell=0}^{|x|-1} \frac{1}{\tilde{q}(\ell)^\alpha}, \quad x \in \mathcal{X}.$$

We assume  $\tilde{q}(\ell) \geq 2$  for all  $\ell$ , so  $\mu$  is a finite measure. The measure spaces that one obtains by making  $\alpha = 0$  or  $\alpha = 1$ , which do not fall under the scope of our work, correspond to the well known cases of the counting measure and the canonical flow measure, respectively. The case of  $\alpha > 1$  has also been considered before, albeit most often in a special situation described above, the homogeneous one. The measures introduced above are indeed a generalization of a class of the measures introduced in [9], where harmonic Bergman spaces on homogeneous trees were first considered. In turn, those are the natural discrete counterpart of the natural Bergman measures in the continuous hyperbolic disk, with the case of  $\alpha = 2$  corresponding to the Lebesgue measure.

We aim to understand the behavior of certain integral operators on  $(\mathcal{X}, \mu)$  and its associated  $L^p$ -spaces. An outstanding example of such an operator of interest in the hyperbolic disk—and other hyperbolic manifolds—is the Bergman projector  $\mathcal{P}$ , the projection from  $L^2$  to the subspace of holomorphic functions. In our real variable, discrete context,  $\mathcal{P}$  is the orthogonal projection from  $L^2(\mathcal{X})$  to the Bergman space  $\mathcal{A}^2(\mathcal{X})$ , the subspace of functions which are harmonic with respect to the combinatorial Laplacian on  $\mathcal{X}$ . We are interested in the  $L^p$  and endpoint behavior of  $\mathcal{P}$ . In addition, understanding weighted inequalities for  $\mathcal{P}$  is important in the study of Toeplitz operators in the continuous case, and we shall make a contribution in that direction in the discrete case. The work [13] already considers  $L^p$ -boundedness for  $\mathcal{P}$  in the homogeneous case, but weighted inequalities do not seem to follow easily from the approach taken there. In addition, our mild assumptions on  $\mathcal{X}$  imply that its geometry may be unbounded. In terms of harmonic analysis, this means that the measure  $\mu$  is not only nondoubling with respect to the natural graph metric, but it is also nondoubling—even non-locally doubling—with respect to the Gromov metric [30, p. 245], which is the right one to study  $\mathcal{P}$  as indicated by the results in [23]. The lack of the doubling property of  $\mu$  is one of the main challenges that we have to overcome in our setting.

The last few years have seen the development of a new technique within harmonic analysis that entails (usually sharp) weighted inequalities: sparse domination. Initially developed to provide a simpler proof of the so-called  $A_2$  theorem [19, 20], it has later evolved to become a widespread tool to prove sharp estimates for many operators of interest. One can consult the survey [24] and the references therein for an extensive account of the developments in the area. A collection of sets  $\mathcal{S}$  is sparse if for each  $Q \in \mathcal{S}$  there exists  $E_Q \subseteq Q$  such that  $\mu(E_Q) \geq \frac{1}{2}\mu(Q)$ , and the family  $\{E_Q : Q \in \mathcal{S}\}$  is pairwise disjoint. The sparse operator associated with  $\mathcal{S}$  is an averaging operator

$$\mathcal{A}_{\mathcal{S}}f(x) = \sum_{Q \in \mathcal{S}} \langle f \rangle_Q 1_Q(x),$$

where  $\langle f \rangle_Q$  denotes the integral average of  $f$  over the set  $Q$ . A sparse domination result for an operator  $T$  —in the dual sense— is an inequality like

$$\langle Tf_1, f_2 \rangle \lesssim \langle \mathcal{A}_{\mathcal{S}}f_1, f_2 \rangle,$$

where the sparse collection of sets  $\mathcal{S}$  may depend on  $f_1$  and  $f_2$ . We intend to follow that trend and look for an appropriate sparse domination result for  $\mathcal{P}$ . Like we said, the main obstacle that we shall encounter is that, even with the right choice of metric structure on  $\mathcal{X}$ ,  $\mu$  fails to be doubling over the relevant family of sets that are candidates to belong to a sparse family  $\mathcal{S}$ . In the nondoubling context, the existing sparse literature is either unsatisfactory [6, 29] or requires a substantial modification of the concept of sparse form [7]. This modification consists in adding an extra bilinear form to the right hand side of the inequality which accounts for certain off-diagonal interactions between the two functions  $f_1$  and  $f_2$ . Despite the fact that this extra form cannot be bounded by a sparse one —for the original operator does not admit sparse domination itself—, one can still recover useful quantitative information, including  $L^p$  boundedness results, from the resulting domination. We follow this and term our extra form  $\mathcal{E}_{\mathcal{S}}$ .

**Theorem A.** Let  $f_1, f_2 \in L^1(\mathcal{X})$ . There exists a sparse collection  $\mathcal{S}$  such that

$$|\langle \mathcal{P}f_1, f_2 \rangle| \lesssim \langle \mathcal{A}_{\mathcal{S}}f_1, f_2 \rangle + \mathcal{E}_{\mathcal{S}}(f_1, f_2).$$

We postpone the definition of  $\mathcal{E}_{\mathcal{S}}$  to Section 2 (see formula (2.1)). It captures the local behavior of the operator, accounting for interactions between vertices in  $\mathcal{X}$  —and hence sets in  $\mathcal{S}$ — that are close to one another.  $\mathcal{E}_{\mathcal{S}}$  cannot be avoided in the general case, but it is not needed when  $q$  is a bounded function. In that case, our argument can be easily modified to yield a domination of  $\mathcal{P}$  involving only a usual sparse form. It should be noted that in that case the result can be achieved using one of the existing sparse domination principles, like that in [22]. However, as we already explained, the nondoubling context poses significant difficulties making all previously existing approaches unapplicable. In order to prove Theorem A, we avoid writing down an explicit formula for the kernel of  $\mathcal{P}$ , and instead we look at it as a variant of a Haar shift. The analogy is far from perfect and we still need to account for tail issues and lack of localization, but it allows us to loosely follow an approach reminiscent of the beautiful proof in [11].

Our sparse domination result for  $\mathcal{P}$  is of course tailored to the operator under analysis. However, a more general principle lies behind it that can be applied to any discrete nondoubling setting, where small sets and local singularities do not play any role. Indeed, given a general linear operator  $T$ , our sparse domination procedure can always be carried out, leading to the identification of the right analogue of  $\mathcal{E}_S$  if it is needed. One can see the dyadic results of [7] and Theorem A above as instances of this general strategy, of which more details are given in Section 2.

We next identify the weighted inequalities for  $\mathcal{P}$  that follow from our sparse domination result. As we explained above, we are forced to include the form  $\mathcal{E}_S$  as part of our dominating term. This enters the definition of our classes of weights: the right one to dominate sparse operators is the Muckenhoupt class, while we necessarily need a larger one. For each  $1 < p < \infty$ , we denote the Muckenhoupt-like class of weights associated to the sets that form the families  $\mathcal{S}$  by  $\mathcal{B}_p(\mu)$ . Additionally, a —generally larger— quantity  $[w]_{\tilde{\mathcal{B}}_p(\mu)}$  is introduced in order to bound the term  $\mathcal{E}_S$  of the sparse domination. Then, our weighted inequalities for  $\mathcal{P}$  read as follows:

**Theorem B.** Let  $1 < p < \infty$ . If  $w \in \tilde{\mathcal{B}}_p(\mu)$ , then

$$\|\mathcal{P}f\|_{L^p(\mathcal{X},w)} \lesssim_{[w]_{\tilde{\mathcal{B}}_p(\mu)}} \|f\|_{L^p(\mathcal{X},w)}.$$

The dependence on  $[w]_{\tilde{\mathcal{B}}_p(\mu)}$  in the inequality above can be traced. For example, if  $p = 2$  one gets  $[w]_{\mathcal{B}_2(\mu)}[w]_{\tilde{\mathcal{B}}_2(\mu)}^{1/2}$ , which does not reduce to the optimal one in the classical case, as we explain below. In our proof, the extra power comes from the estimate of the term  $\mathcal{E}_S$  and does not seem to be avoidable in the nondoubling setting. We have decided to term  $\tilde{\mathcal{B}}_p(\mu)$  the class of weights that appears in Theorem B, because the class of sets over which weights  $w$  are tested in the definition of  $[w]_{\mathcal{B}_p(\mu)}$  makes it the exact discrete analogue of the Bekollé-Bonami class of weights [1], which is known to characterize the weighted bounds for the Bergman projection on the disk. In addition, our quantitative results are in line with their continuous counterpart from [26] —in the doubling setting—. The proof of Theorem B follows the standard argument that deduces weighted bounds from sparse domination (see, for instance, [10]), that we adapt to deal with the new term  $\mathcal{E}_S$  in our sparse domination result. We emphasize again that our scheme of proof for Theorems A and B has a wider applicability: we identify conditions under which an operator defined on a discrete structure that is equipped with a martingale filtration admits a sparse domination result and a family of nontrivial weighted inequalities.

We finally study the endpoint behavior of  $\mathcal{P}$ . At this point, we prefer to formulate our results in two separate ways. First, we study the endpoint behavior of general integral operators  $T$  whose kernels  $K$  satisfy adequate Hörmander-like and integral size conditions, that we term respectively (HörmCond) and (SizeCond). These should be compared to those naturally arising in nonhomogeneous Calderón-Zygmund theory [28, 27]. Our results hold for any tree equipped with any measure, and can be thought of as a general Calderón-Zygmund theory for non-locally doubling trees.

**Theorem C.** Let  $(\mathcal{X}, \mu)$  be a tree equipped with any measure. Let  $T$  be an  $L^2$ -bounded integral operator with symmetric kernel  $K$ . If  $K$  satisfies conditions (SizeCond) and (HörmCond), then  $T$  satisfies the following endpoint bounds:

- (a)  $T$  maps  $L^1(\mathcal{X})$  to  $L^{1,\infty}(\mathcal{X})$ .
- (b)  $T$  maps  $H^1(\mathcal{X})$  to  $L^1(\mathcal{X})$ .
- (c)  $T$  maps  $L^\infty(\mathcal{X})$  to  $BMO(\mathcal{X})$ .

The spaces  $BMO(\mathcal{X})$  and  $H^1(\mathcal{X})$  that appear in the statement of Theorem C are the martingale ones adapted to the natural filtration induced by the Gromov metric, and are introduced in Section 1. In particular, we have  $H^1(\mathcal{X})^* = BMO(\mathcal{X})$  and interpolation with the  $L^p(\mathcal{X})$  scale. In our radial setting, the kernel of  $\mathcal{P}$  satisfies (SizeCond) and (HörmCond), and so we can apply Theorem C to it and we obtain its  $L^p$ -boundedness as a consequence. We also recover the weak  $L^1$ -estimates obtained in [13, 12] in the homogeneous case and extend it to the unbounded branching setting. We expect that Theorem C will be applicable to other integral operators on —non-necessarily radial— trees with unbounded branching behavior. Our nondoubling results do not fit the framework of [16], because  $\mu$  does not satisfy an appropriate growth condition. Some elements in the proof of Theorem C are reminiscent of the arguments that naturally appear in nonhomogeneous Calderón-Zygmund theory [28], adapted to our specific discrete setting.

Our study of  $\mathcal{P}$  allows us to prove pointwise estimates for its kernel that resemble the classical Lipschitz smoothness ones for Calderón-Zygmund operators. This and the fact that  $\mathcal{P}1_{\mathcal{X}} = 1_{\mathcal{X}}$  suggests that the  $H^1$  and  $BMO$  endpoint estimates are suboptimal.

**Theorem D.** The following endpoint results hold for  $\mathcal{P}$ :

- (a)  $\mathcal{P}$  maps  $H^1(\mathcal{X})$  to  $H^1(\mathcal{X})$ .
- (b)  $\mathcal{P}$  maps  $BMO(\mathcal{X})$  to  $BMO(\mathcal{X})$ .

The  $BMO$  estimate for  $\mathcal{P}$  is new even in the homogeneous setting, although the slightly weaker  $L^\infty - BMO$  one had already been noted explicitly and could be deduced from the known results of Calderón-Zygmund theory in homogeneous spaces. The scheme of proof of the  $H^1 - H^1$  estimate seems entirely new. We emphasize once more that our framework is significantly different from all those previously studied in nonhomogeneous harmonic analysis: the measure  $\mu$  does not exhibit polynomial (or upper doubling) growth (so [15] cannot be followed for the  $BMO$  estimate), it is not balanced —in the sense of [21]—, and even if it clearly satisfies the isoperimetric property, it is not locally doubling, ruling out the techniques of [3, 4].

The rest of the paper is organized as follows: in Section 1 we explain our setting, the martingale filtration that we will be working with and some examples of operators that fit our framework. We also collect there some useful kernel estimates for  $\mathcal{P}$  that we use later in the paper. These are deduced from a novel decomposition of the operator into localizing projections. Section 2 contains our sparse domination result for  $\mathcal{P}$  and the

weighted inequalities that descend from it, in terms of the appropriate Bekollé-Bonami-like weight classes. Finally, in Section 3 we develop our discrete non-locally doubling Calderón-Zygmund theory and prove Theorems C and D.

## 1. TREES WITH UNBOUNDED GEOMETRY

**1.1. Radial trees of (possibly) unbounded degree.** A tree  $\mathcal{X}$  is a connected graph without loops. We denote the natural graph distance on  $\mathcal{X}$  by  $d$ . Two vertices  $x, y \in \mathcal{X}$  are called adjacent if  $d(x, y) = 1$ . We choose a vertex in  $\mathcal{X}$  that we call origin and denote by 0. For each  $x \in \mathcal{X}$ , we write

$$|x| := d(x, 0).$$

For each  $x \in \mathcal{X} \setminus \{0\}$ , we denote

$$q(x) + 1 = \#\{y \in \mathcal{X} : d(x, y) = 1\},$$

and we declare  $q(0) = \#\{y \in \mathcal{X} : |y| = 1\}$ . In that way,  $q(x)$  denotes the number of children of  $x \in \mathcal{X}$ , i.e. the vertices adjacent to  $x$  whose distance from 0 is  $|x| + 1$ . As we stated in the Introduction, we work with radial trees. This means that we postulate the existence of a function  $\tilde{q} : \mathbb{Z}_+ \rightarrow \{2, 3, 4, \dots\}$  such that

$$q(x) = \tilde{q}(|x|) \quad \text{for all } x \in \mathcal{X}.$$

Given  $0 \neq x \in \mathcal{X}$  with  $|x| = \ell$  and  $1 \leq k \leq \ell$ , we denote by  $x^{(k)}$  the only vertex  $y$  in the geodesic merging 0 and  $x$  such that  $|y| = \ell - k$ . Conversely, the family of vertices  $y$  such that  $y^{(1)} = x$  is denoted  $\mathcal{D}_1(x)$ . Fix  $\alpha > 1$ . We consider the radially decreasing measure on  $\mathcal{X}$  defined by

$$\mu(\{x\}) := \prod_{k=1}^{|x|} \frac{1}{q(x^{(k)})^\alpha} = \prod_{\ell=0}^{|x|-1} \frac{1}{\tilde{q}(\ell)^\alpha}.$$

Often, we abuse notation and write  $\mu(x) = \mu(\{x\})$  for measures  $\mu$  on  $\mathcal{X}$ . Without loss of generality, we may assume that  $\mu$  has been normalized to be a probability measure. The sector generated by  $x \in \mathcal{X}$  is the set

$$S_x = \{y \in \mathcal{X} : y^{(k)} = x \text{ for some } k \geq 1\} \cup \{x\}.$$

Due to the fast decay of  $\mu$ , we have, and will repeatedly use, the fact that  $\mu(S_x) \sim \mu(x)$ . The point  $x$  is called the vertex of the sector  $S_x$ . Given a sector  $S$ , we will sometimes use the notation  $x_S$  to denote the vertex of  $S$ . Given  $x, y \in \mathcal{X}$ , their confluent  $x \wedge y$  is the point  $z$  that belongs both to the chain  $(x^{(k)})_k$  and  $(y^{(\ell)})_\ell$  such that  $|z|$  is largest.

*Remark 1.1.* An alternative construction of a tree can be carried over without distinguishing an origin 0, by hanging  $\mathcal{X}$  by (a point in) its boundary  $\partial\mathcal{X}$ .  $\partial\mathcal{X}$  can be constructed as follows: let  $\mathcal{G}(\mathcal{X})$  be the space of infinite geodesics—infinite sequences  $(x_n)_n$  of adjacent points such that  $x_{n+2} \neq x_n$  for every  $n$ —. On  $\mathcal{G}(\mathcal{X})$ , we define the equivalence relation  $(x_n)_n \sim (y_m)_m$  if and only if there exists  $N \in \mathbb{Z}$  and  $n_0 \in \mathbb{N}$  such that  $x_n = y_{n+N}$  for every  $n > n_0$ . Then

$$\partial\mathcal{X} := \mathcal{G}(\mathcal{X}) / \sim.$$

Once we fix a hanging point  $\omega \in \partial\mathcal{X}$ , we can define a level structure and then a non-finite,  $\sigma$ -finite measure adapted to it on  $\mathcal{X}$ . We proceed along the lines of [12], but we need to make some modifications because our tree is no longer homogeneous, and the property of being radial is (can be) conditional on the choice of the boundary point over which  $\mathcal{X}$  hangs. Indeed, assume  $\omega \in \partial\mathcal{X}$  has been picked and for each  $y \in \mathcal{X}$  denote by  $(y_n)_n$  the (unique) representative of  $\omega$  such that  $y_0 = y$ . We then set  $y^{(n)} = y_n$  for all  $y$ . Fix  $x_0 \in \mathcal{X}$ . We proclaim that  $x_0$  belongs to the 0-th horocycle and  $x_n$  belongs to the  $(-n)$ -th horocycle, for  $n \geq 1$ . All other points  $y$  are assigned a horocycle as follows: if  $n_1$  and  $n_2$  are the smallest numbers such that  $y^{(n_1)} = x^{(n_2)}$  then  $y$  belongs to the  $(n_1 - n_2)$ -th horocycle. Having fixed  $\omega$ ,  $\mathcal{X}$  is radial if for each  $k \in \mathbb{Z}$ , the following holds: for all  $y$  in the  $k$ -th horocycle the set  $\{z \in \mathcal{X} : z^{(1)} = y\}$  has the same number  $q(y) = q_k$  of elements. Then we may define a measure  $\mu$  on  $\mathcal{X}$  as follows: we set  $\mu(\{x_0\}) = 1$ , and then we impose the relation

$$\mu(\{y\}) = \frac{\mu(\{y^{(1)}\})}{q(y^{(1)})^\alpha},$$

which defines  $\mu$  over the whole  $\mathcal{X}$ .

The measure  $\mu$  is not finite, but all sectors have finite measure and are very similar to the trees we consider below. In fact, any tree with origin can be seen as a sector in a non-finite measure tree as the ones constructed above. All of our results below can be generalized to radial trees of infinite measure writing  $\mathcal{X}$  as an increasing limit of sectors. We leave the concrete details and adjustments needed in each estimate to the interested reader.

**1.2. Martingale structure.  $H^1$  and BMO.** We construct a (one-sided) martingale filtration induced by the Gromov metric in  $\mathcal{X}$ . Each  $\sigma$ -algebra is generated by a partition  $\mathcal{D}_k$ , which is trivial for  $k = 0$  and is constructed as follows for  $k \geq 1$ :

$$\mathcal{D}_k = \left\{ \{x\} : |x| < k \right\} \cup \left\{ S_y : |y| = k \right\}.$$

We use notation inherited from that of the usual dyadic system: we use the letters  $Q, R, S$  to denote elements of  $\mathcal{D}$ , and the only  $S \in \mathcal{D}_k$  that contains a given  $Q \in \mathcal{D}_{k+s}$  is denoted by  $Q^{(s)}$ . When  $Q = \{x\}$  is a singleton, we sometimes write  $x = x_Q$ . Clearly, if  $\{x\} \in \mathcal{D}_k$  for some  $k$ , then  $\{x\} \in \mathcal{D}_\ell$  for all  $\ell > k$ . It will be useful for us to always view singletons as points elements in the dyadic generation  $\mathcal{D}_\ell$  with the smallest possible value of  $\ell$ . This means that we will always think that  $\{x\} \in \mathcal{D}_k$  if

$$\min \{ \ell : \{x\} \in \mathcal{D}_\ell \} = k.$$

Also, when needed, we will denote

$$\mathcal{D}_t(Q) = \{ R \in \mathcal{D} : R^{(t)} = Q \},$$

and

$$\mathcal{D}_{\leq t}(Q) = \bigcup_{k=0}^t \mathcal{D}_k(Q).$$

The conditional expectations induced by  $\mathcal{D} = \{\mathcal{D}_k\}_{k \geq 0}$  are denoted

$$\mathbf{E}_{\mathcal{D}_k} f = \mathbf{E}_k f := \sum_{Q \in \mathcal{D}_k} \langle f \rangle_Q 1_Q,$$

while martingale differences are

$$\mathbf{D}_k f = \mathbf{E}_k f - \mathbf{E}_{k-1} f.$$

In our endpoint estimates we use the spaces  $\mathbf{H}^1(\mathcal{X})$  and  $\mathbf{BMO}(\mathcal{X})$  defined with respect to the filtration generated by  $\mathcal{D}$ . The norm in the latter is given by

$$\|f\|_{\mathbf{BMO}} = \|\mathbf{E}_0 f\|_{\infty} + \sup_{k \geq 1} \|\mathbf{E}_k |f - \mathbf{E}_{k-1} f|\|_{\infty},$$

We will use the following useful expression to compute BMO norms in the sequel:

$$\|f\|_{\mathbf{BMO}} \sim \sup_{Q \in \mathcal{D}} \frac{1}{\mu(Q)} \sum_{x \in Q} |f(x) - \langle f \rangle_Q| \mu(x) + \sup_{\substack{Q \in \mathcal{D} \\ Q \neq \mathcal{X}}} |\langle f \rangle_Q - \langle f \rangle_{Q^{(1)}}| + \left| \sum_{x \in \mathcal{X}} f(x) \mu(x) \right|. \quad (1.1)$$

The norm in  $\mathbf{H}^1(\mathcal{X})$  is defined via the martingale square function:

$$\|f\|_{\mathbf{H}^1} = \left\| \left( \sum_{k \geq 1} |\mathbf{D}_k f|^2 \right)^{\frac{1}{2}} \right\|_1.$$

It will be useful to work with atomic descriptions of  $\mathbf{H}^1(\mathcal{X})$ . In what follows, an algebraic atom is a function  $b \in L^1(\mathcal{X})$  which is supported on some  $Q^{(1)} \in \mathcal{D}_k$  for some  $k$ —which cannot be a single point—and has vanishing integral. Following [5], an algebraic atom is an atomic block if  $b = \sum_i \lambda_i a_i$ , where  $\text{supp}(a_i) \subseteq Q_i$  for some  $Q_i \in \mathcal{D}_{k_i}$ ,  $k_i \geq k$  and

$$\|a_i\|_{\infty} \leq \mu(Q_i)^{-1} \frac{1}{k_i - k + 1}.$$

Every  $a_i$  is said to be a subatom. For every atomic block we define

$$|b|^1 = \inf_{\substack{b = \sum_i \lambda_i a_i \\ a_i \text{ subatoms}}} \sum_{i=1}^{\infty} |\lambda_i|.$$

Then  $f \in \mathbf{H}^1(\mathcal{X})$  if and only if  $f = \sum_j b_j$ , where each  $b_j$  is an atomic block, and

$$\|f\|_{\mathbf{H}^1} = \inf_{\substack{f = \sum_j b_j \\ b_j \text{ atomic block}}} \sum_{j=1}^{\infty} |b_j|^1 = \inf_{\substack{f = \sum_j b_j \\ b_j = \sum_i \lambda_{j,i} a_i}} \sum_{j,i=1}^{\infty} |\lambda_{j,i}|.$$

Atomic blocks can be assumed to adopt a simpler form than those, as we show next. An algebraic atom  $s$  is called a simple algebraic atom if there exists  $a : \mathcal{X} \rightarrow \mathbb{R}$  such that  $\text{supp}(a) \subseteq Q$  and

$$s = a - \langle a \rangle_{Q^{(1)}} 1_{Q^{(1)}}.$$

**Lemma 1.2.** *Every function  $f \in H^1(\mathcal{X})$  can be written as*

$$f = \sum_j \lambda_j s_j,$$

where each simple atomic block  $s_j = a_j - \langle a_j \rangle_{Q_j^{(1)}} 1_{Q_j^{(1)}}$  is such that

$$\|a_j\|_\infty \leq \mu(Q_j)^{-1}.$$

Moreover, we have

$$\|f\|_{H^1} \sim \inf_{\substack{f = \sum_j \lambda_j s_j \\ s_j \text{ simple atomic block}}} \sum_{j=1}^{\infty} |\lambda_j| =: \|f\|_{H_{\text{sb}}^1}.$$

*Proof.* Clearly,  $\|f\|_{H^1} \lesssim \|f\|_{H_{\text{sb}}^1}$ , for simple atomic blocks are atomic blocks. Indeed, if  $b = a - \langle a \rangle_Q 1_Q$  is a simple atomic block then  $a_1 = a$  and  $a_2 = -\langle a \rangle_Q 1_Q$  are subatoms satisfying the right size condition. Therefore, the infimum defining  $\|\cdot\|_{H^1}$  ranges over a larger set than that defining  $\|\cdot\|_{H_{\text{sb}}^1}$ .

Conversely, let  $b$  be an atomic block supported on  $Q \in \mathcal{D}_k$  with  $|b|^1 = 1$ . It is enough to show that  $\|b\|_{H_{\text{sb}}^1} \lesssim 1$ . Let  $b = \sum_j \lambda_j a_j$  be a decomposition of  $b$  into subatoms such that  $\sum_j |\lambda_j| \sim 1$ . For each  $j$ , denote the support of  $a_j$  by  $Q_j \in \mathcal{D}_{k_j}$  and write

$$a_j = a_j - \langle a_j \rangle_{Q_j^{(1)}} 1_{Q_j^{(1)}} + \sum_{\ell=1}^{k_j-k-1} \left[ \langle a_j \rangle_{Q_j^{(\ell)}} 1_{Q_j^{(\ell)}} - \langle a_j \rangle_{Q_j^{(\ell+1)}} 1_{Q_j^{(\ell+1)}} \right] + \langle a_j \rangle_Q 1_Q.$$

Since

$$\sum_j \lambda_j \langle a_j \rangle_Q = 0$$

because of the mean 0 of  $b$ , we can write

$$\begin{aligned} b &= \sum_j \lambda_j \left[ a_j - \langle a_j \rangle_{Q_j^{(1)}} 1_{Q_j^{(1)}} \right] \\ &\quad + \sum_j \lambda_j \frac{1}{k_j - k + 1} \sum_{\ell=1}^{k_j-k-1} (k_j - k + 1) \left[ \langle a_j \rangle_{Q_j^{(\ell)}} 1_{Q_j^{(\ell)}} - \langle a_j \rangle_{Q_j^{(\ell+1)}} 1_{Q_j^{(\ell+1)}} \right], \end{aligned}$$

which we claim is a decomposition of  $b$  into simple atomic blocks. Indeed, for each  $j$  and  $1 \leq \ell \leq k_j - k + 1$ ,

$$\left\| (k_j - k + 1) \langle a_j \rangle_{Q_j^{(\ell)}} 1_{Q_j^{(\ell)}} \right\|_\infty \leq \frac{(k_j - k + 1) \mu(Q_j)}{\mu(Q_j^{(\ell)})} \|a_j\|_\infty \leq \frac{1}{\mu(Q_j^{(\ell)})},$$

as required. An entirely similar estimate holds for the terms of the form  $a_j - \langle a_j \rangle_{Q_j^{(1)}} 1_{Q_j^{(1)}}$ . Therefore,

$$\|b\|_{\mathbb{H}_{\text{sb}}^1} \lesssim \sum_j |\lambda_j| \left( 1 + \sum_{\ell=1}^{k_j-k-1} \frac{1}{k_j - k + 1} \right) \lesssim \sum_j |\lambda_j| \sim 1,$$

which concludes the proof.  $\square$

*Remark 1.3.* The proof of Lemma 1.2 can clearly be done in any probability space, and not only in one whose associated filtration is atomic. We do not need that level of generality here and so we omit the details.

*Remark 1.4.* Above, we have dealt with atoms satisfying an  $L^\infty$  size condition. Equivalent norms and notions of both atomic blocks and simple atomic blocks can be constructed for each  $1 < p < \infty$  replacing said size condition by

$$\|a_j\|_p \leq \frac{1}{\mu(Q_j)^{\frac{1}{p}}}.$$

Lemma 1.2 holds as well for  $L^p$ -atomic blocks. We use the case  $p = 2$  below in Section 3.

**1.3. Operators acting on  $\mathcal{X}$ : Bergman spaces and projection-like operators.** The combinatorial Laplacian on  $\mathcal{X}$  is the operator defined by

$$\Delta f(x) = f(x) - \frac{1}{q(x) + 1} \sum_{y:d(x,y)=1} f(y).$$

A function  $f$  is called harmonic at a point  $x$  if  $\Delta f(x) = 0$ , and is called harmonic if it is harmonic at all  $x \in \mathcal{X}$ . The (real) Bergman space is

$$\mathcal{A}^2(\mathcal{X}) = \{f \in L^2(\mathcal{X}) : f \text{ is harmonic}\}.$$

This is a real-variable version of the classical Bergman space.  $\mathcal{A}^2(\mathcal{X})$  is a reproducing kernel Hilbert space for which an explicit orthonormal basis can be exhibited. To do so, we start defining a family of auxiliary functions. For each  $k \geq 0$ , we denote by  $\varphi_k$  the only function on  $\mathcal{X}$  such that  $\varphi_k(x) = 0$  if  $|x| < k$ ,  $\varphi_k(x) = 1$  if  $|x| = k$ , that is harmonic at all points  $x$  such that  $|x| > k$ , and is radial:  $\varphi_k(x) = \varphi_k(y)$  if  $|x| = |y|$ . The function  $\varphi_k$  encodes the decay properties of the kernels studied below, in terms of the following quantity:

$$\nu_j^k := \prod_{\ell=j}^k \frac{1}{\tilde{q}(\ell)}, \quad k \geq j$$

with the convention that  $\nu_j^k = 1$  if  $j > k$ . Many times, the trivial estimate  $\nu_j^k \leq 2^{-(k-j+1)}$  will be enough for our purposes, but the unboundedness of  $q$  will often force us to use the larger decay captured by  $\nu_j^k$ . We will sometimes use the multiplicative identity  $\nu_j^k =$

$\nu_j^m \cdot \nu_{m+1}^k$  for  $j \leq m \leq k$ . The following explicit formula, which can be checked by induction, allows one to compute the values of  $\varphi_k$ :

$$\varphi_k(x) = \sum_{\ell=1}^{|x|-k+1} \nu_k^{|x|-\ell}, \quad |x| \geq k. \quad (1.2)$$

Formula (1.2) immediately implies that  $1 \leq \varphi_k(x) \leq 2$  for all  $x \in \text{supp}(\varphi_k)$ , a fact that we will use repeatedly. The following more precise estimate will be useful in the sequel:

**Lemma 1.5.** *Let  $k \geq 0$ . If  $x, y$  are such that  $|x| \geq |y| = m \geq k$ , then*

$$\varphi_k(x) - \varphi_k(y) \lesssim \nu_k^m.$$

*Proof.* Indeed,

$$\begin{aligned} \varphi_k(x) - \varphi_k(y) &= \sum_{\ell=1}^{|x|-k+1} \nu_k^{|x|-\ell} - \sum_{\ell=1}^{|y|-k+1} \nu_k^{|y|-\ell} \\ &= \sum_{\ell=|x|-|y|+1}^{|x|-k+1} \nu_k^{|x|-\ell} - \sum_{\ell=1}^{|y|-k+1} \nu_k^{|y|-\ell} + \sum_{\ell=1}^{|x|-|y|} \nu_k^{|x|-\ell} \\ &= \nu_k^m \sum_{\ell=1}^{|x|-|y|} \nu_{m+1}^{|x|-\ell} \lesssim \nu_k^m, \end{aligned}$$

using  $\nu_{m+1}^{|x|-\ell} \lesssim 2^{-|x|+\ell+m}$ . □

Given a sector  $Q \in \mathcal{D}_{k-1}$  and a function  $e : Q \cap \{|x| = k\} \rightarrow \mathbb{R}$  with mean zero, the function

$$E_Q(e)(x) = \begin{cases} \varphi_k(x)e(y) & x \in S_y \text{ for some } y \in \mathcal{D}_1(x_Q), \\ 0 & \text{otherwise,} \end{cases}$$

is harmonic on  $\mathcal{X}$  and we call it the radial harmonic extension of  $e$ . For each  $x \in \mathcal{X}$ , denote  $Q = S_x$  and pick an orthonormal basis  $\{e^\ell\}_{\ell=1}^{q(x)-1}$  of the hyperplane

$$W_Q = \left\{ v = (v_j) \in \mathbb{R}^{q(x)} : \sum_j v_j = 0 \right\}.$$

Let  $\mathcal{D}_1(x) = \{y_j : j = 1, \dots, q(x)\}$ . For each  $\ell$ , define the function  $e_Q^\ell$  as the one supported in the successors of  $x$  that satisfies  $e_Q^\ell(y_j) = e_j^\ell$ . Finally, for each  $\ell$  and  $Q$  set

$$h_Q^\ell(z) = \frac{E_Q(e_Q^\ell)(z)}{\|E_Q(e_Q^\ell)\|_2}, \quad z \in \mathcal{X}.$$

When  $Q \in \mathcal{D}$  is a singleton, by convention we define  $h_Q \equiv 0$ . The next result uses a key property of the functions  $h_Q^\ell$ : they are not only orthogonal to constants, but moreover they

are orthogonal to all radial functions (and they themselves behave like radial functions in almost all sectors).

**Proposition 1.6.** The family

$$\mathcal{U} = \{h_Q^\ell : Q \in \mathcal{D}, 1 \leq \ell \leq q(x_Q) - 1\} \cup \{1_{\mathcal{X}}\}$$

is an orthonormal basis of  $\mathcal{A}^2(\mathcal{X})$ .

*Proof.* The functions under consideration are  $L^2$ -normalized, so we have to check orthogonality and completeness. First, if  $Q \in \mathcal{D}_{k-1}$  and  $\ell \neq \ell'$  then

$$\begin{aligned} \langle h_Q^\ell, h_Q^{\ell'} \rangle &= \sum_{x \in \mathcal{D}_1(x_Q)} \sum_{y \in S_x} e_Q^\ell(x) e_Q^{\ell'}(x) \varphi_k(y)^2 \mu(y) \\ &= \sum_{x \in \mathcal{D}_1(x_Q)} e_Q^\ell(x) e_Q^{\ell'}(x) \sum_{y \in S_x} \varphi_k(y)^2 \mu(y) \\ &= C_{k,Q} \sum_{x \in \mathcal{D}_1(x_Q)} e_Q^\ell(x) e_Q^{\ell'}(x) = C_{k,Q} \sum_{j=1}^{q(x_Q)} e_j^\ell e_j^{\ell'} = 0, \end{aligned}$$

by the orthogonality of the vectors that generate  $h_Q^\ell$  and  $h_Q^{\ell'}$ . Second, since  $\text{supp}(h_Q^\ell) \subseteq Q$  for all  $Q$ , the following —similar— computation for  $\mathcal{D}_{r-1} \ni Q \subsetneq R \in \mathcal{D}_{k-1}$  deals with all the remaining cases:

$$\begin{aligned} \langle h_Q^\ell, h_R^{\ell'} \rangle &= \sum_{x \in \mathcal{D}_1(x_Q)} \sum_{y \in S_x} e_Q^\ell(x) e_R^{\ell'}(x_0) \varphi_r(y) \varphi_k(y) \mu(y) \\ &= e_R^{\ell'}(x_0) \sum_{x \in \mathcal{D}_1(x_Q)} e_Q^\ell(x) \sum_{y \in S_x} \varphi_r(y) \varphi_k(y) \mu(y) \\ &= C_{k,r,Q} e_R^{\ell'}(x_0) \sum_{x \in \mathcal{D}_1(x_Q)} e_Q^\ell(x) = 0. \end{aligned}$$

Above, the point  $x_0 \in \mathcal{D}_1(x_R)$  is the only one such that  $Q \subseteq S_{x_0}$ . Finally, all functions  $h_Q^\ell$  have mean zero so they are orthogonal to constants.

It remains to show that  $\mathcal{U}$  generates the whole space. Suppose  $f \in \mathcal{A}^2(\mathcal{X})$  is orthogonal to every element of  $\mathcal{U}$ . In particular,  $f$  is orthogonal to constants and it can then be checked that  $f(0) = 0$ . We assume by induction that  $f(x) = 0$  if  $|x| < n$  and we prove that  $f(y) = 0$  for all  $y$  with  $|y| = n$ . Fix  $x$  with  $|x| = n - 1$  and denote  $Q = S_x$ . Since  $f(x) = f(x^{(1)}) = 0$ , we have

$$\sum_{y \in \mathcal{D}_1(x)} f(y) = 0.$$

But then, for all  $\ell$

$$0 = \langle h_Q^\ell, f \rangle = \sum_{y \in \mathcal{D}_1(x)} \sum_{z \in S_y} e_Q^\ell(y) f(z) \varphi_n(z) \mu(z)$$

$$\begin{aligned}
&= \sum_{y \in \mathcal{D}_1(x)} e_Q^\ell(y) \sum_{z \in S_y} f(z) \varphi_n(z) \mu(z) \\
&= \sum_{y \in \mathcal{D}_1(x)} e_Q^\ell(y) \sum_{j=1}^{\infty} \sum_{\substack{z \in S_y \\ |z|=|x|+j}} f(z) \varphi_n(z) \mu(z) \\
&= \sum_{y \in \mathcal{D}_1(x)} e_Q^\ell(y) \sum_{j=1}^{\infty} C_{j,n} f(y) \\
&= C_n \sum_{y \in \mathcal{D}_1(x)} e_Q^\ell(y) f(y),
\end{aligned}$$

using the radially of  $\mu$  and  $\varphi_n$ . This implies that  $f(y) = 0$  for all  $y \in \mathcal{D}_1(x)$ .  $\square$

Since pointwise evaluation is a bounded functional,  $\mathcal{A}^2(\mathcal{X})$  is a closed subspace of  $L^2(\mathcal{X})$ , as in the continuous case. The Bergman projector  $\mathcal{P}$  is the orthogonal projection  $L^2(\mathcal{X}) \mapsto \mathcal{A}^2(\mathcal{X})$ . Because of Proposition 1.6,  $\mathcal{P}$  has the explicit expression

$$\mathcal{P}f(x) = \langle f \rangle_{\mathcal{X}} 1_{\mathcal{X}}(x) + \sum_{Q \in \mathcal{D}} \sum_{\ell=1}^{q(x_Q)-1} \langle f, h_Q^\ell \rangle h_Q^\ell(x).$$

$\mathcal{P}$  is an integral operator with a kernel that can be written explicitly. But as we said in the Introduction, it is better to work with it adopting a different point of view. To that end, the following definition does not only ease our notation in the sequel, but it also encodes the right way of manipulating the pieces of the kernel of  $\mathcal{P}$ . For each  $Q \in \mathcal{D}$  which is not a singleton, we set

$$\mathcal{D}_Q f(x) := \sum_{\ell=1}^{q(x_Q)-1} \langle f, h_Q^\ell \rangle h_Q^\ell(x).$$

When  $Q$  is a singleton we set  $\mathcal{D}_Q f = 0$ . Clearly, for any  $f$  we have  $\text{supp}(\mathcal{D}_Q f) \subseteq Q \setminus \{x_Q\}$  and  $\mathcal{D}_Q f$  has vanishing mean. We will sometimes work with truncated versions of  $\mathcal{P}$ , which we denote

$$\mathcal{P}^R f := \sum_{\substack{T \in \mathcal{D} \\ T \subseteq R}} \mathcal{D}_T f,$$

in the case of the localized truncation inside  $R \in \mathcal{D}$ , and for a disjoint family  $\mathcal{F} \subseteq \mathcal{D}$

$$\mathcal{P}_{\mathcal{F}} f := \langle f \rangle_{\mathcal{X}} 1_{\mathcal{X}} + \sum_{\substack{T \in \mathcal{D} \\ T \not\subseteq R \text{ if } R \in \mathcal{F}}} \mathcal{D}_T f.$$

*Remark 1.7.* Following the classical definition of Haar shifts, given a pair  $(r, s) \in \mathbb{N}_+^2$  we define

$$\mathcal{P}_{\gamma}^{r,s} f(x) = \sum_{Q \in \mathcal{D}} \sum_{R \in \mathcal{D}_r(Q)} \sum_{S \in \mathcal{D}_s(Q)} \sum_{k=1}^{q(v_R)-1} \sum_{\ell=1}^{q(v_S)-1} \gamma_{R,S,k,\ell}^Q \langle f, h_R^k \rangle h_S^\ell(x).$$

with  $\gamma \in \ell^\infty$ . We call  $\mathcal{P}_\gamma^{r,s}$  a Bergman shift of complexity  $(r, s)$ . The Bergman projector corresponds to the case  $\mathcal{P}_{\gamma_0}^{0,0}$ , where

$$\gamma_0 = \gamma_{Q,Q,k,\ell}^Q = \delta_{k,\ell}.$$

It should be observed that, unless  $r = s = 0$ ,  $\mathcal{P}_\gamma^{r,s}$  cannot be expressed as a sum involving only the operators  $\mathcal{D}_Q$ . This is one way to see that, when  $q$  is unbounded, one should not expect to prove  $L^p$ -boundedness results for  $\mathcal{P}_\gamma^{r,s}$  if  $p \neq 2$ . The heuristic reason comes from [21], where the family of measures such that a Haar shift and its adjoint are  $L^p$ -bounded is characterized. The measure  $\mu$  does not belong to that class, and in an analogous way one should expect  $L^p$ -boundedness to fail. It is possible to prove positive boundedness results for  $\mathcal{P}_\gamma^{r,s}$  when  $q$  is bounded.

**1.4. Kernel estimates for  $\mathcal{P}$ .** We next collect two key estimates involving the operators  $\mathcal{D}_R$  that will be useful in the following sections. By definition,  $\mathcal{D}_R$  is an integral operator with kernel  $\mathcal{K}_R$  given by

$$\mathcal{K}_R(x, y) = \sum_{j=1}^{q(x_R)-1} h_R^j(x) h_R^j(y).$$

On the other hand, the space  $W_R$  introduced above is a reproducing kernel Hilbert space. Denoting by  $k_R(w, w') = \sum_{j=1}^{q(x_R)-1} e_R^j(w) e_R^j(w')$  its reproducing kernel, for  $R \in \mathcal{D}_{k-1}$  we have

$$\begin{aligned} \mathcal{K}_R(x, y) &= \sum_{j=1}^{q(x_R)-1} h_R^j(x) h_R^j(y) = \frac{q(x_R) \varphi_k(x) \varphi_k(y)}{\|\varphi_k 1_R\|_2^2} \sum_{j=1}^{q(x_R)-1} e_R^j(x^{|x|-|x_R|-1}) e_R^j(y^{|y|-|x_R|-1}) \\ &= \frac{q(x_R) \varphi_k(x) \varphi_k(y)}{\|\varphi_k 1_R\|_2^2} k_R(x^{|x|-|x_R|-1}, y^{|y|-|x_R|-1}). \end{aligned} \quad (1.3)$$

Finally, a direct computation shows that  $k_R$  is uniquely determined by

$$k_R(w, w') = \begin{cases} -\frac{1}{q(x_R)}, & \text{if } w \neq w', \\ 1 - \frac{1}{q(x_R)}, & \text{if } w = w'. \end{cases} \quad (w, w') \in W_R \times W_R.$$

**Lemma 1.8.** *If  $Q \in \mathcal{D}_k$  is not a single point, the following two estimates hold:*

(1) *For all  $\ell \geq 1$  and every  $z \in Q^{(\ell)}$  we have*

$$\sup_{x, y \in Q} |\mathcal{K}_{Q^{(\ell)}}(z, x) - \mathcal{K}_{Q^{(\ell)}}(z, y)| \lesssim \begin{cases} \frac{1}{\mu(Q^{(\ell-1)})} \nu_{k-\ell+1}^k, & \text{if } z \in Q^{(\ell-1)}, \\ \frac{1}{\mu(Q^{(\ell-1)})} \nu_{k-\ell}^k, & \text{otherwise.} \end{cases}$$

(2) For all  $x, y \in Q^{(1)} \setminus \{x_Q^{(1)}\}$ ,

$$|\mathcal{K}_{Q^{(1)}}(x, y)| \lesssim \begin{cases} \frac{1}{\mu(Q)}, & \text{if } x \wedge y \neq x_Q^{(1)}, \\ \frac{1}{q(x_Q^{(1)})\mu(Q)}, & \text{otherwise.} \end{cases}$$

*Proof.* (1) By the triangle inequality, it is enough to prove the assertion for  $y = x_Q$ . Since  $1 \leq \varphi_{k-\ell+1} < 2$ , using (1.3) and Lemma 1.5 yields

$$\begin{aligned} & \mathcal{K}_{Q^{(\ell)}}(z, x) - \mathcal{K}_{Q^{(\ell)}}(z, x_Q) \\ &= \frac{q(x_Q^{(\ell)})\varphi_{k-\ell+1}(z)(\varphi_{k-\ell+1}(x) - \varphi_{k-\ell+1}(x_Q))}{\|\varphi_{k-\ell+1}1_{Q^{(\ell)}}\|_2^2} k_{Q^{(\ell)}}(z^{|z|-k+\ell-1}, x_Q^{(\ell-1)}) \\ &\sim \nu_{k-\ell+1}^k \frac{1}{\mu(Q^{(\ell-1)})} k_{Q^{(\ell)}}(z^{|z|-k+\ell-1}, x_Q^{(\ell-1)}) \\ &\sim \begin{cases} \frac{1}{\mu(Q^{(\ell-1)})} \nu_{k-\ell+1}^k, & \text{if } z \in Q^{(\ell-1)}, \\ -\frac{1}{q(x_Q^{(\ell)})\mu(Q^{(\ell-1)})} \nu_{k-\ell+1}^k, & \text{otherwise.} \end{cases} \end{aligned}$$

(2) Again by (1.3),

$$\begin{aligned} \mathcal{K}_{Q^{(1)}}(x, y) &= \frac{q(x_Q^{(1)})\varphi_k(x)\varphi_k(y)}{\|\varphi_k 1_{Q^{(1)}}\|_2^2} k_{Q^{(1)}}(x^{|x|-k}, y^{|y|-k}) \\ &\sim \begin{cases} \frac{1}{\mu(Q)}, & \text{if } x \wedge y \neq x_Q^{(1)}, \\ -\frac{1}{q(x_Q^{(1)})\mu(Q)}, & \text{otherwise.} \end{cases} \end{aligned}$$

□

We next prove an estimate for truncations of  $\mathcal{P}$ . Given a family  $\mathcal{F}$  of disjoint sets in  $\mathcal{D}$  contained in a sector  $Q_0$ , we denote the kernel of the truncation  $\mathcal{P}_{\mathcal{F}}^{Q_0}$  by  $\mathcal{K}_{\mathcal{F}}^{Q_0}$ , which clearly can be written as

$$\mathcal{K}_{\mathcal{F}}^{Q_0} = \sum_{\substack{Q \in \mathcal{D}, Q \subseteq Q_0 \\ Q \not\subseteq R \in \mathcal{F}}} \mathcal{K}_Q. \quad (1.4)$$

In the particular case when  $Q_0 = \mathcal{X}$ , we slightly abuse notation and exclude  $\mathcal{K}_{\mathcal{X}}$  from the sum —and will deal with it separately below—.

**Proposition 1.9.** Let  $Q \in \mathcal{D}_k$  be such that  $Q \subseteq Q_0$  and  $Q \not\subseteq R$  for any  $R \in \mathcal{F}$ . For every  $z \in Q^{(\ell)} \setminus Q^{(\ell-1)} \subseteq Q_0$ , we have

$$\sup_{x, y \in Q} |\mathcal{K}_{\mathcal{F}}^{Q_0}(z, x) - \mathcal{K}_{\mathcal{F}}^{Q_0}(z, y)| \lesssim \begin{cases} \frac{1}{\mu(Q^{(\ell)})} \nu_{k-\ell}^k, & \text{if } z = x_Q^{(\ell)}, \\ \frac{1}{\mu(Q^{(\ell-1)})} \nu_{k-\ell}^k, & \text{if } z \neq x_Q^{(\ell)}. \end{cases}$$

*Proof.* We have

$$|\mathcal{K}_{\mathcal{F}}^{Q_0}(z, x) - \mathcal{K}_{\mathcal{F}}^{Q_0}(z, y)| \leq \sum_{j=0}^{k-\ell} |\mathcal{K}_{Q^{(\ell+j)}}(z, x) - \mathcal{K}_{Q^{(\ell+j)}}(z, y)|,$$

where the first term of the sum vanishes whenever  $z = x_Q^{(\ell)}$ . By part (1) of Lemma 1.8, one has

$$|\mathcal{K}_{\mathcal{F}}^{Q_0}(z, x) - \mathcal{K}_{\mathcal{F}}^{Q_0}(z, y)| \lesssim \begin{cases} \sum_{j=1}^{k-\ell} \frac{1}{\mu(Q^{(\ell+j-1)})} \nu_{k-\ell-j+1}^k, & \text{if } z = x_Q^{(\ell)} \\ \frac{1}{\mu(Q^{(\ell-1)})} \nu_{k-\ell}^k + \sum_{j=1}^{k-\ell} \frac{1}{\mu(Q^{(\ell+j-1)})} \nu_{k-\ell-j+1}^k, & \text{if } z \neq x_Q^{(\ell)}. \end{cases}$$

The statement follows from the fact that both sums are geometric and hence controlled by their first term, and the one appearing in the term below is in turn controlled by the one on its left.  $\square$

The following notation will sometimes be used in the sequel: if  $Q \in \mathcal{D}$ ,  $\overline{Q}$  is the smallest sector that contains it—it is only different from  $Q$  if the latter is a singleton—. Our last estimate is particularly useful when controlling local terms.

**Proposition 1.10.** Let  $b$  be a simple algebraic atom associated to  $Q \in \mathcal{D}_k$ . For every  $1 \leq \ell \leq |x_Q|$ ,

$$\mathcal{D}_{\overline{Q}^{(\ell)}}(b)(x) \lesssim \begin{cases} \frac{\nu_{k-\ell+1}^k}{\mu(Q^{(\ell-1)})} \|b\|_1, & \text{if } x \in \overline{Q}^{(\ell-1)}, \\ \frac{\nu_{k-\ell}^k}{\mu(Q^{(\ell-1)})} \|b\|_1, & \text{if } x \in \overline{Q}^{(\ell)} \setminus \overline{Q}^{(\ell-1)}. \end{cases} \quad (1.5a)$$

$$\mathcal{D}_{\overline{Q}^{(\ell)}}(b)(x) \lesssim \begin{cases} \frac{\nu_{k-\ell+1}^k}{\mu(Q^{(\ell-1)})} \|b\|_1, & \text{if } x \in \overline{Q}^{(\ell-1)}, \\ \frac{\nu_{k-\ell}^k}{\mu(Q^{(\ell-1)})} \|b\|_1, & \text{if } x \in \overline{Q}^{(\ell)} \setminus \overline{Q}^{(\ell-1)}. \end{cases} \quad (1.5b)$$

Furthermore, if  $Q$  is a single point, then  $\mathcal{D}_{\overline{Q}}b \equiv 0$ .

*Proof.* The case  $\ell > 1$  follows from the fact that  $b$  has null mean and Lemma 1.8 applied to  $\overline{Q}^{(1)}$ . Indeed,

$$|\mathcal{D}_{\overline{Q}^{(\ell)}}b(x)| = \left| \sum_{y \in Q^{(1)}} \left( \mathcal{K}_{\overline{Q}^{(\ell)}}(x, y) - \mathcal{K}_{\overline{Q}^{(\ell)}}(x, x_{\overline{Q}^{(1)}}) \right) b(y) \mu(y) \right|$$

$$\begin{aligned}
&\leq \sup_{y \in Q^{(1)}} \left| \mathcal{K}_{\overline{Q}^{(\ell)}}(x, y) - \mathcal{K}_{\overline{Q}^{(\ell)}}(x, x_{\overline{Q}^{(1)}}) \right| \|b\|_1 \\
&\leq \begin{cases} \frac{\nu_{k-\ell+1}^k}{\mu(Q^{(\ell-1)})} \|b\|_1, & \text{if } x \in \overline{Q}^{(\ell-1)}, \\ \frac{\nu_{k-\ell}^k}{\mu(Q^{(\ell-1)})} \|b\|_1, & \text{if } x \in \overline{Q}^{(\ell)} \setminus \overline{Q}^{(\ell-1)}, \end{cases}
\end{aligned}$$

regardless of whether  $Q = \overline{Q}$  or not. For  $\ell = 1$ , if  $Q$  is a single point one can argue similarly as above. Otherwise, denote by  $R$  the element so that  $x \in R \in \mathcal{D}_1(Q^{(1)})$ . By (2) of Lemma 1.8, one has

$$\begin{aligned}
|\mathcal{D}_{Q^{(1)}}b(x)| &\leq \sum_{y \in Q^{(1)} \setminus \{x_{Q^{(1)}}\}} |\mathcal{K}_{Q^{(1)}}(x, y)| |b(y)| \mu(y) \\
&\lesssim \frac{1}{\mu(Q)} \left[ \|b1_R\|_1 + \frac{1}{q(x_{Q^{(1)}})} \|b1_{Q^{(1)} \setminus (R \cup \{x_{Q^{(1)}}\})}\|_1 \right].
\end{aligned}$$

If  $R = Q$ , then clearly  $|\mathcal{D}_{Q^{(1)}}b(x)| \lesssim \|b\|_1 \mu(Q)^{-1}$ . Otherwise, since  $b = a - \langle a \rangle_{Q^{(1)}} 1_{Q^{(1)}}$  we have that

$$|\mathcal{D}_{Q^{(1)}}b(x)| \lesssim \frac{1}{\mu(Q)} \left[ \mu(R) |\langle a \rangle_{Q^{(1)}}| + \frac{1}{q(x_{Q^{(1)}})} \|b\|_1 \right] \lesssim \frac{\|b\|_1}{q(x_{Q^{(1)}}) \mu(Q)}.$$

Finally, the fact that  $\mathcal{D}_{\overline{Q}}b \equiv 0$  when  $Q = \{v\}$  follows from

$$\begin{aligned}
\mathcal{D}_{\overline{Q}}b(x) &= -\frac{a(v)}{\mu(\overline{Q})} \sum_{y \in \overline{Q} \setminus \{v\}} \mathcal{K}_{\overline{Q}}(x, y) \mu(y) \\
&= -\frac{a(v)}{\mu(\overline{Q})} \sum_{z \in \mathcal{D}_1(v)} k_{\overline{Q}}(x^{|x|-|v|-1}, z) \sum_{y \in S_z} \frac{\varphi_{|v|+1}(x) \varphi_{|v|+1}(y)}{\|\varphi_{|v|+1} 1_{\overline{Q}}\|_2^2} \mu(y) = 0, \quad (1.6)
\end{aligned}$$

by the definition of  $k_{\overline{Q}}$  and the radially of  $\varphi_{|v|+1}$ .  $\square$

*Remark 1.11.* The key to obtaining the estimates above is to exploit the joint cancellation enjoyed by the family  $\{h_Q^\ell\}_\ell$  for each fixed  $Q$ . We do that by studying the operators  $\mathcal{D}_Q$  as a whole instead of considering the individual products  $\langle f, h_Q^\ell \rangle h_Q^\ell$ , of which there can be arbitrarily many within each cube  $Q$ , and for which size estimates are not precise enough. This is a key difference with respect to the usual point of view that one adopts in dyadic harmonic analysis, in which the action of each function in the Haar basis is dealt with separately. In a sense, one can interpret our point of view as a way to dealing with the fact that  $\mathcal{X}$  may be infinite-dimensional.

2. SPARSE FORMS ASSOCIATED WITH  $\mathcal{P}$ 

This section is devoted to the proof of Theorems A and B. We start defining the form  $\mathcal{E}_{\mathcal{S}}$  that appears in the statement of Theorem A. If  $\mathcal{S} \subseteq \mathcal{D}$  is a sparse family we put

$$\mathcal{E}_{\mathcal{S}}(f_1, f_2) := \sum_{Q=\bar{Q} \in \mathcal{S}} \langle f_1 \rangle_Q \left( \frac{1}{q(x_Q^{(1)})} \sum_{\substack{R=\bar{R} \in \mathcal{S} \\ R^{(1)}=Q^{(1)}}} \langle f_2 \rangle_R \right) \mu(Q). \quad (2.1)$$

We need to use the following variant of the Calderón-Zygmund decomposition for general measures, both in the next Subsection and in Section 3.

**Lemma 2.1.** *Let  $R \in \mathcal{D}$ ,  $0 \leq f \in L^1(R)$  and  $\lambda > \mu(R)^{-1} \|f\|_{L^1(R)}$ . Let  $\{Q_j\}_j \subseteq \mathcal{D}$  be a disjoint family such that*

$$\{\mathcal{M}_{\mathcal{D}} f > \lambda\} \subseteq \bigcup_j Q_j,$$

and  $Q_j^{(1)} \not\subseteq \bigcup_{\ell} Q_{\ell}$  for any  $j$ . Then, if we define  $b = \sum_j b_{Q_j}$ , where

$$b_{Q_j} := f 1_{Q_j} - \langle f 1_{Q_j} \rangle_{Q_j^{(1)}} 1_{Q_j^{(1)}},$$

and  $f = g + b$ , the following properties are satisfied:

(1) *Localization:*

$$\text{supp}(b_{Q_j}) \subseteq Q_j^{(1)}, \quad \sum_{x \in Q_j^{(1)}} b_{Q_j}(x) \mu(x) = 0.$$

(2)  *$L^1$ -bounds:*

$$\|g\|_1 \lesssim \|f\|_1, \quad \sum_j \|b_{Q_j}\|_1 \lesssim \|f\|_1.$$

(3) *Higher integrability for  $g$ :*

$$\|g\|_2^2 \lesssim \lambda \|f\|_1, \quad \|g\|_{\text{BMO}} \lesssim \lambda.$$

The statement above is different from the usual one in that the Calderón-Zygmund cubes can be taken to be as large as desired, so long as they are still disjoint (and so smallness of the measure of the exceptional set is not claimed). We take advantage of this additional feature below in the proof of Theorem A, as is usually done in sparse domination proofs. The proof of Lemma 2.1 is a straightforward adaptation from [6, 2], and is hence omitted. The key BMO estimate is proven in [7].

**2.1. Sparse domination.** In the proof of Theorem A we will need to use the uniform  $L^2$ -boundedness of truncations of  $\mathcal{P}$ , that we record below:

$$\sup_{Q_0 \in \mathcal{D}, \mathcal{F} \subseteq \mathcal{D}} \left\| \mathcal{P}_{\mathcal{F}}^{Q_0} \right\|_{L^2(\mathcal{X}) \rightarrow L^2(\mathcal{X})} < \infty. \quad (2.2)$$

The proof of (2.2) follows immediately from the orthogonality and normalization of the operators  $\mathcal{D}_Q$  and we omit it.

*Proof of Theorem A.* Since we will have  $\mathcal{X} \in \mathcal{S}$  for all pairs  $(f_1, f_2)$ , it is enough to prove sparse domination for the operator  $\mathcal{P} - \langle f \rangle_{\mathcal{X}} 1_{\mathcal{X}}$ . Let  $f_1, f_2 \in L^1(\mathcal{X})$  be nonnegative, and set  $Q_0 = \mathcal{X}$ . We have

$$\langle \mathcal{P} f_1, f_2 \rangle = \langle f_1 \rangle_{Q_0} \|f_2\|_1 + \left\langle \sum_{Q \in \mathcal{D}} \mathcal{D}_Q f_1, f_2 \right\rangle = \langle f_1 \rangle_{Q_0} \langle f_2 \rangle_{Q_0} \mu(Q_0) + \langle \mathcal{P}^{Q_0} f_1, f_2 \rangle.$$

Since  $Q_0 \in \mathcal{S}$ , the first term in the right hand side above is acceptable, and we focus on estimating the second for the remainder of the proof. We denote

$$\Omega_i^0 = \{ \mathcal{M}_{\mathcal{D}} f_i > 4 \| \mathcal{M}_{\mathcal{D}} \|_{L^1(\mathcal{X}) \rightarrow L^{1,\infty}(\mathcal{X})} \langle f_i \rangle_{Q_0} \}, \quad i \in \{1, 2\},$$

and  $\Omega^0 = \Omega_1^0 \cup \Omega_2^0$ . We cover  $\Omega^0$  with a disjoint family  $\mathcal{F}^1 \subseteq \mathcal{D}$ , maximal with respect to inclusion. For  $i = 1, 2$ , we apply the Calderón-Zygmund decomposition from Lemma 2.1 to  $f_i$  at height  $\lambda_i = 4 \| \mathcal{M}_{\mathcal{D}} \|_{L^1(\mathcal{X}) \rightarrow L^{1,\infty}(\mathcal{X})} \langle f_i \rangle_{Q_0}$ . We next estimate the non localizing part of  $\mathcal{P}$ , namely the truncation

$$\mathcal{P}_{\mathcal{F}^1}^{Q_0} f = \mathcal{P}^{Q_0} f - \sum_{Q \in \mathcal{F}^1} \mathcal{P}^Q f.$$

We claim

$$\begin{aligned} |\langle \mathcal{P}_{\mathcal{F}^1}^{Q_0} f_1, f_2 \rangle| &\lesssim \langle f_1 \rangle_{Q_0} \langle f_2 \rangle_{Q_0} \mu(Q_0) + \sum_{Q \in \mathcal{F}^1} \langle f_1 \rangle_Q \langle f_2 \rangle_Q \mu(Q) \\ &\quad + \sum_{Q = \bar{Q} \in \mathcal{F}^1} \langle f_1 \rangle_Q \left( \frac{1}{q(x_Q^{(1)})} \sum_{\substack{R = \bar{R} \in \mathcal{F}^1 \\ R^{(1)} = Q^{(1)}}} \langle f_2 \rangle_R \right) \mu(Q). \end{aligned} \quad (2.3)$$

We split according to the Calderón-Zygmund pieces:

$$\begin{aligned} |\langle \mathcal{P}_{\mathcal{F}^1}^{Q_0} f_1, f_2 \rangle| &\leq |\langle \mathcal{P}_{\mathcal{F}^1}^{Q_0} g_1, g_2 \rangle| + |\langle \mathcal{P}_{\mathcal{F}^1}^{Q_0} b_1, g_2 \rangle| + |\langle \mathcal{P}_{\mathcal{F}^1}^{Q_0} g_1, b_2 \rangle| + |\langle \mathcal{P}_{\mathcal{F}^1}^{Q_0} b_1, b_2 \rangle| \\ &=: \text{I} + \text{II} + \text{III} + \text{IV}, \end{aligned}$$

and we analyze separately the four cases. For term I, Cauchy-Schwarz's inequality and (2.2) yield

$$\text{I} \leq \| \mathcal{P}_{\mathcal{F}^1}^{Q_0} g_1 \|_2 \| g_2 \|_2 \lesssim \| g_1 \|_2 \| g_2 \|_2 \lesssim (\langle f_1 \rangle_{Q_0} \| f_1 \|_1 \langle f_2 \rangle_{Q_0} \| f_2 \|_1)^{1/2} = \langle f_1 \rangle_{Q_0} \langle f_2 \rangle_{Q_0} \mu(Q_0).$$

To analyze the terms involving the bad parts in the Calderón-Zygmund decomposition, we write

$$\begin{aligned} \mathcal{P}_{\mathcal{F}^1}^{Q_0}(b_{1,Q}) &= \sum_{\substack{D=\bar{D}\in\mathcal{D} \\ D\cap Q^{(1)}\neq\emptyset}} \mathcal{D}_D b_{1,Q} - \sum_{\substack{R=\bar{R}\in\mathcal{F}^1 \\ R\cap Q^{(1)}\neq\emptyset}} \sum_{\substack{T=\bar{T}\in\mathcal{D} \\ T\subseteq R \\ T\cap Q^{(1)}\neq\emptyset}} \mathcal{D}_T b_{1,Q} \\ &= \sum_{k=1}^{|x_Q|} \mathcal{D}_{Q^{(k)}} b_{1,Q} + \sum_{\substack{T=\bar{T}\in\mathcal{D} \\ T\subseteq Q^{(1)}\setminus\Omega^0}} \mathcal{D}_T b_{1,Q}. \end{aligned} \quad (2.4)$$

For every  $T \subseteq Q^{(1)} \setminus \Omega^0$ , we have  $b_{1,Q}|_T \equiv \langle f1_Q \rangle_{Q^{(1)}}$ , and so the last sum in (2.4) vanishes. Using that  $\mathcal{D}_{\bar{Q}} b_{1,Q} \equiv 0$  when  $Q$  is a singleton, the above formula reduces to

$$\mathcal{P}_{\mathcal{F}^1}^{Q_0}(b_{1,Q}) = \sum_{k=1}^{|x_Q|} \mathcal{D}_{\bar{Q}^{(k)}} b_{1,Q}. \quad (2.5)$$

Since  $\mathcal{P}_{\mathcal{F}^1}^{Q_0}$  is self-adjoint, it is enough to study II and the estimate for III will follow. For every  $1 \leq k \leq |x_Q|$ , we put  $g_2^{(k)} = g_2 - \langle g_2 \rangle_{\bar{Q}^{(k)}}$ . By (1.1), we have  $\langle g_2^{(k)} \rangle_R \lesssim \|g_2\|_{\text{BMO}}$  for every  $R \in \mathcal{D}^1(\bar{Q}^{(k)})$ . Then, by the vanishing mean of  $\mathcal{D}_{\bar{Q}^{(k)}}(b_{1,Q})$  and Proposition 1.10 we have

$$\begin{aligned} \langle \mathcal{D}_{\bar{Q}^{(k)}}(b_{1,Q}), g_2 \rangle &= \langle \mathcal{D}_{\bar{Q}^{(k)}}(b_{1,Q}), g_2 - \langle g_2 \rangle_{\bar{Q}^{(k)}} \rangle \\ &= \|\mathcal{D}_{\bar{Q}^{(k)}}(b_{1,Q})1_{\bar{Q}^{(k-1)}}\|_\infty \|g_2^{(k)}1_{\bar{Q}^{(k-1)}}\|_1 + \sum_{\substack{z \in \mathcal{D}^1(x_Q^{(k)}) \\ z \neq x_Q^{(k-1)}}} \|\mathcal{D}_{\bar{Q}^{(k)}}(b_{1,Q})1_{S_z}\|_\infty \|g_2^{(k)}1_{S_z}\|_1 \\ &\lesssim \nu_{|x_Q|-k+1} \|b_{1,Q}\|_1 \|g_2\|_{\text{BMO}} \left( 1 + \frac{1}{q(x_Q^{(k)})} \sum_{\substack{z \in \mathcal{D}^1(x_Q^{(k)}) \\ z \neq x_Q^{(k-1)}}} 1 \right) \\ &\lesssim \nu_{|x_Q|-k+1} \|b_{1,Q}\|_1 \|g_2\|_{\text{BMO}}. \end{aligned}$$

Summing over  $k$  yields

$$\sum_{Q \in \mathcal{F}^1} \langle \mathcal{P}_{\mathcal{F}^1}^{Q_0}(b_{1,Q}), g_2 \rangle \lesssim \sum_{Q \in \mathcal{F}^1} \|b_{1,Q}\|_1 \|g_2\|_{\text{BMO}} \sum_{k=1}^{|x_Q|} \nu_{|x_Q|-k+1} \lesssim \|f_1\|_{1,Q_0} \|f_2\|_{1,Q_0} \mu(Q_0).$$

We are left with term IV, that we need to further split as follows:

$$\langle \mathcal{P}_{\mathcal{F}^1}^{Q_0} b_1, b_2 \rangle = \sum_{Q,R \in \mathcal{F}^1} \langle \mathcal{P}_{\mathcal{F}^1}^{Q_0} b_{1,Q}, b_{2,R} \rangle \leq 2 \sum_{Q \in \mathcal{F}^1} \sum_{\substack{R \in \mathcal{F}^1 \\ \mu(R) \leq \mu(Q)}} \langle \mathcal{P}_{\mathcal{F}^1}^{Q_0} b_{1,Q}, b_{2,R} \rangle = 2 \sum_{Q \in \mathcal{F}^1} \text{IV}_Q,$$

where

$$\begin{aligned}
IV_Q &= \sum_{x \in \mathcal{X}} \mathcal{P}_{\mathcal{F}^1}^{Q_0}(b_{1,Q})(x) \sum_{\substack{R \in \mathcal{F}^1 \\ \mu(R) \leq \mu(Q)}} b_{2,R}(x) \mu(x) \\
&= \sum_{x \in Q^{(1)}} \mathcal{P}_{\mathcal{F}^1}^{Q_0}(b_{1,Q})(x) \sum_{\substack{R \in \mathcal{F}^1 \\ \mu(R) \leq \mu(Q)}} b_{2,R}(x) \mu(x) \\
&\quad + \sum_{k=2}^{|x_Q|} \sum_{x \in Q^{(k)} \setminus Q^{(k-1)}} \mathcal{P}_{\mathcal{F}^1}^{Q_0}(b_{1,Q})(x) \sum_{\substack{R \in \mathcal{F}^1 \\ \mu(R) \leq \mu(Q)}} b_{2,R}(x) \mu(x) \\
&=: IV_{Q,1} + \sum_{k=2}^{|x_Q|} IV_{Q,k}.
\end{aligned}$$

For  $k \geq 2$  we have  $b_{2,R}(x_Q^{(k)}) = 0$  because  $\mu(R) \leq \mu(Q)$ . Therefore,

$$\begin{aligned}
|IV_{Q,k}| &= \left| \sum_{x \in Q^{(k)} \setminus Q^{(k-1)}} \mathcal{P}_{\mathcal{F}^1}^{Q_0}(b_{1,Q})(x) \sum_{\substack{R \in \mathcal{F}^1 \\ \mu(R) \leq \mu(Q)}} b_{2,R}(x) \mu(x) \right| \\
&= \left| \sum_{x \in Q^{(k)} \setminus (Q^{(k-1)} \cup \{x_Q^{(k)}\})} \sum_{y \in Q^{(1)}} (\mathcal{K}_{\mathcal{F}^1}^{Q_0}(x, y) - \mathcal{K}_{\mathcal{F}^1}^{Q_0}(x, x_Q)) b_{1,Q}(y) \mu(y) \sum_{\substack{R \in \mathcal{F}^1 \\ \mu(R) \leq \mu(Q)}} b_{2,R}(x) \mu(x) \right| \\
&\leq \|b_{1,Q}\|_1 \sup_{y \in Q^{(1)}} \sum_{x \in Q^{(k)} \setminus (Q^{(k-1)} \cup \{x_Q^{(k)}\})} |\mathcal{K}_{\mathcal{F}^1}^{Q_0}(x, y) - \mathcal{K}_{\mathcal{F}^1}^{Q_0}(x, x_Q)| \sum_{\substack{R \in \mathcal{F}^1 \\ \mu(R) \leq \mu(Q)}} |b_{2,R}(x)| \mu(x) \\
&\leq \|b_{1,Q}\|_1 \left( \sum_{\substack{Q^{(k-1)} \neq S = \bar{S} \in \mathcal{D} \\ S^{(1)} = Q^{(k)}}} \sum_{x \in S} \sup_{y \in Q^{(1)}} |\mathcal{K}_{\mathcal{F}^1}^{Q_0}(x, y) - \mathcal{K}_{\mathcal{F}^1}^{Q_0}(x, x_Q)| \sum_{\substack{R \in \mathcal{F}^1, R \subseteq S \\ \mu(R) \leq \mu(Q)}} |b_{2,R}(x)| \mu(x) \right) \\
&\lesssim \|b_{1,Q}\|_1 \left( \sum_{\substack{Q^{(k-1)} \neq S = \bar{S} \in \mathcal{D} \\ S^{(1)} = Q^{(k)}}} \frac{1}{\mu(Q^{(k-1)})} \nu_{|x_Q|}^{|x_Q|} \sum_{x \in S} \sum_{\substack{R \in \mathcal{F}^1, R \subseteq S \\ \mu(R) \leq \mu(Q)}} |b_{2,R}(x)| \mu(x) \right) \\
&\leq \|b_{1,Q}\|_1 \nu_{|x_Q|}^{|x_Q|} \left( \sum_{\substack{Q^{(k-1)} \neq S = \bar{S} \in \mathcal{D} \\ S^{(1)} = Q^{(k)}}} \langle f_2 \rangle_S \right) \lesssim \|b_{1,Q}\|_1 \nu_{|x_Q|}^{|x_Q|} \langle f_2 \rangle_{Q_0},
\end{aligned}$$

by Proposition 1.9 and the maximality of the Calderón-Zygmund cubes  $Q$ . For  $IV_{Q,1}$  we use (2.5) to get

$$IV_{Q,1} = \sum_{x \in Q^{(1)}} \mathcal{P}_{\mathcal{F}^1}^{Q_0}(b_{1,Q})(x) \sum_{\substack{R \in \mathcal{F}^1 \\ \mu(R) \leq \mu(Q)}} b_{2,R}(x) \mu(x)$$

$$\begin{aligned}
&= \sum_{x \in Q^{(1)}} \sum_{k=1}^{|x_Q|} \mathcal{D}_{\overline{Q}^{(k)}} b_{1,Q}(x) \sum_{\substack{R \in \mathcal{F}^1 \\ \mu(R) \leq \mu(Q)}} b_{2,R}(x) \mu(x) \\
&= \sum_{x \in Q^{(1)}} \mathcal{D}_{\overline{Q}^{(1)}} b_{1,Q}(x) \sum_{\substack{R \in \mathcal{F}^1 \\ \mu(R) \leq \mu(Q)}} b_{2,R}(x) \mu(x) \\
&\quad + \sum_{x \in Q^{(1)}} \sum_{k=2}^{|x_Q|} \mathcal{D}_{\overline{Q}^{(k)}} b_{1,Q}(x) \sum_{\substack{R \in \mathcal{F}^1 \\ \mu(R) \leq \mu(Q)}} b_{2,R}(x) \mu(x) \\
&=: \text{IV}_{Q,1}^a + \text{IV}_{Q,1}^b.
\end{aligned}$$

We next use Proposition 1.10 to get

$$\begin{aligned}
|\text{IV}_{Q,1}^b| &\leq \|b_{1,Q}\|_1 \sum_{k=2}^{|x_Q|} \frac{\nu_{|x_Q|^{-k+1}}}{\mu(Q^{(k-1)})} \sum_{x \in Q^{(1)}} \sum_{\substack{R \in \mathcal{F}^1 \\ \mu(R) \leq \mu(Q)}} b_{2,R}(x) \mu(x) \\
&\leq \|b_{1,Q}\|_1 \sum_{k=2}^{|x_Q|} \nu_{|x_Q|^{-k+1}} \langle f_2 \rangle_{Q^{(1)}} \\
&\lesssim \|b_{1,Q}\|_1 \langle f_2 \rangle_{Q_0} \sum_{k=2}^{|x_Q|} \nu_{|x_Q|^{-k+1}} \\
&\lesssim \|b_{1,Q}\|_1 \langle f_2 \rangle_{Q_0}.
\end{aligned}$$

To estimate  $\text{IV}_{Q,1}^a$  we proceed differently depending on whether  $Q = \overline{Q}$  or not. In the latter case, we use Proposition 1.10 and the fact that  $\mu(Q) \sim \mu(Q^{(1)})$  to get

$$\begin{aligned}
|\text{IV}_{Q,1}^a| &\leq \frac{\nu_{|x_Q|}}{\mu(Q)} \|b_{1,Q}\|_1 \sum_{x \in Q^{(1)}} \sum_{\substack{R \in \mathcal{F}^1 \\ \mu(R) \leq \mu(Q)}} |b_{2,R}(x)| \mu(x) \\
&\lesssim \|b_{1,Q}\|_1 \frac{1}{\mu(Q^{(1)})} \sum_{\substack{R \in \mathcal{F}^1 \\ R \subsetneq Q^{(1)}}} \|b_{2,R}\|_1 \\
&\lesssim \|b_{1,Q}\|_1 \langle f_2 \rangle_{Q^{(1)}} \lesssim \|b_{1,Q}\|_1 \langle f_2 \rangle_{Q_0}.
\end{aligned}$$

If  $Q = \overline{Q}$ , then we split again as

$$\begin{aligned}
\text{IV}_{Q,1}^a &= \sum_{x \in Q} \mathcal{D}_{\overline{Q}^{(1)}} b_{1,Q}(x) \sum_{\substack{R \in \mathcal{F}^1 \\ \mu(R) \leq \mu(Q)}} b_{2,R}(x) \mu(x) + \sum_{\substack{x \in Q^{(1)} \setminus Q \\ x \neq x_Q^{(1)}}} \mathcal{D}_{\overline{Q}^{(1)}} b_{1,Q}(x) \sum_{\substack{R \in \mathcal{F}^1 \\ \mu(R) \leq \mu(Q)}} b_{2,R}(x) \mu(x) \\
&=: \text{IV}_{Q,1}^{a,1} + \text{IV}_{Q,1}^{a,2}.
\end{aligned}$$

Hence Proposition 1.10 yields

$$\begin{aligned}
|IV_{Q,1}^{a,1}| &\leq \sum_{x \in Q} |\mathcal{D}_{\overline{Q}^{(1)}} b_{1,Q}(x)| \sum_{\substack{R \in \mathcal{F}^1 \\ R^{(1)} = Q^{(1)}}} |b_{2,R}(x)| \mu(x) \\
&\lesssim \|b_{1,Q}\|_1 \frac{1}{\mu(Q)} \sum_{x \in Q} \sum_{\substack{R \in \mathcal{F}^1 \\ R^{(1)} = Q^{(1)}}} |b_{2,R}(x)| \mu(x) \\
&\leq \|b_{1,Q}\|_1 \left( \sum_{\substack{S \in \mathcal{D}_1(Q^{(1)}) \\ S \neq Q}} \langle f_2 1_S \rangle_{Q^{(1)}} + \langle f_2 \rangle_Q \right) \\
&\leq \|b_{1,Q}\|_1 (\langle f_2 \rangle_{Q^{(1)}} + \langle f_2 \rangle_Q) \lesssim \|b_{1,Q}\|_1 (\langle f_2 \rangle_{Q_0} + \langle f_2 \rangle_Q).
\end{aligned}$$

Finally, for the last term, again by applying Proposition 1.10 and using the fact that if  $\mu(R) \leq \mu(Q)$  and  $Q^{(1)} = R^{(1)}$  then  $R = \overline{R}$ ,

$$\begin{aligned}
|IV_{Q,1}^{a,2}| &\leq \sum_{\substack{x \in Q^{(1)} \setminus Q \\ x \neq x_Q^{(1)}}} |\mathcal{D}_{\overline{Q}^{(1)}} b_{1,Q}(x)| \sum_{\substack{R \in \mathcal{F}^1 \\ \mu(R) \leq \mu(Q)}} |b_{2,R}(x)| \mu(x) \\
&\lesssim \|b_{1,Q}\|_1 \frac{1}{\mu(Q)} \frac{1}{q(x_Q^{(1)})} \sum_{\substack{x \in Q^{(1)} \setminus Q \\ x \neq x_Q^{(1)}}} \sum_{\substack{R \in \mathcal{F}^1 \\ \mu(R) \leq \mu(Q)}} |b_{2,R}(x)| \mu(x) \\
&\leq \|b_{1,Q}\|_1 \frac{1}{\mu(Q)} \frac{1}{q(x_Q^{(1)})} \left( \sum_{\substack{S \in \mathcal{D}_1(Q^{(1)}) \\ S \notin \mathcal{F}^1}} \sum_{\substack{R \in \mathcal{F}^1 \\ R^{(1)} \subseteq S}} \sum_{x \in S} |b_{2,R}(x)| \mu(x) + \sum_{\substack{R = \overline{R} \in \mathcal{F}^1 \\ R^{(1)} = Q^{(1)}}} \sum_{x \in Q^{(1)} \setminus Q} |b_{2,R}(x)| \mu(x) \right) \\
&\leq \|b_{1,Q}\|_1 \left( \frac{1}{q(x_Q^{(1)})} \sum_{\substack{S \in \mathcal{D}_1(Q^{(1)}) \\ S \notin \mathcal{F}^1}} \langle \sum_{\substack{R \in \mathcal{F}^1 \\ R^{(1)} \subseteq S}} |b_{2,R}| \rangle_S + \sum_{\substack{R = \overline{R} \in \mathcal{F}^1 \\ R^{(1)} = Q^{(1)}}} \frac{1}{q(x_Q^{(1)})} \frac{\|b_{2,R}\|_1}{\mu(Q)} \right) \\
&\lesssim \|b_{1,Q}\|_1 \langle f_2 \rangle_{Q_0} \left( \frac{1}{q(x_Q^{(1)})} \sum_{\substack{S \in \mathcal{D}_1(Q^{(1)}) \\ S \notin \mathcal{F}^1}} 1 \right) + \|b_{1,Q}\|_1 \frac{1}{q(x_Q^{(1)})} \sum_{\substack{R = \overline{R} \in \mathcal{F}^1 \\ R^{(1)} = Q^{(1)}}} \langle f_2 \rangle_R \\
&\leq \|b_{1,Q}\|_1 \left( \langle f_2 \rangle_{Q_0} + \frac{1}{q(x_Q^{(1)})} \sum_{\substack{R = \overline{R} \in \mathcal{F}^1 \\ R^{(1)} = Q^{(1)}}} \langle f_2 \rangle_R \right).
\end{aligned}$$

So, when  $Q$  is a sector, we control the term  $IV_{Q,1}$  by

$$|IV_{Q,1}| \lesssim \|b_{1,Q}\|_1 \left( \langle f_2 \rangle_{Q_0} + \langle f_2 \rangle_Q + \frac{1}{q(x_Q^{(1)})} \sum_{\substack{R = \overline{R} \in \mathcal{F}^1 \\ R^{(1)} = Q^{(1)}}} \langle f_2 \rangle_R \right).$$

Collecting all the estimates together yields

$$\begin{aligned}
|IV_Q| &\leq |IV_{Q,1}| + \sum_{k=2}^{|x_Q|} |IV_{Q,k}| \\
&\lesssim \|b_{1,Q}\|_1 \left( \langle f_2 \rangle_{Q_0} \left( 1 + \sum_{k=2}^{|x_Q|} \nu_{|x_Q|-k+1}^{|x_Q|} \right) + \langle f_2 \rangle_Q + \frac{1}{q(x_Q^{(1)})} \sum_{\substack{R=\bar{R} \in \mathcal{F}^1 \\ R^{(1)}=Q^{(1)}}} \langle f_2 \rangle_R \right) \\
&\lesssim \|b_{1,Q}\|_1 \langle f_2 \rangle_{Q_0} + \|b_{1,Q}\|_1 \langle f_2 \rangle_Q + \|b_{1,Q}\|_1 \frac{1}{q(x_Q^{(1)})} \sum_{\substack{R=\bar{R} \in \mathcal{F}^1 \\ R^{(1)}=Q^{(1)}}} \langle f_2 \rangle_R,
\end{aligned}$$

whenever  $Q = \bar{Q}$ , and

$$|IV_Q| \lesssim \|b_{1,Q}\|_1 \langle f_2 \rangle_{Q_0}$$

in the singleton case. Finally, we obtain

$$\begin{aligned}
|IV| &\lesssim \sum_{Q \in \mathcal{F}^1} |IV_Q| \\
&\lesssim \sum_{Q \in \mathcal{F}^1} \|b_{1,Q}\|_1 (\langle f_2 \rangle_{Q_0} + \langle f_2 \rangle_Q) + \sum_{Q=\bar{Q} \in \mathcal{F}^1} \|b_{1,Q}\|_1 \frac{1}{q(x_Q^{(1)})} \sum_{\substack{R=\bar{R} \in \mathcal{F}^1 \\ R^{(1)}=Q^{(1)}}} \langle f_2 \rangle_R \\
&\leq \langle f_1 \rangle_{Q_0} \langle f_2 \rangle_{Q_0} \mu(Q_0) + \sum_{Q \in \mathcal{F}^1} \langle f_1 \rangle_Q \langle f_2 \rangle_Q \mu(Q) \\
&\quad + \sum_{Q=\bar{Q} \in \mathcal{F}^1} \langle f_1 \rangle_Q \left( \frac{1}{q(x_Q^{(1)})} \sum_{\substack{R=\bar{R} \in \mathcal{F}^1 \\ R^{(1)}=Q^{(1)}}} \langle f_2 \rangle_R \right) \mu(Q).
\end{aligned}$$

This concludes the estimate for IV and yields (2.3), which in turn implies

$$\begin{aligned}
|\langle \mathcal{P}^{Q_0} f_1, f_2 \rangle| &\leq C \left( \langle f_1 \rangle_{Q_0} \langle f_2 \rangle_{Q_0} \mu(Q_0) + \sum_{Q \in \mathcal{F}^1} \langle f_1 \rangle_Q \langle f_2 \rangle_Q \mu(Q) \right. \\
&\quad \left. + \sum_{Q=\bar{Q} \in \mathcal{F}^1} \langle f_1 \rangle_Q \left( \frac{1}{q(x_Q^{(1)})} \sum_{\substack{R=\bar{R} \in \mathcal{F}^1 \\ R^{(1)}=Q^{(1)}}} \langle f_2 \rangle_R \right) \mu(Q) \right) + \sum_{R \in \mathcal{F}^1} \langle \mathcal{P}^R(f_1 1_R), f_2 1_R \rangle,
\end{aligned}$$

because  $\text{supp}(\mathcal{P}^R f) \subseteq R$  for each  $R \in \mathcal{D}$ . On the other hand,

$$\begin{aligned}
\mu(\Omega^0) &\leq \mu(\Omega_1^0) + \mu(\Omega_2^0) \\
&\leq \|\mathcal{M}_{\mathcal{D}}\|_{L^1(\mathcal{X}) \rightarrow L^{1,\infty}(\mathcal{X})} \left( \frac{\|f_1\|_1}{4\|\mathcal{M}_{\mathcal{D}}\|_{L^1(\mathcal{X}) \rightarrow L^{1,\infty}(\mathcal{X})} \langle f_1 \rangle_{Q_0}} + \frac{\|f_2\|_1}{4\|\mathcal{M}_{\mathcal{D}}\|_{L^1(\mathcal{X}) \rightarrow L^{1,\infty}(\mathcal{X})} \langle f_2 \rangle_{Q_0}} \right) \\
&= \frac{\mu(Q_0)}{2}.
\end{aligned}$$

We may then declare that  $Q_0 \in \mathcal{S}$  and choose  $E_{Q_0} = Q_0 \setminus \Omega^0$ , because we have just proved that  $\mu(E_{Q_0}) \geq \frac{1}{2}\mu(Q_0)$ . The proof then ends by iteration. Indeed, for every  $R \in \mathcal{F}^1$  we can now estimate

$$\begin{aligned} |\langle \mathcal{P}^R(f_1 1_R), f_2 1_R \rangle| &\leq C \left( \langle f_1 \rangle_R \langle f_2 \rangle_R \mu(R) + \sum_{T \in \mathcal{F}^1(R)} \langle f_1 \rangle_T \langle f_2 \rangle_T \mu(T) \right. \\ &\quad \left. + \sum_{T = \bar{T} \in \mathcal{F}^2(R)} \langle f_1 \rangle_T \left( \frac{1}{q(x_T^{(1)})} \sum_{\substack{S = \bar{S} \in \mathcal{F}^2(R) \\ S^{(1)} = T^{(1)}}} \langle f_2 \rangle_S \right) \mu(T) \right) + \sum_{T \in \mathcal{F}^2(R)} \langle \mathcal{P}^T(f_1 1_T), f_2 1_T \rangle, \end{aligned}$$

applying (2.3). We then proclaim that  $\mathcal{F}^1 \subseteq \mathcal{S}$  and construct the corresponding sets  $E_R$  for each  $R \in \mathcal{F}^1$  as before. The procedure can be iterated and yields the assertion after summing all the estimates for each cube.  $\square$

*Remark 2.2.* We have chosen to write down the proof of Theorem A specifically for  $\mathcal{P}$ , but we could have formulated our result in a much more general way. Clearly, the iterative scheme of proof and the construction of  $\mathcal{S}$  do not rely on the specific definition of  $\mathcal{P}$  and could have been carried over with any operator. Then, the core of the proof consists in estimating the action of the truncations  $\mathcal{P}_{\mathcal{F}^1}^{Q_0}$  over the parts of the Calderón-Zygmund decomposition of  $f_1$  and  $f_2$ . Those estimates in turn rely on uniform  $L^2$ -boundedness of truncations (good-good term), which can be formulated in a general way, and kernel estimates (good-bad and bad-bad terms). The kernel estimates and the form  $\mathcal{E}_{\mathcal{S}}$  are what is specific of  $\mathcal{P}$  and of the way in which we choose to split our operator into scales. It is absolutely conceivable that the same scheme of proof can be applied to the study of other operators whose structure is tied to that of a martingale filtration in any filtered space, finding wider applicability. In the nondoubling situation, however, one should expect that additional terms like  $\mathcal{E}_{\mathcal{S}}$  play a role in the study of most sparse domination results to compensate the lack of regularity of the underlying filtration.

*Remark 2.3.* Theorem A recovers the standard sparse domination result for  $\mathcal{P}$  when  $q : \mathcal{X} \rightarrow \mathbb{N}$  is bounded. Indeed, in that case  $\mu(Q^{(1)}) \leq C_{\text{doub}}\mu(Q)$  for all  $Q$  and so

$$\begin{aligned} \mathcal{E}_{\mathcal{S}}(f_1, f_2) &= \sum_{Q = \bar{Q} \in \mathcal{S}} \langle f_1 \rangle_Q \frac{1}{q(x_Q^{(1)})} \left( \sum_{\substack{R = \bar{R} \in \mathcal{S} \\ R^{(1)} = Q^{(1)}}} \langle f_2 \rangle_R \right) \mu(Q) \\ &\leq \sum_{Q = \bar{Q} \in \mathcal{S}} \langle f_1 \rangle_Q \frac{1}{q(x_Q^{(1)})} \langle f_2 \rangle_{Q^{(1)}} \mu(Q^{(1)}) \\ &\lesssim \sum_{Q = \bar{Q} \in \mathcal{S}} \langle f_1 \rangle_{Q^{(1)}} \langle f_2 \rangle_{Q^{(1)}} \mu(Q^{(1)}). \end{aligned}$$

But

$$\mu(E_Q) \geq \frac{1}{2}\mu(Q) \geq \frac{1}{2C_{\text{doub}}}\mu(Q^{(1)}),$$

which means that if  $\mathcal{S}$  is sparse, then the family  $\tilde{\mathcal{S}} = \{Q^{(1)} : Q \in \mathcal{S}\}$  is  $1/(2C_{\text{doub}})$ -sparse—we have called  $1/2$ -sparse families just sparse in the rest of the paper—and that the bound given by Theorem A becomes

$$\langle \mathcal{P}f_1, f_2 \rangle \lesssim \langle \mathcal{A}_S f_1, f_2 \rangle + \langle \mathcal{A}_{\tilde{\mathcal{S}}} f_1, f_2 \rangle.$$

This is a sparse estimate, because the union of two sparse families is sparse.

**2.2. Discrete Bekollé-Bonami weights.** We now study the weighted inequalities that follow from Theorem A. In the doubling setting, the natural class of weights is the (dyadic) Muckenhoupt one: for  $1 < p < \infty$ , set

$$\mathcal{B}_p(\mu) := \left\{ w : \mathcal{X} \rightarrow (0, \infty) : [w]_{\mathcal{B}_p(\mu)} := \sup_{Q \in \mathcal{D}} \langle w \rangle_Q \langle w^{-\frac{p'}{p}} \rangle_Q^{\frac{p}{p'}} < \infty \right\}.$$

$\mathcal{A}_S$  satisfies weighted inequalities for weights belonging to the Muckenhoupt class. Indeed, for  $w \in \mathcal{B}_p(\mu)$ ,  $\mathcal{A}_S$  is bounded on  $L^p(\mathcal{X}, w)$  and (see, for example, [10])

$$\|\mathcal{A}_S\|_{L^p(\mathcal{X}, w) \rightarrow L^p(\mathcal{X}, w)} \lesssim [w]_{\mathcal{B}_p(\mu)}^{\max\{1, \frac{p'}{p}\}}.$$

Due to the presence of  $\mathcal{E}_S$  in the sparse domination, the Muckenhoupt class is too large and so we introduce

$$\tilde{\mathcal{B}}_p(\mu) := \left\{ w : \mathcal{X} \rightarrow (0, \infty) : [w]_{\tilde{\mathcal{B}}_p(\mu)} := \sup_{\substack{Q, R \in \mathcal{D} \\ Q^{(1)} = R^{(1)}}} \langle w \rangle_Q \langle w^{-\frac{p'}{p}} \rangle_R^{\frac{p}{p'}} < \infty \right\}.$$

It is clear that  $[w]_{\mathcal{B}_p(\mu)} \leq [w]_{\tilde{\mathcal{B}}_p(\mu)}$ , and so  $\tilde{\mathcal{B}}_p(\mu) \subseteq \mathcal{B}_p(\mu)$ . In the proof of Theorem B, we will use the following fact—satisfied by  $\tilde{\mathcal{B}}_p(\mu)$  weights—: if  $E \subseteq S$  and  $\mu(E) \sim \mu(S)$ , then

$$w(S) \lesssim [w]_{\mathcal{B}_p(\mu)} w(E). \quad (2.6)$$

*Proof of Theorem B.* Let  $1 < p < \infty$ ,  $f \in L^p(\mathcal{X}, w)$  and  $g \in L^{p'}(\mathcal{X}, w)$ . By Theorem A we have that there exists a sparse collection  $\mathcal{S}$  such that

$$\langle \mathcal{P}f, gw \rangle \lesssim \langle \mathcal{A}_S f, gw \rangle + \mathcal{E}_S(f, gw).$$

Furthermore, we already know that  $\mathcal{A}_S$  is bounded on  $L^p(\mathcal{X}, w)$ , so that we only have to deal with  $\mathcal{E}_S(f, gw)$ :

$$\begin{aligned} \mathcal{E}_S(f, gw) &= \sum_{Q=\bar{Q} \in \mathcal{S}} \langle f \rangle_Q \left( \frac{1}{q(x_Q^{(1)})} \sum_{\substack{R=\bar{R} \in \mathcal{S} \\ R^{(1)}=Q^{(1)}}} \langle gw \rangle_R \right) \mu(Q) \\ &= \sum_{Q=\bar{Q} \in \mathcal{S}} \langle f w^{\frac{p'}{p}} \rangle_{Q, w^{-\frac{p'}{p}}} (w^{-\frac{p'}{p}}(Q))^{\frac{1}{p}} \left( \frac{1}{q(x_Q^{(1)})} (w^{-\frac{p'}{p}}(Q))^{1-\frac{1}{p}} \sum_{\substack{R=\bar{R} \in \mathcal{S} \\ R^{(1)}=Q^{(1)}}} \langle gw \rangle_R \right) \end{aligned}$$

$$\leq \left( \sum_{Q=\bar{Q} \in \mathcal{S}} \langle fw^{\frac{p'}{p}} \rangle_{Q, w^{-\frac{p'}{p}}}^p w^{-\frac{p'}{p}}(Q) \right)^{\frac{1}{p}} \left( \sum_{Q=\bar{Q} \in \mathcal{S}} w^{-\frac{p'}{p}}(Q) \left( \frac{1}{q(x_Q^{(1)})} \sum_{\substack{R=\bar{R} \in \mathcal{S} \\ R^{(1)}=Q^{(1)}}} \langle gw \rangle_R \right)^{p'} \right)^{\frac{1}{p'}}.$$

Concerning the first factor, by (2.6) and the fact that  $[w^{-\frac{p'}{p}}]_{\mathcal{B}_{p'}(\mu)} = [w]_{\mathcal{B}_p(\mu)}^{\frac{p'}{p}}$ , we obtain

$$\begin{aligned} \left( \sum_{Q=\bar{Q} \in \mathcal{S}} \langle fw^{\frac{p'}{p}} \rangle_{Q, w^{-\frac{p'}{p}}}^p w^{-\frac{p'}{p}}(Q) \right)^{\frac{1}{p}} &\lesssim [w^{-\frac{p'}{p}}]_{\mathcal{B}_{p'}(\mu)}^{\frac{1}{p}} \left( \sum_{Q=\bar{Q} \in \mathcal{S}} \langle fw^{\frac{p'}{p}} \rangle_{Q, w^{-\frac{p'}{p}}}^p w^{-\frac{p'}{p}}(E_Q) \right)^{\frac{1}{p}} \\ &\leq [w]_{\mathcal{B}_p(\mu)}^{\frac{p'}{p^2}} \left( \sum_{x \in \mathcal{X}} \left( \mathcal{M}_{\mathcal{D}, w^{-\frac{p'}{p}}}(fw^{\frac{p'}{p}}) \right)^p(x) w^{-\frac{p'}{p}}(x) \mu(x) \right)^{\frac{1}{p}} \\ &\leq [w]_{\mathcal{B}_p(\mu)}^{\frac{p'}{p^2}} \left( \sum_{x \in \mathcal{X}} f^p(x) w^{p' - \frac{p'}{p}}(x) \mu(x) \right)^{\frac{1}{p}} \\ &= [w]_{\mathcal{B}_p(\mu)}^{\frac{p'}{p^2}} \|f\|_{L^p(\mathcal{X}, w)}. \end{aligned}$$

For the second one we use Jensen's inequality to get

$$\begin{aligned} &\left( \sum_{Q=\bar{Q} \in \mathcal{S}} w^{-\frac{p'}{p}}(Q) \left( \frac{1}{q(x_Q^{(1)})} \sum_{\substack{R=\bar{R} \in \mathcal{S} \\ R^{(1)}=Q^{(1)}}} \langle gw \rangle_R \right)^{p'} \right)^{\frac{1}{p'}} \\ &= \left( \sum_{Q=\bar{Q} \in \mathcal{S}} \frac{w^{-\frac{p'}{p}}(Q)}{\mu(Q)^{p'}} \left( \frac{1}{q(x_Q^{(1)})} \sum_{\substack{R=\bar{R} \in \mathcal{S} \\ R^{(1)}=Q^{(1)}}} \langle g \rangle_{R, w} w(R) \right)^{p'} \right)^{\frac{1}{p'}} \\ &\leq \left( \sum_{Q=\bar{Q} \in \mathcal{S}} \frac{w^{-\frac{p'}{p}}(Q)}{\mu(Q)^{p'}} \frac{1}{q(x_Q^{(1)})} \sum_{\substack{R=\bar{R} \in \mathcal{S} \\ R^{(1)}=Q^{(1)}}} \langle g \rangle_{R, w}^{p'} w(R) w(R)^{p'-1} \right)^{\frac{1}{p'}} \\ &\lesssim \left( \sup_{\substack{Q, R \in \mathcal{D} \\ Q^{(1)}=R^{(1)}}} \frac{w^{-\frac{p'}{p}}(Q)}{\mu(Q)} \frac{w(R)^{p'-1}}{\mu(R)^{p'-1}} \right)^{\frac{1}{p'}} \\ &\quad \times \left( \sum_{Q=\bar{Q} \in \mathcal{S}} \frac{1}{q(x_Q^{(1)})} \sum_{\substack{R=\bar{R} \in \mathcal{S} \\ R^{(1)}=Q^{(1)}}} \langle g \rangle_{R, w}^{p'} w(R) \right)^{\frac{1}{p'}} \\ &\leq \left( \sup_{\substack{Q, R \in \mathcal{D} \\ Q^{(1)}=R^{(1)}}} \langle w^{-\frac{p'}{p}} \rangle_Q^{\frac{1}{p'}} \langle w \rangle_R^{\frac{1}{p}} \right) \left( \sum_{Q=\bar{Q} \in \mathcal{S}} \langle g \rangle_{Q, w}^{p'} w(Q) \right)^{\frac{1}{p'}} \end{aligned}$$

$$\lesssim [w]_{\tilde{\mathcal{B}}_p(\mu)}^{\frac{1}{p}} [w]_{\mathcal{B}_p(\mu)}^{\frac{1}{p'}} \|g\|_{L^{p'}(\mathcal{X}, w)},$$

where in the last step we proceeded as we did for the first factor. Finally, gathering all the estimates together

$$\mathcal{E}_{\mathcal{S}}(f, gw) \lesssim [w]_{\tilde{\mathcal{B}}_p(\mu)}^{\frac{1}{p}} [w]_{\mathcal{B}_p(\mu)}^{\frac{p'}{p^2} + \frac{1}{p'}} \|f\|_{L^p(\mathcal{X}, w)} \|g\|_{L^{p'}(\mathcal{X}, w)}$$

and then

$$\langle \mathcal{P}f, gw \rangle \lesssim [w]_{\tilde{\mathcal{B}}_p(\mu)}^{\frac{1}{p}} [w]_{\mathcal{B}_p(\mu)}^{\frac{p'}{p^2} + \frac{1}{p'}} \|f\|_{L^p(\mathcal{X}, w)} \|g\|_{L^{p'}(\mathcal{X}, w)}.$$

□

*Remark 2.4.* Our Muckenhoupt-type conditions can be viewed as natural counterparts of the conditions found by Bekollé and Bonami [1] in the continuous setting. Indeed, there the only balls that appear in the condition on the weights are those whose radii are comparable to their distance to the boundary. Using the usual embedding of  $\mathcal{X}$  in the hyperbolic disk (as in [8]), the sets that correspond to those balls in the tree are precisely the elements of  $\mathcal{D}$  which are not singletons, which are the ones that appear in the definition of the class  $\tilde{\mathcal{B}}_p(\mu)$ .

*Remark 2.5.* When  $q$  is bounded,  $\mathcal{P}$  admits sparse domination in the usual sense and we recover the linear dependence on  $[w]_{\mathcal{B}_2(\mu)}$  in its  $L^2$ -bound. This would be in line with the sharp result in [26]. Our proof does not provide that because of  $\mathcal{E}_{\mathcal{S}}$ , which gives a power of  $3/2$  in the  $L^2$ -case. We suspect that this is not an artifact of the proof but we leave it as an interesting open problem.

### 3. ENDPOINT ESTIMATES

**3.1. General Calderón-Zygmund theory in  $\mathcal{X}$ .** In this section we work with  $L^2$ -bounded integral operators  $T$  of the form

$$Tf(x) = \sum_{y \in \mathcal{X}} K(x, y) f(y) \mu(y),$$

for some kernel  $K: \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{C}$  which satisfies  $K(x, y) = K(y, x)$  for all  $x, y \in \mathcal{X}$ .  $K$  is said to satisfy Hörmander's condition if

$$\sup_{Q \in \mathcal{D}} \sup_{x, y \in Q} \sum_{z \in \mathcal{X} \setminus Q} |K(z, x) - K(z, y)| \mu(z) < \infty. \quad (\text{HörmCond})$$

The above is just the classical Hörmander's condition rewritten in our setting. The following estimate can be viewed as an integral size condition in the spirit of the one in [6]—already present in [27], albeit in a more implicit fashion—:

$$\sup_{Q \in \mathcal{D}} \sup_{x \in Q} \sum_{z \in Q^{(1)} \setminus Q} |K(x, z)| \mu(z) < \infty. \quad (\text{SizeCond})$$

*Remark 3.1.* The symmetry condition  $K(x, y) = K(y, x)$  for the kernel can be dropped but then (HörmCond) and (SizeCond) have to be modified accordingly to adjust the proof below. In particular,  $H^1$  and  $L^1$  estimates require the integration to run over the first variable of the kernel, while BMO ones require that it runs over the second.

Throughout this subsection, we do not need to assume that  $\mathcal{X}$  is radial, nor do we need any assumption on the measure  $\mu$ .

*Proof of Theorem C.* We proof the three claimed estimates in turn.

**Weak-type bound.** Let  $0 \leq f \in L^1$ . If  $0 < \lambda \leq \|f\|_1$ , then

$$\lambda \mu(\{x \in \mathcal{X} : |Tf(x)| > \lambda\}) \leq \lambda \leq \|f\|_1,$$

and there is nothing to prove. Otherwise, fix  $\lambda > \|f\|_1$ . We apply Lemma 2.1 to the Calderón-Zygmund cubes  $\{Q_j\}_j$ —that is, to the smallest disjoint cubes in  $\mathcal{D}$  that cover  $\{\mathcal{M}_{\mathcal{D}}f > \lambda\}$ —. Denote  $\Omega_\lambda = \bigcup_j Q_j$  and write

$$\begin{aligned} \mu\left(\{x \in \mathcal{X} : |Tf(x)| > \lambda\}\right) &\leq \mu(\Omega_\lambda) + \mu\left(\{x \in \mathcal{X} : |Tg(x)| > \frac{\lambda}{2}\}\right) \\ &\quad + \mu\left(\{x \in \mathcal{X} \setminus \Omega_\lambda : |Tb(x)| > \frac{\lambda}{2}\}\right) \\ &=: \text{I} + \text{II} + \text{III}. \end{aligned}$$

The weak-type of  $\mathcal{M}_{\mathcal{D}}$  immediately gives

$$\text{I} \lesssim \frac{\|f\|_1}{\lambda}.$$

Chebyshev's inequality, the  $L^2$ -boundedness of  $T$ , and property (3) of Lemma 2.1 yield

$$\text{II} \lesssim \frac{1}{\lambda^2} \|Tg\|_2^2 \lesssim \frac{1}{\lambda^2} \|g\|_2^2 \lesssim \frac{1}{\lambda} \|f\|_1.$$

Finally, by Chebyshev's inequality again,

$$\text{III} \leq \frac{1}{\lambda} \sum_{x \in \mathcal{X} \setminus \Omega_\lambda} |Tb(x)| \mu(x) \leq \frac{1}{\lambda} \sum_j \sum_{x \in \mathcal{X} \setminus Q_j} |Tb_j(x)| \mu(x) =: \frac{1}{\lambda} \sum_j \text{III}_j.$$

To estimate each term, we first split into two regions:

$$\begin{aligned} \text{III}_j &= \sum_{x \in \mathcal{X} \setminus Q_j} \left| \sum_{y \in Q_j^{(1)}} K(x, y) b_j(y) \mu(y) \right| \mu(x) \\ &= \sum_{x \in \mathcal{X} \setminus Q_j^{(1)}} \left| \sum_{y \in Q_j^{(1)}} K(x, y) b_j(y) \mu(y) \right| \mu(x) + \sum_{x \in Q_j^{(1)} \setminus Q_j} \left| \sum_{y \in Q_j^{(1)}} K(x, y) b_j(y) \mu(y) \right| \mu(x). \end{aligned}$$

For the first one, we use the mean zero of  $b_j$  and (HörmCond):

$$\begin{aligned} \sum_{x \in \mathcal{X} \setminus Q_j^{(1)}} \left| \sum_{y \in Q_j^{(1)}} K(x, y) b_j(y) \mu(y) \right| \mu(x) &= \sum_{x \in \mathcal{X} \setminus Q_j^{(1)}} \left| \sum_{y \in Q_j^{(1)}} (K(x, y) - K(x, x_{Q_j})) b_j(y) \mu(y) \right| \mu(x) \\ &\leq \sum_{y \in Q_j^{(1)}} |b_j(y)| \left( \sup_{z \in Q_j^{(1)}} \sum_{x \in \mathcal{X} \setminus Q_j^{(1)}} |K(x, y) - K(x, z)| \mu(x) \right) \mu(y) \lesssim \|b_j\|_1. \end{aligned}$$

For the remaining term, we further use the definition of  $b_j$ :

$$\begin{aligned} \sum_{x \in Q_j^{(1)} \setminus Q_j} \left| \sum_{y \in Q_j^{(1)}} K(x, y) b_j(y) \mu(y) \right| \mu(x) \\ \leq \sum_{x \in Q_j^{(1)} \setminus Q_j} \left| \sum_{y \in Q_j} K(x, y) b_j(y) \mu(y) \right| \mu(x) + \sum_{x \in Q_j^{(1)}} \left| T \left( \langle f 1_{Q_j} \rangle_{Q_j^{(1)}} 1_{Q_j^{(1)}} \right) (x) \right| \mu(x) \\ =: \alpha_j + \beta_j. \end{aligned}$$

We deal with  $\beta_j$  using Hölder's inequality and the  $L^2$ -bound:

$$\beta_j \leq \mu(Q_j^{(1)})^{\frac{1}{2}} \left\| T \left( \langle f 1_{Q_j} \rangle_{Q_j^{(1)}} 1_{Q_j^{(1)}} \right) \right\|_2 \lesssim \mu(Q_j^{(1)})^{\frac{1}{2}} \left\| \langle f 1_{Q_j} \rangle_{Q_j^{(1)}} 1_{Q_j^{(1)}} \right\|_2 = \|f 1_{Q_j}\|_1.$$

For  $\alpha_j$  we use Fubini's theorem and (SizeCond):

$$\begin{aligned} \alpha_j &\leq \sum_{x \in Q_j^{(1)} \setminus Q_j} \sum_{y \in Q_j} |K(x, y)| |f(y)| \mu(y) \mu(x) \\ &\leq \sum_{y \in Q_j} |f(y)| \mu(y) \sup_{y' \in Q_j} \sum_{x \in Q_j^{(1)} \setminus Q_j} |K(x, y')| \mu(x) \lesssim \|f 1_{Q_j}\|_1. \end{aligned}$$

Finally, we collect all estimates together and by property (2) of Lemma 2.1,

$$\text{III} \lesssim \frac{1}{\lambda} \sum_j (\|b_j\|_1 + \|f 1_{Q_j}\|_1) \lesssim \frac{1}{\lambda} \|f\|_1,$$

as desired.

**$L^\infty$  – BMO bound.** We use the characterization of the BMO norm given in (1.1). Let  $f \in L^\infty(\mathcal{X})$ . The estimate

$$\sup_{Q \in \mathcal{D}} \frac{1}{\mu(Q)} \sum_{x \in Q} |Tf(x) - \langle Tf \rangle_Q| d\mu \lesssim \|f\|_\infty$$

is standard and we omit its proof. On the other hand, for each  $Q \in \mathcal{D}$  we use the kernel representation to write

$$\left| \langle Tf \rangle_Q - \langle Tf \rangle_{Q^{(1)}} \right| = \frac{1}{\mu(Q) \mu(Q^{(1)})} \left| \mu(Q^{(1)}) \sum_{x \in Q} Tf(x) \mu(x) - \mu(Q) \sum_{z \in Q^{(1)}} Tf(z) \mu(z) \right|$$

$$\begin{aligned}
&= \frac{1}{\mu(Q)\mu(Q^{(1)})} \left| \sum_{\substack{x \in Q \\ z \in Q^{(1)}}} \sum_{y \in \mathcal{X}} (K(x, y) - K(z, y)) f(y) \mu(y) \mu(x) \mu(z) \right| \\
&\leq \sup_{\substack{x \in Q \\ z \in Q^{(1)}}} \sum_{y \in \mathcal{X} \setminus Q^{(1)}} |K(x, y) - K(z, y)| |f(y)| \mu(y) \\
&\quad + \langle |T(f1_Q)| \rangle_Q + \langle |T(f1_{Q^{(1)}})| \rangle_{Q^{(1)}} + \sup_{x \in Q} \sum_{y \in Q^{(1)} \setminus Q} |K(x, y)| |f(y)| \mu(y) \\
&=: \text{I} + \text{II} + \text{III} + \text{IV}.
\end{aligned}$$

The estimate I  $\lesssim \|f\|_\infty$  follows directly from (HörmCond). Next, by the  $L^2$ -bound for  $T$

$$\text{II} \leq \mu(Q)^{-\frac{1}{2}} \|Tf\|_{L^2(Q)} \lesssim \mu(Q)^{-\frac{1}{2}} \|f\|_{L^2(Q)} \lesssim \|f\|_\infty,$$

and a similar estimate holds for III. Finally, (SizeCond) implies IV  $\lesssim \|f\|_\infty$  and completes the proof.

**$H^1 - L^1$  bound.** It is enough to prove that  $\|Tb\|_1 \lesssim 1$  for every simple atomic block  $b$ . Fix one such  $b$ , so that

$$b = a - \langle a \rangle_Q 1_Q,$$

where  $\text{supp}(a) \subseteq R \in \mathcal{D}_1(Q)$  and  $\|a\|_\infty \leq \mu(R)^{-1}$ . We write

$$\|Tb\|_1 \leq \|Ta\|_{L^1(Q)} + \|Tb\|_{L^1(\mathcal{X} \setminus Q)},$$

and handle each term in turn. For the first one, we further split

$$\|Ta\|_{L^1(Q)} \leq \|Ta\|_{L^1(R)} + \|Ta\|_{L^1(Q \setminus R)} + |\langle a \rangle_Q| \|T1_Q\|_{L^1(Q)}.$$

The first term on the right hand side above is dealt with using the  $L^2$ -bound of  $T$  and the size estimate on  $a$ :

$$\|Ta\|_{L^1(R)} \leq \mu(R)^{\frac{1}{2}} \|Ta\|_{L^2(R)} \lesssim \mu(R)^{\frac{1}{2}} \|a\|_{L^2(R)} \leq 1.$$

Similarly, for the third term on the right hand side above we obtain

$$|\langle a \rangle_Q| \|T1_Q\|_{L^1(Q)} \leq \|a\|_\infty \frac{\mu(R)}{\mu(Q)} \|T1_Q\|_{L^2(Q)} \mu(Q)^{\frac{1}{2}} \lesssim 1,$$

using  $\text{supp}(a) \subseteq R$ . The second term on the right hand side above requires (SizeCond):

$$\begin{aligned}
\|Ta\|_{L^1(Q \setminus R)} &\leq \sum_{x \in Q \setminus R} \sum_{y \in R} |K(x, y)| |a(y)| \mu(y) \mu(x) \\
&\leq \|a\|_1 \sup_{y \in R} \sum_{x \in Q \setminus R} |K(x, y)| \mu(x) \\
&\lesssim \|a\|_1 \leq \|a\|_\infty \mu(R) \leq 1.
\end{aligned}$$

Finally, using (HörmCond) and the mean zero of  $b$  in  $Q$  yields

$$\|Tb1_{\mathcal{X} \setminus Q}\|_1 = \sum_{x \in \mathcal{X} \setminus Q} \left| \sum_{y \in Q} K(x, y) b(y) \mu(y) \right| \mu(x)$$

$$\begin{aligned}
&= \sum_{x \in \mathcal{X} \setminus Q} \left| \sum_{y \in Q} (K(x, y) - K(x, x_Q)) b(y) \mu(y) \right| \mu(x) \\
&\leq \|b\|_1 \sup_{y \in Q} \sum_{x \in \mathcal{X} \setminus Q} \left| K(x, y) - K(x, x_Q) \right| \mu(x) \\
&\lesssim \|b\|_1 \leq \|b\|_{\mathbb{H}^1} \leq 1,
\end{aligned}$$

which finishes the proof.  $\square$

**3.2. Endpoint results for  $\mathcal{P}$ .** We start checking the Hörmander and size conditions for the kernel of  $\mathcal{P}$ . As one could have expected, they are weaker than the estimates in Section 1 and can be deduced from them. For instance, (HörmCond) can be checked as follows: fix a non-singleton  $Q \in \mathcal{D}$  —otherwise, there is nothing to prove— and  $x \neq y \in Q$ . Summing in annuli and using Proposition 1.9 yields

$$\begin{aligned}
\sum_{z \in \mathcal{X} \setminus Q} |\mathcal{K}(z, x) - \mathcal{K}(z, y)| \mu(z) &= \sum_{\ell=1}^{|x_Q|} \sum_{z \in Q^{(\ell)} \setminus Q^{(\ell-1)}} |\mathcal{K}(z, x) - \mathcal{K}(z, y)| \mu(z) \\
&\leq \sum_{\ell=1}^{|x_Q|} \mu(x_Q^{(\ell)}) \frac{\nu_{|x_Q|}^{|x_Q|-\ell+1}}{\mu(Q^{(\ell)})} + \sum_{\ell=1}^{|x_Q|} \mu(Q^{(\ell)} \setminus (Q^{(\ell-1)} \cup \{x_Q^{(\ell)}\})) \frac{\nu_{|x_Q|}^{|x_Q|-\ell}}{\mu(Q^{(\ell-1)})} \\
&\lesssim \sum_{\ell=1}^{|x_Q|} \nu_{|x_Q|}^{|x_Q|-\ell+1} + \sum_{\ell=1}^{|x_Q|} \mu(Q^{(\ell-1)}) q(x_Q^{(\ell)}) \frac{\nu_{|x_Q|}^{|x_Q|-\ell+1}}{q(x_Q^{(\ell)}) \mu(Q^{(\ell-1)})} \lesssim 1.
\end{aligned}$$

A similar computation using Lemma 1.8, part (2) gives (SizeCond) for  $\mathcal{K}$ . As a consequence, we can apply Theorem C to  $\mathcal{P}$  and deduce the corresponding endpoint estimates. We finish proving the stronger  $\mathbb{H}^1 - \mathbb{H}^1$  and BMO – BMO endpoint bounds.

*Proof of Theorem D.* Even if the  $\mathbb{H}^1$  estimate implies the BMO one, we prove both directly, without using duality.

**BMO  $\rightarrow$  BMO estimate:** Let  $f \in \text{BMO}$ . Since  $\mathcal{P}1_{\mathcal{X}} = 1_{\mathcal{X}}$

$$\left| \sum_{x \in \mathcal{X}} \mathcal{P}f(x) \mu(x) \right| = \left| \sum_{z \in \mathcal{X}} f(z) \mu(z) \right| \leq \|f\|_{\text{BMO}},$$

and we just have to bound the two suprema in (1.1). Fix  $Q \in \mathcal{D}$  and split  $f$  as  $f = f - \langle f \rangle_{Q^{(1)}} + \langle f \rangle_{Q^{(1)}} =: \tilde{f} + \langle f \rangle_{Q^{(1)}}$ . Since  $\mathcal{P}$  sends constants to constants, it is enough estimate the BMO norm of  $\mathcal{P}\tilde{f}$ . Since  $\tilde{f}$  has mean 0 over  $Q^{(1)}$ , we have

$$|\langle \tilde{f} \rangle_{Q^{(\ell)}}| + \sup_{R \in \mathcal{D}_\ell(Q)} |\langle \tilde{f} \rangle_R| \lesssim (\ell + 1) \|f\|_{\text{BMO}}, \quad \ell \geq 0. \quad (3.1)$$

To estimate the norm of  $\mathcal{P}\tilde{f}$ , we study separately the two suprema in (1.1). For the first, we fix  $Q \in \mathcal{D}$  and write  $\tilde{f} = \tilde{f}1_Q + \tilde{f}1_{\mathcal{X}\setminus Q}$  and we denote by I and II the corresponding associated terms after applying triangle inequality. Concerning the first part, John-Nirenberg inequality and the  $L^2$ -boundedness of  $\mathcal{P}$  give

$$\begin{aligned} \text{I} &= \frac{1}{\mu(Q)} \sum_{x \in Q} |\mathcal{P}(\tilde{f}1_Q)(x) - \langle \mathcal{P}(\tilde{f}1_Q) \rangle_Q| \mu(x) \\ &\lesssim \left( \frac{1}{\mu(Q)} \sum_{x \in Q} |\mathcal{P}(\tilde{f}1_Q)(x)|^2 \mu(x) \right)^{\frac{1}{2}} \lesssim \left( \frac{1}{\mu(Q)} \sum_{x \in \mathcal{X}} |\tilde{f}1_Q(x)|^2 \mu(x) \right)^{\frac{1}{2}} \\ &= \left( \frac{1}{\mu(Q)} \sum_{x \in Q} |f(x) - \langle f \rangle_{Q^{(1)}}|^2 \mu(x) \right)^{\frac{1}{2}} \leq \|f\|_{\text{BMO}}. \end{aligned}$$

For II, if  $Q$  is a single point there is nothing to do. Otherwise, we decompose it into annuli:

$$\begin{aligned} \text{II} &= \frac{1}{\mu(Q)} \sum_{x \in Q} |\mathcal{P}(\tilde{f}1_{\mathcal{X}\setminus Q})(x) - \langle \mathcal{P}(\tilde{f}1_{\mathcal{X}\setminus Q}) \rangle_Q| \mu(x) \\ &= \frac{1}{\mu(Q)} \sum_{x \in Q} \left| \sum_{z \in \mathcal{X}\setminus Q} \mathcal{K}(z, x) \tilde{f}(z) \mu(z) - \frac{1}{\mu(Q)} \sum_{y \in Q} \sum_{z \in \mathcal{X}\setminus Q} \mathcal{K}(z, y) \tilde{f}(z) \mu(z) \mu(y) \right| \mu(x) \\ &\leq \frac{1}{\mu(Q)^2} \sum_{x \in Q} \sum_{y \in Q} \sum_{z \in \mathcal{X}\setminus Q} |\mathcal{K}(z, x) - \mathcal{K}(z, y)| |\tilde{f}(z)| \mu(z) \mu(y) \mu(x) \\ &\leq \sup_{x, y \in Q} \sum_{z \in \mathcal{X}\setminus Q} |\mathcal{K}(z, x) - \mathcal{K}(z, y)| |\tilde{f}(z)| \mu(z) \\ &\leq \sum_{\ell=1}^{|x_Q|} \sup_{x, y \in Q} \sum_{z \in Q^{(\ell)} \setminus Q^{(\ell-1)}} |\mathcal{K}(z, x) - \mathcal{K}(z, y)| |\tilde{f}(z)| \mu(z) =: \sum_{\ell=1}^{|x_Q|} \text{II}_\ell. \end{aligned}$$

We now estimate each term using Lemma 1.9 and (3.1):

$$\begin{aligned} \text{II}_\ell &\lesssim |\tilde{f}(x_Q^{(\ell)})| \mu(x_Q^{(\ell)}) \frac{\nu_{|x_Q|-\ell+1}^{|x_Q|}}{\mu(Q^{(\ell)})} + \frac{\nu_{|x_Q|-\ell}^{|x_Q|}}{\mu(Q^{(\ell-1)})} \sum_{z \in Q^{(\ell)} \setminus (Q^{(\ell-1)} \cup \{x_Q^{(\ell)}\})} |\tilde{f}(z)| \mu(z) \\ &\leq \nu_{|x_Q|-\ell+1} \langle |\tilde{f}| \rangle_{Q^{(\ell)}} + \nu_{|x_Q|-\ell+1} \frac{1}{q(x_Q^{(\ell)})} \sum_{R \in \mathcal{D}_1(Q^{(\ell)})} \langle |\tilde{f}| \rangle_R \\ &\leq 2 \cdot \nu_{|x_Q|-\ell+1} \langle |\tilde{f}| \rangle_{Q^{(\ell)}} + \nu_{|x_Q|-\ell+1} \frac{1}{q(x_Q^{(\ell)})} \sum_{R \in \mathcal{D}_1(Q^{(\ell)})} \langle |\tilde{f}| \rangle_R - \langle |\tilde{f}| \rangle_{Q^{(\ell)}} \\ &\lesssim \ell \cdot \nu_{|x_Q|-\ell+1} \| |\tilde{f}| \|_{\text{BMO}} \lesssim \ell \cdot \nu_{|x_Q|-\ell+1} \|f\|_{\text{BMO}}. \end{aligned}$$

The resulting quantity is acceptable due to the geometric decay of  $\nu_{|x_Q|}^{|x_Q|-\ell}$ . To bound the second term in (1.1), we split

$$\begin{aligned} |\langle \mathcal{P}\tilde{f} \rangle_Q - \langle \mathcal{P}\tilde{f} \rangle_{Q^{(1)}}| &\leq |\langle \mathcal{P}(\tilde{f}1_Q) \rangle_Q| + |\langle \mathcal{P}(\tilde{f}1_{Q^{(1)}}) \rangle_{Q^{(1)}}| \\ &\quad + |\langle \mathcal{P}(\tilde{f}1_{Q^{(1)} \setminus Q}) \rangle_Q| + |\langle \mathcal{P}(\tilde{f}1_{\mathcal{X} \setminus Q^{(1)}}) \rangle_Q - \langle \mathcal{P}(\tilde{f}1_{\mathcal{X} \setminus Q^{(1)}}) \rangle_{Q^{(1)}}|. \end{aligned}$$

The first two terms above are estimated similarly. Indeed, by the  $L^2$ -boundedness of  $\mathcal{P}$ ,

$$|\langle \mathcal{P}(\tilde{f}1_{Q^{(1)}}) \rangle_{Q^{(1)}}| \lesssim \left( \langle |\tilde{f}|^2 \rangle_{Q^{(1)}} \right)^{\frac{1}{2}} = \left( \langle |\tilde{f} - \langle \tilde{f} \rangle_{Q^{(1)}}|^2 \rangle_{Q^{(1)}} \right)^{\frac{1}{2}} \leq \|\tilde{f}\|_{\text{BMO}},$$

and a similar computation applies to  $|\langle \mathcal{P}(\tilde{f}1_Q) \rangle_Q|$ . The third term above is estimated using Lemma 1.8, part (2), as follows:

$$\begin{aligned} |\langle \mathcal{P}(\tilde{f}1_{Q^{(1)} \setminus Q}) \rangle_Q| &\leq \frac{1}{\mu(Q)} \sum_{x \in Q} \sum_{y \in Q^{(1)} \setminus Q} |\mathcal{K}(x, y)| |\tilde{f}(y)| \mu(y) \\ &\leq \sup_{x \in Q} \sum_{y \in Q^{(1)} \setminus Q} |\mathcal{K}(x, y)| |\tilde{f}(y)| \mu(y) \lesssim \sum_{y \in Q^{(1)} \setminus Q} |\mathcal{K}(x_Q, y)| |\tilde{f}(y)| \mu(y) \\ &\leq \sum_{y \in Q^{(1)} \setminus Q} \sum_{\ell=1}^{|x_Q|} |\mathcal{K}_{Q^{(\ell)}}(x_Q, y)| |\tilde{f}(y)| \mu(y) \\ &\leq \left( \sum_{\ell=1}^{|x_Q|} \frac{1}{\mu(Q^{(\ell)})} \right) f(x_Q^{(1)}) \mu(x_Q^{(1)}) \\ &\quad + \left( \frac{1}{q(x_Q^{(1)}) \mu(Q)} + \sum_{\ell=1}^{|x_Q|} \frac{1}{\mu(Q^{(\ell)})} \right) \sum_{y \in Q^{(1)} \setminus (Q \cup \{x_Q^{(1)}\})} |\tilde{f}(y)| \mu(y) \\ &\lesssim \langle |\tilde{f}| \rangle_{Q^{(1)}} + \frac{1}{q(x_Q^{(1)})} \sum_{R \in \mathcal{D}_1(Q^{(1)})} \langle |\tilde{f}| \rangle_R \lesssim \|f\|_{\text{BMO}}, \end{aligned}$$

finishing the computation as with  $\text{II}_\ell$  before. Finally, for the last term we write

$$\begin{aligned} &|\langle \mathcal{P}(\tilde{f}1_{\mathcal{X} \setminus Q^{(1)}}) \rangle_Q - \langle \mathcal{P}(\tilde{f}1_{\mathcal{X} \setminus Q^{(1)}}) \rangle_{Q^{(1)}}| \\ &\leq \frac{1}{\mu(Q)} \frac{1}{\mu(Q^{(1)})} \sum_{x \in Q} \sum_{y \in Q^{(1)}} |\mathcal{P}(\tilde{f}1_{\mathcal{X} \setminus Q^{(1)}})(x) - \mathcal{P}(\tilde{f}1_{\mathcal{X} \setminus Q^{(1)}})(y)| \mu(y) \mu(x) \\ &\leq \frac{1}{\mu(Q)} \frac{1}{\mu(Q^{(1)})} \sum_{x \in Q} \sum_{y \in Q^{(1)}} \sum_{z \in \mathcal{X} \setminus Q^{(1)}} |\tilde{f}(z)| |\mathcal{K}(z, x) - \mathcal{K}(z, y)| \mu(z) \mu(y) \mu(x) \\ &\leq \sup_{x, y \in Q^{(1)}} \sum_{z \in \mathcal{X} \setminus Q^{(1)}} |\tilde{f}(z)| |\mathcal{K}(z, x) - \mathcal{K}(z, y)| \mu(z), \end{aligned}$$

and we split into annuli as with term I above to conclude.

$H^1 \rightarrow H^1$  **estimate:** Let  $b \in H^1$  be a simple atomic block supported on  $Q^{(1)} \in \mathcal{D}$ , i.e.  $b = a - \langle a \rangle_{Q^{(1)}} 1_{Q^{(1)}}$ , with  $a$  supported in  $Q$  satisfying the size condition. It is enough to check that  $\|\mathcal{P}b\|_{H^1} \lesssim 1$ , and we do it by finding a decomposition of it into atomic blocks that need not be simple, by the equivalence in Lemma 1.2. Because of cancellation, we have

$$\begin{aligned} \mathcal{P}b &= \sum_{R \in \mathcal{D}} \mathcal{D}_R b = \sum_{\substack{R \in \mathcal{D} \\ R \subsetneq Q^{(1)}}} \mathcal{D}_R b + \sum_{\substack{R \in \mathcal{D} \\ R \supseteq Q^{(1)}}} \mathcal{D}_R b \\ &= \sum_{\substack{R \in \mathcal{D} \\ R \subsetneq Q^{(1)}}} \sum_{\ell} \langle a - \langle a \rangle_{Q^{(1)}}, h_R^\ell \rangle h_R^\ell + \sum_{k=1}^{|Q|} \mathcal{D}_{Q^{(k)}} b = \mathcal{P}^Q a + \sum_{k=1}^{|Q|} \mathcal{D}_{Q^{(k)}} b. \end{aligned}$$

Clearly,  $\text{supp}(\mathcal{P}^Q a) \subseteq Q$  and has mean zero, so it is an atomic block. By the  $L^2$ -boundedness,  $\|\mathcal{P}^Q a\|_2 \lesssim \|a\|_2 \leq \mu(Q)^{-1/2}$ , so  $\|\mathcal{P}^Q a\|_{H^1} \lesssim 1$ .

We next claim that, for each  $k$ ,  $\mathcal{D}_{Q^{(k)}} b$  is an atomic block. It is supported on  $Q^{(k)}$ , where it has zero mean, and we can write it as

$$\mathcal{D}_{Q^{(k)}} b = 1_{Q^{(k-1)}} \mathcal{D}_{Q^{(k)}} b + \sum_{\substack{T: T^{(1)} = Q^{(k)} \\ T \neq Q^{(k-1)}}} 1_T \mathcal{D}_{Q^{(k)}} b.$$

We have

$$\begin{aligned} \|1_{Q^{(k-1)}} \mathcal{D}_{Q^{(k)}} b\|_2 &\leq \mu(Q^{(k-1)})^{\frac{1}{2}} \|1_{Q^{(k-1)}} \mathcal{D}_{Q^{(k)}} b\|_\infty \\ &\lesssim \frac{\mu(Q^{(k-1)})^{\frac{1}{2}} \nu_{|x_Q|^{-k+1}}^{|x_Q|}}{\mu(Q^{(k-1)})} \|b\|_1 \\ &\leq \frac{\nu_{|x_Q|^{-k+1}}^{|x_Q|}}{\mu(Q^{(k-1)})^{\frac{1}{2}}}, \end{aligned}$$

by Hölder's inequality, Proposition 1.10 and the fact that  $\|b\|_1 \leq \|b\|_{H^1} \leq 1$ . Analogously we get

$$\|1_T \mathcal{D}_{Q^{(k)}} b\|_2 \lesssim \frac{\nu_{|x_Q|^{-k}}^{|x_Q|}}{\mu(Q^{(k-1)})^{\frac{1}{2}}}$$

for every  $T \in \mathcal{D}$  such that  $T^{(1)} = Q^{(k)}$ ,  $T \neq Q^{(k-1)}$ . It follows that  $\mathcal{D}_{Q^{(k)}} b$  is an atomic block for every  $k = 1, \dots, |Q|$  and

$$\|\mathcal{D}_{Q^{(k)}} b\|_2 \lesssim \frac{\nu_{|x_Q|^{-k+1}}^{|x_Q|}}{\mu(Q^{(k-1)})^{\frac{1}{2}}} + (q(x_Q^{(k)}) - 1) \frac{\nu_{|x_Q|^{-k}}^{|x_Q|}}{\mu(Q^{(k-1)})^{\frac{1}{2}}} \sim \frac{\nu_{|x_Q|^{-k+1}}^{|x_Q|}}{\mu(Q^{(k-1)})^{\frac{1}{2}}}.$$

Finally, we get

$$\|\mathcal{P}b\|_{H^1} \lesssim 1 + \sum_{k=1}^{|Q|} \nu_{|x_Q|^{-k+1}}^{|x_Q|} \lesssim 1.$$

□

*Remark 3.2.* It is natural to ask whether the above estimates apply to the Bergman shifts introduced in Remark 1.7. It is not difficult to see that this is the case if  $q$  is bounded, because in that case the corresponding kernels satisfy (HörmCond) and (SizeCond). However, in the general radial case  $\mu$  fails to be balanced (in the sense of [7]), so one cannot expect a general result to hold, and therefore there is no hope of a richer Calderón-Zygmund theory via representation as in [17, 14].

## REFERENCES

- [1] D. Békollé and A. Bonami, *Inégalités à poids pour le noyau de Bergman*, C. R. Acad. Sci., Paris, Sér. A **286** (1978), 775–778.
- [2] L. Cadilhac and J.M. Conde-Alonso and J. Parcet, *Spectral multipliers in group algebras and non-commutative Calderón-Zygmund theory*, J. Math. Pures Appl. **163** (2022), 450–472.
- [3] A. Carbonaro and G. Mauceri and S. Meda,  *$H^1$  and BMO for certain locally doubling metric measure spaces*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) **8** (2009), 543–582.
- [4] A. Carbonaro and G. Mauceri and S. Meda,  *$H^1$  and BMO for certain locally doubling metric measure spaces of finite measure*, Colloq. Math. **118** (2010), 13–41.
- [5] J.M. Conde-Alonso and J. Parcet, *Atomic blocks for noncommutative martingales*, Indiana Univ. Math. J. **65** (2016), 1425–1443.
- [6] J.M. Conde-Alonso and J. Parcet, *Nondoubling Calderón-Zygmund theory: a dyadic approach*, J. Fourier Anal. Appl. **25** (2019), 1267–1292.
- [7] J.M. Conde-Alonso and J. Pipher and N. Wagner, *Balanced measures, sparse domination and complexity-dependent weight classes*, Math. Ann. **391** (2025), 2209–2253.
- [8] J.M. Cohen and F. Colonna, *Embeddings of trees in the hyperbolic disk*, Complex Variables Theory Appl. **24** (1994), 311–335.
- [9] J.M. Cohen and F. Colonna and M.A. Picardello and D. Singman, *Bergman spaces and Carleson measures on homogeneous isotropic trees*, Potential Anal. **44** (2016), 745–766.
- [10] D. Cruz-Uribe and J.M. Martell and C. Pérez, *Sharp weighted estimates for classical operators*, Adv. Math. **229** (2012), 408–441.
- [11] A. Culiuc and F. Di Plinio and Y. Ou, *Uniform sparse domination of singular integrals via dyadic shifts*, Math. Res. Lett. **25** (2018), 21–42.
- [12] F. De Mari and M. Monti and E. Rizzo, *Horocyclic harmonic Bergman spaces on homogeneous trees*, Anal. Appl. **23** (2025), 447–474.
- [13] F. De Mari and M. Monti and M. Vallarino, *Harmonic Bergman projectors on homogeneous trees*, Potential Anal. **61** (2024), 153–182.
- [14] F. Di Plinio and B.D. Wick and T. Williams, *Wavelet representation of singular integral operators*, Math. Ann. **386** (2023), 1829–1889.
- [15] E. Doubtsov and A.V. Vasin, *Calderón-Zygmund operators on RBMO*, Proc. Amer. Math. Soc. **151** (2023), 595–610.
- [16] T. Hytönen, *A framework for non-homogeneous analysis on metric spaces, and the RBMO space of Tolsa*, Publ. Mat. **54** (2010), 485–504.
- [17] T.P. Hytönen, *The sharp weighted bound for general Calderón-Zygmund operators*, Ann. of Math. (2) **175** (2012), 1473–1506.
- [18] M. Keller and D. Lenz and R. Wojciechowski, “Graphs and discrete Dirichlet spaces”, Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 358, Springer, Cham, 2021.

- [19] A.K. Lerner, *On an estimate of Calderón-Zygmund operators by dyadic positive operators*, J. Anal. Math. **121** (2013), 141–161.
- [20] A.K. Lerner, *A simple proof of the  $A_2$  conjecture*, Int. Math. Res. Not. IMRN (2013), no. 14, 3159–3170.
- [21] L.D. López-Sánchez and J.M. Martell, and J. Parcet, *Dyadic harmonic analysis beyond doubling measures*, Adv. Math. **267** (2014), 44–93.
- [22] E. Lorist, *On pointwise  $\ell^r$ -sparse domination in a space of homogeneous type*, J. Geom. Anal. **31** (2021), 9366–9405.
- [23] M. Monti,  *$H^1$  and BMO spaces for exponentially decreasing measures on homogeneous trees*, to appear in: N. Arcozzi, M. M. Peloso and A. Tabacco, (New Trends in) Complex and Fourier Analysis, Springer INdAM Series, Springer. Cambridge Tracts in Mathematics, arXiv:2301.07600,
- [24] M.C. Pereyra, *Dyadic harmonic analysis and weighted inequalities: the sparse revolution*, in “New trends in applied harmonic analysis. Vol. 2—harmonic analysis, geometric measure theory, and applications”, Appl. Numer. Harmon. Anal., 159–239, Birkhäuser/Springer, Cham, 2019.
- [25] P. Petersen, “Riemannian geometry”, Graduate Texts in Mathematics, 171, Springer, New York, 2006.
- [26] S. Pott and M.C. Reguera, *Sharp Békollé estimates for the Bergman projection*, J. Funct. Anal. **265** (2013), 3233–3244.
- [27] X. Tolsa, *A proof of the weak  $(1, 1)$  inequality for singular integrals with non doubling measures based on a Calderón-Zygmund decomposition*, Publ. Mat. **45** (2001), 163–174.
- [28] X. Tolsa, *BMO,  $H^1$ , and Calderón-Zygmund operators for non doubling measures*, Math. Ann. **319** (2001), 89–149.
- [29] A. Volberg and P. Zorin-Kranich, *Sparse domination on non-homogeneous spaces with an application to  $A_p$  weights*, Rev. Mat. Iberoam. **34** (2018), 1401–1414.
- [30] W. Woess, “Random walks on infinite graphs and groups”, Cambridge Tracts in Mathematics, **138**, Cambridge University Press, Cambridge, 2000.

**Jose M. Conde-Alonso**

jose.conde@uam.es

Dept. of Mathematics, Universidad Autónoma de Madrid, 7 Francisco Tomás y Valiente, 28049 Madrid, Spain.

**Filippo De Mari**

filippo.demari@unige.it

Dipartimento di Matematica and MaLGa Center, Università di Genova, Via Dodecaneso 35, 16146 Genova, Italy.

**Matteo Monti**

matteo.monti@unimib.it

Dipartimento di Matematica e Applicazioni, Università degli studi di Milano-Bicocca, Via Roberto Cozzi, 55, 20126 Milano, Italy.

**Elena Rizzo**

elena.rizzo@unimi.it

Dipartimento di Matematica “Federigo Enriques”, Via Cesare Saldini, 50, 20133 Milano,

Italy.

**Maria Vallarino**

`maria.vallarino@polito.it`

Dipartimento di Scienze Matematiche “Giuseppe Luigi Lagrange”, Politecnico di Torino,  
Corso Duca degli Abruzzi 24, 10129 Torino, Italy.