SHARP AND OPTIMAL INEQUALITIES IN HARMONIC ANALYSIS

by

GIANMARCO BROCCHI

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

School of Mathematics
College of Engineering and Physical Sciences
University of Birmingham
June 2021

Abstract

This thesis deals with three topics in Harmonic Analysis:

1. Sharp restriction theory;
2. Sparse domination for square function operators;
3. Two weight theory for the Bergman projection.

In the first part we study some sharp inequalities that arise composing a k-plane transform with the square of the Fourier extension operator from the paraboloid. We study the sharp form of these inequalities. We compute the optimal constants and characterise maximisers.

The second and main part of this thesis develops on sparse domination for square function operators. In particular we derive a sparse domination in form under minimal testing conditions. We called this domination a "quadratic" as it dominates the non-linear operator $(Sf)^2$ rather than Sf. This produces optimal weighted estimates for the dominated square functions.

We show that a quadratic domination holds also for non-integral square functions associated with a general elliptic operator L. This refines and improves the domination in [BFP16] when the operator is a square function.

The last part of the thesis studies the Bergman projection P on the complex unit ball $\mathbb{B}^d$ in $\mathbb{C}^d$. We derive sufficient conditions for two weight estimates for P via sparse domination. These conditions are given in terms of "bumped" Orlicz averages of the two weights. On the way, we also derive mixed B_2–B_∞ estimates for the Bergman projection on $L^2(\mathbb{B}^d)$.

To nonno Giovanni

Acknowledgements

No mathematician is an island.

I surely would not be such without the support of my advisors: Maria Carmen and Jonathan. I would like to personally thank Maria for the incredible support she gave me during incredible times. I wish to stay more to see Balbino and Santiago growing. I would also like to thank Jon for being the kind of mathematician who I aspire to be. Thanks to both for engaging me at the blackboards.

Funding During my PhD, I was supported by an EPSRC Scholarship for the University of Birmingham. I would like to acknowledge Maria's support through her grant which allowed me to attend more conferences outside Europe.

These years would have been completely different without many wonderful people:

Giuseppe, thanks for your invaluable company and for hosting me (remotely) in the final month; Sebastiano, thanks for always reading something to us; thanks to Gian Maria, for making maths such a great story to tell.

Thanks to Shohei for all the sushi and suggestions, to Julian for all his corrections. Thanks to Diogo, for cheering up all our lunches at the Staff House; thanks to Marco and Gennady for reaching out and sharing questions and ideas. Thanks for all the endless maths discussions.

Thanks to Riccardo for running and discussing science at the same time.

Thanks to Yahya, for visiting me every week during the first lockdown and beyond; to Alex, for organising all the surprise trips and for his secret salad; to Soodeh, for always being so excited about going everywhere. Thanks to Tejas, for being so spontaneous and for all the board games! Thanks Constantin for the unique chats over countless flat white, and for reading so much news. Thanks to Francesco, for showing me how to `git rebase` a house. Thanks to Jen for reading Stein with me; and to Andy for the Scotch eggs and all the pints. Thanks to Davide, for his great taste in food.

Thanks to the friends of the Swiss house: Ettore, Marianna, Edo, Geggio, Ceci, il Casti, Lorenzo. Thanks to the friends of the Belgian house: Michele, for being so welcoming and for all the puns; Angelos, *ευχαριστώ* for the all the movies at the Astro-kine-polis; Daniela, for winning the first match of every game; thanks to Shivangee, Francesco and Luca.

Thanks zii e nonna, for the walks together when I return; thanks Giò, Lore e Marco for all the video calls, for the great holiday trips. Thanks to Fr. Dries for all our meetings at the Manresa House, to Kensy for the online masses in hard times, and to Don Ale for all the (always online) SV meetings.

Thanks Ludo, for always making my Monday nights — and Tuesday afternoons, sometimes. Thanks Fede, for reading all my drafts and being so patient with my English. Grazie Catherine, per farmi sempre ridere.

Thanks to the friends who are with me in every house: Pippo, Ettore, Luigi, Ale. Thanks for showing that distance doesn't count.
Thanks mamma e babbo, for always saying "you can count on us".

CONTENTS

Notation in this thesis

The results that are contributions by author are indicated with letters: A, B, C, ... Other results are indicated with numbers, as the equations, using (Chapter.section.#).

For two positive quantities X and Y we will write $X \lesssim Y$ to mean that there exists a constant $C > 0$ such that $X \leq CY$. We write $X \approx Y$ when also the reverse inequality $Y \lesssim X$ holds, and so the quantities X and Y are equal up to constants.

Symbols

$\mathscr{D}$ dyadic system

$\langle f\rangle_Q := \fint_Q f$ average of the function f over the set Q

$\langle f\rangle_Q^\sigma$ weighted average of the function f over the set Q with respect to the measure $\sigma\,\mathrm{d}x$

$\mathscr{S}$ sparse family

$\mathbb{1}_Q$ indicator function on the set Q

CHAPTER 1

INTRODUCTION

Dove c'è gusto, non c'è perdenza.

1.1 Sharp restriction theory

One of the main tools used in harmonic analysis is the Fourier transform. It is known, by the Hausdorff–Young theorem, that it maps L^p to $L^{p'}$, for $1 \leq p \leq 2$, where $p' = \frac{p}{p-1}$. However, the Fourier transform is not surjective on these spaces, and a function in its image is more regular than a generic element in $L^{p'}$. Indeed, unlike a generic function in $L^{p'}$, the Fourier transform of a function in L^p can be meaningfully restricted to curved hypersurfaces, although these have zero Lebesgue measure. The so-called "restriction estimates" quantify this phenomenon.

These estimates have important applications to dispersive PDEs, such as Schrödinger and wave equations. Solutions to these equations can be seen as Fourier transforms of functions supported on characteristic hypersurfaces, thus enabling the use of restriction estimates to obtain meaningful bounds. This procedure is called "Fourier extension".

Sharp restriction theory aims to compute the norms of these Fourier extension operators. It also seeks to characterise functions that achieve maximal norm; such functions are called maximisers.

Example 1.1.1. The Fourier extension from the parabola $\tau = \xi^2$ corresponds to the time evolution of the Schrödinger equation $i\partial_t u = \partial_x^2 u$. Gaussians are maximisers in dimension $d = 1, 2$, and it is conjectured that maximisers are gaussians in every dimension.

Despite all the efforts made so far, this conjecture remains open: maximisers are known only for the simplest surfaces in low dimension. The latest research directions try to find new methods to obtain sharp inequalities. Among these avenues, there are the heat flow techniques [BBCH09] and new tomography bounds for Fourier extensions [BBF+18; BV20; BN21]. These recent works investigate bounds for the Fourier extension operator composed with a k-plane transform.

Our first result is a sharp inequality of this kind, which holds in any dimension $d \geq 3$. It involves the Fourier extension from the paraboloid composed with the $(d-2)$-plane transform T_{d-2}, which is the operator that averages a function on a given $(d-2)$-dimensional plane. To state it, we denote by $\mathcal{A}_{d-2,d}$ the collection of all affine subspaces of $\mathbb{R}^d$ of codimension 2, which is the domain of T_{d-2}; while $e^{-it\Delta}$ is the solution operator of the free Schrödinger equation.

Theorem A. *Let $d \geq 3$. The following estimate:*

$$\|T_{d-2}(|e^{-it\Delta}f|^2)\|_{L^2(\mathbb{R}\times\mathcal{A}_{d-2,d})} \leq C_d \|f\|^2_{L^2(\mathbb{R}^d)}$$

is saturated only by gaussians and the optimal constant is

$$C_d = \left((d-2)\frac{\pi^{d/2}}{\Gamma(d/2)}\right)^{1/2}.$$

The inequality in Theorem A follows by composing Strichartz estimates for the Fourier extension operator and known L^p bounds for the operator T_{d-2}. The novelty is the sharp form of the inequality and the characterisation of maximisers.

Remark 1.1.2. Some maximisers for the L^p inequality for T_{d-2} are known [Dro14] and they are not gaussian, while the Fourier extension inequality from the paraboloid in higher dimension is not known in sharp form, and its maximisers have not been characterised.

We give a proof of Theorem A in Chapter 2.

Other k-plane Strichartz estimates in [BBF+18] have a weight in the L^2 norm on the right hand side. The investigation of these weights in Fourier extension inequalities has contributed in leading the author's attention to more general weighted estimates. From a broader point of view, weighted estimates — with optimal dependence on the characteristic of the weight — can be obtained via a powerful method today popular as *sparse domination*, which is having an incredible impact on harmonic analysis.

1.2 Background on weights and sparse domination

This section presents some background material in order to state our main results.

Weights appear in a variety of situations: for example, on a bounded domain, weights may arise as the Jacobian of a transformation or perturbation of the domain itself. They have many applications to PDEs [FKP91], approximation theory, quasiconformal theory [AIS01; PV02], complex analysis and operator theory [APR17].

In this thesis we call *weight* a positive, locally integrable function. We are interested in understanding how the norm of the operator depends on the weight in the underlying measure. It was known [HMW73] that for certain singular integral operators, like the Hilbert transform, for $1 < p < \infty$ the finiteness of the Muckenhoupt characteristic

$$[w]_{A_p} := \sup_Q \left(\frac{1}{|Q|} \int_Q w \right) \left(\frac{1}{|Q|} \int_Q w^{-\frac{1}{p-1}} \right)^{p-1} \tag{A_p}$$

is a necessary and sufficient condition for the boundedness on $L^p(w)$. Since the qualitative problem was settled, the quantitative problem attracted interest. It consists in the following question:

Given a bounded (sub)linear operator T from $L^p(w)$ to itself, what is the smallest power $\alpha \geq 0$ such that

$$\|T\|_{L^p(w)\to L^p(w)} \leq c_p\,[w]_{A_p}^{\alpha} \quad ?$$

The search for the optimal dependence on the characteristic of the weight was also motivated by a problem in quasi-conformal theory about regularity of solutions of the Beltrami equation

$$\partial f(z) = \mu(z)\,\bar{\partial} f(z) \tag{1.2.1}$$

where $f, \mu\colon \mathbb{C} \to \mathbb{C}$ and $\|\mu\|_{L^\infty} < 1$. The open question was:

What is the minimal q such that any solution $f \in W^{1,q}_{\mathrm{loc}}$ to (1.2.1) is continuous?

The condition $q > 1 + \|\mu\|_{L^\infty}$ was known to be sufficient [AIS01], while there are counterexamples for $q < 1 + \|\mu\|_{L^\infty}$. The critical value $q = 1 + \|\mu\|_{L^\infty}$ was shown to be sufficient by Petermichl and Volberg [PV02]. Their result follows from a sharp weighted estimate for a singular integral operator: the Ahlfors–Beurling operator, the complex analogue of the Hilbert transform.

The same question about optimal dependence on the weight can be asked for more general singular integral operators.

1.2.1 Integral operators

In the following, C will denote a positive constant which may change from line to line.

Singular Integral Operators

We consider the class of operators named after A. Calderón and A. Zygmund.

Definition 1.2.1 (Calderón–Zygmund kernel)**.** We say that a function $K(x, y)$ on $\mathbb{R}^d \times \mathbb{R}^d \setminus \{x = y\}$ is a Calderón–Zygmund kernel if there exists $C > 0$ and $\alpha \in (0, 1]$ such that

K satisfies the following size and regularity conditions:

$$|K(x,y)| \leq C|x-y|^{-d},$$

$$|K(x+h,y)-K(x,y)|+|K(x,y+h)-K(x,y)| \leq C\frac{|h|^{\alpha}}{|x-y|^{d+\alpha}}$$

for all $h \in \mathbb{R}^d$ such that $|x-y| > 2|h|$.

We denoted by C_c^∞ the space of smooth, compactly supported functions, and by $(C_c^\infty)'$ its dual. If the underlying measure is doubling, the space C_c^∞ is dense in any L^p space, for $1 \leq p < \infty$.

Definition 1.2.2 (Singular Integral Operator)**.** We say that a linear map $T\colon C_c^\infty \to (C_c^\infty)'$ associated with a Calderón–Zygmund kernel K is a Singular Integral Operator if for all $f, g \in C_c^\infty$ with disjoint supports one has the following integral representation

$$\langle Tf, g\rangle = \iint_{\mathbb{R}^d\times\mathbb{R}^d} K(x,y)f(y)g(x)\,\mathrm{d}y\,\mathrm{d}x.$$

Square function operators

In most of this thesis, and in particular in Chapter 3, we will focus on general square function operators, which we now introduce.

Definition 1.2.3 (Littlewood–Paley kernels)**.** A collection of functions $\{k_t(x,y)\}_{t>0}$ is a family of Littlewood–Paley kernels if there exists positive constants C_1, C_2 and $\alpha \in (0,1]$ such that the kernels k_t satisfy the following size and regularity conditions for all $x, y \in \mathbb{R}^d$:

$$|k_t(x,y)| \leq C_1\frac{t^{\alpha}}{(t+|x-y|)^{\alpha+d}}, \tag{C1}$$

$$|k_t(x+h,y)-k_t(x,y)|+|k_t(x,y+h)-k_t(x,y)| \leq C_2\frac{|h|^{\alpha}}{(t+|x-y|)^{d+\alpha}} \tag{C2}$$

for all $h \in \mathbb{R}^d$ and $t > |h|$.

Let $\{\theta_t\}_{t>0}$ be the family of integral operators $\theta_t f(x) = \int_{\mathbb{R}^d} k_t(x,y) f(y)\,\mathrm{d}y$. We consider the vertical square function

$$Sf(x) \coloneqq \left(\int_0^\infty |\theta_t f(x)|^2 \frac{\mathrm{d}t}{t} \right)^{1/2}. \tag{1.2.2}$$

Example 1.2.4 (Littlewood–Paley square function). A standard example for which the size and regularity conditions hold is $\theta_t f = f * \psi_t$, where $\psi_t(x) = t^{-d}\psi(t^{-1}x)$ and ψ is a mean zero Schwartz function which gives rise to the Littlewood–Paley square function [Gra14, §6.1]. In particular, conditions (C1) and (C2) are off-diagonal conditions compatible with the scaling.

1.2.2 Quantitative weighted estimates

The dependence of the operator norm $\|T\|^p_{L^p(w)\to L^p(w)}$ on the Muckenhoupt characteristic (A_p) has been first investigated by Buckley [Buc93, Theorem 2.5] for the Hardy–Littlewood maximal function

$$Mf(x) \coloneqq \sup_{B \ni x} ⨍_B |f(y)|\,\mathrm{d}y$$

where the supremum is taken over all balls B containing x and $⨍_B f \coloneqq |B|^{-1} \int_B f$.

Theorem 1.2.5 (Buckley 1993). *For $p > 1$ and for all weights $w \in A_p$ it holds that*

$$\|Mf\|^p_{L^p(w)} \leq C[w]^{p'}_{A_p} \|f\|^p_{L^p(w)}$$

and the power of the Muckenhoupt characteristic is the best possible.

A decade later, quantitative weighted estimates (optimal in terms of the Muckenhoupt characteristic $[w]_{A_p}$) have been obtained for the Hilbert transform [Pet07], the Riesz transform [Pet08], Haar shift [LPR10] and for general Calderón–Zygmund operators [Hyt12]:

Theorem 1.2.6 (Hytönen 2012)**.** *Let T be a Singular Integral Operator with Calderón–Zygmund kernel. For any $1 < p < \infty$ and weight $w \in A_p$ it holds that*

$$\|T\|^p_{L^p(w)\to L^p(w)} \leq c_p [w]^{\max\{p',p\}}_{A_p}.$$

Remark 1.2.7. The power in the characteristic of the weight is sharp and it can be matched with power weights. The sharpness had already been shown by Buckley [Buc93, Theorem 2.14] for some singular integral operator T and its maximal truncation $T^{\#}$ given by

$$T^{\#}f(x) := \sup_{\epsilon>0} |(K \cdot \mathbb{1}_{\mathbb{R}^d\setminus B(0,\epsilon)}) * f(x)|.$$

Theorem 1.2.6 is known as the "A_2 theorem", as the estimates for general $p \in (1,\infty)$ can be extrapolated [Gra14, Theorem 7.5.3] from the one with $p = 2$. This result has been extended and simplified by many authors [HLP13; HLM+12; HRT17; Hyt14], especially via sparse domination techniques [Ler13a; Lac17; Ler16], which we now introduce.

1.2.3 Sparse domination

Sparse domination consists in controlling non-local operators by a sum of positive averages. An operator T might be dominated either pointwise:

$$|Tf(x)| \leq C \sum_{Q\in\mathscr{S}} \left(\frac{1}{|Q|}\int_Q |f|\right) \mathbb{1}_Q(x) \tag{1.2.3}$$

for all x in a fixed cube Q_0; or in form:

$$\left|\int_{Q_0} Tf \cdot g \,\mathrm{d}x\right| \leq C \sum_{Q\in\mathscr{S}} \left(\frac{1}{|Q|}\int_Q |f|\right)\left(\frac{1}{|Q|}\int_Q |g|\right)|Q|. \tag{1.2.4}$$

In both cases, the constant C in (1.2.3) and (1.2.4) does not depend on the input functions, while the collections $\mathscr{S}$ do and they are *sparse* in the following sense

Definition 1.2.8 (Sparse collection)**.** A collection of cubes $\mathscr{S}$ is $\frac{1}{\tau}$-sparse, for $\tau \geq 1$, if for any $Q \in \mathscr{S}$ there exists a subset $E_Q \subseteq Q$ such that $\{E_Q\}_{Q\in\mathscr{S}}$ are pairwise disjoint and $|Q| \leq \tau |E_Q|$.

Roughly speaking, a collection of sets $\mathscr{S}$ is sparse if it contains a disjoint subcollection of sets that are not too small. For simplicity, the reader can think of $\mathscr{S}$ to be a collection of dyadic cubes, although sparse families can be defined for general Borel sets [Hän18].

The thrust of sparse domination consists in the fact that sparse expressions (the right hand sides in (1.2.3) and (1.2.4)) enjoy the same boundedness properties of several interesting operators, but are much simpler to deal with. As a consequence, sparse domination produces – in an unified manner – plenty of unweighted, weighted, and vector valued estimates, for the dominated operator. Moreover these estimates are often optimal in the dependence on the weight.

A growing list of operators sparsely dominated includes: Calderón–Zygmund operators [CCDO17; CR16; Ler16; CDO18b], bilinear Hilbert transform and multilinear singular integrals [CDO18a], variational Carleson operators [DDU18], oscillatory and random singular integrals [LS17], pseudodifferential operators [BC20], Stein's square function [CD17], and singular Radon transforms [Obe19].

Despite the fact that having a sparse domination is not always possible [BCOR19], the sparse paradigm has been extended beyond the classical theory to obtain weighted estimates for Bochner–Riesz multipliers [LMR19; BBL17], singular integral operators on spaces of non-homogeneous type [VZ18], and more general non-integral operators [BFP16] which will be discussed in §1.4.

Sparse domination can be seen as a technique to deduce properties of an operator from its action on a single function, or on a class of functions. For this reason sparse domination is particularly suited for improving classical $T(1)$ theorems.

1.2.4 Sparse $T1$ theorems

In the '80s David and Journé [DJ84] showed that L^2-boundedness of singular integral operators follows from the uniform boundedness on indicator functions. Let Q be a cube and let $\mathbb{1}_Q$ be the indicator function on Q taking values 1 on Q and 0 otherwise. The result in [DJ84] can be rephrased as follow.

Theorem 1.2.9 (David & Journé 1984). *Let T be a Singular Integral Operator with Calderón–Zygmund kernel and let $T^\star$ be its adjoint. If there exists $C > 0$ such that*

$$\langle |T\mathbb{1}_Q|, \mathbb{1}_Q\rangle + \langle |T^\star \mathbb{1}_Q|, \mathbb{1}_Q\rangle \leq C|Q| \tag{1.2.5}$$

holds for all cubes $Q \subseteq \mathbb{R}^d$, then $\|T\|_{L^2\to L^2} < \infty$.

This kind of results are known as "$T(1)$ theorems", as the operator is tested on constant functions. This classical result has been recast by Lacey and Mena [LM17b, Theorem 1.1] to a sparse domination under minimal assumptions.

Theorem 1.2.10 (Lacey & Mena 2016). *Let T be a singular integral operator with Calderón–Zygmund kernel that satisfies the testing condition (1.2.5). Then for any pair of compactly supported functions $f, g \in C_c^\infty$ there exists a sparse collection $\mathscr{S}$ such that*

$$|\langle Tf, g\rangle| \leq C \sum_{Q\in\mathscr{S}} \left(\fint_Q |f|\right)\left(\fint_Q |g|\right)|Q|$$

where $C > 0$ is a positive constant independent of f and g, and the symbol $\fint_Q f$ denotes the average of f over Q.

This theorem, instead of just L^2 boundedness, implies:

- weak $(1,1)$ bound [CCDO17, Appendix B], [BB18, Prop 3.1];
- strong L^p-bounds for $1 < p < \infty$;

- strong weighted bounds on $L^p(w)$ with optimal dependence on $[w]_{A_p}$ as in Theorem 1.2.6:

$$\|T\|_{L^p(w)\to L^p(w)} \le c_p [w]_{A_p}^{\max\{\frac{1}{p-1},1\}};$$

- weak weighted estimates at the endpoint for $w \in A_1$ [FN19, Theorem 1.4]:

$$\|T\|_{L^1(w)\to L^{1,\infty}(w)} \le c\,[w]_{A_1} \log(e + [w]_{A_\infty}),$$

where $[w]_{A_1} := \sup_Q \left(\fint_Q w\right) \|w^{-1}\|_{L^\infty(Q)}$ and $[w]_{A_\infty}$ is the Wilson characteristic

$$[w]_{A_\infty} := \sup_Q \frac{1}{w(Q)} \int_Q M(w\mathbb{1}_Q).$$

- Upper bound on γ_1, γ_2 for the asymptotic behaviour of the unweighted norm at the endpoints [FN19, Prop. 5.4]:

$$\lim_{p\to 1^+} \|T\|_{L^p\to L^p} \simeq (p-1)^{-\gamma_1}, \qquad \lim_{p\to\infty} \|T\|_{L^p\to L^p} \simeq p^{\gamma_2};$$

- vector valued estimates [LN20].

Again, we emphasise that only the testing condition (1.2.5) is assumed.

In the spirit of Lacey and Mena, we will derive a sparse $T1$ theorem for square functions.

1.3 $T1$ theorem for square functions

In this section we introduce the main result of Chapter 3.

The first $T(1)$ theorem for square functions is by Christ and Journé [CJ87]: they showed that a square function S is bounded on $L^2(\mathbb{R}^d)$ if θ_t applied to the constant function 1 gives rise to a Carleson measure $\nu := |\theta_t 1(x)|^2 \, \mathrm{d}t/t \, \mathrm{d}x$ on the upper half space $\mathbb{R}^{d+1}_+$.

A Carleson measure on $\mathbb{R}^{d+1}_+$ is a measure which acts like a d-dimensional measure in the following sense. Let Q be a cube in $\mathbb{R}^d$ with sides parallel to the coordinate axes. Denote by ℓQ and $|Q|$ the side length and the Lebesgue measure of Q, so that $(\ell Q)^d = |Q|$. Consider the Carleson box $B_Q := Q \times (0, \ell Q)$. Then ν is a Carleson measure if $\nu(B_Q)/|Q|$ is finite for any cube Q.

Successively, it has been shown [Hof08; Hof10; LM17a] that S is bounded in $L^2(\mathbb{R}^d)$ if there exists a constant $C_{\mathsf{T}} > 0$ such that the following *local* testing condition holds for any cube Q:

$$\int_Q \int_0^{\ell Q} |\theta_t \mathbb{1}_Q(x)|^2 \frac{\mathrm{d}t}{t} \, \mathrm{d}x \leq C_{\mathsf{T}} |Q|. \tag{T}$$

Example. For the case of a Littlewood–Paley square function in Example 1.2.4, where $\theta_t f = f * \psi_t$ and ψ is a mean zero Schwartz function rescaled as $\psi_t(x) = t^{-d}\psi(t^{-1}x)$, condition (T) is the cancellation condition $\int \psi = 0$.

The testing condition (T) is equivalent to the following testing condition: there exists $C > 0$ such that $\langle S(\mathbb{1}_Q)^2, \mathbb{1}_Q \rangle \leq C|Q|$ for all cubes Q. The reader can find the details in §3.1.7.

We have the following local $T1$ theorem for square functions.

Theorem 1.3.1 (Christ & Journé 1987, Auscher, Hofmann, Lacey, et al. 2002)**.** *Let S be a square function associated to a family of Littlewood–Paley operators $\{\theta_t\}_{t>0}$. If there exists $C > 0$ such that*

$$\langle S(\mathbb{1}_Q)^2, \mathbb{1}_Q \rangle \leq C|Q| \tag{T'}$$

holds for all cubes $Q \subseteq \mathbb{R}^d$, then $\|S\|_{L^2 \to L^2} < \infty$.

We present the main result from Chapter 3, also in [Bro20], where the theorem above has been recast to a *quadratic* sparse domination.

Theorem B (B. 2020). *Let S be a square function associated to $\{\theta_t\}_{t>0}$ satisfying conditions* (C1), (C2) *and* (T). *Then for any $f, g \in C_c^\infty$ there exists a sparse collection $\mathscr{S}$ such that*

$$|\langle (Sf)^2, g\rangle| \le C(C_1 + C_2 + C_{\mathsf{T}}) \sum_{Q\in\mathscr{S}} \left(\fint_Q |f| \right)^2 \left(\fint_Q |g| \right) |Q| \tag{1.3.1}$$

where $C = C(\alpha, d)$ is a positive constant independent of f and g.

The reader can compare (1.3.1) with the sparse form in Theorem 1.2.10.

The domination (1.3.1) implies the L^2-boundedness of S, which in turn implies the Carleson condition (T), see also [MM14, Remark 1.6]. So we have the following

Corollary B. *Let S be a vertical square function in* (1.2.2) *associated to a family of Littlewood–Paley operators $\{\theta_t\}_{t>0}$. Then S admits a sparse domination if and only if the Carleson condition* (T) *holds.*

Previous sharp weighted inequalities for square functions

Under condition (T) the square function S was known to be bounded on the weighted space $L^p(w)$ for $p \in (1, \infty)$, provided that w belongs to the Muckenhoupt class for which the quantity in (A_p) is finite.

For $p \in (1, \infty)$ and w in A_p, let $\alpha(p)$ be the best exponent in the inequality

$$\sup_{f\neq 0} \frac{\|Sf\|_{L^p(w)}}{\|f\|_{L^p(w)}} \le C(S, p)[w]_{A_p}^{\alpha(p)}. \tag{1.3.2}$$

When $p = 2$, Buckley [Buc93] showed the upper bound $\alpha(2) \le 3/2$. Later Wittwer improved it to $\alpha(2) = 1$ and showed that it is sharp for the dyadic and the continuous square functions [Wit02, Theorem 3.1–3.2]. The same result was obtained independently by Hukovic, Treil and Volberg using Bellman functions [HTV00, Theorem 0.1–0.4].

Lerner was the first to prove that $\alpha(p) = \max\{\frac{1}{2}, \frac{1}{p-1}\}$ cannot be improved [Ler06, Theorem 1.2] and to conjecture estimate (1.3.2) for Littlewood–Paley square functions. After improving the best known exponent for $p > 2$ [Ler08b, Corollary 1.3], Lerner proved the estimate

$$\|Sf\|_{L^3(w)} \le C[w]_{A_3}^{1/2}\|f\|_{L^3(w)} \tag{1.3.3}$$

for Littlewood–Paley square functions pointwise controlled by the intrinsic square function [Ler11, Theorem 1.1]. Lerner achieved this by applying the local mean oscillation formula to a dyadic variant of the Wilson intrinsic square function [Wil07]. The sharp estimate (1.3.2) for all $1 < p < \infty$ follows from (1.3.3) by the sharp extrapolation theorem [DGPP05], see also [Gra14, Theorem 7.5.3]. A proof of the sharp bound (1.3.2) for the dyadic square function using local mean oscillation can be found in [CMP12, Theorem 1.8].

Lerner's result relies on a pointwise control of the square function S, and it exploits the local behaviour of the Wilson intrinsic square function. Instead Theorem B implies the weighted estimate (1.3.3) by duality, and so the estimate (1.3.2) in the full range with optimal dependence on the A_p characteristic. As for the list after Theorem 1.2.10, other estimates follow from the sparse domination in Theorem B. In particular, see [LS12] and [HL18] for weak type estimates and [LL16; DLR16] for mixed A_p–A_∞ estimates.

Our different approach can dispense with the extra "locality" assumption used in [Ler11; Zor16]. As a consequence, we can allow for square functions with general kernel. A similar approach, where the input function is decomposed using wavelets, has been used in [DWW20]: the basis of wavelets used there allows to derive new $T(1)$ theorems on weighted Sobolev spaces, see in particular [DWW20, §8 and Theorem C].

The fact that sparse domination for $(Sf)^2$ gives better estimates than the one for Sf is true also for more general square functions.

1.4 Beyond classical square functions

In this section we present the main result of Chapter 4.

Classical operators in harmonic analysis come with an integral representation and a kernel. On the other hand, many operators coming from elliptic PDEs are "non-integral", in the sense that they do not possess such an explicit representation.

In contrast to the classical Calderón–Zygmund theory, which cannot be applied in this situation, sparse domination has proven to be more flexible and it has yielded many results in this context [BFP16; BCDH17; CDO18a; BD20b; BD20a]. In the case of operators associated to an elliptic operator L, the usual assumptions on the kernel are replaced by hypotheses on the action of the operator on the semigroup e^{-tL} generated by L.

Example 1.4.1. Prominent examples are operators attached to the elliptic operator $L = -\mathrm{div}(A\nabla)$, where A is bounded and elliptic with complex coefficients. For example, the Riesz transforms $\nabla L^{-1/2}$ or the constituent operators $\{\sqrt{t}\nabla e^{-tL}\}_{t>0}$ of the square function

$$G_L f = \left(\int_0^\infty |\sqrt{t}\nabla e^{-tL} f|^2 \frac{\mathrm{d}t}{t}\right)^{1/2}$$

might not possess integral kernels.

In contrast to the classical setting where L is the Laplacian, these operators are in general not bounded on $L^p(\mathbb{R}^d)$ for every $p \in (1, \infty)$. As proved in [Aus07], boundedness might occur only in a restricted range $(p_0, q_0) \subsetneq (1, \infty)$. This interval depends on the perturbation A in $L = -\mathrm{div}(A\nabla)$, see also [BK03] and [HM03]. Weighted estimates have been introduced in this setting by Auscher and Martell in the seminal series of papers [AM07a; AM07b; AM06; AM08].

Operators bounded on L^p only for a limited range of p can be bounded on the weighted space $L^r(w)$, for $r \in (p_0, q_0)$, only for a restricted class of weights. This is a consequence of the extrapolation theorem [Gra14, Theorem 7.5.3], as the finiteness of the quantity $\|T\|_{L^r(w)\to L^r(w)}$ for all $w \in A_r$ would imply the boundedness of T on unweighted L^p spaces for all $1 < p < \infty$.

For this reason, we consider a subclass of A_p weights which additionally satisfy a reverse Hölder property: there is $q > 1$ such that

$$[w]_{RH_q} := \sup_{Q \text{ cube}} \left(\fint_Q w^q \right)^{1/q} \left(\fint_Q w \right)^{-1} \tag{RH_q}$$

is finite. This subclass of A_p will be denoted by $A_p \cap RH_q$, as in [AM07a].

Aiming to a sparse bound, note that the pointwise sparse domination

$$|Tf(x)| \leq C \sum_{Q \in \mathscr{S}} \left(\fint_Q |f|^{p_0} \right)^{1/p_0} \mathbb{1}_Q(x)$$

would yield (weighted) estimates in the open range $p \in (p_0, \infty)$. A way to restrict further the range to (p_0, q_0) is to consider a sparse domination in form of the following kind

$$\left| \int_{\mathbb{R}^d} Tf \cdot g \, \mathrm{d}x \right| \leq C \sum_{Q \in \mathscr{S}} \left(\fint_Q |f|^{p_0} \right)^{1/p_0} \left(\fint_Q |g|^{q_0'} \right)^{1/q_0'} |Q|.$$

This has been obtained in [BFP16] for a large class of non-integral operators, including the Riesz transform and the square function above.

As we will see in Chapter 4, in the case of non-integral square functions a sparse domination for $(G_L f)^2$ yields better estimates. This is true for a large class of square functions, as shown in the following result.

Theorem C (Bailey, B., Reguera 2020)**.** *Let S be a vertical square function as defined in Chapter 4, which is bounded on L^p for $p \in (p_0, q_0)$, $p_0 < 2 < q_0$. For any f and g in $C_c^\infty(\mathbb{R}^d)$ there exists a sparse family $\mathscr{S}$ such that*

$$\left| \int_{\mathbb{R}^d} (Sf)^2 \cdot g \, \mathrm{d}x \right| \leq C \sum_{Q \in \mathscr{S}} \left(\fint_Q |f|^{p_0} \right)^{2/p_0} \left(\fint_Q |g|^{\left(\frac{q_0}{2}\right)'} \right)^{1/\left(\frac{q_0}{2}\right)'} |Q|.$$

where C is a positive constant independent of f and g.

The sparse domination in Theorem C implies the following weighted estimates.

Corollary C. *Let S be a vertical square function as defined in Chapter 4, which is bounded on L^p for $p \in (p_0, q_0)$, $p_0 < 2 < q_0$. For a weight $w \in A_{p/p_0} \cap RH_{(q_0/p)'}$ the square function S is bounded on $L^p(w)$ with*

$$\|S\|_{L^p(w) \to L^p(w)} \leq C \left([w]_{A_{p/p_0}} [w]_{RH_{(q_0/p)'}} \right)^{\max\left(\frac{1}{p-p_0}, \frac{q_0-2}{q_0-p}\right)}$$

where C is positive constant independent of the weight.

The power in the characteristic of the weights is sharp for the sparse form in Theorem C, see Proposition C.

It remains open to understand:

(a) if the sparse bound in Theorem C and the weighted estimates in Corollary C are sharp for *all* the square functions considered;

(b) what are the minimal assumptions on the operator S for the sparse bound.

Example 1.4.2. The domination from Theorem C yields weighted estimates for square functions associated with divergence forms and Laplace–Beltrami operators on Riemannian manifolds. These are presented in §4.3.

Theorem C and its corollary are presented in Chapter 4.

We move to another application of sparse domination in a completely different context.

1.5 Two weight theory for the Bergman projection

In this section we introduce the results from Chapter 5.

An active direction of research aims to better understand the projections onto spaces of holomorphic functions in terms of weighted estimates.

To fix some notation, let $\Omega \subset \mathbb{C}^d$ be a bounded domain (open, connected set) with smooth boundary, and let $\mathrm{d}\nu$ be the normalised Lebesgue measure on Ω. Let $\mathcal{H}(\Omega)$ be the space of holomorphic functions on Ω. The Bergman space is the subspace of holomorphic functions of $L^2(\Omega)$

$$A^2(\Omega) := \{f \in \mathcal{H}(\Omega) \cap L^2(\Omega)\}.$$

The L^2 inner product makes $A^2(\Omega)$ a Hilbert space. The evaluation at any point $z_0 \in \Omega$ is a continuous functional on $A^2(\Omega)$. This follows from the mean value theorem for holomorphic functions and an application of Hölder's inequality: for any ball $B(z_0, r) \subseteq \Omega$ we have

$$|f(z_0)| = \left| \fint_{B(z_0,r)} f(\zeta)\,\mathrm{d}\nu(\zeta) \right| \le \left(\fint_{B(z_0,r)} |f(\zeta)|^2\,\mathrm{d}\nu(\zeta) \right)^{1/2} \le \frac{c_{2d}}{r^d} \|f\|_{A^2(\Omega)}.$$

By the Riesz representation theorem, the evaluation at z_0 can be written as

$$f(z_0) = \int_\Omega K(z_0, \zeta) f(\zeta)\,\mathrm{d}\nu(\zeta) \tag{1.5.1}$$

where $K(z_0, \cdot)$ is a function in $A^2(\Omega)$ called Bergman kernel. The associated operator

$$Pf(z) := \int_\Omega K(z, \zeta) f(\zeta)\,\mathrm{d}\nu(\zeta)$$

is the identity on A^2, in view of (1.5.1). Moreover P is self-adjoint and idempotent, see [Kra01, §1.4]. It follows that P is a projection from $L^2(\Omega)$ to $A^2(\Omega)$, so we call it the Bergman projection. When Ω is the unit ball in $\mathbb{C}^d$, the Bergman kernel can be written explicitly and a sparse domination for P is available [PR13; APR17; RTW17]. The one-weight theory for P is then well understood. On the other hand, the *two weight* theory is still at its early stages and is becoming increasingly important in connection with the resolution of the Sarason conjecture, see §1.5.1.

1.5.1 Connection with Operator Theory

Given f in the Bergman space $A^2 \subset L^2$, one can consider the Toeplitz operator $T_f h := (P \circ m_f)(h) = P(f \cdot h)$, where m_f denotes the pointwise multiplication by f.

In [ACS78] Sarason showed sufficient conditions on f, g for the composition of Toeplitz operators to be bounded. The question about necessity of these conditions was left open. We state the conjecture on the Bergman space $A^2(\mathbb{D})$.

Conjecture 1.5.1 (Sarason)**.** *Given $f, g \in A^2(\mathbb{D})$, the operator $T_f T_{\bar{g}}$ is bounded on $A^2(\mathbb{D})$ if and only if*

$$\|B(|f|^2)B(|g|^2)\|_{L^\infty(\mathbb{D})} < \infty,$$

where the operator B is the Berezin transform

$$Bf(z) := \int_{\mathbb{D}} f(\zeta) \frac{(1-|z|^2)^2}{|1-\bar{\zeta}z|^4} \, \mathrm{d}\nu(\zeta).$$

Cruz-Uribe [Cru94] showed that the Sarason conjecture on Bergman spaces is related to the two weight boundedness of the Bergman projection. Indeed, we have the following abstract diagram, where f and g are general functions and $A^p(\Omega) := \{f \in \mathcal{H}(\Omega) \cap L^p(\Omega)\}$..

$$\begin{tikzcd}
A^p(\Omega) \arrow[r, "T_f \circ T_{\bar{g}}"] \arrow[d, "m_{\bar{g}}"'] & A^p(\Omega) \\
L^p(\Omega, |g|^{-p}) \arrow[r, "P"] & A^p(\Omega, |f|^p) \arrow[u, "T_f"']
\end{tikzcd}$$

Figure 1.1: Relation between composition of Toeplitz operators and weighted estimates for the Bergman projection P.

The original diagram in [Cru94] considered the Hardy space $H^2(\mathbb{T})$ and the corresponding Riesz projection. The case of the Bergman space $A^2(\mathbb{D})$ has been studied in [APR17].

Remark 1.5.2. When $f \in A^p(\Omega)$, the Toeplitz operator on the right of the diagram can be replaced by m_f, since for a holomorphic function h we have $T_f(h) = P(f \cdot h) = f \cdot h$.

1.5.2 Bergman projection on the unit ball

In this thesis we consider the Bergman projection on the unit ball $\mathbb{B}^d \subset \mathbb{C}^d$ given by

$$Pf(z) := \int_{\mathbb{B}^d} \frac{f(\zeta)}{(1 - z\bar{\zeta})^{d+1}} \, \mathrm{d}\nu(\zeta).$$

In Chapter 5 we address the following question

> What are the sufficient conditions on two weight u, ω for the boundedness of the Bergman projection $P : L^2(u) \to A^2(\omega)$?

Remark 1.5.3. For two weights u and ω, we have

$$\|P\|_{L^2(u)\to L^2(\omega)} \eqsim \|P(\sigma\, \cdot)\|_{L^2(\sigma)\to L^2(\omega)} \tag{1.5.2}$$

where $\sigma = u^{-1}$. This equivalent formulation is due to Sawyer and it holds for general domains other than the unit ball, see Appendix C.

For the complex ball, the weights for which any of the two quantities in (1.5.2) is finite are the Békollé–Bonamí weights. These weights satisfy a (A_p) condition where the role of cubes is played by Carleson tents.

Definition 1.5.4 (Carleson tent on the unit ball)**.** Given a point $z \in \mathbb{B}^d \setminus \{0\}$, consider the following set

$$T_z := \left\{ \zeta \in \mathbb{B}^d \; : \; |1 - \langle \zeta, \tfrac{z}{|z|} \rangle| \leq 1 - |z| \right\}.$$

The set T_z is the intersection of $\mathbb{B}^d$ with the ball centred at $z/|z|$ with radius $1-|z|$, whose boundary contains the point z. For $z = 0$, set $T_0 = \mathbb{B}^d$.

Definition 1.5.5 (Békollé–Bonamí weights)**.** Given two weights w, σ on $\mathbb{B}^d$, we define their joint $\mathcal{B}_2$ characteristic:

$$[w,\sigma]_{\mathcal{B}_2} := \sup_{z\in\mathbb{B}^d} \langle w\rangle_{T_z} \langle \sigma\rangle_{T_z}$$

where $\langle w\rangle_{T_z} := |T_z|^{-1}\int_{T_z} w$. For general $1<p<\infty$, we define the quantity

$$[w,\sigma]_{\mathcal{B}_p} := \sup_{z\in\mathbb{B}^d} \langle w\rangle_{T_z} \langle \sigma\rangle_{T_z}^{p-1}.$$

When $\sigma = w^{1-p'}$ is the dual weight of w, the quantity $[w]_{\mathcal{B}_p} := [w, w^{1-p'}]_{\mathcal{B}_p}$ is the $\mathcal{B}_p$ characteristic of w. We say that w is a Békollé–Bonamí weight if $[w]_{\mathcal{B}_p}$ is finite and we write $w\in\mathcal{B}_p$.

In order to state the main results of this section, we introduce bump conditions.

1.5.3 Bumps conditions

Since the '70s it is known that the joint Muckenhoupt condition for two weight w,σ

$$[w,\sigma]_{A_p} := \sup_{Q} \langle w^{1/p}\rangle_{p,Q} \langle \sigma^{-1/p}\rangle_{p',Q} \; < \; +\infty \tag{1.5.3}$$

is necessary but not sufficient for the boundedness of singular integral operators T from $L^p(\sigma)$ to $L^p(w)$. Aiming to find suitable sufficient conditions, researchers have replaced the averages in (1.5.3) with smaller averages, and so assuming a stronger condition on the weights w,σ.

A way to generalise the L^p averages is to consider Orlicz averages, which we now introduce.

Orlicz bumps

Definition 1.5.6 (Young function)**.** Let $\Phi\colon [0,\infty) \to [0,\infty)$ be a continuous, convex, strictly increasing function such that

$$\Phi(0) = 0 \quad \text{and} \quad \lim_{t\to+\infty} \frac{\Phi(t)}{t} = +\infty.$$

Given a set Q and a Young function Φ we denote by $\langle f\rangle_{\Phi,Q}$ the Orlicz average defined via the Luxembourg norm

$$\langle f\rangle_{\Phi,Q} \coloneqq \inf\{\lambda > 0 \,:\, \langle \Phi(f/\lambda)\rangle_Q \le 1\}.$$

In [Pér95, Theorem 1.7] Pérez characterised the Young functions for which the associated maximal function

$$M_\Phi f \coloneqq \sup_Q \langle |f|\rangle_{\Phi,Q} \mathbb{1}_Q$$

is bounded on L^p.

Theorem 1.5.7 (Pérez 1995)**.** *Given a Young function Φ, the associated maximal function M_Φ maps L^p to L^p, for $1 < p < \infty$, if and only if*

$$\int_1^\infty \frac{\Phi(t)}{t^p}\frac{\mathrm{d}t}{t} < +\infty. \tag{B_p}$$

Note that the operator M_Φ is also bounded on L^∞ [And15, Lemma 3.2]. We say that a Young function Φ belongs to B_p if the condition (B_p) holds.

Using the dyadic structure on the ball introduced in §5.1 we can define the Orlicz averages of two weights.

Definition 1.5.8. Given two weight w, σ and two Young functions $\Phi, \Psi \in B_2$, the joint Orlicz bumps condition reads

$$[w,\sigma]_{\Phi,\Psi} \coloneqq \sup_{\widehat{K}\in\mathcal{T}} \frac{\langle w\rangle_{\widehat{K}}}{\langle w^{1/2}\rangle_{\Phi,\widehat{K}}} \frac{\langle \sigma\rangle_{\widehat{K}}}{\langle \sigma^{1/2}\rangle_{\Psi,\widehat{K}}}$$

where $\langle\,\cdot\,\rangle_{\Phi,\widehat{K}}$ denotes the Orlicz average on the dyadic tent $\widehat{K}$ with respect to Φ.

The first result of Chapter 5 is a sufficient bump condition in terms of Orlicz averages. This will be deduced from the sparse operator dominating P. In particular, once a dyadic structure on $\mathbb{B}^d$ has been constructed in §5.1, we have the following.

Theorem D. *Let σ, ω be two weights*[1] *on $\mathbb{B}^d$ and let Φ, Ψ be two Young functions such that the associated maximal function is bounded on L^2. Then the Bergman projection P on $L^2(\mathbb{B}^d)$ satisfies the following bound*

$$\|P(\sigma\,\cdot)\|_{L^2(\sigma)\to L^2(\omega)} \leq C\,[\sigma,\omega]_{\Phi,\Psi}$$

where C is a positive constant independent of σ, ω.

This result is deduced by combining the domination in [RTW17] with the known estimates for sparse forms [Li17]. Nevertheless, to the best of our knowledge, these estimates have not appeared in the context of Bergman spaces.

In Chapter 5 we derive the bump conditions in the theorem via testing conditions on sparse operators. We will also obtain the following mixed estimates in terms of the B_∞ characteristic.

Definition 1.5.9. We consider the quantity

$$[\sigma]_{B_\infty} \coloneqq \sup_{\widehat{K}\in\mathcal{T}} \frac{1}{\sigma(\widehat{K})}\int_{\widehat{K}} M(\sigma \mathbb{1}_{\widehat{K}})$$

where M is the maximal operator

$$Mf(z) \coloneqq \sup_{\widehat{K}\in\mathcal{T}} \langle |f|\rangle_{\widehat{K}} \mathbb{1}_{\widehat{K}}(z)$$

over the collection of sets $\mathcal{T}$ introduced in §5.1. We say that a weigth σ belongs to B_∞ if the quantity $[\sigma]_{B_\infty}$ is finite.

[1]Weight will always mean: positive, measurable function.

Theorem E. *Let σ, ω be two weights on $\mathbb{B}^d$ in the class B_∞ such that their joint $\mathcal{B}_2$ characteristic $[\omega, \sigma]_{\mathcal{B}_2}$ is finite. The Bergman projection P on $L^2(\mathbb{B}^d)$ satisfies the following bound*

$$\|P(\sigma\,\cdot)\|_{L^2(\sigma)\to L^2(\omega)} \leq C\,[\omega,\sigma]_{\mathcal{B}_2}^{1/2}\big([\sigma]_{B_\infty}^{1/2} + [\omega]_{B_\infty}^{1/2}\big)$$

where C is a positive constant independent of σ and ω.

CHAPTER 2

SHARP k-PLANE INEQUALITIES

The more I do this, the more I think Analysis is a mistake.

C. B.

2.1 Strichartz estimates for Schrödinger equation

We study a family of Strichartz estimates for the solution of the free Schrödinger equation

$$\begin{cases} i\partial_t u(x,t) - \Delta u(x,t) = 0 \\ u(x,0) = f(x) \in L^2(\mathbb{R}^d). \end{cases}$$

In quantum mechanics, the quantity $|u(x,t)|^2$ represents the probability to find a particle at a point x at time t. Bounds for this probability function in time and space are the content of a family of estimates introduced by Strichartz [Str77]:

$$\|e^{-it\Delta} f\|_{L^q_t(\mathbb{R})L^p_x(\mathbb{R}^d)} \leq C\|f\|_{L^2(\mathbb{R}^d)} \tag{2.1.1}$$

where the constant C depends on exponents and dimension, which are related by the following scaling condition:

$$\frac{2}{q} + \frac{d}{p} = \frac{d}{2} \quad \text{with} \quad \begin{cases} p \in [2, \infty] & \text{if } d = 1 \\ p \in [2, \infty) & \text{if } d = 2 \\ p \in \left[2, \frac{2d}{d-2}\right] & \text{if } d \geq 3 \end{cases}$$

For a given dimension d, a pair (q, p) satisfying the above relation is called *admissible*.

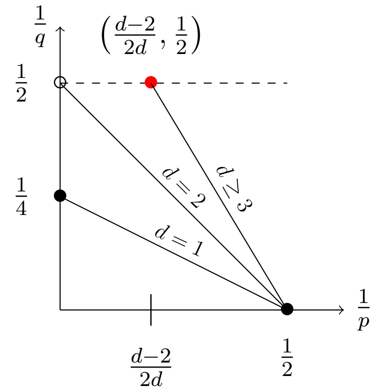

Figure 2.1: Admissible exponents.

We are interested in the sharp form of these inequalities. We will focus on the symmetric case $q = p$. Moreover, we consider the inequality for the modulus of the wave function u:

$$\||u|^2\|_{L^q(\mathbb{R}\times\mathbb{R}^d)} \leq C\|f\|^2_{L^2(\mathbb{R}^d)}, \quad \text{with} \quad q = 1 + \frac{2}{d}. \tag{2.1.2}$$

When the exponent in the L^q-norm is an integer, we can write

$$\||u|^2\|^q_{L^q(\mathbb{R}^{1+d})} = \iint \underbrace{|u|^2 \ldots |u|^2}_{q \text{ times}} \, dxdt$$

and study the inequalities in sharp form.

Characterisation of extremisers in (2.1.2) and the value of the best constant have been computed only when $d \in \{1, 2\}$, exploiting the above representation for integer exponents in the norm.

In [BBF+18], Bennett, Bez, Flock, Gutierrez, and Iliopoulou obtained the inequality

$$\|X(|u|^2)\|_{L^3_{t,\ell}} \leq \left(\frac{\pi}{2}\right)^{\frac{1}{3}} \|f\|^2_{L^2(\mathbb{R}^2)}, \tag{2.1.3}$$

where X is the (spatial) X-ray transform on $\mathbb{R}^2$. The constant in (2.1.3) is sharp, and extremisers are Gaussians. The authors also established similar results in higher dimensions for general k-plane transform in $\mathbb{R}^d$ by considering the L^{d+1} norms and a weight on the right hand side. We consider instead a family of estimates which arise by applying the k-plane transform to the function $|u|^2$. By restricting ourselves on integer exponents, we can study the sharp form of these inequalities also in higher dimensions.

2.2 Sharp k-plane estimates for Schrödinger equation

We study sharp inequalities obtained in the same spirit of (2.1.3), namely composing an L^p-bound of a k-plane transform with a Strichartz estimate (2.1.2).

Let us introduce some notations. With k we indicate an integer between 1 and $d-1$. Given an affine k-dimensional plane U in $\mathbb{R}^d$ this is identified by $(\omega_1, \dots, \omega_k)$, $\omega_j \in \mathbb{S}^{d-1}$ and a vector $b \in \mathbb{R}^d$, such that

$$U = \{x_1\omega_1 + \cdots + x_k\omega_k + b \,:\, (x_1, \dots, x_k) \in \mathbb{R}^k\} = U_0 + b$$

where U_0 is a k-dimensional subspaces in $\mathbb{R}^d$. Let $\mathcal{A}_{k,d}$ be the bundle of affine k-dimensional subspace of $\mathbb{R}^d$. There is a natural projection from $\mathcal{A}_{k,d}$ onto the Grassmannian $\mathrm{Gr}_k(\mathbb{R}^d)$ of all k-dimensional subspaces in $\mathbb{R}^d$: the one that maps an affine k-plane U to its translated copy U_0 passing through the origin.

$$\begin{aligned} p : \mathcal{A}_{k,d} &\to \mathrm{Gr}_k(\mathbb{R}^d) \\ U &\mapsto U_0 \end{aligned}$$

The projection p is a fibration that turns $\mathcal{A}_{k,d}$ into a vector bundle over $\mathrm{Gr}_k(\mathbb{R}^d)$.

The manifold $\mathcal{A}_{k,d}$ can be endowed with a measure μ given by the product of the uniform measure μ_{Gr} on the Grassmannian times the Lebesgue measure on the orthogonal $(d-k)$-plane.

The k-plane transform $T_{k,d}$ maps measurable functions on $\mathbb{R}^d$ to functions on $\mathcal{A}_{k,d}$:

given an affine k-plane $U \in \mathcal{A}_{k,d}$, the value of $T_{k,d}(g)$ at U is the average of the function g over U, namely

$$T_{k,d}(g)(U) = \int_U g = \int_{U_0} g(x_1\omega_1 + \cdots + x_k\omega_k + b)\, d\mathcal{L}^k(x)$$

where $\mathcal{L}^k$ is the k-dimensional Lebesgue measure. The k-plane transform $T_{k,d}(g)$ is then a function on $\mathcal{A}_{k,d}$.

Composing the L^p bounds for the k-plane transform with the Strichartz estimates in (2.1.1) gives raise to a family of inequalities:

$$\|T_k(|u(t,\cdot)|^2)\|_{L^r_t(\mathbb{R};L^s(\mathcal{A}_{k,d}))} \leq C\,\|f\|^2_{L^2(\mathbb{R}^d)} \tag{2.2.1}$$

for $1 \leq r, s \leq \infty$ and $k \in \{1, \dots, d-1\}$ satisfying the following scaling conditions

$$\frac{1}{r} = \frac{d-k}{2}\left(1 - \frac{1}{s}\right) \tag{2.2.2}$$

with

$$1 \leq s \leq d+1, \quad r \geq 1.$$

The optimal constant $C = C(r,s,d,k)$ depends on k and d, as well as on the exponents r, s.

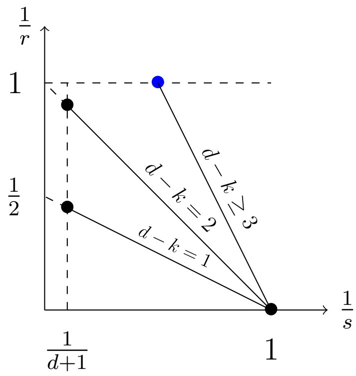

Figure 2.2: Exponents satisfying (2.2.2).

Aiming to study the sharp form of the inequality in (2.2.1), we focus on the symmetric case, when $r = s$. We have the following inequalities:

$$\|T_k(|u(t,\cdot)|^2)\|_{L^q(\mathbb{R}\times\mathcal{A}_{k,d})} \leq C\,\|f\|^2_{L^2(\mathbb{R}^d)}, \quad \text{with} \quad q = 1 + \frac{2}{d-k}. \tag{2.2.3}$$

One should compare this inequality with the one in (2.1.2).

Remark 2.2.1. When $k = 0$ the exponent q in (2.2.3) coincides with the Strichartz exponent in (2.1.2). Indeed, the classical Strichartz estimates in (2.1.2) can be seen as a special case of (2.2.3), T_0 being just evaluation at a point.

Remark 2.2.2. For $k = d-1$ the L^3 norm on the left hand side of (2.2.3) is

$$\|T_{d-1}(|u(t,\cdot)|^2)\|^3_{L^3(\mathbb{R}\times\mathcal{A}_{d-1,d})} = \frac{d-2}{4\pi}\int_{\mathbb{R}}\int_{(\mathbb{R}^d)^3}\frac{|u(t,x_0)|^2|u(t,x_1)|^2|u(t,x_2)|^2}{\Delta_2(x_0,x_1,x_2)}dxdt \quad (2.2.4)$$

where $\Delta_2(x_0,x_1,x_2)$ is the area of the triangle in $\mathbb{R}^d$ with vertices (x_0,x_1,x_2). It can be written as $\det(x^Tx)$ where x denotes the $d\times 3$ matrix with columns x_0,x_1,x_2. The expression in (2.2.4) is a special case of Drury's identity in [Rub18, eq (3.8)].

When $d = 2$ the Radon transform T_{d-1} is an X-ray transform and the inequality corresponds to (2.1.3), which is the only case known in sharp form.

Applying a high dimensional k-plane transform allows us to regain integer exponents in any dimension. When $k = d-2$ we characterise extremisers and compute the optimal constant. Our approach uses Drury's identity for the k-plane transform from [Rub18].

2.3 Main result

We consider inequality (2.2.3) in the case $q = d-k = 2$. We have the following theorem.

Theorem A. *Let $d \geq 3$. The following estimate holds:*

$$\|T_{d-2}(|u(t,\cdot)|^2)\|_{L^2(\mathbb{R}\times\mathcal{A}_{d-2,d})} \leq C_d\|f\|^2_{L^2(\mathbb{R}^d)} \quad (2.3.1)$$

where the sharp constant is

$$C_d = \left(\frac{(d-2)\pi^{d/2}}{\Gamma(d/2)}\right)^{1/2}$$

The equality in (2.3.1) is achieved if and only if $\widehat{f}(\xi) = \exp(a|\xi|^2 + b\cdot\xi + c)$ with $\Re(a) < 0$, $b\in\mathbb{C}^d$ and $c\in\mathbb{C}$.

We give a proof of the theorem.

2.3.1 Proof of Theorem A

After applying the Drury's identity [Rub18, eq (3.8)] with $k = d-2$, we have

$$\|T_{d-2}(|u(t,\cdot)|^2)\|^2_{L^2(\mathbb{R}\times\mathcal{A}_{d-2,d})} = \frac{d-2}{2\pi}\int_{\mathbb{R}}\int_{(\mathbb{R}^d)^2}\frac{|u(t,x)|^2|u(t,y)|^2}{|x-y|^2}dxdydt. \tag{2.3.2}$$

Notation comparison

To help the reader comparing the identities (2.3.2) and [Rub18, eq (3.8)], we provide some information for the notation used by Rubin in [Rub18]. In the article, $\mathfrak{M}_{d,k}$ denotes the space of real matrices with d rows and k columns which can be identified with $\mathbb{R}^{d\times k}$. The identity (2.3.2) above corresponds to [Rub18, eq (3.8)] with $k = d-2$ and $q = 1$. Given $x = (x_1, \dots, x_k) \in (\mathbb{R}^d)^k$, the quantity $\Delta_k(x)$ appearing in [Rub18] denotes the volume of the convex hull of $\{0, x_1, \dots, x_k\}$. So $\Delta_1(x_0, x_1)$ is the length of the vector $x_1 - x_0$. The identity is justified since the integrand $|u(\cdot,x)|^2|u(\cdot,y)|^2|x-y|^{-2}$ belongs to $L^1(\mathbb{R}^{2d}) = L^1(\mathfrak{M}_{d,2})$ as required, by (2.3.1).

We write $|u|^2$ as $u\bar{u}$ and expand using the expression for the solution to the free Schrödinger equation:

$$u(x,t) = (2\pi)^{-d}\int_{\mathbb{R}^d} e^{ix\cdot\xi}e^{it|\xi|^2}\widehat{f}(\xi)d\xi.$$

After one application of Fubini, and computing the inverse Fourier transform of $|x|^{-2}$, we obtain

$$\int_{\mathbb{R}}\int_{(\mathbb{R}^d)^2}\frac{|u(t,x)|^2|u(t,y)|^2}{|x-y|^2}\,dxdydt = (2\pi)^{1-2d}\int_{\mathbb{R}^{2d}_\xi}\int_{\mathbb{R}^{2d}_\eta}\widehat{F}(\xi)\overline{\widehat{F}(\eta)}\,d\Sigma_\xi(\eta)d\xi \tag{2.3.3}$$

where $\xi = (\xi_0, \xi_1) \in \mathbb{R}^{2d}$, $F = f \otimes f$ and the measure $d\Sigma_\xi(\eta)$ is given by

$$d\Sigma_\xi(\eta) = \frac{\boldsymbol{\delta}\Big(|\xi|^2 - |\eta|^2\Big)\,\boldsymbol{\delta}\Big((\xi - \eta)\cdot(1,1)\Big)}{|\xi_0 - \eta_0|^{d-2}}d\eta. \tag{2.3.4}$$

Remark 2.3.1. The quantity in the left hand side of (2.3.3) is a real number. Also, the product $F(\xi)\overline{F}(\eta)$ can be written as

$$F(\xi)\overline{F}(\eta) = \frac{1}{2}\left(F\overline{F} + \overline{F}F\right) = \frac{1}{2}\left(|F(\xi)|^2 + |F(\eta)|^2 - |F(\xi) - F(\eta)|^2\right). \tag{2.3.5}$$

Note that the measure $d\Sigma_\xi(\eta)$ is also real valued, and it is symmetric in ξ and η, in the sense that $d\Sigma_\xi(\eta)d\xi = d\Sigma_\eta(\xi)d\eta$.

Using the symmetry of the measure and the expression (2.3.5), the integral in (2.3.3) can be written as

$$\int_{\mathbb{R}^{2d}_\xi}\int_{\mathbb{R}^{2d}_\eta} \widehat{F}(\xi)\overline{\widehat{F}(\eta)}\, d\Sigma_\xi(\eta)d\xi = \int |\widehat{F}(\eta)|^2 \int d\Sigma_\eta(\xi)d\xi - \frac{1}{2}\iint |\widehat{F}(\xi) - \widehat{F}(\eta)|^2\, d\Sigma_\xi(\eta)d\xi.$$

If the measure $d\Sigma_\xi(\eta)$ is finite almost everywhere, so has to be each of the two terms. In this case the first term on the right hand side equals the right and side of (2.3.1): the inequality is maximised when the difference $|\widehat{F}(\xi) - \widehat{F}(\eta)|^2$ vanishes. The functional equation $\widehat{F}(\xi) = \widehat{F}(\eta)$ is used to characterise extremisers.

In the follows lemma, we show that this is the case: the measure $d\Sigma_\xi(\eta)$ is finite.

Lemma 2.3.2. *Let $\xi, \eta \in \mathbb{R}^{2d}$. For any $d \geq 2$, the measure $d\Sigma_\eta(\xi)$ defined in (2.3.4) is independent of η. Moreover its total mass is finite and equals*

$$\int_{\mathbb{R}^{2d}} d\Sigma_\eta(\xi)d\xi = \frac{\pi^{d/2}}{\Gamma(d/2)} = \frac{1}{2}|\mathbb{S}^{d-1}|.$$

Proof. We make the change of variables: $\xi - \eta \mapsto (v, w)$, so that $v = \xi_0 - \eta_0$ and $w = \xi_1 - \eta_1$.

Then

$$\int_{\mathbb{R}^{2d}} d\Sigma_\eta(\xi)\, d\xi = \iint_{(\mathbb{R}^d)^2} \frac{\boldsymbol{\delta}\left(\left|\begin{pmatrix} v \\ w \end{pmatrix} + \eta\right|^2 - |\eta|^2\right) \boldsymbol{\delta}\Big(v + w\Big)}{|v|^{d-2}} dv dw.$$

On the support of the second Dirac delta we have $v = -w$. We complete the square in the argument of the first Dirac delta

$$\begin{aligned} \left|\begin{pmatrix} v \\ w \end{pmatrix} + \eta\right|^2 - |\eta|^2 &= |v|^2 + |w|^2 + 2\langle \begin{pmatrix} v \\ -v \end{pmatrix}, \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} \rangle \\ &= 2\left(|v|^2 + v \cdot (\eta_1 - \eta_2)\right) = 2\left(|v - \zeta|^2 - |\zeta|^2\right) \end{aligned}$$

where we use the new variable $\zeta = \frac{1}{2}(\eta_2 - \eta_1)$. Integrating the other Dirac delta we obtain:

$$\iint_{(\mathbb{R}^d)^2} \frac{\boldsymbol{\delta}\left(\left|\begin{pmatrix} v \\ w \end{pmatrix} + \eta\right|^2 - |\eta|^2\right) \boldsymbol{\delta}\Big(v + w\Big)}{|v|^{d-2}} dv dw = \int_{\mathbb{R}^d} \frac{\boldsymbol{\delta}\Big(|v - \zeta|^2 - |\zeta|^2\Big)}{|v|^{d-2}} dv.$$

We write $v = r\omega + \zeta$, for $\omega \in \mathbb{S}^{d-1}$ and $r \in \mathbb{R}_+$, so that $|v - \zeta|^2 = r^2$. Then in spherical coordinates, after using the variable $s = r^2$, the integral equals

$$\begin{aligned} \int_{\mathbb{S}^{d-1}} \int_0^\infty \frac{\boldsymbol{\delta}\Big(r^2 - |\zeta|^2\Big)}{|r\omega + \zeta|^{d-2}} r^{d-1} dr\, d\sigma(\omega) &= \frac{1}{2} \int_{\mathbb{S}^{d-1}} \int_0^\infty \frac{\boldsymbol{\delta}\Big(s - |\zeta|^2\Big)}{|\omega + \zeta/\sqrt{s}|^{d-2}}\, ds\, d\sigma(\omega) \\ &= \frac{1}{2} \int_{\mathbb{S}^{d-1}} \frac{d\sigma(\omega)}{|\omega + \omega'|^{d-2}} \end{aligned} \tag{2.3.6}$$

where $\omega' := \zeta/|\zeta|$. The denominator is a function of ω which is rotations invariant. We can choose a rotation ρ such that $\rho(\omega') = -e_d$. We have

$$\int_{\mathbb{S}^{d-1}} \frac{1}{|\omega + \omega'|^{d-2}}\, d\sigma(\omega) = \int_{\mathbb{S}^{d-1}} \frac{1}{|\omega + \rho(\omega')|^{d-2}}\, d\sigma(\omega) = \int_{\mathbb{S}^{d-1}} \frac{1}{|\omega - e_d|^{d-2}}\, d\sigma(\omega).$$

This integral is finite, since in a chart centered at the singularity e_d the integrand equals

$|x|^{-(d-2)}$, which is in $L^1_{\rm loc}(\mathbb{B}^{d-1})$.

In order to compute the exact value of the integral, consider spherical coordinates $(\theta_1, \ldots, \theta_{d-1})$ on $\mathbb{S}^{d-1}$. These coordinates can be chosen such that $\omega_d = \langle \omega, e_d \rangle = \cos(\theta_{d-1})$. After writing $|\omega - e_d|^{d-2}$ as $(2 - 2\langle \omega, e_d \rangle)^{\frac{d-2}{2}}$, the integral in (2.3.6) becomes

$$
\begin{aligned}
&\frac{1}{2}\int_{\mathbb{S}^{d-1}} \frac{d\sigma(\omega)}{|\omega - e_d|^{d-2}} = \\
&\quad = \frac{1}{2}\int_0^{2\pi}\int_{[0,\pi]^{d-2}} \frac{\sin^{d-2}(\theta_{d-1})}{(2-2\cos\theta_{d-1})^{\frac{d-2}{2}}} \prod_{k=1}^{d-2} (\sin(\theta_{d-1-k}))^{d-2-k}\, d\theta_{d-1-k}\, d\theta_{d-1} \\
&\qquad\qquad = \frac{1}{2}\left|\mathbb{S}^{d-2}\right| \int_0^{\pi} \frac{(\sin\theta)^{d-2}}{(2-2\cos\theta)^{\frac{d-2}{2}}} d\theta = \frac{\pi^{\frac{d-1}{2}}}{2^{\frac{d-2}{2}}\Gamma\left(\frac{d-1}{2}\right)} \int_0^{\pi} \frac{(\sin\theta)^{d-2}}{(1-\cos\theta)^{\frac{d-2}{2}}} d\theta
\end{aligned}
$$

where we have renamed $\theta := \theta_{d-1}$. The last integral can be computed exactly via the expression

$$
\int_0^{\pi} \left(\frac{\sin^2(\theta)}{1-\cos\theta}\right)^{\frac{n}{2}} d\theta = \int_0^{\pi} (1+\cos\theta)^{\frac{n}{2}} d\theta = \frac{2^{\frac{n}{2}}\sqrt{\pi}\,\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2}+1)}.
$$

for any $n \in \mathbb{N}$. The case $n = d-2$ gives the desired result. $\square$

Now we turn to the characterisation of extremisers:

Theorem 2.3.3. *Extremisers for* (2.3.1) *are Gaussians.*

Proof. Following the approach in [BBF+18], we have equality in (2.3.1) if and only if $\widehat{F}(\xi) = \widehat{F}(\eta)$ for almost every ξ and η in the support of the measure $d\Sigma_\xi(\eta)d\xi$. The functional equation

$$
\widehat{f}(\xi_1)\widehat{f}(\xi_2) = \widehat{f}(\eta_1)\widehat{f}(\eta_2) \tag{2.3.7}
$$

is known as the Maxwell-Boltzmann functional equation, and it has been previously studied in [BBJP17; BBF+18]. The solution of (2.3.7), when $\widehat{f}$ is locally integrable, is given by $\widehat{f}(\xi) = \exp(a|\xi|^2 + b\cdot\xi + c)$ with $\Re(a) < 0$, $b \in \mathbb{C}^d$ and $c \in \mathbb{C}$.

We need to check that $\widehat{f}$ is locally integrable, and this is the case since $\widehat{f} \in L^2$. See also [BBJP17, Remark p. 471] and [BBF+18, Proof of Theorem 2.1]. $\square$

CHAPTER 3

A SPARSE QUADRATIC $T(1)$ THEOREM

Sometimes science is more art than science, a lot of people don't get that.

R. & M.

This chapter is based on the paper

G. Brocchi. A sparse quadratic $T1$ theorem, *New York Journal of Mathematics*, **26** (2020), 1232–1272.

The only addition is Section 3.1.7 on the equivalence of the testing conditions:

1. There exists a constant $C_{\mathsf{T}} > 0$ such that

$$\int_Q \int_0^{\ell Q} |\theta_t \mathbb{1}_Q(x)|^2 \frac{\mathrm{d}t}{t} \,\mathrm{d}x \leq C_{\mathsf{T}} |Q|. \tag{T}$$

holds uniformly for any cube Q.

2. There exists a constant $C > 0$ such that

$$\langle S(\mathbb{1}_Q)^2, \mathbb{1}_Q \rangle \leq C|Q| \tag{T'}$$

holds uniformly for any cube Q.

Remark 3.0.1. It is possible to relax the L^2 testing condition (T) to the following weak L^0 condition: there is a non-increasing function $\varphi\colon (0,\infty) \to [0,1)$ vanishing at infinity such that for all cubes Q and $\lambda > 0$ it holds that

$$\frac{|\{x \in Q : S(\mathbb{1}_Q)(x) > \lambda\}|}{|Q|} \leq \varphi(\lambda). \tag{T_0}$$

See [MMV19, Theorem 1.6] and the remarks after it, where the measure ν_Q appearing there can be replaced by $\mathbb{1}_Q \, \mathrm{d}x$. In particular, conditions (T_0) and (T) are equivalent, as each one is a necessary and sufficient condition for the L^2 boundedness of S.

We prove the following theorem.

Theorem B. *Let S be the vertical square function defined in (1.2.2) satisfying conditions (C1), (C2) and any of the equivalent testing conditions: (T), (T'), (T_0). Then for any pair of compactly supported functions $f, g \in C_c^\infty(\mathbb{R}^d)$ there exists a sparse collection $\mathscr{S}$ such that*

$$\Big| \int_{\mathbb{R}^d} (Sf)^2 \cdot g \,\mathrm{d}x \Big| \leq C(C_1 + C_2 + C_{\mathsf{T}}) \sum_{Q \in \mathscr{S}} \left(\frac{1}{|Q|} \int_Q |f| \right)^2 \left(\frac{1}{|Q|} \int_Q |g| \right) |Q|$$

where $C = C(\alpha, d)$ is a positive constant independent of f and g.

Guide to this chapter

In §3.1 we introduce shifted random dyadic grids and the associated Haar basis. Furthermore we use the classical reduction to good cubes. In §3.2 we start the proof of Theorem B by decomposing the operator into off-diagonal and diagonal parts. These are split further each one into two terms

$$\langle (Sf)^2, g \rangle \lesssim \underbrace{(\mathrm{I}) + (\mathrm{II})}_{\text{off-diagonal}} + \underbrace{(\mathrm{III}_a) + (\mathrm{III}_b)}_{\text{diagonal}}.$$

The off-diagonal part is bounded by a dyadic form using standard techniques in §3.3 and off-diagonal estimates in §3.4. The dyadic form is dominated by a sparse form in §3.7.

Terms (III_a) and (III_b) come from a Calderón–Zygmund decomposition $g = a + b$, where a is the average part and b is the bad part of g.

In §3.5 we introduce the stopping cubes used to control the diagonal part. We reduce (III_a) to a telescopic sum on stopping cubes plus off-diagonal terms. We remark that the stopping family depends only on the functions f and g. Furthermore, the testing condition (T) is used only in this section and only once.

In §3.6 we deal with (III_b). We exploit the zero average property of b together with the regularity of the kernel ($\mathsf{C2}$) to restore a setting in which off-diagonal estimates can be applied as in the previous sections, see §3.6.1.

In §3.8 we collect some of the proofs postponed to ease the reading. In Appendix B we recall some known results about conditional expectations and Haar projections which are used in §3.6.

Notation

For two positive quantities X and Y the notation $X \lesssim Y$ means that there exists a constant $C > 0$ such that $X \leq CY$. The dependence of C on other parameters will be indicated by subscripts $X \lesssim_{d,r,\alpha} Y$ when appropriate.

Given a cube Q in $\mathbb{R}^d$, the quantities ∂Q, ℓQ and $|Q|$ denote, respectively, boundary, size length, and the Lebesgue measure of Q. We also denote by $3Q$ the (non-dyadic) cube with the same centre of Q and side length $3\ell Q$.

The average of a function f over a cube Q will be denoted by

$$\langle f \rangle_Q := \fint_Q f := \frac{1}{|Q|} \int_Q f(y)\,\mathrm{d}y.$$

We consider $\mathbb{R}^d$ with the ℓ^∞ metric, so that $\operatorname{dist}(x, 0) = \max_i |x_i|$. The distance between two cubes P and R will be denoted by $\mathrm{d}(P, R)$, while

$$D(P, R) := \ell P + \mathrm{d}(P, R) + \ell R$$

is the “long distance”, as defined in [NTV03, Definition 6.3].

3.1 Preliminaries

3.1.1 Dyadic cubes

The standard dyadic grid $\mathscr{D}$ on $\mathbb{R}^d$ is a collection of nested cubes organised in generations

$$\mathscr{D}_j := \{2^{-j}([0,1)^d + m), m \in \mathbb{Z}^d\}.$$

Each generation $\mathscr{D}_j$ is a partition of the whole space and $\mathscr{D} = \cup_{j\in\mathbb{Z}}\mathscr{D}_j$. Any cube $Q \in \mathscr{D}_j$ has 2^d children in $\mathscr{D}_{j+1}$ and one parent in $\mathscr{D}_{j-1}$. For $k \in \mathbb{N}$ we denote by $Q^{(k)}$ the k-ancestor of Q, that is the unique cube R in the same grid $\mathscr{D}$ such that $R \supset Q$ and $\ell R = 2^k \ell Q$. We also denote by $\mathsf{ch}_k(Q)$ the set of the k-grandchildren of Q, so that if $P \in \mathsf{ch}_k(Q)$ then $P^{(k)} = Q$.

3.1.2 Haar functions

Classical Haar functions are an orthonormal basis of $L^2(\mathbb{R}^d)$ given by linear combinations of indicator functions supported on cubes in $\mathscr{D}$.

On $\mathbb{R}$, for a given interval $I \in \mathscr{D}$ let I^- and I^+ be the left and the right dyadic child of I. Consider the functions $h_I^0 := |I|^{-1/2}\mathbb{1}_I$ and $h_I^1 := (\mathbb{1}_{I^-} - \mathbb{1}_{I^+})|I|^{-1/2}$. Then $\{h_I^1\}_{I\in\mathscr{D}}$ is an orthonormal complete system of $L^2(\mathbb{R})$. In higher dimensions, as a cube I is the product of intervals $I_1 \times \cdots \times I_d$, a non-constant Haar function h_I^ϵ is the product $h_{I_1}^{\epsilon_1} \times \cdots \times h_{I_d}^{\epsilon_d}$, where $\epsilon = (\epsilon_i)_i \in \{0,1\}^d \setminus \{0\}^d$.

A function f in L^2 can be written in the Haar basis:

$$f = \sum_{I\in\mathscr{D}} \sum_{\epsilon\in\{0,1\}^d\setminus\{0\}^d} \langle f, h_I^\epsilon\rangle h_I^\epsilon$$

$$= \sum_{I \in \mathscr{D}} \sum_{J \in \mathsf{ch}_1(I)} (\langle f \rangle_J - \langle f \rangle_I)\, \mathbb{1}_J =: \sum_{I \in \mathscr{D}} \Delta_I f.$$

In this chapter the sum over ϵ is not important, so both the superscript and the sum will be omitted and h_I will denote a non-constant Haar function. Two bounds that will be used are

$$\begin{aligned} \|\Delta_I f\|_{L^1} &\le |\langle f, h_I \rangle| |I|^{1/2} \le \int_I |f|, \\ \|\Delta_I f\|_{L^\infty} &\le |\langle f, h_I \rangle| |I|^{-1/2} \le \fint_I |f|. \end{aligned} \tag{3.1.1}$$

3.1.3 Good and bad cubes

A cube is called *good* if it is distant from the boundary of any much larger cube. More precisely, we have the following

Definition 3.1.1 (Good cubes)**.** Given two parameters $r \in \mathbb{N}$ and $\gamma \in (0, \frac{1}{2})$, a cube $R \in \mathscr{D}$ is r-good if $\mathrm{d}(R, \partial P) > (\ell R)^\gamma (\ell P)^{1-\gamma}$ for any $P \in \mathscr{D}$ with $\ell P \ge 2^r \ell R$.

A cube which is not good is a bad cube.

Recall that a family of Littlewood–Paley kernels comes with a parameter $\alpha \in (0, 1]$, see Definition 1.2.3. It is useful to fix $\gamma = \alpha/(4\alpha + 4d)$. This is just a convenient choice and any other value of γ strictly between 0 and $\alpha/(2\alpha + 2d)$ would work as well.

3.1.4 Shifted dyadic cubes

Given a sequence $\omega = \{\omega_i\}_{i \in \mathbb{Z}} \in (\{0,1\}^d)^{\mathbb{Z}}$ and a cube $R \in \mathscr{D}_j$ of length 2^{-j}, the translation of R by ω is defined by

$$R \dot{+} \omega := R + x_j \qquad \text{where} \qquad x_j := \sum_{i > j} \omega_i 2^{-i}.$$

For a fixed ω, let $\mathscr{D}^\omega$ be the collection of dyadic cubes in $\mathscr{D}$ translated by ω. The standard dyadic grid corresponds to $\mathscr{D}^0$ where $\omega_i = 0$ for all $i \in \mathbb{Z}$. Shifted dyadic grids enjoy the same nested properties of the standard grid $\mathscr{D}^0$, together with other properties that will be useful later, see Remark 3.3.5. For more on dyadic grids, we refer the reader to the beautiful survey [Per19, §3].

3.1.5 Random shifts

Let $\mathbb{P}$ be the unique probability measure on $\Omega := (\{0,1\}^d)^{\mathbb{Z}}$ such that the coordinate projections are independent and uniformly distributed. Fix $R \in \mathscr{D}^0$ with $\ell R = 2^{-j}$ and consider $J \in \mathscr{D}^0$ with $\ell J > \ell R$. The translated cube $J \dot{+} \omega$ is

$$\begin{aligned} J \dot{+} \omega &= J + \sum_{2^{-i} < \ell R} \omega_i 2^{-i} + \sum_{\ell R \leq 2^{-i} < \ell J} \omega_i 2^{-i}, \\ R \dot{+} \omega &= R + \sum_{2^{-i} < \ell R} \omega_i 2^{-i}. \end{aligned}$$

The position of $R \dot{+} \omega$ depends on the i such that $2^{-i} < \ell R$ while the goodness of $R \dot{+} \omega$, since R and J are translated by the same ω, depends on the i such that $2^{-i} \geq \ell R$. Then position and goodness of a cube are independent random variables, see [Hyt12].

Let $\mathbb{1}_{\text{good}}$ be the function on $\mathscr{D}^\omega$ which takes value 0 on bad cubes and 1 on good cubes. The probability of a cube R to be good is $\pi_{\text{good}} = \mathbb{P}(R \dot{+} \omega \text{ is good}) = \mathbb{E}_\omega[\mathbb{1}_{\text{good}}(R \dot{+} \omega)]$, where $\mathbb{E}_\omega$ is the expectation with respect to $\mathbb{P}$. The probability $\pi_{\text{good}} > 0$ provided to choose r large enough in Definition 3.1.1, see [Hyt17, Lemma 2.3]. The indicator function $\mathbb{1}_{R \dot{+} \omega}(\,\cdot\,)$ depends only on the position of $R \dot{+} \omega$, so by the independence of goodness and position, for any cube $R \in \mathscr{D}^0$ we have

$$\mathbb{E}_\omega[\mathbb{1}_{\text{good}}(R \dot{+} \omega)] \cdot \mathbb{E}_\omega[\mathbb{1}_{R \dot{+} \omega}(\,\cdot\,)] = \mathbb{E}_\omega[\mathbb{1}_{\{R \dot{+} \omega \text{ good}\}}(\,\cdot\,)]. \tag{3.1.2}$$

3.1.6 Calderón–Zygmund decomposition on dyadic grandchildren

Let R be a dyadic cube. For $r \in \mathbb{N}$ we denote by R_r a r-dyadic child of R in $\mathsf{ch}_r(R)$, so that $R_r^{(r)} = R$.

Proposition 3.1.2 (Calderón–Zygmund decomposition on r-grandchildren)**.** *Let $r \in \mathbb{N}$ and f be a function in $L^1(\mathbb{R}^d)$. For any $\lambda > 0$ there exists a collection of maximal dyadic cubes $\mathscr{L}$ and two functions a and b such that $f = a + b$, with $\|a\|_{L^\infty} \le 2^{d(r+1)}\lambda$ and*

$$b := \sum_{L \in \mathscr{L}} \sum_{L_r \in \mathsf{ch}_r(L)} b_{L_r}, \qquad \textit{where} \qquad b_{L_r} := \Big(f - \langle f \rangle_{L_r}\Big) \mathbb{1}_{L_r}.$$

Remark 3.1.3. When $r = 0$, this is the usual Calderón–Zygmund decomposition of f, see [Gra14, Theorem 5.3.1].

Proof. Given $\lambda > 0$, let $\mathscr{L}$ be the collection of maximal dyadic cubes L covering the set

$$E := \Big\{x \in \mathbb{R}^d \,:\, \sup_{Q \in \mathscr{D}} \langle |f| \rangle_Q \mathbb{1}_Q(x) > \lambda\Big\} = \bigcup_{L \in \mathscr{L}} L$$

so that $\langle |f| \rangle_L \in (\lambda, 2^d\lambda]$ for each $L \in \mathscr{L}$. Let

$$a := f\mathbb{1}_{E^\complement} + \sum_{L \in \mathscr{L}} \sum_{L_r \in \mathsf{ch}_r(L)} \langle f \rangle_{L_r} \mathbb{1}_{L_r}, \qquad b := f - a.$$

The cubes in $\mathsf{ch}_r(L)$ are a partition of L. Since the cubes L in $\mathscr{L}$ are disjoint, we have

$$\|a\|_{L^\infty} \le \lambda + \sup_{L \in \mathscr{L}} \sup_{L_r \in \mathsf{ch}_r(L)} |\langle f \rangle_{L_r}|.$$

Let $L^{(1)}$ be the dyadic parent of L. Then the average of f is controlled by

$$\Big|\frac{1}{|L_r|}\int_{L_r} f\Big| \le \frac{|L^{(1)}|}{|L_r|} \fint_{L^{(1)}} |f| \le 2^{d(r+1)}\lambda.$$

$\square$

3.1.7 Equivalence of testing conditions

The testing condition in (T) can also be expressed as testing against indicator functions on cubes.

Compare the following proposition with [AT98, Lemma 6 (Localization)].

Proposition 3.1.4. *Let S be a Littlewood–Paley square function. Then the two testing conditions are equivalent:*

1. *there exists $C > 0$ such that*

$$\langle S(\mathbb{1}_Q)^2, \mathbb{1}_Q \rangle \le C|Q|$$

holds for all dyadic cubes Q;

2. *there exists $C > 0$ such that*

$$\int_Q \int_0^{\ell Q} |\theta_t \mathbb{1}_Q|^2 \frac{\mathrm{d}t}{t} \,\mathrm{d}x \le C|Q|$$

holds for all dyadic cubes Q.

To prove Proposition 3.1.4, we use the following general reduction for large scales exploiting the decay of the kernel (condition (C1)).

Proposition 3.1.5. *Let f, g be functions supported on a fixed $Q \in \mathscr{D}$ in $\mathbb{R}^d$, then*

$$\int_Q \int_{\ell Q}^{\infty} |\theta_t f(x)|^2 g(x) \frac{\mathrm{d}t}{t} \,\mathrm{d}x \le \frac{1}{2d} \langle |f| \rangle_Q^2 \langle |g| \rangle_Q |Q|.$$

Proof. Bound $k_t(u, v)$ by $t^{\alpha}(t + |u - v|)^{-(\alpha+d)}$, then by dropping the term $|u - v|$ in the denominator, we estimate

$$\begin{aligned}\int_Q \int_{\ell Q}^{\infty} |\theta_t f(x)|^2 g(x) \frac{\mathrm{d}t}{t} \,\mathrm{d}x &\le \int_Q \int_{\ell(Q)}^{\infty} |t^{-d} \|f\|_{L^1}|^2 g(x) \frac{\mathrm{d}t}{t} \,\mathrm{d}x \\ &= \left(|Q| \fint_Q |f| \right)^2 \int_{\ell(Q)}^{\infty} t^{-2d-1} \,\mathrm{d}t \left(\int_Q g \right) \\ &= \left(\fint_Q |f| \right)^2 |Q|^2 \frac{\ell(Q)^{-2d}}{2d} |Q| \left(\fint_Q g \right) \\ &\lesssim_d \langle |f| \rangle_Q^2 \langle |g| \rangle_Q |Q|.\end{aligned}$$

□

Proof of Proposition 3.1.4. Since the integrand is non-negative, we can bound with

$$\int_Q \int_0^{\ell Q} |\theta_t \mathbb{1}_Q|^2 \frac{\mathrm{d}t}{t} \,\mathrm{d}x \le \int_Q \int_0^{\infty} |\theta_t \mathbb{1}_Q|^2 \frac{\mathrm{d}t}{t} \,\mathrm{d}x \le C|Q|.$$

For the opposite implication, we decompose the integral in t in two ranges, then we estimate:

$$\begin{aligned}\langle S(\mathbb{1}_Q)^2, \mathbb{1}_Q\rangle &= \int_Q \int_0^{\ell Q} |\theta_t \mathbb{1}_Q|^2 \frac{\mathrm{d}t}{t} \,\mathrm{d}x \\ &\quad + \int_Q \int_{\ell Q}^{\infty} |\theta_t \mathbb{1}_Q|^2 \frac{\mathrm{d}t}{t} \,\mathrm{d}x \le (C_{\mathsf{T}} + (2d)^{-1})|Q|\end{aligned}$$

where we applied the testing condition (T) to the fist term and Proposition 3.1.5 to the second. □

3.2 Proof of Theorem B

We start by decomposing the dual form $\langle (Sf)^2, g\rangle$.

3.2.1 Decomposition

For any fixed $\omega \in \Omega = (\{0,1\}^d)^{\mathbb{Z}}$ the upper half space $\mathbb{R}^{d+1}_+$ can be decomposed in the Whitney regions

$$W_R := R \times \left[\frac{\ell R}{2}, \ell R\right), \quad R \in \mathscr{D}^{\omega}.$$

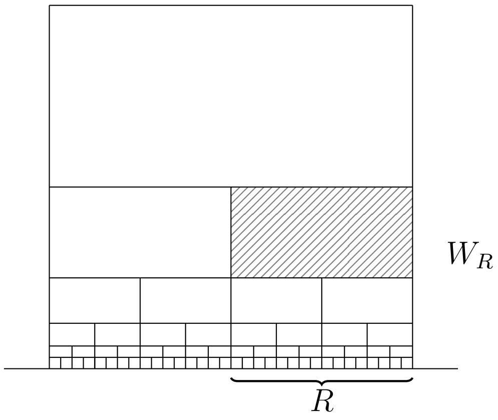

Figure 3.1: Decomposition of $\mathbb{R}^2_+$ in Whitney regions. The region W_R and the corresponding interval R are shown.

Thus we can write

$$\langle (Sf)^2, g\rangle = \iint_{\mathbb{R}^{d+1}_+} |\theta_t f(x)|^2 \frac{\mathrm{d}t}{t} g(x)\,\mathrm{d}x = \sum_{R\in\mathscr{D}^\omega} \iint_{W_R} |\theta_t f(x)|^2 \frac{\mathrm{d}t}{t} g(x)\,\mathrm{d}x.$$

Then we decompose $f = \sum_{P\in\mathscr{D}^\omega} \Delta_P f$. Given $R \in \mathscr{D}^\omega$, we distinguish two collections of P:

$$\mathscr{P}^\omega_R \coloneqq \{P \in \mathscr{D}^\omega \,:\, P \supset R^{(r)}\}, \qquad \text{and} \qquad \mathscr{D}^\omega \setminus \mathscr{P}^\omega_R.$$

We shall sometimes omit the superscript ω in the following. Bound the operator:

$$\begin{aligned}\sum_{R\in\mathscr{D}} \iint_{W_R} &|\theta_t f(x)|^2 \frac{\mathrm{d}t}{t} g(x)\,\mathrm{d}x \\ &\le 2 \sum_{R\in\mathscr{D}} \iint_{W_R} \Big(\big|\sum_{P\in\mathscr{D}\setminus\mathscr{P}_R} \theta_t \Delta_P f(x)\big|^2 + \big|\sum_{P\in\mathscr{P}_R} \theta_t \Delta_P f(x)\big|^2\Big) |g(x)| \frac{\mathrm{d}t}{t}\,\mathrm{d}x.\end{aligned} \tag{3.2.1}$$

Consider the second term in (3.2.1). Let P_R be the dyadic child of P containing R. Then $\Delta_P f \mathbb{1}_P = \Delta_P f \mathbb{1}_{P\setminus P_R} + \langle \Delta_P f\rangle_{P_R} \mathbb{1}_{P_R}$ and we split the operator accordingly as before to obtain:

$$\sum_{R\in\mathscr{D}} \iint_{W_R} |\theta_t f(x)|^2 \frac{\mathrm{d}t}{t} g(x)\,\mathrm{d}x \lesssim$$

$$\lesssim \sum_{R\in\mathscr{D}} \iint_{W_R} \Big| \sum_{P\in\mathscr{D}\setminus\mathscr{P}_R} \theta_t \Delta_P f(x)\Big|^2 |g(x)| \frac{\mathrm{d}t}{t}\,\mathrm{d}x \tag{I}$$

$$+ \sum_{R\in\mathscr{D}} \iint_{W_R} \Big| \sum_{P\in\mathscr{P}_R} \theta_t (\Delta_P f \mathbb{1}_{P\setminus P_R})(x)\Big|^2 |g(x)| \frac{\mathrm{d}t}{t}\,\mathrm{d}x \tag{II}$$

$$+ \sum_{R\in\mathscr{D}} \iint_{W_R} \Big| \sum_{P\in\mathscr{P}_R} \theta_t (\langle\Delta_P f\rangle_{P_R} \mathbb{1}_{P_R})(x)\Big|^2 |g(x)| \frac{\mathrm{d}t}{t}\,\mathrm{d}x. \tag{III}$$

In each term, without loss of generality, we can assume g to be supported on R. We write $|g| = a + b$ using the Calderón–Zygmund decomposition in Proposition 3.1.2 at height $\lambda = A\langle |g|\rangle_R$ for $A > 1$. Then the bad part b is decomposed in the Haar basis. We split term (III) in two parts: (III_a) and (III_b), defined below.

$$(\mathrm{III}) = \sum_{R\in\mathscr{D}} \iint_{W_R} \Big| \sum_{P\in\mathscr{P}_R} \theta_t (\langle\Delta_P f\rangle_{P_R} \mathbb{1}_{P_R})(x)\Big|^2 \frac{\mathrm{d}t}{t} a(x)\,\mathrm{d}x \tag{III_a}$$

$$+ \sum_{R\in\mathscr{D}} \iint_{W_R} \Big| \sum_{P\in\mathscr{P}_R} \theta_t (\langle\Delta_P f\rangle_{P_R} \mathbb{1}_{P_R})(x)\Big|^2 \sum_{\substack{Q\in\mathscr{D}\\ Q\subset R}} \Delta_Q b(x) \frac{\mathrm{d}t}{t}\,\mathrm{d}x. \tag{III_b}$$

3.2.2 Good reduction

Averaging over all dyadic grids $\mathscr{D}^\omega$ we have

$$\begin{aligned}\iint_{\mathbb{R}^{d+1}_+} |\theta_t f(x)|^2 |g(x)| \frac{\mathrm{d}t}{t}\,\mathrm{d}x &= \mathbb{E}_\omega \sum_{R\in\mathscr{D}^\omega} \iint_{W_R} |\theta_t f(x)|^2 |g(x)| \frac{\mathrm{d}t}{t}\,\mathrm{d}x \\ &\lesssim \mathbb{E}_\omega\big[\mathrm{I}+\mathrm{II}+\mathrm{III}\big] = \mathbb{E}_\omega\big[\mathrm{I}+\mathrm{II}+\mathrm{III}_a\big] + \mathbb{E}_\omega\big[\mathrm{III}_b\big]\end{aligned}$$

because all the integrands are non-negative and the expectation $\mathbb{E}_\omega$ is linear.

By using the identity (3.1.2) and writing 1 as $\pi_{\mathrm{good}}^{-1}\mathbb{E}_\omega[\mathbb{1}_{\mathrm{good}}(\,\cdot\dot{+}\omega)]$, one can turn a sum over all cubes in $\mathscr{D}^\omega$ into a sum over good cubes, in particular:

$$\begin{aligned}\mathbb{E}_\omega\big[\mathrm{I}+\mathrm{II}+\mathrm{III}_a\big] &= \pi_{\mathrm{good}}^{-1}\mathbb{E}_\omega\big[\mathbb{1}_{\mathrm{good}}(R\dot{+}\omega)\big(\mathrm{I}+\mathrm{II}+\mathrm{III}_a\big)\big], \\ \mathbb{E}_\omega\big[\mathrm{III}_b\big] &= \pi_{\mathrm{good}}^{-1}\mathbb{E}_\omega\big[\mathbb{1}_{\mathrm{good}}(Q\dot{+}\omega)\big(\mathrm{III}_b\big)\big].\end{aligned} \tag{3.2.2}$$

We refer the reader to [MM14, §2.2] for an expanded version of (3.2.2) with $g \equiv 1$.

From now on, the cubes Q in (III_b) and the cubes R in all other cases are considered to be good cubes. The superscript in $\mathscr{D}^\omega$, as well as the expectation $\mathbb{E}_\omega$ and the probability π_{good} will be omitted.

3.3 Reduction of (I) to a dyadic form

We start by showing that

$$(\mathrm{I}) = \sum_{\substack{R\in\mathscr{D} \\ R \text{ good}}} \iint_{W_R} \Big|\sum_{P\in\mathscr{D}\setminus\mathscr{P}_R} \theta_t \Delta_P f\Big|^2 |g| \frac{\mathrm{d}t}{t}\,\mathrm{d}x \lesssim \sum_{j\in\mathbb{N}} 2^{-cj} B_j^{\mathscr{D}}(g,f)$$

for $c > 0$, where $B_j^{\mathscr{D}}(g,f)$ is the dyadic form given by

$$B_j^{\mathscr{D}}(g,f) \coloneqq \sum_{K\in\mathscr{D}} \langle |g|\rangle_{3K} \sum_{\substack{P\in\mathscr{D} \\ P\subset 3K \\ \ell P = 2^{-j}\ell K}} \langle f, h_P\rangle^2. \tag{3.3.1}$$

We remark that the function g barely plays any role in this section.

3.3.1 Different cases for P

Given $R \in \mathscr{D}$, the cubes P are grouped according to their length and position with respect to R. This decomposition also appeared in [MM15, §4.3].

Table 3.3.1: Different cases for P given R according to their lengths (first row) and position.

$\ell P \geq 2^{r+1}\ell R$			$\ell R \leq \ell P \leq 2^r \ell R$		$\ell P < \ell R$	
$P \supset R$	$P \not\supset R$					
					$\mathscr{P}_{\text{subscale}}$	
	$3P \setminus P \supset R$	$3P \not\supset R$	$3P \not\supset R$	$3P \supset R$	$P \subset 3R$	$P \not\subset 3R$
	$\mathscr{P}_{\text{near}}$	$\mathscr{P}_{\text{far}}$		$\mathscr{P}_{\text{close}}$	inside	far
$\mathscr{P}_R$	$\mathscr{D} \setminus \mathscr{P}_R$					

Remark 3.3.1. Since $3P$ is the union of 3^d cubes in $\mathscr{D}$, the condition $3P \not\supset R$ is equivalent to $3P \cap R = \emptyset$, which implies that $\mathrm{d}(P, R) > \ell P$. The condition $\ell P \geq 2^{r+1}\ell R$ allows to exploit the goodness of R also with dyadic children of P.

We decompose the sum over $P \in \mathscr{D} \setminus \mathscr{P}_R$ in four terms.

$$
\begin{aligned}
&\sum_{R \in \mathscr{D}} \iint_{W_R} \Big| \sum_{P \in \mathscr{D} \setminus \mathscr{P}_R} \theta_t(\Delta_P f) \Big|^2 |g| \frac{\mathrm{d}t}{t} \,\mathrm{d}x \\
&\quad \lesssim \sum_{R \in \mathscr{D}} \iint_{W_R} \Big| \sum_{\substack{P \,:\, \ell P > 2^r \ell R \\ 3P \setminus P \supset R}} \theta_t(\Delta_P f) \Big|^2 |g| \frac{\mathrm{d}t}{t} \,\mathrm{d}x && \text{(near)} \\
&\qquad + \sum_{R \in \mathscr{D}} \iint_{W_R} \Big| \sum_{\substack{P \,:\, \ell P \geq \ell R \\ \mathrm{d}(P,R) > \ell P}} \theta_t(\Delta_P f) \Big|^2 |g| \frac{\mathrm{d}t}{t} \,\mathrm{d}x && \text{(far)} \\
&\qquad + \sum_{R \in \mathscr{D}} \iint_{W_R} \Big| \sum_{\substack{P \,:\, 3P \supset R \\ \ell R \leq \ell P \leq 2^r \ell R}} \theta_t(\Delta_P f) \Big|^2 |g| \frac{\mathrm{d}t}{t} \,\mathrm{d}x && \text{(close)} \\
&\qquad + \sum_{R \in \mathscr{D}} \iint_{W_R} \Big| \sum_{P \,:\, \ell P < \ell R} \theta_t(\Delta_P f) \Big|^2 |g| \frac{\mathrm{d}t}{t} \,\mathrm{d}x. && \text{(subscale)}
\end{aligned}
$$

3.3.2 Estimates case by case

We start with a well–known bound.

Lemma 3.3.2. *Let $P, R \in \mathscr{D}$ with R good. If one of the following conditions holds*

(1) $\ell P \geq \ell R$ and P and R are disjoint;

(2) $\ell P < \ell R$;

then for $(x,t) \in W_R$ we have

$$|\theta_t(\Delta_P f)(x)| \lesssim_{\alpha,d} (C_1 + C_2)\frac{(\sqrt{\ell R \ell P})^\alpha}{D(R,P)^{\alpha+d}}\|\Delta_P f\|_{L^1}.$$

The proof uses the goodness of R in case (1) and the zero average of $\Delta_P f$ in case (2), see also [LM17a, §5],[MM14, §2.4]. Details of the proof are deferred to §3.8.
We apply Lemma 3.3.2 for P in $\mathscr{P}_i$ with $i \in \{\text{near, far, close, subscale}\}$ and estimate $\|\Delta_P f\|_{L^1}$ as in (3.1.1). Then we apply Cauchy–Schwarz in ℓ^2.

$$\begin{aligned}
&\sum_{R\in\mathscr{D}} \iint_{W_R} \Big|\sum_{P\in\mathscr{P}_i} \theta_t(\Delta_P f)(x)\Big|^2 |g(x)| \frac{\mathrm{d}t}{t}\,\mathrm{d}x \\
&\quad\lesssim \sum_{R\in\mathscr{D}} \iint_{W_R} \left(\sum_{P\in\mathscr{P}_i} |\langle f, h_P\rangle| \frac{(\sqrt{\ell R \ell P})^\alpha}{D(R,P)^{\alpha+d}} |P|^{1/2}\right)^2 |g(x)|\frac{\mathrm{d}t}{t}\,\mathrm{d}x \\
&\quad\leq \sum_{R\in\mathscr{D}} \iint_{W_R} \left(\sum_{P\in\mathscr{P}_i} \langle f, h_P\rangle^2 \frac{(\sqrt{\ell R \ell P})^\alpha}{D(R,P)^{\alpha+d}} \cdot \sum_{P\in\mathscr{P}_i} \frac{(\sqrt{\ell R \ell P})^\alpha}{D(R,P)^{\alpha+d}}|P|\right) |g(x)|\frac{\mathrm{d}t}{t}\,\mathrm{d}x. \qquad (3.3.2)
\end{aligned}$$

The quantity in parenthesis in (3.3.2) does not depend on t, so we bound

$$\int_{\ell R/2}^{\ell R} \frac{\mathrm{d}t}{t} \leq \frac{2}{\ell R}\int_{\ell R/2}^{\ell R} \mathrm{d}t = 1.$$

The second sum in parenthesis in (3.3.2) is finite in all cases.

Lemma 3.3.3. *Let $i \in \{$near, far, close, subscale$\}$, then*

$$\sum_{P\in\mathscr{P}_i} \frac{(\sqrt{\ell R \ell P})^\alpha}{D(R,P)^{\alpha+d}}|P| \lesssim 1.$$

Details of the proof are in §3.8. We proceed with studying

$$\sum_{R\in\mathscr{D}} \int_R \left(\sum_{P\in\mathscr{P}_i} \langle f, h_P\rangle^2 \frac{(\sqrt{\ell R \ell P})^\alpha}{D(R,P)^{\alpha+d}}\right) |g(x)|\,\mathrm{d}x$$

for $i \in \{\text{near, far, close, subscale}\}$. When P and R are disjoint, it is useful to rearrange the sums using a common ancestor of P and R.

Lemma 3.3.4 (Common ancestor). *Let $R, P \in \mathscr{D}$ be disjoint cubes with R good. If $\mathrm{d}(R,P) > \max(\ell R, \ell P)^{1-\gamma} \min(\ell R, \ell P)^{\gamma}$ then there exists $K \supseteq P \cup R$ such that*

$$\ell K \left(\frac{\min(\ell P, \ell R)}{\ell K} \right)^{\gamma} \leq 2^r \, \mathrm{d}(R,P).$$

A proof in the case $\ell P \geq \ell R$ can be found in [Hyt17, Lemma 3.7]. When $\ell P < \ell R$, the same ideas carry over, see §3.8 for a proof of this case.

Remark 3.3.5. For any $P, R \in \mathscr{D}^{\omega}$ there exists (almost surely) a common ancestor $K \in \mathscr{D}^{\omega}$. Indeed, dyadic grids (like the standard grid $\mathscr{D}^0$) without this property have zero measure in the probability space $(\Omega, \mathbb{P})$, see [Per19, §3.1.1 and Example 3.2].

3.3.3 P far from R

In this case $\mathrm{d}(P,R) > \ell P$ and $\ell P = \max(\ell P, \ell R)$, so the hypotheses of Lemma 3.3.4 are satisfied. Let K be the common ancestor of P and R given by Lemma 3.3.4. Since $\ell P \geq 2^{r+1} \ell R$, let $\ell P = 2^{-j} \ell K$ and $\ell R = 2^{-i-j} \ell K$ for some $i, j \in \mathbb{Z}_+$, with $i \geq r+1$. We have

$$\begin{aligned} &\sum_{R \in \mathscr{D}} \int_R g \left(\sum_{P \in \mathscr{P}_{\mathrm{far}}} \langle f, h_P \rangle^2 \frac{(\sqrt{\ell R \ell P})^{\alpha}}{\mathrm{d}(R,P)^{\alpha+d}} \right) \\ &\quad = \sum_{K \in \mathscr{D}} \sum_{i,j} \sum_{\substack{R \,:\, R \subset K \\ \ell R = 2^{-i-j} \ell K}} \int_R g \sum_{\substack{P \,:\, P \subset K \\ \ell P = 2^{-j} \ell K \\ \mathrm{d}(P,R) > \ell P}} \langle f, h_P \rangle^2 \frac{(\sqrt{\ell R \ell P})^{\alpha}}{\mathrm{d}(R,P)^{\alpha+d}}. \end{aligned}$$

By using the lower bound $\mathrm{d}(P,R) \gtrsim_r (\ell K)^{1-\gamma} (\ell R)^{\gamma}$ with $\gamma = \alpha/(4\alpha + 4d)$, we estimate

$$\frac{\sqrt{\ell P \ell R}}{\mathrm{d}(P,R)} \lesssim_r \frac{2^{-j-i/2} \ell K}{\ell K 2^{-(i+j)\gamma}}$$

so that

$$\frac{(\sqrt{\ell P \ell R})^{\alpha}}{\mathrm{d}(P,R)^{\alpha+d}} \lesssim_{r,\alpha,d} \frac{2^{-(j+i/2)\alpha}}{2^{-(i+j)\gamma(\alpha+d)}|K|} = \frac{2^{-(3j+i)\alpha/4}}{|K|}. \tag{3.3.3}$$

For any fixed integer m, the set $\{R \subset K : \ell R = 2^{-m}\ell K\}$ is a partition of K, so we bound

$$\begin{aligned}
&\sum_{K\in\mathscr{D}}\sum_{i,j}\sum_{\substack{R\,:\,R\subset K\\ \ell R=2^{-i-j}\ell K}}\int_R g \sum_{\substack{P\,:\,P\subset K\\ \ell P=2^{-j}\ell K\\ \mathrm{d}(P,R)>\ell P}} \langle f,h_P\rangle^2 \frac{(\sqrt{\ell R\ell P})^{\alpha}}{\mathrm{d}(R,P)^{\alpha+d}}\\
&\lesssim \sum_{j\in\mathbb{N}} 2^{-3j\alpha/4} \sum_{i\geq r+1} 2^{-i\alpha/4} \sum_{K\in\mathscr{D}} \fint_K |g| \sum_{\substack{P\,:\,P\subset K\\ \ell P=2^{-j}\ell K}} \langle f,h_P\rangle^2.
\end{aligned}$$

We can sum in i, then we have

$$\begin{aligned}
\sum_{j\in\mathbb{N}} 2^{-3j\alpha/4} \sum_{K\in\mathscr{D}} \langle |g|\rangle_K \sum_{\substack{P\subset K\\ \ell P=2^{-j}\ell K}} \langle f,h_P\rangle^2
&\leq 3^d \sum_{j\in\mathbb{N}} 2^{-3j\alpha/4} \sum_{K\in\mathscr{D}} \langle |g|\rangle_{3K} \sum_{\substack{P\subset 3K\\ \ell P=2^{-j}\ell K}} \langle f,h_P\rangle^2\\
&= \sum_{j\in\mathbb{N}} 2^{-3j\alpha/4} B_j^{\mathscr{D}}(g,f).
\end{aligned}$$

A sparse domination of $B_j^{\mathscr{D}}(g,f)$ is proved in §3.7.

3.3.4 P near R

Recall that $P \in \mathscr{P}_{\mathrm{near}}$ if $3P \setminus P \supset R$ and $\ell P \geq 2^{r+1}\ell R$. By the goodness of R, we have $\mathrm{d}(P,R) > (\ell P)^{1-\gamma}(\ell R)^{\gamma}$. So the hypotheses of Lemma 3.3.4 are satisfied and there exists $K \supseteq P \cup R$ such that $\mathrm{d}(P,R) \gtrsim_r (\ell K)^{1-\gamma}(\ell R)^{\gamma}$. Arguing as in the far term leads to

$$\sum_{R\in\mathscr{D}} \int_R g \left(\sum_{P\in\mathscr{P}_{\mathrm{near}}} \langle f,h_P\rangle^2 \frac{(\sqrt{\ell R\ell P})^{\alpha}}{\mathrm{d}(R,P)^{\alpha+d}} \right) \lesssim \sum_{j\in\mathbb{N}} 2^{-3j\alpha/4} B_j^{\mathscr{D}}(g,f).$$

3.3.5 P comparable and close to R

In this case $\ell R \leq \ell P \leq \ell R^{(r)}$ and $3P \supset R$. Using the trivial bound $D(P,R) \geq \ell R$ we have

$$\sum_{R\in\mathscr{D}} \int_R |g| \left(\sum_{P\in\mathscr{P}_{\text{close}}} \langle f, h_P\rangle^2 \frac{(\sqrt{\ell R \ell P})^\alpha}{D(R,P)^{\alpha+d}} \right) \lesssim_{r,\alpha} \sum_{R\in\mathscr{D}} \int_R |g| \sum_{\substack{P\,:\,3P\supset R\\ \ell R\leq \ell P\leq 2^r \ell R}} \frac{\langle f, h_P\rangle^2}{|R|}.$$

Rearrange the sum in groups of P such that $\ell P = 2^k \ell R$ for $k \in \{0,\dots,r\}$. Then

$$\begin{aligned}
\sum_{R\in\mathscr{D}} \int_R |g| \sum_{k=0}^{r} \sum_{\substack{P\,:\,3P\supset R\\ \ell P=2^k\ell R}} \langle f, h_P\rangle^2 \frac{1}{|R|} &= \sum_{k=0}^{r} \sum_{P\in\mathscr{D}} \langle f, h_P\rangle^2 \frac{2^{kd}}{|P|} \sum_{\substack{R\subset 3P\\ \ell R=2^{-k}\ell P}} \int_R |g| \\
&\leq \sum_{k=0}^{r} \sum_{P\in\mathscr{D}} \langle f, h_P\rangle^2 \frac{2^{kd}}{|P|} \int_{3P} |g| \\
&\lesssim_{r,d} \sum_{P\in\mathscr{D}} \langle f, h_P\rangle^2 \frac{3^d}{|3P|} \int_{3P} |g| \\
&= 3^d \sum_{P\in\mathscr{D}} \langle f, h_P\rangle^2 \langle |g| \rangle_{3P}.
\end{aligned}$$

We define

$$B_0^{\mathscr{D}}(g,f) \coloneqq \sum_{P\in\mathscr{D}} \langle f, h_P\rangle^2 \langle |g| \rangle_{3P}. \tag{3.3.4}$$

Then $B_0^{\mathscr{D}}(g,f)$ is bounded by a sparse form in §3.7.

3.3.6 Subscale

When $\ell P < \ell R$ we distinguish two subcases, as shown in Table 3.3.1.

Inside : $P \subset 3R$

The leading term in the long-distance $D(R,P)$ is ℓR, so we bound

$$\sum_{R\in\mathscr{D}} \int_R |g| \left(\sum_{\substack{P\,:\,\ell P<\ell R\\ P\subset 3R}} \langle f, h_P\rangle^2 \frac{(\sqrt{\ell R \ell P})^\alpha}{D(R,P)^{\alpha+d}} \right)$$

$$\leq \sum_{R\in\mathscr{D}} \fint_R |g| \sum_{\substack{P\,:\,\ell P<\ell R\\ P\subset 3R}} \langle f,h_P\rangle^2 \left(\frac{\ell P}{\ell R}\right)^{\alpha/2}$$

$$= \sum_{j\in\mathbb{N}} 2^{-j\alpha/2} \sum_{R\in\mathscr{D}} \langle |g|\rangle_R \sum_{\substack{P\,:\,P\subset 3R\\ \ell P=2^{-j}\ell R}} \langle f,h_P\rangle^2$$

$$\lesssim_d \sum_{j\in\mathbb{N}} 2^{-j\alpha/2} B_j^{\mathscr{D}}(g,f).$$

See §3.7 for the sparse domination of $B_j^{\mathscr{D}}(g,f)$.

Far : $P \not\subset 3R$

In this case $\mathrm{d}(P,R) > \ell R > \ell P$, so the hypotheses of Lemma 3.3.4 are satisfied. After Cauchy–Schwarz, rearrange the sum using the common ancestor K, then let $\ell P = 2^{-m}\ell R = 2^{-m-i}\ell K$ and estimate the decay factor as in (3.3.3):

$$\sum_{R\in\mathscr{D}} \int_R |g| \sum_{\substack{P:\ell P<\ell R\\ \mathrm{d}(P,R)>\ell R}} \langle f,h_P\rangle^2 \frac{(\sqrt{\ell P\ell R})^\alpha}{D(P,R)^{\alpha+d}}$$

$$\leq \sum_{i,m}\sum_{K\in\mathscr{D}} \sum_{\substack{R\subset K\\ \ell R=2^{-i}\ell K}} \int_R |g| \sum_{\substack{P\subset K\\ \ell P=2^{-m-i}\ell K}} \langle f,h_P\rangle^2 \frac{(\sqrt{\ell P\ell R})^\alpha}{\mathrm{d}(P,R)^{\alpha+d}}$$

$$\lesssim_r \sum_{i\in\mathbb{N}} 2^{-i\alpha/2} \sum_{m\in\mathbb{N}}\sum_{K\in\mathscr{D}} \int_K |g| \sum_{\substack{P\subset K\\ \ell P=2^{-m-i}\ell K}} \langle f,h_P\rangle^2 \frac{2^{-(m+i)\alpha/4}2^{-i\alpha/2}}{|K|}$$

$$\leq \sum_{i\in\mathbb{N}} 2^{-i\alpha/2} \sum_{j\in\mathbb{N}} 2^{-j\alpha/4} \sum_{K\in\mathscr{D}} \fint_K |g| \sum_{\substack{P\subset K\\ \ell P=2^{-j}\ell K}} \langle f,h_P\rangle^2$$

where $j \coloneqq m+i$ and we bounded by the sum over all $j \geq 0$, since all terms are nonnegative. After summing in i, what is left is bounded by $B_j^{\mathscr{D}}(g,f)$. This concludes this case and the reduction of (I) to a dyadic form. $\square$

3.4 Reduction of (II) to a dyadic form

In this section we prove the following bound

$$\sum_{R\in\mathscr{D}} \iint_{W_R} |g| \Big| \sum_{P\,:\,P\supset R^{(r)}} \theta_t(\Delta_P f \mathbb{1}_{P\setminus P_R}) \Big|^2 \frac{\mathrm{d}t}{t}\,\mathrm{d}x \lesssim B_0^{\mathscr{D}}(g,f). \tag{3.4.1}$$

The dyadic form $B_0^{\mathscr{D}}(g,f)$ defined in (3.3.4) is controlled by a sparse form in §3.7.

Remark 3.4.1. The goodness of R gives the lower bound on the distance $\mathrm{d}(R,\partial P) > (\ell P)^{1-\gamma}(\ell R)^{\gamma}$.

As will be clear from the proof, inequality (3.4.1) holds if one replaces the indicator $\mathbb{1}_{P\setminus P_R}$ with $\mathbb{1}_{K\setminus P_R}$ where K is $\mathbb{R}^d$ or any other larger cube containing P.

To prove (3.4.1), we use a classical estimate for the Poisson kernel.

Lemma 3.4.2 (Poisson off-diagonal estimates). *Let $\beta \in (0,1], r \in \mathbb{N}$ and γ as in the introduction and let $Q, P \in \mathscr{D}$ such that $Q^{(r)} \subset P$ and Q is r-good. Then*

$$\int_{\mathbb{R}^d\setminus P} \frac{(\ell Q)^{\beta}}{\mathrm{d}(y,Q)^{\beta+d}}\,\mathrm{d}y \lesssim \left(\frac{\ell Q}{\ell P}\right)^{\eta}$$

where $\eta = \beta - \gamma(\beta+d)$.

Proof. Decompose $\mathbb{R}^d \setminus P$ in annuli $A_k = 3^{k+1}P \setminus 3^k P$ for $k \in \mathbb{N}$. Then on each annulus $\mathrm{d}(y,Q) > \mathrm{d}(\partial(3^k P), Q)$. Since $\ell P > 2^r \ell Q$, use the goodness of Q to obtain the bound. □

Proof of (3.4.1). When $(x,t) \in W_R$ the size condition (C1) and Lemma 3.4.2 give

$$\begin{aligned}\theta_t(\Delta_P f \mathbb{1}_{P\setminus P_R})(x) &\lesssim \|\Delta_P f\|_{L^\infty} \int_{P\setminus P_R} \frac{(\ell R)^{\alpha}}{(\ell R + \mathrm{d}(y,R))^{\alpha+d}}\,\mathrm{d}y \\ &\lesssim \frac{|\langle f, h_P\rangle|}{|P|^{1/2}} \left(\frac{\ell R}{\ell P_R}\right)^{\eta}\end{aligned}$$

where $\eta = \alpha - \gamma(\alpha+d) > 0$. The sum $\sum_{P\supset R^{(r)}} (\ell R/\ell P_R)^{\eta}$ is a geometric series. An application of the Cauchy–Schwarz inequality gives

$$\begin{aligned}
\sum_{R\in\mathscr{D}} \iint_{W_R} |g| \left| \sum_{P\supset R^{(r)}} \frac{|\langle f, h_P\rangle|}{|P|^{1/2}} \left(\frac{\ell R}{\ell P_R}\right)^{\eta} \right|^2 \frac{\mathrm{d}t}{t}\,\mathrm{d}x \\
&\leq \sum_{R\in\mathscr{D}} \sum_{P\supset R^{(r)}} \frac{\langle f, h_P\rangle^2}{|P|} \left(\frac{\ell R}{\ell P_R}\right)^{\eta} \int_R |g(x)|\,\mathrm{d}x \\
&\lesssim \sum_{i\geq r+1} 2^{-i\eta} \sum_{P\in\mathscr{D}} \frac{\langle f, h_P\rangle^2}{|P|} \sum_{\substack{R\subset P\\ \ell R = 2^{-i}\ell P}} \int_R |g| \\
&= \sum_{i\geq r+1} 2^{-i\eta} \sum_{P\in\mathscr{D}} \frac{\langle f, h_P\rangle^2}{|P|} \int_P |g|.
\end{aligned}$$

We sum in i and then we bound by the dyadic form $B_0^{\mathscr{D}}(g,f)$. □

3.5 Reduction of (III$_a$) to a sparse form

In this section we prove that there exists $c>0$ and a sparse family $\mathscr{S}\subseteq\mathscr{D}$ such that

$$\begin{aligned}
(\mathrm{III}_a) &\lesssim \sum_{R\in\mathscr{D}} \langle |g|\rangle_R \iint_{W_R} \Big| \sum_{P\in\mathscr{P}_R} \langle \Delta_P f\rangle_{P_R} \theta_t \mathbb{1}_{P_R} \Big|^2 \frac{\mathrm{d}t}{t}\,\mathrm{d}x \\
&\lesssim \sum_{j\in\mathbb{N}} 2^{-cj} B_j^{\mathscr{D}}(g,f) + \Lambda_{\mathscr{S}}(g,f)
\end{aligned} \tag{3.5.1}$$

where $\Lambda_{\mathscr{S}}(g,f) = \sum_{S\in\mathscr{S}} \langle |g|\rangle_S \langle |f|\rangle_S^2 |S|$. We remind the reader that P_R is the dyadic child of P which contains R, and $\mathscr{P}_R$ is the collection of P containing $R^{(r)}$.

Remark 3.5.1 (Bound on a). Recall that a is the good part of g in the Calderón–Zygmund decomposition of Proposition 3.1.2 with $\lambda = A\langle |g|\rangle_R$. So $\|a\|_\infty \leq 2^{d(r+1)} A\langle |g|\rangle_R$ and the first inequality in (3.5.1) follows.

3.5.1 Stopping cubes

Given two functions f and g and a cube $Q \subseteq \mathbb{R}^d$, consider the collections:

$$\mathscr{A}_f(Q) = \{S \in \mathscr{D}, S \subset Q \,:\, \langle |f| \rangle_S > A \langle |f| \rangle_Q\},$$
$$\mathscr{A}_g(Q) = \{S \in \mathscr{D}, S \subset Q \,:\, \langle |g| \rangle_S > A \langle |g| \rangle_Q\}.$$

Let $\mathscr{A}^\star(Q)$ be the maximal dyadic components of the set

$$\mathscr{A}(Q) = \mathscr{A}_f(Q) \cup \mathscr{A}_g(Q).$$

The weak $(1,1)$ bound for the dyadic maximal function ensures that there exists a constant $A > 1$ such that $|\mathscr{A}(Q)| \leq \frac{1}{2}|Q|$ and so

$$\Big| \bigcup_{S \in \mathscr{A}^\star(Q)} S \Big| = \sum_{S \in \mathscr{A}^\star(Q)} |S| \leq \frac{1}{2}|Q|.$$

Fix Q_0 in $\mathscr{D}$ containing the support of f and g. The stopping family $\mathscr{S}$ is defined iteratively:

$$\mathscr{S}_0 \coloneqq \{Q_0\}, \qquad \mathscr{S}_{n+1} \coloneqq \bigcup_{Q \in \mathscr{S}_n} \mathscr{A}^\star(Q), \qquad \mathscr{S} \coloneqq \bigcup_{n \in \mathbb{N}} \mathscr{S}_n.$$

Remark 3.5.2. The family $\mathscr{S}$ is $\frac{1}{2}$-sparse, since for any $S \in \mathscr{S}$ the set $E_S \coloneqq S \backslash \bigcup_{S' \in \mathscr{A}^\star(S)} S'$ has measure $|E_S| > \frac{1}{2}|S|$ and $\{E_S\}_{S \in \mathscr{S}}$ are disjoint.

In the same way, taking $\mathscr{A}^\star(Q)$ to be the maximal dyadic components of $\mathscr{A}_g(Q)$ produces a sparse family that we denote with $\mathscr{S}_g$. It will be used later when only the stopping cubes related to g are needed.

For a given $Q \in \mathscr{D}$, denote by $\widehat{Q}$ the minimal stopping cube $S \in \mathscr{S}$ such that $S \supseteq Q$.

For $S \in \mathscr{S}$ let $\mathsf{Tree}(S)$ be the family of dyadic cubes contained in S, but not in any $S' \in \mathscr{A}^\star(S)$

$$\mathsf{Tree}(S) \coloneqq \{R \in \mathscr{D} \,:\, \widehat{R} = S\}.$$

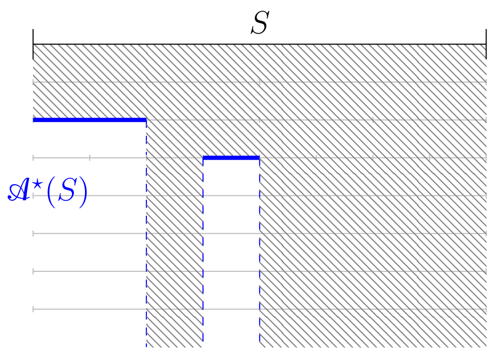

Figure 3.2: Example of $\mathsf{Tree}(S)$.

Also, we define $\mathsf{Tree}_r(S) \coloneqq \{R \in \mathscr{D} \,:\, \widehat{R^{(r)}} = S\}$. Note that the maximal cubes in $\mathsf{Tree}_r(S)$ are the r-grandchildren of S. See Figure B.1 in the appendix.

3.5.2 Reduction to a telescoping sum

We follow the decomposition in [LM17a; MM14] where the sum $\sum_{P \in \mathscr{P}_R} \langle \Delta_P f \rangle_{P_R} \mathbb{1}_{P_R}$ is decomposed in a telescopic sum plus off-diagonal terms. The off-diagonal terms are then bounded by a sum of the dyadic forms $B_j^{\mathscr{D}}(g, f)$ or directly by a sparse form.

Given $S \in \mathscr{S}$ such that $S \supset P_R$, the indicator function $\mathbb{1}_{P_R}$ can be written as $\mathbb{1}_S - \mathbb{1}_{S \setminus P_R}$. Recall that $\widehat{P_R}$ is the minimal stopping cube containing P_R. Then

$$\langle \Delta_P f \rangle_{P_R} \mathbb{1}_{P_R} = \begin{cases} \langle \Delta_P f \rangle_{P_R} \mathbb{1}_{\widehat{P_R}} - \langle \Delta_P f \rangle_{P_R} \mathbb{1}_{\widehat{P_R} \setminus P_R} & \text{if } P_R \notin \mathscr{S} \qquad (3.5.2) \\ \langle \Delta_P f \rangle_{P_R} \mathbb{1}_{\widehat{P_R}} & \text{if } P_R \in \mathscr{S} \qquad (3.5.3) \end{cases}$$

and in the latter case we have

$$\langle \Delta_P f \rangle_{P_R} \mathbb{1}_{\widehat{P_R}} = \mathbb{1}_{\widehat{P_R}} \langle f \rangle_{P_R} - \mathbb{1}_{\widehat{P_R}} \langle f \rangle_P = (\mathbb{1}_{\widehat{P_R}} \langle f \rangle_{P_R} - \mathbb{1}_{\widehat{P}} \langle f \rangle_P) + \mathbb{1}_{\widehat{P} \setminus \widehat{P_R}} \langle f \rangle_P.$$

The term $\langle \Delta_P f \rangle_{P_R} \mathbb{1}_{\widehat{P_R} \setminus P_R}$ is supported away from R, so one can use off-diagonal estimates as in (3.4.1). Also notice that in the bound (3.4.1) and in its proof one can replace $|g|$ by $\langle |g| \rangle_R$. In the same way, off-diagonal estimates are used for $\mathbb{1}_{\widehat{P} \setminus \widehat{P_R}} \langle f \rangle_P$ as shown in Lemma 3.5.4 below.

The terms $\langle \Delta_P f\rangle_{P_R} \mathbb{1}_{\widehat{P_R}}$ and $\mathbb{1}_{\widehat{P_R}}\langle f\rangle_{P_R} - \mathbb{1}_{\widehat{P}}\langle f\rangle_P$ left from (3.5.2) and (3.5.3) are rearranged to obtain a telescopic series. We have

$$\begin{aligned} \mathbb{1}_{\widehat{P_R}}\langle \Delta_P f\rangle_{P_R} = \mathbb{1}_{\widehat{P_R}}\langle f\rangle_{P_R} - \mathbb{1}_{\widehat{P_R}}\langle f\rangle_P \qquad & \text{when } P_R \notin \mathscr{S} \\ \text{and} \quad \mathbb{1}_{\widehat{P_R}}\langle f\rangle_{P_R} - \mathbb{1}_{\widehat{P}}\langle f\rangle_P \qquad & \text{when } P_R \in \mathscr{S}. \end{aligned}$$

If $P_R \notin \mathscr{S}$ then P and P_R are contained in the same minimal stopping cube $\widehat{P}$. So $\widehat{P_R} = \widehat{P}$ and the two cases add up to $2(\mathbb{1}_{\widehat{P_R}}\langle f\rangle_{P_R} - \mathbb{1}_{\widehat{P}}\langle f\rangle_P)$ which leads to the telescopic sum

$$\sum_{\substack{P\in\mathscr{D} \\ R^{(r)}\subset P\subseteq Q_0}} \mathbb{1}_{\widehat{P_R}}\langle f\rangle_{P_R} - \mathbb{1}_{\widehat{P}}\langle f\rangle_P = \mathbb{1}_{\widehat{R^{(r)}}}\langle f\rangle_{R^{(r)}} - \mathbb{1}_{\widehat{Q_0}}\langle f\rangle_{Q_0}. \tag{3.5.4}$$

Since f is supported on a fixed Q_0, the average on larger cubes $Q_0^{(n)}$ containing Q_0 decreases:

$$\langle f\rangle_{Q_0^{(n)}} = \frac{1}{|Q_0^{(n)}|}\int_{Q_0} f \leq \frac{1}{|Q_0^{(n)}|}\|f\|_{L^1} \to 0 \quad \text{ as } n\to\infty.$$

Thus when the sum in (3.5.4) extends to all $P \supset R^{(r)}$, the term $\mathbb{1}_{\widehat{R^{(r)}}}\langle f\rangle_{R^{(r)}}$ is the only one remaining.

We have then identified three terms

$$\sum_{\substack{P\in\mathscr{D} \\ P\supset R^{(r)}}} \langle \Delta_P f\rangle_{P_R}\mathbb{1}_{P_R} = \sum_{\text{telescopic}} - \sum_{\text{far}} + \sum_{\text{sparse}}$$

where

$$\sum_{\text{far}} := \sum_{P\,:\,P\supset R^{(r)}} \langle \Delta_P f\rangle_{P_R}\mathbb{1}_{\widehat{P_R}\setminus P_R}, \qquad \sum_{\text{sparse}} := \sum_{\substack{P\,:\,P\supset R^{(r)} \\ P_R\in\mathscr{S}}} \mathbb{1}_{\widehat{P}\setminus P_R}\langle f\rangle_P$$

$$\text{and} \quad \sum_{\text{telescopic}} := \sum_{P\supset R^{(r)}} 2(\mathbb{1}_{\widehat{P_R}}\langle f\rangle_{P_R} - \mathbb{1}_{\widehat{P}}\langle f\rangle_P) = \mathbb{1}_{\widehat{R^{(r)}}}\langle f\rangle_{R^{(r)}}.$$

Since the case with $\sum_{\text{far}}$ is done in (3.4.1), we show how to deal with the remaining two cases.

3.5.3 Bound by a sparse form

We bound the operator applied to $\mathbb{1}_{\widehat{R^{(r)}}}\langle f\rangle_{R^{(r)}}$ and $\mathbb{1}_{\widehat{P}\setminus P_R}\langle f\rangle_P$.

Lemma 3.5.3. *Let $\mathscr{S}$ be the sparse collection defined in §3.5.1, then*

$$\sum_{R\in\mathscr{D}}\iint_{W_R}\langle|g|\rangle_R|\theta_t\mathbb{1}_{\widehat{R^{(r)}}}(x)|^2\langle f\rangle_{R^{(r)}}^2\frac{\mathrm{d}t}{t}\,\mathrm{d}x\lesssim_{r,d}C_{\mathsf{T}}\sum_{S\in\mathscr{S}}\langle|g|\rangle_S\langle|f|\rangle_S^2|S|.$$

Proof. The set $\{\mathsf{Tree}_r(S)\colon S\in\mathscr{S}\}$ is a partition of $\mathscr{D}$, so we write

$$\begin{aligned}
&\sum_{R\in\mathscr{D}}\iint_{W_R}\langle|g|\rangle_R|\theta_t\mathbb{1}_{\widehat{R^{(r)}}}(x)|^2\langle f\rangle_{R^{(r)}}^2\frac{\mathrm{d}t}{t}\,\mathrm{d}x\\
&\qquad=\sum_{S\in\mathscr{S}}\sum_{R\,:\,\widehat{R^{(r)}}=S}2^{rd}\langle|g|\rangle_{R^{(r)}}\langle f\rangle_{R^{(r)}}^2\iint_{W_R}|\theta_t\mathbb{1}_S(x)|^2\frac{\mathrm{d}t}{t}\,\mathrm{d}x\\
&\qquad\lesssim_{r,d}\sum_{S\in\mathscr{S}}\langle|g|\rangle_S\langle|f|\rangle_S^2\sum_{R\,:\,R\subset S}\iint_{W_R}|\theta_t\mathbb{1}_S(x)|^2\frac{\mathrm{d}t}{t}\,\mathrm{d}x\\
&\qquad=\sum_{S\in\mathscr{S}}\langle|g|\rangle_S\langle|f|\rangle_S^2\int_S\int_0^{\ell S}|\theta_t\mathbb{1}_S(x)|^2\frac{\mathrm{d}t}{t}\,\mathrm{d}x\\
&\qquad\le C_{\mathsf{T}}\sum_{S\in\mathscr{S}}\langle|g|\rangle_S\langle|f|\rangle_S^2|S|
\end{aligned}$$

where we used the stopping conditions for f and g, and the testing condition (T). □

Lemma 3.5.4. *Let $\mathscr{S}$ be the sparse collection defined in §3.5.1, then*

$$\sum_{R\in\mathscr{D}}\iint_{W_R}\left|\sum_{\substack{P:P\supset R^{(r)}\\P_R\in\mathscr{S}}}\theta_t(\mathbb{1}_{\widehat{P}\setminus P_R})\langle f\rangle_P\right|^2|g|\frac{\mathrm{d}t}{t}\,\mathrm{d}x\lesssim\sum_{S\in\mathscr{S}'}\langle|f|\rangle_S^2\langle|g|\rangle_S|S|\tag{3.5.5}$$

where $\mathscr{S}'$ is the sparse collection of dyadic parents of $\mathscr{S}$.

Proof. Since $P\supset R^{(r)}$, the dyadic child $P_R=R^{(k)}$ for some integer $k\ge r$. For $(x,t)\in W_R$, an application of Poisson off-diagonal estimates (Lemma 3.4.2) gives

$$\theta_t(\mathbb{1}_{\widehat{P}\setminus P_R})(x)=\theta_t(\mathbb{1}_{\widehat{R^{(k+1)}}\setminus R^{(k)}})(x)\lesssim(\ell R/\ell R^{(k)})^\eta=2^{-k\eta}.$$

After applying Cauchy–Schwarz the sums are rearranged using P as the common ancestor:

$$\begin{aligned}\sum_{R\in\mathscr{D}}\int_R|g|\sum_{\substack{P\,:\,P\supset R^{(r)}\\ \widehat{P_R}=P_R}}\langle f\rangle_P^2\left(\frac{\ell R}{\ell P_R}\right)^\eta &= \sum_{k\geq r}2^{-k\eta}\sum_{\substack{P\in\mathscr{D}\\ \text{with } P_R\in\mathscr{S}}}\langle f\rangle_P^2\sum_{\substack{R\,:\,R\subset P\\ \ell R=2^{-k-1}\ell P}}\int_R|g|\\ &= \sum_{k\geq r}2^{-k\eta}\sum_{P:P_R\in\mathscr{S}}\langle f\rangle_P^2\int_P|g|\\ &\leq \sum_{P:P_R\in\mathscr{S}}\langle|f|\rangle_P^2\int_P|g|.\end{aligned}$$

Let $\mathscr{S}'$ be the collection $\{P\in\mathscr{D}\,:\,P\supset S,\,\ell P=2\ell S$ for some $S\in\mathscr{S}\}$. If $\mathscr{S}$ is τ-sparse, then $\mathscr{S}'$ is $\tau 2^{-d}$-sparse. This establishes (3.5.5) and concludes the proof. □

The sparse collection in (3.5.1) can be taken as the union of $\mathscr{S}'$ and the stopping family in §3.5.1.

3.6 Reduction of (III$_b$) to a sparse form

In this section we show that there exists $c>0$ and a sparse family $\widetilde{\mathscr{S}}$ such that

$$\begin{aligned}\sum_{R\in\mathscr{D}}\iint_{W_R}\Big|\sum_{P\in\mathscr{P}_R}\theta_t\langle\Delta_P f\rangle_{P_R}\mathbb{1}_{P_R}\Big|^2\sum_{\substack{Q\in\mathscr{D}\\ Q\text{ good},Q\subset R}}\Delta_Q b(x)\frac{\mathrm{d}t}{t}\,\mathrm{d}x\\ \lesssim\sum_{j\in\mathbb{N}}2^{-cj}B_j^{\mathscr{D}}(g,f)+\Lambda_{\widetilde{\mathscr{S}}}(g,f).\end{aligned}$$

In order to exploit the goodness of Q, for example via Poisson off-diagonal estimates as in Lemma 3.4.2, we need a gap of at least r generations between Q and P_R. This motivates the Calderón–Zygmund decomposition in Proposition 3.1.2. In particular, since b is the bad part of g at height $\lambda=A\langle|g|\rangle_R$ given by Proposition 3.1.2, we have

$$\sum_{\substack{Q\in\mathscr{D}\\ Q\subset R}}\Delta_Q b=\sum_{L\in\mathscr{L}}\sum_{L_r\in\mathsf{ch}_r(L)}\sum_{\substack{Q\in\mathscr{D}\\ Q\subseteq L_r}}\Delta_Q b_{L_r}.$$

Since $A > 1$, the cubes in $\mathscr{L}$ are strictly contained in R. If we choose the constant A as in the construction of the stopping family in §3.5.1, then the cubes in $\mathscr{L}$ are also stopping cubes in $\mathscr{S}_g$. We can regroup the dyadic cubes $Q \subseteq L_r$ in the stopping trees $\mathsf{Tree}_r(S)$ for all $S \in \mathscr{S}_g$ inside R.

$$\sum_{L\in\mathscr{L}} \sum_{L_r\in\mathsf{ch}_r(L)} \sum_{\substack{Q\in\mathscr{D}\\ Q\subseteq L_r}} \Delta_Q b_{L_r} = \sum_{\substack{S\in\mathscr{S}_g\\ S\subset R}} \sum_{S_r\in\mathsf{ch}_r(S)} \sum_{\substack{Q\in\mathsf{Tree}_r(S)\\ Q\subseteq S_r}} \Delta_Q b_{S_r}.$$

The last sum is the Haar projection of b on $\operatorname{Span}\{h_Q : Q \in \mathsf{Tree}_r(S), Q \subseteq S_r\}$. We denote this quantity by

$$\mathcal{P}_{S_r}(b) := \sum_{\substack{Q\in\mathsf{Tree}_r(S)\\ Q\subseteq S_r}} \Delta_Q b_{S_r}.$$

Remark 3.6.1. The Haar projection $\mathcal{P}_{S_r}b$ is supported on S_r and equals $\mathcal{P}_{S_r}(|g|)$. Indeed $b_{S_r} = \mathbb{1}_{S_r}(|g| - \langle|g|\rangle_{S_r})$ and for $Q \subseteq S_r$ the Haar coefficient $\langle b_{S_r}, h_Q\rangle = \langle|g|, h_Q\rangle$.

We have then proved the following identity

$$(\mathrm{III}_b) = \sum_{S\in\mathscr{S}_g} \sum_{\substack{S_r\in\mathsf{ch}_r(S)\\ S_r \text{ good}}} \sum_{R:R\supset S} \iint_{W_R} \Big|\sum_{P\in\mathscr{P}_R} \theta_t\langle\Delta_P f\rangle_{P_R}\mathbb{1}_{P_R}\Big|^2 \mathcal{P}_{S_r}(|g|)\frac{\mathrm{d}t}{t}\,\mathrm{d}x.$$

With a slight abuse of notation, we omit the subscript in the stopping family $\mathscr{S}_g$ in the following.

Remark 3.6.2 (Estimates for $\mathcal{P}_{S_r}$). The Haar projection $\mathcal{P}_{S_r}(|g|)$ has zero average and satisfies the following bound

$$\|\mathcal{P}_{S_r}g\|_{L^1} \lesssim |S_r|\langle|g|\rangle_S. \tag{3.6.1}$$

A proof of (3.6.1) is in §B.0.4. In particular, summing over all $S_r \in \mathsf{ch}_r(S)$ gives

$$\sum_{S_r\in\mathsf{ch}_r(S)} \|\mathcal{P}_{S_r}g\|_{L^1} \lesssim \sum_{S_r\in\mathsf{ch}_r(S)} |S_r|\langle|g|\rangle_S \le \int_S |g|. \tag{3.6.2}$$

3.6.1 Recover decay and telescopic sum

Let P_S be the dyadic child of P containing S. Then

$$\sum_{P:P\supset R^{(r)}} \langle \Delta_P f\rangle_{P_R} \mathbb{1}_{P_R} = \sum_{P:P\supset S} \langle \Delta_P f\rangle_{P_S} \mathbb{1}_{P_S} - \sum_{P:S\subset P\subseteq R^{(r)}} \langle \Delta_P f\rangle_{P_S} \mathbb{1}_{P_S}.$$

The second term can be handled as in the subscale case (§3.3.6), while the first can be reduced to a telescopic sum which equals $\langle f\rangle_S \mathbb{1}_S$.

If one tries to reduce $\langle \Delta_P f\rangle_{P_S} \mathbb{1}_{P_S}$ to a telescopic term plus off-diagonal terms as in §3.5.2, the off-diagonal factor which should provide decay is the quantity

$$\int_{\mathbb{R}^d\setminus P_S} \frac{(\ell R)^\alpha}{\mathrm{d}(y, S_r)^{\alpha+d}} \,\mathrm{d}y.$$

Here the scale (numerator) and the distance (denominator) don't match and Lemma 3.4.2 seems unable to provide enough decay in order to handle the integral *and* the sum over R. But the zero average property of $\mathcal{P}_{S_r}(g)$ comes to the rescue bringing a factor $(\ell S_r)^{\alpha/2}$ at the numerator by exploiting the smoothness condition of the kernel. We will explain how.

Let x_{S_r} be the centre of the S_r and consider the sublinear operator

$$K_t^{S_r} f(x) := \int_{\mathbb{R}^d} \frac{(t|x - x_{S_r}|)^{\alpha/2}}{(t + |x-y|)^{\alpha+d}} |f(y)| \,\mathrm{d}y$$

Since the Haar projection $\mathcal{P}_{S_r}(g)$ is supported on S_r, we have the following bound.

Lemma 3.6.3. *Let S_r and R be dyadic cubes with $S_r \subset R$, then*

$$\iint_{W_R} |\theta_t f(x)|^2 \frac{\mathrm{d}t}{t} \mathcal{P}_{S_r}(g)(x) \,\mathrm{d}x \lesssim \iint_{W_R} \left(K_t^{S_r} f(x)\right)^2 \frac{\mathrm{d}t}{t} |\mathcal{P}_{S_r}(g)(x)| \,\mathrm{d}x. \tag{3.6.3}$$

Proof. The idea is to use the zero average of $\mathcal{P}_{S_r}(g)$ to exploit the smoothness condition

(C2). We recall that $\mathcal{P}_{S_r}(g)$ is supported on $S_r \subset R$. Consider the operator

$$\mathcal{K} f(x) := \int_{\ell R/2}^{\ell R} \left| \int k_t(x,y) f(y) \, \mathrm{d}y \right|^2 \frac{\mathrm{d}t}{t}$$

so that the left hand side of (3.6.3) equals $\int \mathcal{K} f(x) \mathcal{P}_{S_r} g(x) \, \mathrm{d}x$. Let x_{S_r} be the centre of S_r. Then

$$\int \mathcal{K} f(x) \mathcal{P}_{S_r} g(x) \, \mathrm{d}x = \int \big(\mathcal{K} f(x) - \mathcal{K} f(x_{S_r})\big) \mathcal{P}_{S_r} g(x) \, \mathrm{d}x$$

and the difference $\mathcal{K} f(x) - \mathcal{K} f(x_{S_r})$ can be factorised as

$$\begin{aligned}
&\int_{\ell R/2}^{\ell R} \left| \int k_t(x,y) f(y) \, \mathrm{d}y \right|^2 - \left| \int k_t(x_{S_r},y) f(y) \, \mathrm{d}y \right|^2 \frac{\mathrm{d}t}{t} \\
&= \int_{\ell R/2}^{\ell R} \left(\int [k_t(x,y) - k_t(x_{S_r},y)] f(y) \, \mathrm{d}y \right) \left(\int [k_t(x,y) + k_t(x_{S_r},y)] f(y) \, \mathrm{d}y \right) \frac{\mathrm{d}t}{t} \\
&=: \int_{\ell R/2}^{\ell R} \mathcal{K}_{S_r}^{-} f(x) \cdot \mathcal{K}_{S_r}^{+} f(x) \frac{\mathrm{d}t}{t}.
\end{aligned}$$

For $x \in S_r$, since $S_r \subset R$ and $t \in (\ell R/2, \ell R)$, the distance $|x - x_{S_r}| \le \ell S_r/2 < \ell R/2 < t$, so by conditions (C2) and (C1) we have

$$\begin{aligned}
\mathcal{K}_{S_r}^{-} f(x) \cdot \mathcal{K}_{S_r}^{+} f(x) &\lesssim \int \frac{|x - x_{S_r}|^{\alpha}}{(t + |x-y|)^{\alpha+d}} |f(y)| \, \mathrm{d}y \cdot \int \frac{t^{\alpha}}{(t + |x-y|)^{\alpha+d}} |f(y)| \, \mathrm{d}y \\
&= \left(\int \frac{(t|x - x_{S_r}|)^{\alpha/2}}{(t + |x-y|)^{\alpha+d}} |f(y)| \, \mathrm{d}y \right)^2 =: \big(K_t^{S_r} f(x)\big)^2.
\end{aligned}$$

□

The operator $K_t^{S_r}$ satisfies Poisson-like off-diagonal estimates.

Lemma 3.6.4 (Estimates for $K_t^{S_r}$)**.** *Let $x \in S_r \subset R$ and $t \in (\ell R/2, \ell R)$. Let $Q \in \mathcal{D}$ such that $Q \supset S_r$. Then there exists $\eta > 0$ such that the following estimates hold:*

$$K_t^{S_r} \mathbb{1}_{\mathbb{R}^d \setminus Q}(x) \lesssim \left(\frac{\ell S_r}{\max(\ell R^{(r)}, \ell Q)} \right)^{\eta}, \qquad K_t^{S_r} \mathbb{1}_{Q}(x) \lesssim \frac{|Q|}{|R|} \left(\frac{\ell S_r}{\ell R} \right)^{\alpha/2}.$$

Remark 3.6.5. Notice that the first estimate is better than the one in Lemma 3.4.2 on smaller scale (when $\ell Q < \ell R^{(r)}$). For the second one, since $\ell S_r < \ell R$, we can also estimate

$$K_t^{S_r}\mathbb{1}_Q(x) \lesssim \frac{|Q|}{|R|}$$

provided that $x \in S_r$ and $t \in (\ell R/2, \ell R)$.

Proof of Lemma 3.6.4. For the second estimate, by forgetting the distance in the denominator, we simply have

$$K_t^{S_r}(\mathbb{1}_Q)(x) \lesssim \int_Q \frac{(\ell S_r \ell R)^{\alpha/2}}{(\ell R + \mathrm{d}(y, S_r))^{\alpha+d}}\,\mathrm{d}y \leq \frac{|Q|}{|R|}\left(\frac{\ell S_r}{\ell R}\right)^{\alpha/2}.$$

For the first estimate, when $Q \supset R^{(r)}$ use $(a+b)^\alpha = (a+b)^{2\alpha/2} \geq (2ab)^{\alpha/2}$ in order to apply off-diagonal estimates. For $x \in S_r$ and $t \in (\ell R/2, \ell R)$ we bound

$$\begin{aligned} K_t^{S_r}\mathbb{1}_{\mathbb{R}^d\setminus Q}(x) &\lesssim \int_{\mathbb{R}^d\setminus Q} \frac{(\ell S_r \ell R)^{\alpha/2}}{(\ell R + \mathrm{d}(y, S_r))^{\alpha+d}}\,\mathrm{d}y \\ &\lesssim \int_{\mathbb{R}^d\setminus Q} \frac{(\ell S_r \ell R)^{\alpha/2}}{(\ell R \cdot \mathrm{d}(y, S_r))^{\alpha/2}} \frac{\mathrm{d}y}{\mathrm{d}(y, S_r)^d} = \int_{\mathbb{R}^d\setminus Q} \frac{(\ell S_r)^{\alpha/2}}{\mathrm{d}(y, S_r)^{\alpha/2+d}}\,\mathrm{d}y. \end{aligned} \tag{3.6.4}$$

Then apply Lemma 3.4.2 with $\beta = \alpha/2$

$$\int_{\mathbb{R}^d\setminus Q} \frac{(\ell S_r)^{\alpha/2}}{\mathrm{d}(y, S_r)^{\alpha/2+d}}\,\mathrm{d}y \lesssim \left(\frac{\ell S_r}{\ell Q}\right)^\eta.$$

When $S_r \subset Q \subset R^{(r)}$, split $\mathbb{1}_{\mathbb{R}^d\setminus Q}$ as $\mathbb{1}_{\mathbb{R}^d\setminus R^{(r)}} + \mathbb{1}_{R^{(r)}\setminus Q}$. Estimate $K_t^{S_r}(\mathbb{1}_{\mathbb{R}^d\setminus R^{(r)}})$ as in (3.6.4). Then applying Lemma 3.4.2 with $\beta = \alpha/2$ gives

$$K_t^{S_r}(\mathbb{1}_{\mathbb{R}^d\setminus R^{(r)}})(x) \lesssim \left(\frac{\ell S_r}{\ell R^{(r)}}\right)^\eta$$

where η is positive and equals $\frac{\alpha}{2} - \gamma(\frac{\alpha}{2} + d) < \frac{\alpha}{2}$. For $K_t^{S_r}(\mathbb{1}_{R^{(r)}\setminus Q})$ we bound

$$K_t^{S_r}(\mathbb{1}_{R^{(r)}\setminus Q})(x) \lesssim \int_{R^{(r)}} \frac{(\ell S_r \ell R)^{\alpha/2}}{(\ell R)^{\alpha/2}(\ell R)^{\alpha/2+d}}\,\mathrm{d}y$$

$$\leq \frac{|R^{(r)}|}{|R|}\left(\frac{\ell S_r}{\ell R}\right)^{\alpha/2} \lesssim_{r,d} \left(\frac{\ell S_r}{\ell R}\right)^{\alpha/2} = 2^{r\alpha/2}\left(\frac{\ell S_r}{\ell R^{(r)}}\right)^{\alpha/2}.$$

Adding the two bounds gives

$$K_t^{S_r}\mathbb{1}_{\mathbb{R}^d\setminus Q}(x) \lesssim \left(\frac{\ell S_r}{\ell R^{(r)}}\right)^{\eta} + \left(\frac{\ell S_r}{\ell R^{(r)}}\right)^{\alpha/2} \leq 2\left(\frac{\ell S_r}{\ell R^{(r)}}\right)^{\eta}$$

since $\ell S_r < \ell R^{(r)}$ and $\min(\eta, \alpha/2) = \eta$. $\square$

3.6.2 Reduction to telescopic: different terms

Apply Lemma 3.6.3 with $\sum \langle \Delta_P f\rangle_{P_R}\mathbb{1}_{P_R}$ in place of f to obtain

$$(\mathrm{III}_b) \lesssim \sum_{S\in\mathscr{S}} \sum_{\substack{S_r\in\mathsf{ch}_r(S)\\ S_r \text{ good}}} \sum_{R:R\supset S} \iint_{W_R} \left(K_t^{S_r} \sum_{P:P\supset R^{(r)}} \langle \Delta_P f\rangle_{P_R}\mathbb{1}_{P_R}\right)^2 \frac{\mathrm{d}t}{t}|\mathcal{P}_{S_r} g|\,\mathrm{d}x.$$

We split the sum in P to obtain a telescopic sum as in §3.5.2, with an extra subscale term:

$$\sum_{P\supset R^{(r)}} \langle \Delta_P f\rangle_{P_R}\mathbb{1}_{P_R} = \sum_{\text{telescopic}} - \sum_{\text{far}} + \sum_{\text{sparse}} - \sum_{\text{subscale}}$$

where

$$\sum_{\text{telescopic}} := 2\sum_{P:P\supset S} \left(\langle f\rangle_{P_S}\mathbb{1}_{\widehat{P_S}} - \langle f\rangle_{P}\mathbb{1}_{\widehat{P}}\right) = 2\langle f\rangle_S \mathbb{1}_S$$

$$\sum_{\text{sparse}} := \sum_{\substack{P:P\supset S\\ P_S\in\mathscr{S}}} \langle f\rangle_P \mathbb{1}_{\widehat{P}\setminus P_S}$$

$$\sum_{\text{subscale}} := \sum_{\substack{P:P\subseteq R^{(r)}\\ P\supset S}} \langle \Delta_P f\rangle_{P_S}\mathbb{1}_{P_S}$$

$$\sum_{\text{far}} := \sum_{P:P\supset S} \langle \Delta_P f\rangle_{P_S}\mathbb{1}_{\widehat{P_S}\setminus P_S}.$$

Then we bound

$$\Big|\sum_{\text{telescopic}} - \sum_{\text{far}} + \sum_{\text{sparse}} - \sum_{\text{subscale}}\Big| \le \Big|\sum_{\text{telescopic}}\Big| + \Big|\sum_{\text{far}}\Big| + \Big|\sum_{\text{sparse}}\Big| + \Big|\sum_{\text{subscale}}\Big|.$$

We estimate $K_t^{S_r}$ applied to each term by using sublinearity and Lemma 3.6.4. Then take the supremum in t on the Whitney region W_R to bound the remaining integral $\int_{\ell R/2}^{\ell R} \mathrm{d}t/t$ by 1.

We give the details in each case.

3.6.3 Telescopic term

This case is bounded by the sparse form $\Lambda_{\mathscr{S}}(f,g) = \sum_{S\in\mathscr{S}} \langle |f| \rangle_S^2 \int_S |g|$, where $\mathscr{S}$ is the stopping family of g.

Lemma 3.6.6. *It holds that*

$$\sum_{S\in\mathscr{S}} \sum_{S_r \in \mathsf{ch}_r(S)} \sum_{R\,:\,R\supset S} \iint_{W_R} \langle f \rangle_S^2 \left(K_t^{S_r} \mathbb{1}_S\right)^2 \frac{\mathrm{d}t}{t} |\mathcal{P}_{S_r} g| \,\mathrm{d}x \lesssim \Lambda_{\mathscr{S}}(f,g)$$

Proof. For $x \in S_r$ and $t \in (\ell R/2, \ell R)$ we estimate $K_t^{S_r}(\mathbb{1}_S)(x) \lesssim |S|/|R|$ and $\int_{\ell R/2}^{\ell R} \mathrm{d}t/t$ by 1. Then by using (3.6.2) for the Haar projection we have

$$\begin{aligned}
&\sum_{S\in\mathscr{S}} \sum_{S_r \in \mathsf{ch}_r(S)} \langle f \rangle_S^2 \sum_{R:R\supset S} \iint_{W_R} \left(K_t^{S_r} \mathbb{1}_S(x)\right)^2 \frac{\mathrm{d}t}{t} |\mathcal{P}_{S_r} g(x)| \,\mathrm{d}x \\
&\lesssim \sum_{S\in\mathscr{S}} \sum_{S_r \in \mathsf{ch}_r(S)} \langle f \rangle_S^2 \sum_{R:R\supset S} \left(\frac{|S|}{|R|}\right)^2 \|\mathcal{P}_{S_r} g\|_{L^1} \\
&\lesssim_{r,d} \sum_{S\in\mathscr{S}} \langle f \rangle_S^2 \sum_{R:R\supset S} \left(\frac{|S|}{|R|}\right)^2 \int_S |g| \le \sum_{S\in\mathscr{S}} \langle |f| \rangle_S^2 \int_S |g|.
\end{aligned}$$

□

3.6.4 Subscale term

This term is bounded in a similar way as in the subscale case in §3.3.6.

Lemma 3.6.7. *It holds that*

$$\sum_{S\in\mathscr{S}}\sum_{S_r\in\mathsf{ch}_r(S)}\sum_{R:\,R\supset S}\iint_{W_R}\Big(K_t^{S_r}\sum_{\text{subscale}}\Big)^2\frac{\mathrm{d}t}{t}|\mathcal{P}_{S_r}g|\,\mathrm{d}x\lesssim\sum_{j\in\mathbb{N}}2^{-j\alpha/4}B_j^{\mathscr{D}}(g,f).$$

Proof. First, since $K_t^{S_r}$ is sublinear, we bound

$$K_t^{S_r}\Big(\sum_{\text{subscale}}\Big)\leq\sum_{\substack{P:\,P\subseteq R^{(r)}\\P\supset S}}|\langle\Delta_P f\rangle_{P_S}|K_t^{S_r}(\mathbb{1}_{P_S}).$$

Then for $x\in S_r$ and $t\in(\ell R/2,\ell R)$ we estimate $K_t^{S_r}\mathbb{1}_{P_S}$ using Lemma 3.6.4

$$K_t^{S_r}\mathbb{1}_{P_S}(x)\lesssim\left(\frac{\ell P_S}{\ell R}\right)^d\left(\frac{\ell S_r}{\ell R}\right)^{\alpha/2}.$$

Bound $\ell P_S<\ell P$ and $|\langle\Delta_P f\rangle_{P_S}|\leq|\langle f,h_P\rangle||P|^{-1/2}$, then we apply the Cauchy–Schwarz inequality

$$\begin{aligned}
&\sum_{S\in\mathscr{S}}\sum_{S_r\in\mathsf{ch}_r(S)}\sum_{R:\,R\supset S}\int_R\Bigg(\sum_{\substack{P:\,P\subseteq R^{(r)}\\P\supset S}}|\langle f,h_P\rangle|\frac{|P|^{1/2}}{|R|}\Bigg)^2\left(\frac{\ell S_r}{\ell R}\right)^{\alpha}|\mathcal{P}_{S_r}g|\,\mathrm{d}x\\
&\leq\sum_{S\in\mathscr{S}}\sum_{S_r\in\mathsf{ch}_r(S)}\sum_{R\supset S}\Bigg(\sum_{\substack{P:\,P\subseteq R^{(r)}\\P\supset S}}\frac{\langle f,h_P\rangle^2}{|R|}\Bigg)\Bigg(\sum_{P\subseteq R^{(r)}}\frac{|P|}{|R|}\Bigg)\left(\frac{\ell S_r}{\ell R}\right)^{\alpha}\|\mathcal{P}_{S_r}g\|_{L^1}\\
&\leq\sum_{S\in\mathscr{S}}\sum_{S_r\in\mathsf{ch}_r(S)}\sum_{R\supset S}\left(\frac{\ell S}{\ell R}\right)^{\alpha/2}\\
&\qquad\Bigg(\sum_{\substack{P:P\supset S\\P\subseteq R^{(r)}}}\frac{\langle f,h_P\rangle^2}{|R|}\left(\frac{\ell P}{\ell R}\right)^{\alpha/4}\Bigg)\Bigg(\sum_{P\subseteq R^{(r)}}\frac{|P|}{|R|}\left(\frac{\ell P}{\ell R}\right)^{\alpha/4}\Bigg)\|\mathcal{P}_{S_r}g\|_{L^1}.
\end{aligned}$$

The second factor after Cauchy–Schwarz is controlled as in subscale case in Lemma 3.3.3

where $P \subset 3R$. Then bound $\|\mathcal{P}_{S_r} g\|_{L^1}$ as in (3.6.2) to obtain

$$\begin{aligned}
&\sum_{S\in\mathscr{S}} \int_S |g| \sum_{R\,:\,R\supset S} \left(\frac{\ell S}{\ell R}\right)^{\alpha/2} \sum_{P\,:\,S\subset P\subseteq R^{(r)}} \frac{\langle f, h_P\rangle^2}{|R|} \left(\frac{\ell P}{\ell R}\right)^{\alpha/4} \\
&= \sum_{R\in\mathscr{D}} \frac{1}{|R|} \sum_{\substack{S\in\mathscr{S}\\ S\subset R}} \int_S |g| \left(\frac{\ell S}{\ell R}\right)^{\alpha/2} \sum_{P\,:\,S\subset P\subseteq R^{(r)}} \langle f, h_P\rangle^2 \left(\frac{\ell P}{\ell R}\right)^{\alpha/4}.
\end{aligned}$$

For $i, j \in \mathbb{N}$, let $\ell P = 2^{-j}\ell R^{(r)}$ and $\ell S = 2^{-i}\ell R$. Extend the sum over all P such that $P \subseteq R^{(r)}$ and rearrange

$$\begin{aligned}
&\sum_{R\in\mathscr{D}} \frac{1}{|R|} \sum_{\substack{S\in\mathscr{S}\\ S\subset R}} \int_S |g| \left(\frac{\ell S}{\ell R}\right)^{\alpha/2} \sum_{P\,:\,S\subset P\subseteq R^{(r)}} \langle f, h_P\rangle^2 \left(\frac{\ell P}{\ell R}\right)^{\alpha/4} \\
&= \sum_{i,j} \sum_{R\in\mathscr{D}} 2^{-i\alpha/2} \frac{1}{|R|} \sum_{\substack{S\subset R\\ \ell S = 2^{-i}\ell R}} \int_S |g| \sum_{\substack{P\subseteq R^{(r)}\\ \ell P = 2^{-j}\ell R^{(r)}}} \langle f, h_P\rangle^2 \left(\frac{\ell P}{\ell R^{(r)}}\right)^{\alpha/4} 2^{r\alpha/4} \\
&\lesssim_{r,d} \sum_{i,j} 2^{-i\alpha/2} 2^{-j\alpha/4} \sum_{R\in\mathscr{D}} \fint_R |g| \sum_{\substack{P\subseteq R^{(r)}\\ \ell P = 2^{-j}\ell R^{(r)}}} \langle f, h_P\rangle^2 \\
&\lesssim_{r,d} \sum_{j\in\mathbb{N}} 2^{-j\alpha/4} \sum_{R^{(r)}\in\mathscr{D}} \fint_{R^{(r)}} |g| \sum_{\substack{P\subseteq R^{(r)}\\ \ell P = 2^{-j}\ell R^{(r)}}} \langle f, h_P\rangle^2 \\
&\le 3^d \sum_{j\in\mathbb{N}} 2^{-j\alpha/4} B_j^{\mathscr{D}}(g, f).
\end{aligned}$$

□

3.6.5 Far and Sparse terms

In this subsection we show that the quantity

$$\sum_{S\in\mathscr{S}} \sum_{S_r\in\mathsf{ch}_r(S)} \sum_{R\,:\,R\supset S} \iint_{W_R} \left(K_t^{S_r}\Big(\sum_{\text{far}} + \sum_{\text{sparse}}\Big)\right)^2 \frac{\mathrm{d}t}{t} |\mathcal{P}_{S_r} g| \,\mathrm{d}x$$

is bounded, up to a constant, by the sum of $B_0^{\mathscr{D}}(g,f)$ and $\Lambda_{\mathscr{S}}(f,g)$.

Since $K_t^{S_r}$ is sublinear and positive, we bound

$$\begin{aligned} K_t^{S_r}\Big(\sum_{\text{far}}+\sum_{\text{sparse}}\Big) &\leq \sum_{P\,:\,P\supset S} |\langle \Delta_P f\rangle_{P_S}| K_t^{S_r}(\mathbb{1}_{\widehat{P_S}\setminus P_S}) \\ &\quad + \sum_{\substack{P:P\supset S\\ P_S\in\mathscr{S}}} |\langle f\rangle_P| K_t^{S_r}(\mathbb{1}_{\widehat{P}\setminus P_S}) \\ &\leq \sum_{P\,:\,P\supset S}\Big(|\langle \Delta_P f\rangle_{P_S}| + |\langle f\rangle_P|\mathbb{1}_{\{P_S\in\mathscr{S}\}}\Big) K_t^{S_r}(\mathbb{1}_{\mathbb{R}^d\setminus P_S}). \end{aligned}$$

Then split the sum over P and consider the two cases:

$$\sum_{P\,:\,P\supset S} = \sum_{P:P\supset R^{(r)}} + \sum_{\substack{P\,:\,P\subseteq R^{(r)}\\ P\supset S}} =: (i)+(ii).$$

Lemma 3.6.8 (Bound for (i))**.** *Let F_P be either $\langle \Delta_P f\rangle_{P_R}$ or $\langle f\rangle_P \mathbb{1}_{\{P_S\in\mathscr{S}\}}$. Then*

$$\begin{aligned} \sum_{S\in\mathscr{S}}\sum_{S_r\in\mathsf{ch}_r(S)}\sum_{R:R\supset S}\iint_{W_R}\Big(\sum_{P\,:\,P\supset R^{(r)}} |F_P|\cdot K_t^{S_r}\mathbb{1}_{\mathbb{R}^d\setminus P_S}(x)\Big)^2\frac{\mathrm{d}t}{t}|\mathcal{P}_{S_r}g|\,\mathrm{d}x \\ \lesssim \Lambda_{\mathscr{S}}(f,g). \end{aligned}$$

Lemma 3.6.9 (Bound for (ii))**.** *Let F_P be either $\langle \Delta_P f\rangle_{P_S}$ or $\langle f\rangle_P \mathbb{1}_{\{P_S\in\mathscr{S}\}}$. Then*

$$\begin{aligned} \sum_{S\in\mathscr{S}}\sum_{S_r\in\mathsf{ch}_r(S)}\sum_{R\,:\,R\supset S}\iint_{W_R}\Big(\sum_{\substack{P\,:\,P\supset S\\ P\subseteq R^{(r)}}} |F_P|\cdot K_t^{S_r}\mathbb{1}_{\mathbb{R}^d\setminus P_S}(x)\Big)^2\frac{\mathrm{d}t}{t}|\mathcal{P}_{S_r}g|\,\mathrm{d}x \\ \lesssim \Lambda_{\mathscr{S}}(f,g). \end{aligned}$$

Proof of Lemma 3.6.8. In this case $P\supset R\supset S$, so the dyadic child P_S equals P_R. Using Lemma 3.6.4, since $\ell S_r < \ell S$, we have

$$K_t^{S_r}\mathbb{1}_{\mathbb{R}^d\setminus P_R}(x) \lesssim \left(\frac{\ell S_r}{\ell P_R}\right)^\eta \leq \left(\frac{\ell S}{\ell R}\right)^\eta\left(\frac{\ell R}{\ell P_R}\right)^\eta.$$

We bound $\int_{\ell R/2}^{\ell R} \mathrm{d}t/t \leq 1$ and then apply Cauchy–Schwarz

$$\sum_{S\in\mathscr{S}} \sum_{S_r\in\mathsf{ch}_r(S)} \sum_{R:R\supset S} \left(\sum_{P\,:\,P\supset R^{(r)}} |F_P| \left(\frac{\ell S}{\ell R}\right)^{\eta} \left(\frac{\ell R}{\ell P_R}\right)^{\eta} \right)^2 \|\mathcal{P}_{S_r} g\|_{L^1}$$
$$\leq \sum_{S\in\mathscr{S}} \sum_{S_r\in\mathsf{ch}_r(S)} \sum_{R:R\supset S} \sum_{P\,:\,P\supset R^{(r)}} F_P^2 \left(\frac{\ell S}{\ell R}\right)^{2\eta} \left(\frac{\ell R}{\ell P_R}\right)^{\eta} \|\mathcal{P}_{S_r} g\|_{L^1}$$

since $\sum_{P\,:\,P\supset R^{(r)}} (\ell R/\ell P_R)^{\eta} \leq 1$. Bound the sum of Haar projections as in (3.6.2)

$$\sum_{S\in\mathscr{S}} \sum_{S_r\in\mathsf{ch}_r(S)} \sum_{R:R\supset S} \sum_{P:P\supset R^{(r)}} F_P^2 \left(\frac{\ell S}{\ell R}\right)^{2\eta} \left(\frac{\ell R}{\ell P_R}\right)^{\eta} \|\mathcal{P}_{S_r} g\|_{L^1}$$
$$\lesssim \sum_{S\in\mathscr{S}} \sum_{R:R\supset S} \sum_{P:P\supset R^{(r)}} F_P^2 \left(\frac{\ell S}{\ell R}\right)^{2\eta} \left(\frac{\ell R}{\ell P_R}\right)^{\eta} \int_S |g|.$$

Rearrange the sums

$$\sum_{S\in\mathscr{S}} \sum_{\substack{R\in\mathscr{D}\\ R\supset S}} \sum_{\substack{P\in\mathscr{D}\\ P\supset R^{(r)}}} = \sum_{R\in\mathscr{D}} \sum_{P:P\supset R^{(r)}} \sum_{i\in\mathbb{N}} \sum_{\substack{S\in\mathscr{S}\\ S\subset R\\ \ell S=2^{-i}\ell R}}$$

then we continue as in the proof of (3.4.1).

$$\sum_{R\in\mathscr{D}} \sum_{P:P\supset R^{(r)}} F_P^2 \left(\frac{\ell R}{\ell P_R}\right)^{\eta} \sum_{i\in\mathbb{N}} \sum_{\substack{S\in\mathscr{S}\\ S\subset R\\ \ell S=2^{-i}\ell R}} \left(\frac{\ell S}{\ell R}\right)^{2\eta} \int_S |g|$$
$$\leq \sum_{R\in\mathscr{D}} \sum_{P:P\supset R^{(r)}} F_P^2 \left(\frac{\ell R}{\ell P_R}\right)^{\eta} \sum_{i\in\mathbb{N}} 2^{-i\eta} \int_R |g|$$
$$\leq \sum_{P\in\mathscr{D}} F_P^2 \sum_{k\geq r} \sum_{\substack{R\,:\,R\subset P\\ \ell R=2^{-k-1}\ell P}} \left(\frac{\ell R}{\ell P_R}\right)^{\eta} \int_R |g|$$
$$\leq \sum_{P\in\mathscr{D}} F_P^2 \sum_{k\geq r} 2^{-k\eta} \int_P |g| \leq \sum_{P\in\mathscr{D}} F_P^2 \int_P |g|.$$

Now we distinguish the two cases for F_P.

If $F_P = \langle \Delta_P f\rangle_{P_R}$

$$\sum_{P\in\mathscr{D}} F_P^2 \int_P |g| \le \sum_{P\in\mathscr{D}} \frac{\langle f, h_P\rangle^2}{|P|}\int_P |g|$$
$$\le 3^d B_0^{\mathscr{D}}(g,f).$$

Then $B_0^{\mathscr{D}}(g,f)$ is bounded by a sparse form in Lemma 3.7.4.

If $F_P = \langle f\rangle_P \mathbb{1}_{\{P_R\in\mathscr{S}\}}$

$$\sum_{P\in\mathscr{D}} F_P^2 \int_P |g| = \sum_{P\,:\,P_R\in\mathscr{S}} \langle f\rangle_P^2 \int_P |g|$$
$$= \Lambda_{\mathscr{S}'}(f,g)$$

where $\mathscr{S}'$ is the sparse collection of dyadic parents of $\mathscr{S}$.

□

Proof of Lemma 3.6.9. For $x\in S_r$ and $t\in(\ell R/2,\ell R)$, since $S_r\subset S\subset P\subseteq R^{(r)}$, by Lemma 3.6.4

$$K_t^{S_r}(\mathbb{1}_{\mathbb{R}^d\setminus P_S})(x) \lesssim \left(\frac{\ell S_r}{\ell R^{(r)}}\right)^{\eta}.$$

Then we distribute the decay factor which is bounded as following

$$\left(\frac{\ell S_r}{\ell R^{(r)}}\right)^{\eta} \le \left(\frac{\ell S}{\ell R^{(r)}}\right)^{\eta/2}\left(\frac{\ell S}{\ell P}\right)^{\eta/2}\left(\frac{\ell P}{\ell R^{(r)}}\right)^{\eta/2}.$$

Estimate the integral $\int_{\ell R/2}^{\ell R} \mathrm{d}t/t \le 1$ and the sum of Haar projections as in (3.6.2).

$$\sum_{S\in\mathscr{S}}\sum_{S_r\in\mathsf{ch}_r(S)}\sum_{R\supset S}\iint_{W_R}\left(\sum_{\substack{P:P\supset S\\ P\subseteq R^{(r)}}} F_P\left(\frac{\ell S_r}{\ell R^{(r)}}\right)^{\eta}\right)^2\frac{\mathrm{d}t}{t}|\mathcal{P}_{S_r}g|\,\mathrm{d}x$$
$$\le \sum_{S\in\mathscr{S}}\sum_{S_r\in\mathsf{ch}_r(S)}\|\mathcal{P}_{S_r}g\|_{L^1}\sum_{R\,:\,R\supset S}\left(\frac{\ell S}{\ell R^{(r)}}\right)^{\eta}\left(\sum_{\substack{P:P\supset S\\ P\subseteq R^{(r)}}} F_P\left(\frac{\ell S}{\ell P}\right)^{\eta/2}\left(\frac{\ell P}{\ell R^{(r)}}\right)^{\eta/2}\right)^2$$
$$\lesssim \sum_{S\in\mathscr{S}}\int_S |g|\sum_{R\,:\,R\supset S}\left(\frac{\ell S}{\ell R^{(r)}}\right)^{\eta}\left(\sum_{\substack{P:P\supset S\\ P\subseteq R^{(r)}}} F_P\left(\frac{\ell S}{\ell P}\right)^{\eta/2}\left(\frac{\ell P}{\ell R^{(r)}}\right)^{\eta/2}\right)^2.$$

Apply the Cauchy–Schwarz inequality.

$$\sum_{S\in\mathscr{S}}\int_S|g|\sum_{R\,:\,R\supset S}\left(\frac{\ell S}{\ell R^{(r)}}\right)^{\eta}\left(\sum_{\substack{P:P\supset S\\P\subseteq R^{(r)}}}F_P\left(\frac{\ell S}{\ell P}\right)^{\eta/2}\left(\frac{\ell P}{\ell R^{(r)}}\right)^{\eta/2}\right)^2$$
$$\leq \sum_{S\in\mathscr{S}}\int_S|g|\sum_{R\,:\,R\supset S}\left(\frac{\ell S}{\ell R^{(r)}}\right)^{\eta}\sum_{\substack{P:P\supset S\\P\subseteq R^{(r)}}}F_P^2\left(\frac{\ell S}{\ell P}\right)^{\eta}\cdot\sum_{\substack{P:P\supset S\\P\subseteq R^{(r)}}}\left(\frac{\ell P}{\ell R^{(r)}}\right)^{\eta}$$

The last sum is finite: since $P\supset S$ there is only one ancestor for each generation. Since all terms are non-negative, we bound by removing the restriction $P\subset R^{(r)}$ in the sum in P.

$$\begin{aligned}\sum_{S\in\mathscr{S}}\int_S|g|\sum_{R\,:\,R\supset S}\left(\frac{\ell S}{\ell R^{(r)}}\right)^{\eta}\sum_{P:P\supset S}F_P^2\left(\frac{\ell S}{\ell P}\right)^{\eta} &\leq \sum_{S\in\mathscr{S}}\int_S|g|\sum_{P:P\supset S}F_P^2\left(\frac{\ell S}{\ell P}\right)^{\eta}\\ &= \sum_{P\in\mathscr{D}}F_P^2\sum_{i\in\mathbb{N}}2^{-i\eta}\sum_{\substack{S\in\mathscr{S}\\S\subset P\\ \ell S=2^{-i}\ell P}}\int_S|g|\\ &\leq \sum_{P\in\mathscr{D}}F_P^2\int_P|g|\sum_{i\in\mathbb{N}}2^{-i\eta}.\end{aligned}$$

The two cases for F_P are as at the end of the proof of Lemma 3.6.8. □

3.7 Sparse domination of the dyadic form

In this section we prove a sparse domination of the dyadic form $B_j^{\mathscr{D}}(g,f)$ defined in (3.3.1).

Writing 1 as $\langle\mathbb{1}_P\rangle_P$ we have

$$B_j^{\mathscr{D}}(g,f)=\int_{\mathbb{R}^d}\sum_{K\in\mathscr{D}}\langle|g|\rangle_{3K}\sum_{\substack{P\in\mathscr{D}\\P\subset 3K\\ \ell P=2^{-j}\ell K}}\frac{\langle f,h_P\rangle^2}{|P|}\mathbb{1}_P(x)\,\mathrm{d}x.$$

Let Q_0 be a dyadic cube containing the support of f and g. On the complement of Q_0 the form is controlled.

Lemma 3.7.1. *Let $B_j^{\mathscr{D}}\restriction_{Q_0^{\complement}}(g,f)$ be the restriction of $B_j^{\mathscr{D}}(g,f)$ to the complement $(Q_0)^{\complement}$,* *then*

$$B_j^{\mathscr{D}}\restriction_{Q_0^\complement}(g,f) \lesssim_d 2^{-jd}\langle|g|\rangle_{Q_0}\langle|f|\rangle_{Q_0}^2|Q_0|.$$

Proof. Decompose $(Q_0)^\complement$ in the union of $Q_0^{(k+1)} \setminus Q_0^{(k)}$ for $k \in \mathbb{Z}_+$. The non-zero terms in $B_j^{\mathscr{D}}\restriction_{Q_0^\complement}(g,f)$ are the ones where P intersects Q_0 and $(Q_0^{(k)})^\complement$. Then $P \supset Q_0^{(k)}$ and in particular $P = Q_0^{(m)}$ for $m > k$. There is only one ancestor for each m, so we have

$$\begin{aligned}
B_j^{\mathscr{D}}\restriction_{Q_0^{(k+1)}\setminus Q_0^{(k)}}(g,f) &= \int_{Q_0^{(k+1)}\setminus Q_0^{(k)}} \sum_{K\in\mathscr{D}} \langle|g|\rangle_{3K} \sum_{\substack{P\subset 3K \\ \ell P = 2^{-j}\ell K \\ P\supset Q_0^{(k)}}} \left(\frac{\langle f, h_P\rangle}{|P|^{1/2}}\right)^2 \mathbb{1}_P(x)\,\mathrm{d}x \\
&\lesssim \sum_{m=k+1}^{\infty} \langle|g|\rangle_{3Q_0^{(m+j)}} \langle|f|\rangle_{Q_0^{(m)}}^2 |Q_0^{(m)}| \\
&= \sum_{m=k+1}^{\infty} 3^{-d}2^{-(m+j)d}\langle|g|\rangle_{Q_0} 2^{-2md}\langle|f|\rangle_{Q_0}^2 2^{md}|Q_0| \\
&\le 2^{-jd}\langle|g|\rangle_{Q_0}\langle|f|\rangle_{Q_0}^2|Q_0| \sum_{m=k+1}^{\infty} 2^{-2md}.
\end{aligned}$$

The last sum is bounded by 2^{-kd} and summing over $k \in \mathbb{Z}_+$ concludes the proof. $\square$

It is enough to construct a sparse family inside Q_0. Taking the supremum of $\langle|g|\rangle_{3K}$ over all $K \in \mathscr{D}$ we have

$$B_j^{\mathscr{D}}(g,f) \le \int M^{3\mathscr{D}}g(x)\cdot(S_j^{3\mathscr{D}}f(x))^2\,\mathrm{d}x$$

where $M^{3\mathscr{D}}$ and $S_j^{3\mathscr{D}}$ denote the maximal function and the square function given by

$$M^{3\mathscr{D}}f \coloneqq \sup_{Q\in\mathscr{D}}\langle|f|\rangle_{3Q}\mathbb{1}_{3Q}, \qquad \left(S_j^{3\mathscr{D}}f(x)\right)^2 \coloneqq \sum_{R\in\mathscr{D}} \sum_{\substack{P\in\mathscr{D} \\ P\subset 3R \\ \ell P = 2^{-j}\ell R}} \frac{\langle f, h_P\rangle^2}{|P|}\mathbb{1}_P(x). \tag{3.7.1}$$

As we see below, $S_j^{3\mathscr{D}}$ is pointwise controlled by the square function $S_j^{\mathscr{D}}f(x)$ given by

$$\left(S_j^{\mathscr{D}}f(x)\right)^2 \coloneqq \sum_{Q\in\mathscr{D}} \sum_{\substack{P\in\mathscr{D} \\ P\subset Q \\ \ell P = 2^{-j}\ell Q}} \frac{\langle f, h_P\rangle^2}{|P|}\mathbb{1}_P(x).$$

Proposition 3.7.2 (Pointwise control). *Let $f \in L^2(\mathbb{R}^d)$ and $j \in \mathbb{N}_0$. For all $x \in \mathbb{R}^d$ it holds that*

$$S_j^{\mathscr{D}} f(x) \le S_j^{3\mathscr{D}} f(x) \le 3^{d/2} S_j^{\mathscr{D}} f(x)$$

Proof. The enlarged cube $3R$ is the union of 3^d cubes $\{R_a\}_a$ in the same dyadic grid $\mathscr{D}$. So

$$\begin{aligned}
\left(S_j^{3\mathscr{D}} f(x)\right)^2 &= \sum_{R\in\mathscr{D}} \sum_{\substack{P:\,P\subset 3R\\ \ell P = 2^{-j}\ell R}} \langle f, h_P\rangle^2 \frac{\mathbb{1}_P(x)}{|P|} \\
&= \sum_{R\in\mathscr{D}} \sum_{a=1}^{3^d} \sum_{\substack{P:\,P\subset R_a\\ \ell P = 2^{-j}\ell R_a}} \langle f, h_P\rangle^2 \frac{\mathbb{1}_P(x)}{|P|} \\
&= \sum_{a=1}^{3^d} \left(S_j^{\mathscr{D}} f(x)\right)^2 \le 3^d \left(S_j^{\mathscr{D}} f(x)\right)^2.
\end{aligned}$$

□

We show that the square function $S_j^{\mathscr{D}}$ satisfies a weak $(1,1)$ bound. The proof follows the one for dyadic shifts without separation of scales [HPTV14, Theorem 5.2] and [LM17b, Lemma 4.4].

Proposition 3.7.3. *Let $j \in \mathbb{Z}_+$. There exists $C > 0$ such that for any $f \in L^1(\mathbb{R}^d)$ it holds that*

$$\sup_{\lambda>0} \lambda |\{x \in \mathbb{R}^d \,:\, S_j^{\mathscr{D}} f(x) > \lambda\}| \le C(1+j)\|f\|_{L^1}.$$

In particular $\|S_j^{\mathscr{D}}\|_{L^1\to L^{1,\infty}}$ grows at most polynomially in j.

Proof. First, $S_j^{\mathscr{D}}$ is bounded in L^2 with norm independent of j.

We want to show that for any $\lambda > 0$ we have

$$|\{x \in \mathbb{R}^d \,:\, S_j^{\mathscr{D}} f(x) > \lambda\}| \le C\frac{\|f\|_1}{\lambda}.$$

Let $f = g + b$ be the Calderón-Zygmund decomposition of f at height $\lambda > 0$. Then $\|g\|_\infty \le 2^d\lambda$ and in particular $\|g\|_2^2 \lesssim \lambda\|f\|_1$, while $b = \sum_{Q\in\mathscr{L}} b_Q$, where b_Q is supported on

Q and $\int b_Q = 0$. The cubes Q in $\mathscr{L}$ are maximal dyadic cubes such that $\lambda < \langle |f| \rangle_Q \leq 2^d \lambda$.

Let E be the union of the cubes in $\mathscr{L}$. Then $|E| = \sum_{Q \in \mathscr{L}} |Q| \leq \lambda^{-1} \|f\|_1$ so it is enough to estimate the superlevel sets on the complement of E. Using the decomposition of f we have

$$\begin{aligned} |\{x \in E^{\complement} \colon S_j^{\mathscr{D}} f(x) > \lambda\}| &\leq \left| \left\{ x \colon S_j^{\mathscr{D}} g(x) > \frac{\lambda}{2} \right\} \right| + \left| \left\{ x \in E^{\complement} \colon S_j^{\mathscr{D}} b(x) > \frac{\lambda}{2} \right\} \right| \\ &\lesssim_d \frac{\|f\|_1}{\lambda} + \frac{2}{\lambda} \|S_j^{\mathscr{D}} b\|_{L^1(E^{\complement})}. \end{aligned}$$

The last bound follows by using Chebyshev's inequality for the good part:

$$\left| \left\{ S_j^{\mathscr{D}} g > \frac{\lambda}{2} \right\} \right| \leq \frac{4}{\lambda^2} \|S_j^{\mathscr{D}} g\|_2^2 \lesssim \frac{\|g\|_2^2}{\lambda^2} \lesssim \frac{\|f\|_1}{\lambda}$$

and Markov's inequality for the bad part. The sublinearity of $S_j^{\mathscr{D}}$ and the triangle inequality imply that

$$\|S_j^{\mathscr{D}} b\|_{L^1(E^{\complement})} \leq \sum_{Q \in \mathscr{L}} \|S_j^{\mathscr{D}} b_Q\|_{L^1(E^{\complement})}.$$

For each $Q \in \mathscr{L}$, only dyadic cubes $K \supset Q$ contribute to the $\|S_j^{\mathscr{D}} b_Q\|_{L^1(E^{\complement})}$, since if $K \subseteq Q$, then K would be inside E. Thus K is an ancestor of Q, so $K = Q^{(k)}$ for some integer $k \geq 1$. For $k > j$ each j-child $P \subset K$ contains Q, and so $\langle b_Q, h_P \rangle$ vanishes, by the zero average of b_Q. Thus we estimate

$$\begin{aligned} \|S_j^{\mathscr{D}} b_Q\|_{L^1(E^{\complement})} &\leq \int_{E^{\complement}} \sum_{K \in \mathscr{D}} \sum_{\substack{P \subset K \\ \ell P = 2^{-j} \ell K}} |\langle b_Q, h_P \rangle| \frac{\mathbb{1}_P(x)}{|P|^{1/2}} \,\mathrm{d}x \\ &\leq \sum_{k=1}^{j} \sum_{\substack{K \supset Q \\ \ell K = \ell Q^{(k)}}} \sum_{\substack{P \subset K \\ \ell P = 2^{-j} \ell K}} |\langle b_Q, h_P \rangle| |P|^{1/2} \\ &\leq \sum_{k=1}^{j} \sum_{\substack{K \supset Q \\ \ell K = \ell Q^{(k)}}} \sum_{\substack{P \subset K \\ \ell P = 2^{-j} \ell K}} \|b_Q\|_{L^1(P)} \\ &\leq \sum_{k=1}^{j} \sum_{\substack{K \supset Q \\ \ell K = \ell Q^{(k)}}} \|b_Q\|_{L^1(K)} = \sum_{k=1}^{j} \|b_Q\|_{L^1}. \end{aligned}$$

Since $\|b_Q\|_{L^1(K)} = \|b_Q\|_{L^1} \lesssim \lambda|Q| < \int_Q |f|$, and there is only one ancestor of Q for each k, we have

$$\sum_{k=1}^{j} \|b_Q\|_{L^1} \lesssim \sum_{k=1}^{j} \int_Q |f| \leq j \int_Q |f|.$$

Summing over all $Q \in \mathscr{L}$ gives the bound

$$\sum_{Q\in\mathscr{L}} \|S_j^{\mathfrak{D}} b_Q\|_{L^1(E^{\complement})} \lesssim \sum_{Q\in\mathscr{L}} j\|f\|_{L^1(Q)} \leq j\|f\|_{L^1(\mathbb{R}^d)}.$$

□

The operator $M^{3\mathfrak{D}}$ defined in (3.7.1) is also weak $(1,1)$ as it is bounded by the Hardy–Littlewood maximal function, which is weakly bounded.

The following lemma exploits the weak boundedness of the operators $M^{3\mathfrak{D}}$ and $S_j^{\mathfrak{D}}$ to construct a sparse collection $\mathscr{S}$. The proof adapts the one in [LM17b, Lemma 4.5] to our square function. We include the details for the convenience of the reader.

Lemma 3.7.4 (Sparse domination of $B_j^{\mathfrak{D}}$)**.** *Let $j \in \mathbb{Z}_+$. For any pair of compactly supported functions $f, g \in L^\infty(\mathbb{R}^d)$ there exists a sparse collection $\mathscr{S}$ such that*

$$B_j^{\mathfrak{D}}(g,f) \lesssim \int M^{3\mathfrak{D}} g \cdot (S_j^{\mathfrak{D}} f)^2 \lesssim (1+j)^2 \sum_{S\in\mathscr{S}} \langle |f|\rangle_S^2 \langle |g|\rangle_S |S|$$

where the implicit constant does not depend on j.

Proof of Lemma 3.7.4. Fix a cube $Q_0 \in \mathfrak{D}$ containing the union of the supports of f and g. By Lemma 3.7.1 it is enough to construct a sparse family inside Q_0. Consider the set $F(Q_0)$ given by

$$\{x \in Q_0 \,:\, M^{3\mathfrak{D}} g(x) > C\langle |g|\rangle_{Q_0}\} \cup \{x \in Q_0 \,:\, S_j^{\mathfrak{D}} f(x) > C(1+j)\langle |f|\rangle_{Q_0}\}.$$

By the weak boundedness of $M^{3\mathfrak{D}}$ and $S_j^{\mathfrak{D}}$, there exists $C > 0$ such that $|F(Q_0)| \leq \frac{1}{2}|Q_0|$.

Then

$$\int_{Q_0} M^{3\mathscr{D}} g \cdot (S_j^{\mathscr{D}} f)^2 \leq \int_{Q_0 \setminus F(Q_0)} M^{3\mathscr{D}} g \cdot (S_j^{\mathscr{D}} f)^2 + \int_{F(Q_0)} M^{3\mathscr{D}} g \cdot (S_j^{\mathscr{D}} f)^2$$
$$\leq C^3 (1+j)^2 \langle |g| \rangle_{Q_0} \langle |f| \rangle_{Q_0}^2 |Q_0| + \sum_{Q \in \mathscr{F}} \int_Q M^{3\mathscr{D}} g \cdot (S_j^{\mathscr{D}} f)^2$$

where $\mathscr{F}$ is the collection of maximal dyadic cubes covering $F(Q_0)$. Iterating on each $Q \in \mathscr{F}$ produces a sparse family of cubes $\mathscr{S}$, since $\{E_Q \coloneqq Q \setminus F(Q)\}_{Q \in \mathscr{S}}$ are pairwise disjoint and $|E_Q| > \frac{1}{2}|Q|$ for each Q in $\mathscr{S}$. □

3.8 Proofs for the reduction to a dyadic form

Proof of Lemma 3.3.2. We distinguish three cases: $\ell P > 2^r \ell R$, where the goodness is used; $\ell P \in [\ell R, 2^r \ell R]$, where ℓP and ℓR are comparable; and $\ell P < \ell R$, where we use the zero-average of $\Delta_P f$ and the regularity condition (C2).

($\ell P > 2^r \ell R$) Using the size condition (C1) and taking the supremum in $(x,t) \in W_R$

$$\theta_t(\Delta_P f)(x) \leq C_1 \int_P \frac{t^\alpha}{(t+|x-y|)^{\alpha+d}} |\Delta_P f(y)| \, \mathrm{d}y$$
$$\leq C_1 \|\Delta_P f\|_{L^1} \frac{(\ell R)^\alpha}{(\frac{\ell R}{2} + \mathrm{d}(R,P))^{\alpha+d}}. \tag{3.8.1}$$

If $\mathrm{d}(R,P) > \ell P$, since $\ell P > 2^r \ell R$ the conclusion follows. Otherwise, by the goodness of R, we have $\ell P < \mathrm{d}(R,P) \left(\frac{\ell P}{\ell R}\right)^\gamma$. The same bound holds for ℓR, so

$$D(P,R)^{\alpha+d} < 3^{\alpha+d} \, \mathrm{d}(P,R)^{\alpha+d} \left(\frac{\ell P}{\ell R}\right)^{\gamma(\alpha+d)}$$

which implies

$$D(P,R)^{\alpha+d} \left(\frac{\ell R}{\ell P}\right)^{\alpha/2} \lesssim_{\alpha,d} \mathrm{d}(P,R)^{\alpha+d} \left(\frac{\ell R}{\ell P}\right)^{-\gamma(\alpha+d)+\alpha/2}$$
$$\leq \mathrm{d}(P,R)^{\alpha+d}$$

since $\ell R/\ell P < 1$ and $\alpha/2 - \gamma(\alpha+d)$ is non-negative for $\gamma \leq \frac{\alpha}{2(\alpha+d)}$. Then multiply and divide (3.8.1) by $D(P,R)^{\alpha+d}(\ell P)^{-\alpha/2}(\ell R)^{\alpha/2}$ to conclude.

($\ell R \leq \ell P \leq 2^r \ell R$) The lengths of P and R are comparable and the conclusion follows.

($\ell P < \ell R$) Let x_P be the centre of P. Then

$$\begin{aligned}\int k_t(x,y)\Delta_P f(y)\,\mathrm{d}y &= \int (k_t(x,y) - k_t(x,x_P))\Delta_P f(y)\,\mathrm{d}y \\ &\leq C_2 \int \frac{|y-x_P|^\alpha}{(t+|x-y|)^{\alpha+d}} |\Delta_P f(y)|\,\mathrm{d}y\end{aligned}$$

by the smoothness condition (C2), since $|y-x_P| \leq \frac{\ell P}{2} < \frac{\ell R}{2} < t$. To conclude, note that

$$\frac{(\ell P)^\alpha}{(\frac{\ell R}{2} + \mathrm{d}(R,P))^{\alpha+d}} < \frac{(\ell P)^\alpha}{(\frac{\ell R}{4} + \frac{\ell P}{4} + \mathrm{d}(R,P))^{\alpha+d}} \leq 4^{\alpha+d} \frac{(\sqrt{\ell R \ell P})^\alpha}{D(R,P)^{\alpha+d}}.$$

□

3.8.1 Counting close cubes

In both cases "near" and "close", given a fixed R we estimate the number of P such that $3P \supset R$. Given a discrete set A, we denote by $\#A$ its cardinality.

Lemma 3.8.1. *For $k \in \mathbb{N}$ let $\mathscr{P}_k(R) := \{P : 3P \supset R, \ell P = 2^k \ell R\}$. Then $\#\mathscr{P}_k(R) = 3^d$.*

Proof. Let $R^{(k)}$ be the k-ancestor of R. Then $R^{(k)}$ belongs to $\mathscr{P}_k(R)$. There are $3^d - 1$ cubes P adjacent to $R^{(k)}$ with $\ell P = \ell R^{(k)}$. Each of them is such that $3P \supset R^{(k)}$, so in particular $3P \supset R$.

On the other hand, if P is not adjacent to $R^{(k)}$ and $\ell P = \ell R^{(k)}$ then $\mathrm{d}(P, R^{(k)}) \geq \ell P$, so $3P$ does not contain $R^{(k)}$, nor R.

This shows that the P in $\mathscr{P}_k(R)$ are exactly the cubes contained in $3R^{(k)}$ with $\ell P = \ell R^{(k)}$, and there are 3^d of such cubes. □

Proof of Lemma 3.3.3. We present each case separately.

far $\ell P \geq 2^{r+1}\ell R$ and $\mathrm{d}(P,R) > \ell P$. The largest term in $D(P,R)$ is $\mathrm{d}(P,R)$. Fix R and $k \in \mathbb{N}$. Given $m \in \mathbb{N}$ there are at most 2^{md} cubes P with length $2^k \ell R$ such that $2^m \ell P < \mathrm{d}(P,R) \leq 2^{m+1}\ell P$, so rearranging the sum

$$\sum_{\substack{P\,:\,\ell P \geq \ell R \\ \mathrm{d}(R,P) > \ell P}} \frac{(\sqrt{\ell R \ell P})^\alpha}{\mathrm{d}(R,P)^{\alpha+d}}|P| =$$

$$= \sum_{m=1}^{\infty}\sum_{k=r}^{\infty} \sum_{\substack{P\,:\,\ell P = 2^{k+1}\ell R \\ 2^{m+1} \geq \mathrm{d}(P,R)/\ell P > 2^m}} \left(\frac{\sqrt{\ell R \ell P}}{\mathrm{d}(R,P)}\right)^\alpha \left(\frac{\ell P}{\mathrm{d}(P,R)}\right)^d$$

$$\leq \sum_{m=1}^{\infty}\sum_{k=0}^{\infty} 2^{md} 2^{-\alpha(k/2+m)} 2^{-md} \leq \sum_{k,m} 2^{-\alpha(k/2+m)}.$$

near For P such that $3P \setminus P \supset R$ and $\ell P \geq 2^{r+1}\ell R$, the decay comes from $\mathrm{d}(P,R)$, which is bounded below by $\ell P(\ell R/\ell P)^\gamma$, and $\gamma = \alpha/(4\alpha + 4d)$. Then

$$\sum_{\substack{P\,:\,3P\setminus P \supset R \\ \ell P \geq 2^{r+1}\ell R}} \frac{(\sqrt{\ell R \ell P})^\alpha}{\mathrm{d}(R,P)^{\alpha+d}}|P| = \sum_{k=r+1}^{\infty} \sum_{\substack{P\,:\,3P\setminus P \supset R \\ \ell P = 2^k \ell R}} \frac{|P|}{\mathrm{d}(P,R)^d}\left(\frac{\sqrt{\ell P \ell R}}{\mathrm{d}(P,R)}\right)^\alpha$$

$$\lesssim_d \sum_{k=r+1}^{\infty} 2^{-k\alpha/4}$$

where, by Lemma 3.8.1, the P in the sum are at most 3^d for each k.

close For $\ell R \leq \ell P \leq \ell R^{(r)}$ and $3P \supset R$, the leading term in the long-distance is ℓR, so

$$\frac{(\sqrt{\ell R \ell P})^\alpha}{D(R,P)^{\alpha+d}}|P| \leq \frac{(2^{r/2}\ell R)^\alpha}{(\ell R)^\alpha}\frac{|P|}{|R|} \leq 2^{\alpha r/2}\frac{|P|}{|R|}.$$

We fix a scale k for P, such that $0 \leq k \leq r$, then we estimate

$$\sum_{\substack{P\,:\,3P \supset R \\ \ell R \leq \ell P \leq \ell R^{(r)}}} \frac{(\sqrt{\ell R \ell P})^\alpha}{D(R,P)^{\alpha+d}}|P| \lesssim \sum_{\substack{P\,:\,3P \supset R \\ \ell R \leq \ell P \leq \ell R^{(r)}}} \frac{|P|}{|R|} = \sum_{k=0}^{r} \sum_{\substack{P\,:\,3P \supset R \\ \ell P = 2^k \ell R}} 2^{kd}$$

$$\leq 2^{rd} \sum_{k=0}^{r} |\{P \,:\, 3P \supset R, \ell P = 2^k \ell R\}| \leq 2^{rd} 3^d (r+1).$$

Where to estimate the number of P we used Lemma 3.8.1.

subscale, $P \subset 3R$ The leading term in the long-distance $D(R,P)$ is again ℓR. For any $k \in \mathbb{N}$, there are $3^d 2^{kd}$ cubes P such that $P \subset 3R$ and $2^k \ell P = \ell R$, so

$$\sum_{P\,:\,P\subset 3R} \frac{(\sqrt{\ell R \ell P})^\alpha}{D(R,P)^{\alpha+d}}|P| \leq \sum_{P\subset 3R} \left(\frac{\ell P}{\ell R}\right)^{\alpha/2+d}$$
$$= \sum_{k=1}^{\infty} \sum_{\substack{P\subset 3R\\ \ell P = 2^{-k}\ell R}} 2^{-k\frac{\alpha}{2}} 2^{-kd} \lesssim_d \sum_{k=1}^{\infty} 2^{-k\frac{\alpha}{2}} < \infty.$$

subscale, $P \not\subset 3R$ In this case $\mathrm{d}(P,R) > \ell R > \ell P$. Regroup the P according to length and distance:

$$\sum_{\substack{P\,:\,P\not\subset 3R\\ \ell P<\ell R}} \frac{(\sqrt{\ell R \ell P})^\alpha}{D(R,P)^{\alpha+d}}|P| = \sum_{k\in\mathbb{N}} \sum_{\substack{P\,:\,2^k\ell P=\ell R\\ \mathrm{d}(P,R)>\ell R}} 2^{-kd} \left(\frac{\ell R}{D(P,R)}\right)^d 2^{-k\alpha/2} \left(\frac{\ell R}{D(P,R)}\right)^\alpha$$
$$\leq \sum_{k,m} \sum_{\substack{P\,:\,2^k\ell P=\ell R\\ 2^{m+1}\geq \mathrm{d}(P,R)/\ell R > 2^m}} 2^{-k(d+\alpha/2)} 2^{-md} 2^{-m\alpha} \leq \sum_{k,m} 2^{-k\alpha/2} 2^{-m\alpha}.$$

This because there are at most 2^{md} cubes R in the range given by the distance, which means at most $2^{md}\cdot 2^{kd}$ cubes P with $\ell P = 2^{-k}\ell R$.

□

Proof of Lemma 3.3.4 (for $\ell P < \ell R$). Recall that $\gamma \in (0, \frac{1}{2})$. Let K be the minimal cube $K \supset R$ such that $\ell K \geq 2^r \ell R$ and $\mathrm{d}(P,R) \leq \ell K \left(\frac{\ell P}{\ell K}\right)^\gamma$. (The set of such cubes is not empty since $\ell K \left(\frac{\ell P}{\ell K}\right)^\gamma$ equals $\ell P \left(\frac{\ell K}{\ell P}\right)^{1-\gamma}$ which goes to infinity as $\ell K \to \infty$.) First, observe that $P \subset K$. Suppose not, then

$$\ell K \left(\frac{\ell P}{\ell K}\right)^\gamma < \ell K \left(\frac{\ell R}{\ell K}\right)^\gamma < \mathrm{d}(R, \partial K) \overset{P\subset K^\complement}{\leq} \mathrm{d}(R,P)$$

which is absurd because of the second condition on K. It remains to show the upper bound for ℓK. By minimality of K, one of the following conditions holds: either

$$\frac{\ell K}{2} < 2^r \ell R \quad \text{or} \quad \frac{\ell K}{2}\left(\frac{\ell P}{\frac{1}{2}\ell K}\right)^{\gamma} < \mathrm{d}(P,R).$$

Since by hypothesis $\mathrm{d}(P,R) > (\ell R)^{1-\gamma}(\ell P)^{\gamma}$, the first implies

$$\ell K \left(\frac{\ell P}{\ell K}\right)^{\gamma} \leq 2^r \ell R \left(\frac{\ell P}{\ell K}\right)^{\gamma} \leq 2^r \ell R \left(\frac{\ell P}{\ell R}\right)^{\gamma} < 2^r\,\mathrm{d}(P,R).$$

The latter gives: $\ell K(\ell P/\ell K)^{\gamma} < 2\,\mathrm{d}(P,R) \leq 2^r\,\mathrm{d}(P,R)$. $\square$

CHAPTER 4

QUADRATIC SPARSE DOMINATION FOR NON-INTEGRAL SQUARE FUNCTIONS

There is a time in the night in which all theorems are true.

G. M.

This chapter is based on the paper

J. Bailey, G. Brocchi, M. C. Reguera. *Quadratic sparse domination and weighted estimates for non-integral square functions*, arXiv:2007.15928.

We prove a version of Theorem C in a doubling metric measure space $(\mathbb{X}, d, \mu)$.

Theorem C (Bailey, B., Reguera 2020)**.** *Let S be a vertical square function bounded on L^p for $p \in (p_0, q_0)$, $p_0 < 2 < q_0$, given by*

$$Sf(x) \coloneqq \left(\int_0^\infty |\mathcal{Q}_t f(x)|^2 \,\frac{\mathrm{d}t}{t} \right)^{1/2},$$

with $\mathcal{Q}_t$ satisfying Assumption 4.1.2. For any f and g in $C_c^\infty(\mathbb{R}^d)$ there exists a sparse family $\mathscr{S} \subseteq \mathscr{D}$ such that

$$\Big| \int_{\mathbb{X}} (Sf)^2 \cdot g \,\mathrm{d}\mu \Big| \leq C \sum_{Q \in \mathscr{S}} \left(\fint_Q |f|^{p_0} \,\mathrm{d}\mu \right)^{2/p_0} \left(\fint_Q |g|^{\left(\frac{q_0}{2}\right)'} \,\mathrm{d}\mu \right)^{1/\left(\frac{q_0}{2}\right)'} \mu(Q)$$

where C is a positive constant independent of f and g.

The above theorem implies weighted estimates for weights in the intersection of the Muckenhoupt class (A_p) and the reverse Hölder class (RH_q) defined in the introduction.

Corollary C. *Let S be a vertical square function considered in Theorem C. For p in (p_0, q_0), $p_0 < 2 < q_0$, and a weight $w \in A_{p/p_0} \cap RH_{(q_0/p)'}$ the square function S is bounded on $L^p(w)$ with*

$$\|S\|_{L^p(w)\to L^p(w)} \leq C\left([w]_{A_{p/p_0}}[w]_{RH_{(q_0/p)'}}\right)^{\max\left(\frac{1}{p-p_0},\frac{q_0-2}{q_0-p}\right)}$$

where C is positive constant independent of the weight.

The result is sharp for certain square functions, see [Ler06, pag. 488], [Ler11] and [BD20a, Remark 15 (ii)]. Sharpness can be deduced from the asymptotic behaviour of the unweighted estimates [FN19]. Unfortunately, these asymptotics are not easy to exactly compute for our non-integral square functions. However, the estimate in Corollary C implies an upper bound on the asymptotic behaviour of the unweighted norm $\|S\|_{L^p\to L^p}$, see Section 4.5.5. In particular, when such asymptotic behaviour is known to match the upper bound, the weighted estimates in Corollary C are sharp.

The power in the characteristic of the weight is sharp for the sparse form in Theorem C. These weighted estimates follow by applying the sharp quantitative limited range extrapolation by Nieraeth [Nie19, Theorem 2.2]. The sharpness is proved in the same article [Nie19, page 418].

We state the application in our case.

Proposition C. *Let $p \in (p_0, q_0)$, $p > 2$, for $p_0 < 2 < q_0$, and let $w \in A_{p/p_0} \cap RH_{(q_0/p)'}$ and $\sigma := w^{1-p^*}$ be the dual weight of w, where $p^* := (p/2)'$. For any sparse family $\mathscr{S} \subseteq \mathscr{D}$ and functions f, $g \in L^1_{\mathrm{loc}}(\,\mathrm{d}\mu)$ we have*

$$\sum_{Q\in\mathscr{S}}\left(\fint_Q |f|^{p_0}\,\mathrm{d}\mu\right)^{2/p_0}\left(\fint_Q |g|^{q_0^*}\,\mathrm{d}\mu\right)^{1/q_0^*}\mu(Q)$$
$$\leq C\left([w]_{A_{p/p_0}}[w]_{RH_{(q_0/p)'}}\right)^{\max\left(\frac{1}{p-p_0},\frac{q_0-2}{q_0-p}\right)}\|f\|^2_{L^p(w)}\|g\|_{L^{p^*}(\sigma)},$$

where the constant C is independent of the weight. The power

$$\gamma(p) := \max\left(\frac{1}{p-p_0}, \left(\frac{q_0}{p}\right)' \frac{1}{2q_0^*}\right)$$

on the weight characteristic is sharp.

Remark 4.0.1. The weighted estimates in Proposition C and their sharpness can be derived by combining [LN20, Prop. 2.9] in the scalar case (where $X = \mathbb{C}$, $m = 1$, $q = 2$, $r = p_0$, $s = q_0$) with [LN20, Prop. 2.5].

Since $(Sf)^2$ is non-negative, the dual pairing is maximised by non-negative functions. For a non-negative function g, let $h = \sqrt{g}$. One can consider the sparse domination with h^2 in place of g. The $(q_0/2)'$-average of g on a cube Q is

$$\langle g \rangle_{Q,\left(\frac{q_0}{2}\right)'} = \langle h^2 \rangle_{Q,\left(\frac{q_0}{2}\right)'} = \langle h \rangle^2_{Q,2\left(\frac{q_0}{2}\right)'} = \langle h \rangle^2_{Q,\frac{1}{\frac{1}{2}-\frac{1}{q_0}}}$$

which appears in the sparse form considered in [LN20, Prop. 2.5].

Guide to this chapter

We start by describing our framework in Section 4.1. Section 4.2 contains some preliminary results that will be of use later. Section 4.3 discusses the examples that fit the assumptions and that one should keep in mind as references. The proof of Theorem C requires us to understand the boundedness properties of a grand maximal operator associated with the corresponding square functions. These boundedness properties are included in Section 4.4. Section 4.5 is dedicated to the proof of Theorem C.

We do not include the derivation of the weighted estimates in Proposition C nor its sharpness, which can be found in the original paper [BBR20, Section 7] for $p \in (2, q_0)$ or in [Nie19; LN20] for the full range.

4.1 Setting

Motivated by finding a uniform setting that will include several examples of square functions, we consider the following general framework.

The underlying space $(\mathbb{X}, d, \mu)$ is a locally compact separable metric space $(\mathbb{X}, d)$ equipped with a Borel measure μ that is finite on compact sets and strictly positive on any non-empty open set. For a measurable subset $E \subset \mathbb{X}$, we denote $|E| := \mu(E)$, and $\fint_B f \, \mathrm{d}\mu := |B|^{-1} \int_B f \, \mathrm{d}\mu$ for $f \in L^1_{\mathrm{loc}}(\mathbb{X}, \mu)$.

The measure μ will be assumed to satisfy the doubling property,

$$|B(x, 2r)| \lesssim |B(x, r)| \tag{4.1.1}$$

for all $x \in \mathbb{X}$ and $r > 0$, where $B(x, s)$ denotes the ball of radius $s > 0$ centered at a point $x \in \mathbb{X}$ and $X \lesssim Y$ will be used to signify that there exists a constant $C > 0$ such that $X \leq CY$. There will then exist some $\nu > 0$ for which

$$|B(x, r)| \lesssim \left(\frac{r}{s}\right)^{\nu} |B(x, s)| \qquad \forall \, x \in \mathbb{X}, \; r \geq s > 0. \tag{4.1.2}$$

It will be assumed that there exists some non-decreasing function $\varphi : (0, \infty) \to (0, \infty)$ with $\varphi(1) = 1$ for which

$$|B(x, r)| \eqsim \varphi\left(\frac{r}{s}\right) |B(x, s)| \tag{4.1.3}$$

for all $x \in \mathbb{X}$ and r, $s > 0$, where $X \eqsim Y$ means that both $X \lesssim Y$ and $Y \lesssim X$ hold. This technical condition has been imposed in order to prove boundedness of a certain maximal operator that is essential to our proof. This point will be elaborated upon further in Remark 4.1.4 and Section 4.4.

Let $\theta \in [0, \pi/2)$. We say that a linear operator L with dense domain $\mathcal{D}_2(L)$ in $L^2(\mathbb{X}, \mu)$ is θ-accretive if its spectrum is contained in the closed sector $\Sigma_{\theta+} := \{z \in \mathbb{C} : |\arg z| \leq \theta\} \cup \{0\}$ and $\langle Lf, f \rangle \in \Sigma_{\theta+}$ for all f in $\mathcal{D}_2(L)$.

We consider an unbounded operator L on $L^2(\mathbb{X}, \mu)$ satisfying the below assumption.

Assumption 4.1.1. The operator L is a linear, injective, θ-accretive operator with dense domain $\mathcal{D}_2(L)$ in $L^2(\mathbb{X},\mu)$, and there exists some $1 \leq p_0 < 2 < q_0 \leq \infty$ and $c > 0$ such that for all balls B_1, B_2 of radius $\sqrt{t}$,

$$\|e^{-tL}\|_{L^{p_0}(B_1)\to L^{q_0}(B_2)} \lesssim |B_1|^{-\frac{1}{p_0}}|B_2|^{\frac{1}{q_0}} \exp\Big(-\frac{c\,d(B_1,B_2)^2}{t}\Big).$$

From Assumption 4.1.1, it follows that L possesses a bounded holomorphic functional calculus on $L^2(\mathbb{X},\mu)$ and $-L$ is the generator of the analytic semigroup $(e^{-tL})_{t>0}$ on $L^2(\mathbb{X},\mu)$, see [Haa06, §7.1.3].

We consider square function operators associated with L. These will be defined to be operators S that satisfy the following set of assumptions.

Assumption 4.1.2. (a) The operator S is sublinear and bounded on $L^2(\mathbb{X},\mu)$.

(b) *(Off-diagonal estimates for the constituent operators)* The operator S is of the form

$$Sf(x) := \left(\int_0^\infty |\mathcal{Q}_t f(x)|^2 \,\frac{\mathrm{d}t}{t}\right)^{1/2},$$

where $\{\mathcal{Q}_t\}_{t>0}$ is a collection of bounded operators on $L^2(\mathbb{X},\mu)$ which satisfy the property that there exists some $1 \leq p_0 < 2 < q_0 \leq \infty$ such that for all balls B_1, B_2 of radius $\sqrt{t}$,

$$\|\mathcal{Q}_t\|_{L^{p_0}(B_1)\to L^{q_0}(B_2)} \lesssim |B_1|^{-\frac{1}{p_0}}|B_2|^{\frac{1}{q_0}} \left(1+\frac{d(B_1,B_2)^2}{t}\right)^{-(\nu+1)}.$$

(c) *(Cancellation with respect to L)* There exists $A_0 > 0$ and $N_0 \in \mathbb{N}$ such that for all integers $N \geq N_0$,

$$\mathcal{Q}_t (sL)^N e^{-sL} = \frac{t^{A_0} s^N}{(t+s)^{A_0+N}} \Theta^{(N)}_{t+s},$$

where $\{\Theta^{(N)}_r\}_{r>0}$ is a collection of bounded operators on $L^2(\mathbb{X},\mu)$ that satisfies off-diagonal estimates at all scales in the sense that

$$\|\Theta_r^{(N)}\|_{L^{p_0}(B_1)\to L^{q_0}(B_2)} \lesssim |B_{1,\sqrt{r}}|^{-\frac{1}{p_0}} |B_{2,\sqrt{r}}|^{\frac{1}{q_0}} \left(1 + \frac{d(B_1, B_2)^2}{r}\right)^{-\frac{\nu+1}{2}}$$

for all rescaled balls $B_{1,\sqrt{r}}$, $B_{2,\sqrt{r}}$ and $r > 0$, where $B_{\sqrt{r}} := (\sqrt{r}/r(B))B$, $r(B)$ is the radius of B, and tB represents the t-dilate of B, $tB := B(x, tr(B))$ for positive $t > 0$.

(d) *(Cotlar type inequality)* There exists an exponent $p_1 \in [p_0, 2)$ such that for all $x \in \mathbb{X}$ and $r > 0$

$$\left(\fint_{B(x,r)} |Se^{-r^2L} f|^{q_0} \,\mathrm{d}\mu\right)^{1/q_0} \lesssim \inf_{y\in B(x,r)} \mathcal{M}_{p_1}(Sf)(y) + \inf_{y\in B(x,r)} \mathcal{M}_{p_1}(f)(y),$$

where we denote by $\mathcal{M}$ the uncentered Hardy–Littlewood maximal function and $\mathcal{M}_p f := (\mathcal{M}|f|^p)^{1/p}$ for any $p \geq 1$.

Remark 4.1.3. In general, the exponents p_0 and q_0 are determined by the off-diagonal estimates for the constituent operator $\mathcal{Q}_t$, rather than by the off-diagonal estimates for $\{e^{-tL}\}_{t>0}$. For our aim, it is enough to assume that the range in which one has off-diagonal estimates for $\{e^{-tL}\}_{t>0}$ contains the range (p_0, q_0) in the Assumption 4.1.2.

Remark 4.1.4. As our work is intended to build upon the article [BFP16], it will be instructive to compare our assumptions with the hypotheses of [BFP16]. In both articles the assumptions imposed upon the underlying operator L are identical, as well as for the L^2-boundedness and the Cotlar type inequality for the operator S. We have further assumed that S is of the form of a square function with constituent operators $\mathcal{Q}_t$ that satisfy off-diagonal bounds. Also, the cancellative condition of S with respect to L, Assumption (b) of [BFP16], has been replaced by a cancellative condition of the constituent operators $\mathcal{Q}_t$.

This assumption, under condition (4.1.3) on the measure μ, implies the cancellative condition of S with respect to L. The operators under consideration are then a subclass of the operators considered by [BFP16]. As such, we utilise some of the intermediary results from [BFP16]. This will be particularly useful in Section 4.4 to prove the boundedness of a certain maximal function operator.

Example 4.1.5. Square function operators that satisfy the previous set of assumptions are the square functions associated with an elliptic operator $L = -\mathrm{div}(A\nabla)$, such as

$$g_L f := \left(\int_0^\infty |tLe^{-tL} f|^2 \frac{\mathrm{d}t}{t} \right)^{1/2} \quad \text{and} \quad G_L f := \left(\int_0^\infty |\sqrt{t} \nabla e^{-tL} f|^2 \frac{\mathrm{d}t}{t} \right)^{1/2},$$

and some square functions associated with the Laplace–Beltrami operator. We discuss these examples in detail in Section 4.3.

We recall the notion of a sparse family for a system of dyadic cubes $\mathscr{D} = \{\mathscr{D}^b\}_{b=1}^K$.

Definition 4.1.6. A collection $\mathscr{S} \subseteq \mathscr{D}$ is $\frac{1}{2}$-sparse if for each $b \in \{1, \ldots, K\}$ there are pairwise disjoint sets $\{F_Q\}_{Q \in \mathscr{S} \cap \mathscr{D}^b}$ such that $F_Q \subseteq Q$ and $|Q| \leq 2|F_Q|$.

4.2 Preliminaries

In this section we gather a collection of useful results concerning dyadic analysis in metric measure spaces, off-diagonal estimates for a family of operators, and properties of Muckenhoupt and reverse Hölder weight classes.

4.2.1 Dyadic Analysis on a Doubling Metric Space

We recall some well-known definitions and facts from dyadic harmonic analysis as written in [BFP16]. For detailed information on the construction of dyadic systems of cubes in doubling metric spaces, the interested reader is referred to [HK12] and references therein.

Definition 4.2.1. A dyadic system of cubes in a metric measure space $(\mathbb{X}, \mu)$, with parameters $0 < c_0 \leq C_0 < \infty$ and $\delta \in (0, 1)$, is a family of open subsets $\left(Q_\alpha^l\right)_{\alpha \in \mathcal{A}_l, l \in \mathbb{Z}}$ that satisfies the following properties:

- For each $l \in \mathbb{Z}$, there exists a subset Z_l with $\mu(Z_l) = 0$ such that

$$\mathbb{X} = \bigsqcup_{\alpha \in \mathcal{A}_l} Q^l_\alpha \bigsqcup Z_l;$$

- If $l \geq k$, $\alpha \in \mathcal{A}_k$ and $\beta \in \mathcal{A}_l$ then either $Q^l_\beta \subseteq Q^k_\alpha$ or $Q^k_\alpha \cap Q^l_\beta = \emptyset$;

- For every $l \in \mathbb{Z}$ and $\alpha \in \mathcal{A}_l$, there exists a point z^l_α with the property that

$$B(z^l_\alpha, c_0\delta^l) \subseteq Q^l_\alpha \subseteq B(z^l_\alpha, C_0\delta^l).$$

The point z^l_α can be seen as the centre of the cube Q^l_α and the side-length is defined by $\ell(Q^l_\alpha) := \delta^l$.

The below theorem asserts the existence of adjacent systems of dyadic cubes for a doubling metric space. For a proof of this result, refer to [HK12].

Theorem 4.2.2 ([HK12, Thm. 4.1])**.** *Let $(\mathbb{X}, d, \mu)$ be a doubling metric space. There exists $0 < c_0 \leq C_0 < \infty$, $\delta \in (0, 1)$, finite constants $K = K(c_0, C_0, \delta)$ and $C = C(\delta)$, and a finite collection of dyadic systems $\mathscr{D}^b$ with parameters (c_0, C_0, δ), $b = 1, \cdots, K$, that satisfies the following property. For any ball $B = B(x, r) \subseteq \mathbb{X}$, there exists $b \in \{1, \cdots, K\}$ and $Q \in \mathscr{D}^b$ such that*

$$B \subseteq Q \quad \text{and} \quad \operatorname{diam}(Q) \leq Cr.$$

From this point forward we fix a dyadic collection $\mathscr{D} := \cup_{b=1}^K \mathscr{D}^b$ as in the previous theorem. The following covering lemma will be useful in §4.5.

Lemma 4.2.3 ([Lor21, Lemma 2.2])**.** *Let $(\mathbb{X}, d, \mu)$ be a doubling metric space with $\operatorname{diam}(\mathbb{X}) = \infty$ and $\mathscr{D}$ a dyadic system with parameters (c_0, C_0, δ). Let $\alpha \geq 3/\delta$ and $E \subset \mathbb{X}$ with $\operatorname{diam}(E) \in (0, \infty)$. There exists a partition $\mathscr{P} \subseteq \mathscr{D}$ of the space $\mathbb{X}$, made with dyadic cubes, such that*

$$E \subseteq \alpha Q\,, \quad \forall Q \in \mathscr{P}.$$

Let w be a weight on $\mathbb{X}$. The uncentered dyadic maximal function $\mathcal{M}_{p,w}^{\mathscr{D}}$ of exponent $p \in [1, \infty)$ is defined by

$$\mathcal{M}_{p,w}^{\mathscr{D}} f(x) := \sup_{Q \in \mathscr{D}} \left(\frac{1}{w(Q)} \int_Q |f(y)|^p w(y) \, \mathrm{d}y \right)^{1/p} \mathbb{1}_Q(x),$$

where the notation $\mathbb{1}_E$ is used to denote the characteristic function of a set $E \subset \mathbb{X}$ and $w(E) := \int_E w \, \mathrm{d}\mu$. When $w \equiv 1$, $\mathcal{M}_{p,w}^{\mathscr{D}}$ will just be the usual dyadic maximal function of exponent p and the shorthand notation $\mathcal{M}_p^{\mathscr{D}} = \mathcal{M}_{p,1}^{\mathscr{D}}$ will be employed. Similarly, we will also use the notation $\mathcal{M}_w^{\mathscr{D}} = \mathcal{M}_{1,w}^{\mathscr{D}}$. It is known that $\mathcal{M}_p^{\mathscr{D}}$ is of weak-type (p, p) and strong (q, q) for all $q > p$, see [CW71]. Moreover, $\mathcal{M}_w^{\mathscr{D}}$ is bounded on $L^p(w)$ for all $p \in [1, \infty)$ with a constant independent of the weight,

$$\|\mathcal{M}_w^{\mathscr{D}} f\|_{L^p(w)} \leq p' \|f\|_{L^p(w)}. \tag{4.2.1}$$

4.2.2 Off-Diagonal Estimates

In this section, we define three different notions of off-diagonal estimates that will be used throughout this article. For an extensive and detailed account of off-diagonal estimates for operator families, the reader is referred to [AM07b]. Throughout this section, we will consider exponents $1 \leq p_0 < 2 < q_0 \leq \infty$.

Definition 4.2.4 (Off-diagonal estimates at scale $\sqrt{t}$)**.** A family of operators $\{T_t\}_{t>0}$ is said to satisfy (p_0, q_0) off-diagonal estimates at scale $\sqrt{t}$ if for any two balls B_1, B_2 of radius $\sqrt{t}$ we have

$$\left(\fint_{B_2} |T_t(f \mathbb{1}_{B_1})|^{q_0} \, \mathrm{d}\mu \right)^{1/q_0} \lesssim \rho\Big(\frac{d(B_1, B_2)}{\sqrt{t}} \Big) \left(\fint_{B_1} |f|^{p_0} \, \mathrm{d}\mu \right)^{1/p_0},$$

where $\rho \colon [0, \infty) \to (0, 1]$ is a non-increasing function such that $\lim_{x \to \infty} |x|^a \rho(x) = 0$ for some $a \geq 0$, and $\rho(0) = 1$.

Remark 4.2.5. Some comments are in order.

- Examples of ρ that we will use are the Gaussian function $\rho(x) = e^{-c|x|^2}$ and $\rho(x) = \langle x\rangle^{-s}$, where $\langle x\rangle = (1+|x|^2)^{1/2}$ is the Japanese bracket. For the Gaussian case, the positive constant c is not relevant and may change from line to line. See also comments after [AM07b, Def. 2.1]. For our sparse domination, the choice $\rho(x) = \langle x\rangle^{-2(\nu+1)}$ suffices.

- Off-diagonal estimates at scale $\sqrt{t}$ are stable under composition. That is, if T_t satisfies (p_1, p_2) off-diagonal estimates at scale $\sqrt{t}$ and S_t satisfies (p_2, p_3) off-diagonal estimates at scale $\sqrt{t}$ then S_tT_t will satisfy (p_1, p_3) off-diagonal estimates at scale $\sqrt{t}$. It should be noted, however, that the value of c or s in the above examples of ρ may change for the composition.

- For $p_0 \leq p \leq q \leq q_0$, Hölder's inequality implies that if an operator family satisfies (p_0, q_0) off-diagonal estimates at scale $\sqrt{t}$ then it will also satisfy (p, q) estimates.

- Off-diagonal estimates for $p \leq q$ do not imply $L^p - L^q$ boundness of T_t, see [AM07b].

In order to apply off-diagonal estimates, we often need to decompose the support of a function f into finitely overlapping balls with radius to match the scale.

Definition 4.2.6. We say that a collection of balls $\mathcal{B}$ has finite overlap if there exists a finite constant $\Lambda_{\mathcal{B}}$ such that

$$\|\sum_{B\in\mathcal{B}} \mathbb{1}_B\|_{L^\infty} = \Lambda_{\mathcal{B}}.$$

Remark 4.2.7. Let $\mathcal{B}$ be a collection of finite overlapping balls covering a set Ω. Then

$$\sum_{B\in\mathcal{B}} \mu(B) = \int_\Omega \sum_{B\in\mathcal{B}} \mathbb{1}_B \,\mathrm{d}\mu \leq \Lambda_{\mathcal{B}}\, \mu(\Omega).$$

Lemma 4.2.8. *Let $\Omega \subset \mathbb{X}$ be an open set, and let $\mathcal{R}$ be a family of finite overlapping balls, with the same radius, covering Ω. If there exists $m \in \mathbb{N}$ such that $mR \supset \Omega$ for all*

$R \in \mathcal{R}$, *then for any* $f \in L^{p_0}(\Omega)$, $p_0 \geq 1$, *we have*

$$\sum_{R\in\mathcal{R}} \left(\fint_R |f|^{p_0} \,\mathrm{d}\mu \right)^{1/p_0} \lesssim m^\nu \left(\fint_\Omega |f|^{p_0} \,\mathrm{d}\mu \right)^{1/p_0}. \tag{4.2.2}$$

Proof. For $p_0 > 1$, Hölder's inequality implies that

$$\begin{aligned} \sum_{R\in\mathcal{R}} \left(\fint_R |f|^{p_0} \,\mathrm{d}\mu \right)^{\frac{1}{p_0}} &\leq \left(\sup_{R\in\mathcal{R}} \frac{1}{|R|} \right)^{\frac{1}{p_0}} \left(\sum_{R\in\mathcal{R}} \int_R |f|^{p_0} \,\mathrm{d}\mu \right)^{\frac{1}{p_0}} \left(\sum_{R\in\mathcal{R}} 1 \right)^{\frac{1}{p_0'}} \\ &= \left(\sup_{R\in\mathcal{R}} \frac{|\Omega|}{|R|} \right)^{\frac{1}{p_0}} \left(\fint_\Omega |f|^{p_0} \,\mathrm{d}\mu \right)^{\frac{1}{p_0}} (\#\mathcal{R})^{\frac{1}{p_0'}}. \end{aligned}$$

Since $mR \supset \Omega$ for all $R \in \mathcal{R}$, the doubling property implies that

$$\left(\sup_{R\in\mathcal{R}} \frac{|\Omega|}{|R|} \right)^{\frac{1}{p_0}} (\#\mathcal{R})^{\frac{1}{p_0'}} \lesssim \sup_{R\in\mathcal{R}} \frac{|mR|}{|R|} \lesssim m^\nu.$$

The case $p_0 = 1$ is even simpler since it does not require the use of Hölder's inequality nor an estimate on the cardinality $\#\mathcal{R}$. □

Remark 4.2.9. If T_s satisfies (p_0, q_0) off-diagonal estimates at scale $\sqrt{s}$, then it satisfies

$$\left(\fint_{B(r)} |T_s(f \mathbb{1}_{B_1})|^{q_0} \,\mathrm{d}\mu \right)^{1/q_0} \lesssim \rho \left(\frac{d(B_1, B(r))}{\sqrt{s}} \right) \left(\fint_{B_1} |f|^{p_0} \,\mathrm{d}\mu \right)^{1/p_0}, \tag{4.2.3}$$

for balls $B(r)$ of radius $r \geq \sqrt{s}$ and B_1 of radius $\sqrt{s}$.

Proof of (4.2.3)*.* It is enough to cover the larger ball $B(r)$ with a collection $\mathcal{B}$ of smaller, finite overlapping balls of radius $\sqrt{s}$.

$$\begin{aligned} \left(\fint_{B(r)} |T_s(f \mathbb{1}_{B_1})|^{q_0} \,\mathrm{d}\mu \right)^{1/q_0} &= \left(\sum_{B\in\mathcal{B}} \frac{|B|}{|B(r)|} \fint_B |T_s(f \mathbb{1}_{B_1})|^{q_0} \,\mathrm{d}\mu \right)^{1/q_0} \\ &\leq \left(\sum_{B\in\mathcal{B}} \frac{|B|}{|B(r)|} \right)^{1/q_0} \left(\sup_{B\in\mathcal{B}} \fint_B |T_s(f \mathbb{1}_{B_1})|^{q_0} \,\mathrm{d}\mu \right)^{1/q_0} \\ (\text{ by Remark 4.2.7 }) &\leq \Lambda_{\mathcal{B}}^{1/q_0} \sup_{B\in\mathcal{B}} \left(\fint_B |T_s(f \mathbb{1}_{B_1})|^{q_0} \,\mathrm{d}\mu \right)^{1/q_0}. \end{aligned}$$

We can use off-diagonal estimates at scale $\sqrt{s}$ to obtain

$$\sup_{B\in\mathcal{B}} \left(\fint_B |T_s(f\mathbb{1}_{B_1})|^{q_0} \,\mathrm{d}\mu \right)^{1/q_0} \lesssim \sup_{B\in\mathcal{B}} \rho\left(\frac{d(B,B_1)}{\sqrt{s}}\right) \left(\fint_{B_1} |f|^{p_0} \,\mathrm{d}\mu \right)^{1/p_0}.$$

The estimate then follows from the fact that the supremum of $\rho(d(B,B_1)/\sqrt{s})$ over $B \in \mathcal{B}$ is at most $\rho\big(d(B(r),B_1)/\sqrt{s}\big)$. $\square$

We denote the semigroup by $P_t := e^{-tL}$. This is used as an approximation of the identity at scale $\sqrt{t}$, since for any $p \in (p_0,q_0)$ we have

$$\lim_{t\to 0} \|f - e^{-tL}f\|_{L^p} = 0 \quad \text{and} \quad \lim_{t\to\infty} \|e^{-tL}f\|_{L^p} = 0.$$

For $N > 0$, we also consider the family of operators $Q_t^{(N)} := c_N^{-1}(tL)^N e^{-tL}$ with $c_N = \int_0^\infty s^N e^{-s} \frac{\mathrm{d}s}{s}$. These operators will satisfy an adapted Calderón reproducing formula for functions $f \in L^p$ with $p \in (p_0,q_0)$, namely

$$f = \int_0^\infty Q_t^{(N)} f \frac{\mathrm{d}t}{t}.$$

Also define

$$P_t^{(N)} := \int_1^\infty Q_{st}^{(N)} \frac{\mathrm{d}s}{s}.$$

Then $P_t^{(N)}$ is related to the operator $Q_t^{(N)}$ through $t\partial_t P_t^{(N)} = -Q_t^{(N)}$. We also have that as L^p-bounded operators,

$$P_t^{(N)} = \mathrm{Id} + \int_0^t Q_s^{(N)} \frac{\mathrm{d}s}{s}.$$

Remark 4.2.10. It is known that for any integer $N \in \mathbb{N}\setminus\{0\}$ the operators $P_t^{(N)}$ and $Q_t^{(N)}$ satisfy (p,p) off-diagonal estimates at scale $\sqrt{t}$ for all $t > 0$ and all $p \in [p_0,q_0]$ with $p < \infty$ (see the arguments in [HLM+11, Prop 3.1], for instance).

Definition 4.2.11 (Off-diagonal estimates at all scales)**.** A family of operators $\{T_t\}_{t>0}$ is said to satisfy (p_0,q_0) off-diagonal estimates at all scales if for all balls B_1, B_2 of radius

r_1, r_2 we have

$$\|T_t\|_{L^{p_0}(B_1)\to L^{q_0}(B_2)} \lesssim \left|B_{1,\sqrt{t}}\right|^{-\frac{1}{p_0}} \left|B_{2,\sqrt{t}}\right|^{\frac{1}{q_0}} \rho\left(\frac{d(B_1, B_2)}{\sqrt{t}}\right),$$

where $B_{i,\sqrt{t}} := (\sqrt{t}/r_i)B_i$ for $i = 1, 2$ and $\rho\colon [0,\infty) \to (0,1]$ is a non-increasing function such that $\lim_{x\to\infty}|x|^a\rho(x) = 0$ for some $a \geq 0$, and $\rho(0) = 1$.

It is trivial to see that off-diagonal estimates at all scales implies off-diagonal estimates at scale $\sqrt{t}$. This stronger condition is used in our cancellation hypothesis, Assumption 4.1.2(c).

Let $\psi : (0,\infty) \to (0,\infty)$ be a non-decreasing function. A space of homogeneous type $(\mathbb{X}, \mu)$ is said to be of ψ-growth if

$$|B(x,r)| = \mu(B(x,r)) \eqsim \psi(r)$$

uniformly for all $x \in \mathbb{X}$ and $r > 0$. Notice that this condition is stronger than (4.1.3). For spaces of ψ-growth, one encounters another notion of off-diagonal estimate. These types of estimates are studied in [AM07b].

Definition 4.2.12 (Full off-diagonal estimates)**.** Suppose that $(\mathbb{X}, \mu)$ is of ψ-growth. A family of operators $\{T_t\}_{t>0}$ is said to satisfy (p_0, q_0) full off-diagonal estimates if for all closed sets E, F we have

$$\|T_t\|_{L^{p_0}(E)\to L^{q_0}(F)} \lesssim \psi(\sqrt{t})^{\frac{1}{q_0}-\frac{1}{p_0}} \rho\left(\frac{d(E,F)}{\sqrt{t}}\right),$$

where $\rho\colon [0,\infty) \to (0,1]$ is a non-increasing function such that $\lim_{x\to\infty}|x|^a\rho(x) = 0$ for some $a \geq 0$, and $\rho(0) = 1$.

Remark 4.2.13. It is not difficult to show that for spaces of ψ-growth, the three different notions of off-diagonal estimates, Definitions 4.2.4, 4.2.11 and 4.2.12, are all equivalent for a particular choice of ρ.

4.3 Applications

In this section, we consider two distinct applications of our quadratic sparse domination result and Corollary C. For the first application, weighted estimates for square functions associated with divergence form elliptic operators will be proved. For the particular case of the Laplacian operator Δ, this will allow us to recover some estimates from [BD20a]. The second example that we will look at are square functions associated with the Laplace–Beltrami operator on a Riemannian manifold.

4.3.1 Elliptic Operators

Fix $n \in \mathbb{N}\backslash\{0\}$ and consider the Euclidean space $\mathbb{R}^d$ equipped with the Lebesgue measure. This is a space of ψ-growth, so all definitions of off-diagonal estimates are equivalent, see Remark 4.2.13.

Let A be an $n \times n$ matrix-valued function on $\mathbb{R}^d$ that is bounded and elliptic in the sense that

$$\mathrm{Re}\langle A(x)\xi, \xi\rangle_{\mathbb{C}^n} \geq \lambda|\xi|^2,$$

for some $\lambda > 0$, for all ξ, $x \in \mathbb{R}^d$. Consider the divergence form elliptic operator

$$L = -\mathrm{div}A\nabla,$$

defined through its corresponding sesquilinear form as a densely defined and maximally accretive operator on $L^2(\mathbb{R}^d)$. The operator L generates an analytic semigroup $\left\{e^{-zL}\right\}_{z\in\Sigma_{\pi/2-\theta}}$, where

$$\theta := \sup\left\{|\arg\langle Lf, f\rangle| : f \in \mathcal{D}_2(L)\right\}.$$

Let g_L and G_L denote the square function operators associated with L defined by

$$g_L f := \left(\int_0^\infty |tL\, e^{-tL} f|^2 \,\frac{\mathrm{d}t}{t}\right)^{1/2} \quad \text{and} \quad G_L f := \left(\int_0^\infty |\sqrt{t}\nabla e^{-tL} f|^2 \,\frac{\mathrm{d}t}{t}\right)^{1/2}.$$

In the articles [AM07b] and [AM06], off-diagonal estimates for the constituent operators of g_L and G_L were studied in great detail. The below proposition outlines some properties of such off-diagonal estimates that will be required in order to apply Corollary C to these two square functions.

Proposition 4.3.1 ([AM06, Prop. 3.3])**.** *For $m \in \mathbb{N}$ and $0 < \mu < \pi/2 - \theta$, there exists maximal intervals $\mathcal{J}^m(L)$ and $\mathcal{K}^m(L)$ in $[1, \infty]$ satisfying the below properties.*

- *If p_0, $q_0 \in \mathcal{J}^m(L)$ with $p_0 \leq q_0$ then $\left\{(zL)^m e^{-zL}\right\}_{z \in \Sigma_\mu}$ satisfies (p_0, q_0) full off-diagonal estimates.*

- *If p_0, $q_0 \in \mathcal{K}^m(L)$ with $p_0 \leq q_0$ then $\left\{\sqrt{z}\nabla(zL)^m e^{-zL}\right\}_{z \in \Sigma_\mu}$ satisfies (p_0, q_0) full off-diagonal estimates.*

- *The interiors $\operatorname{int} \mathcal{J}^m(L)$ and $\operatorname{int} \mathcal{K}^m(L)$ are independent of m.*

- *The inclusion $\mathcal{K}^m(L) \subseteq \mathcal{J}^m(L)$ is satisfied.*

- *The point $p = 2$ is contained in $\mathcal{K}^m(L)$.*

Remark 4.3.2. Observe that for any $m \geq 1$, $\mathcal{J}^1(L) \subset \mathcal{J}^m(L)$. To see this, let $p_0, q_0 \in \mathcal{J}^1(L)$ with $p_0 \leq q_0$. Then $(tL)e^{-tL/m}$ must satisfy both (p_0, q_0) and (q_0, q_0) off-diagonal estimates. This fact, when combined with the decomposition

$$(tL)^m e^{-tL} = (tL)e^{-tL/m} \cdots (tL)e^{-tL/m}$$

and the property that full off-diagonal estimates are stable under composition (c.f. [AM07b, Thm. 2.3 (b)]) then implies that p_0, $q_0 \in \mathcal{J}^m(L)$.

It is also not difficult to see that $\mathcal{J}^0(L) \subset \mathcal{J}^1(L)$. Indeed, consider the expression

$$tLe^{-tL} = e^{-\frac{t}{3}L} \cdot (tL)e^{-\frac{t}{3}L} \cdot e^{-\frac{t}{3}L}.$$

For p_0, $q_0 \in \mathcal{J}^0(L)$ with $p_0 < 2 < q_0$, Proposition 4.3.1 tells us that the operator $e^{-\frac{t}{3}L}$ will satisfy both $(p_0, 2)$ and $(2, q_0)$ full off-diagonal estimates. It is also well-known that $tLe^{-\frac{t}{3}L}$ satisfies $(2, 2)$ full off-diagonal estimates. The stability of full off-diagonal estimates under composition then implies that tLe^{-tL} satisfies (p_0, q_0) full off-diagonal estimates.

Applying Corollary C to the operators L and g_L will produce the following weighted result.

Proposition 4.3.3. *Let p_0, $q_0 \in \mathcal{J}^0(L)$ with $p_0 < 2 < q_0$. Then, for any $p \in (p_0, q_0)$ and $w \in A_{\frac{p}{p_0}} \cap RH_{(\frac{q_0}{p})'}$,*

$$\|g_L\|_{L^p(w)} \lesssim \left([w]_{A_{\frac{p}{p_0}}} \cdot [w]_{RH_{(\frac{q_0}{p})'}}\right)^{\gamma(p)},$$

where $\gamma(p)$ is as defined in Corollary C.

Proof. To prove the proposition, it is sufficient to check that the hypotheses of Corollary C, namely Assumptions 4.1.1 and 4.1.2, are valid for the operators L and g_L and the indices p_0, q_0. Assumption 4.1.1 is clearly valid since the definition of $\mathcal{J}^0(L)$ implies that the semigroup e^{-tL} will satisfy (p_0, q_0) full off-diagonal estimates.

It remains to prove the validity of Assumption 4.1.2. Part (a), the L^2-boundedness of g_L, follows from the fact that L possesses a bounded holomorphic functional calculus on L^2. Assumption 4.1.2(b), the off-diagonal estimates of the operator family tLe^{-tL} is given by Remark 4.3.2. Assumption 4.1.2(c) follows on observing that

$$\begin{aligned} \mathcal{Q}_s(tL)^N e^{-tL} &= sLe^{-sL}(tL)^N e^{-tL} \\ &= \frac{st^N}{(s+t)^{N+1}}((s+t)L)^{N+1} e^{-(s+t)L} \end{aligned}$$

and that since p_0, $q_0 \in \mathcal{J}^0(L)$ the operator family $\Theta_r^{(N)} = (rL)^{N+1}e^{-rL}$ will possess (p_0, q_0) full off-diagonal bounds for any $N \geq N_0 = 0$ by Remark 4.3.2. Finally, for Assumption

4.1.2(d), in the proof of [AM06, Thm. 7.2 (a)] it was shown that for any ball $B(x,r)$ we have

$$\left(\fint_{B(x,r)} |g_L e^{-r^2 L} f|^{q_0}\, \mathrm{d}\mu\right)^{1/q_0} \lesssim \sum_{j\geq 1} c(j) \left(\fint_{2^{j+1}B(x,r)} |g_L f|^{p_0}\, \mathrm{d}\mu\right)^{1/p_0}, \tag{4.3.1}$$

for some sequence of numbers $c(j) > 0$ that satisfies $\sum_{j\geq 1} c(j) \lesssim 1$. It should be noted that this argument was written for the square function with constituent operators $(tL)^{\frac{1}{2}} e^{-tL}$, but it applies equally well to our choice of square function. This clearly implies that

$$\left(\fint_{B(x,r)} |g_L e^{-r^2 L} f|^{q_0}\, \mathrm{d}\mu\right)^{1/q_0} \lesssim \inf_{y\in B(x,r)} \mathcal{M}_{p_0}(g_L f)(y),$$

and thus Assumption 4.1.2(d) is valid. □

Similarly, Corollary C can be applied to the square function G_L.

Proposition 4.3.4. *Let p_0, $q_0 \in \mathcal{K}^0(L)$ with $p_0 < 2 < q_0$. Then, for any $p \in (p_0, q_0)$ and $w \in A_{\frac{p}{p_0}} \cap RH_{(\frac{q_0}{p})'}$,*

$$\|G_L\|_{L^p(w)} \lesssim \left([w]_{A_{\frac{p}{p_0}}} \cdot [w]_{RH_{(\frac{q_0}{p})'}}\right)^{\gamma(p)}.$$

Proof. In order to apply Corollary C, it is sufficient to show that G_L satisfies Assumptions 4.1.1 and 4.1.2. Assumption 4.1.1 is implied by p_0, $q_0 \in \mathcal{K}^0(L) \subset \mathcal{J}^0(L)$.

Let us now demonstrate the validity of Assumption 4.1.2. The L^2-boundedness of G_L, Assumption 4.1.2(a), follows from the ellipticity condition of A and a straightforward integration by parts argument that can be found in [Aus07, pg. 74]. Assumption 4.1.2(b) is implied by the condition p_0, $q_0 \in \mathcal{K}^0(L)$. For Assumption 4.1.2(c), notice that

$$\begin{aligned}
\mathcal{Q}_s Q_t^{(N)} &= \sqrt{s}\nabla e^{-sL}(tL)^N e^{-tL}\\
&= \frac{s^{\frac{1}{2}} t^N}{(s+t)^{N+\frac{1}{2}}} \sqrt{s+t}\nabla\left((s+t)L\right)^N e^{-(s+t)L}\\
&=: \frac{s^{\frac{1}{2}} t^N}{(s+t)^{N+\frac{1}{2}}} \Theta_{s+t}^{(N)}.
\end{aligned}$$

Also observe that

$$\Theta_r^{(N)} = \sqrt{r}\nabla e^{-rL/2}(rL)^N e^{-rL/2}.$$

As p_0, $q_0 \in \mathcal{K}^0(L)$, Proposition 4.3.1 tells us that operator family $\sqrt{r}\nabla e^{-rL/2}$ will satisfy $(2, q_0)$ full off-diagonal estimates. Similarly, since $\mathcal{K}^0(L) \subset \mathcal{J}^N(L)$ for any $N \geq N_0 = 0$, the family $(rL)^N e^{-rL/2}$ satisfies $(p_0, 2)$ full off-diagonal bounds. It then follows from the stability of full off-diagonal bounds under composition that the family $\Theta_r^{(N)}$ will satisfy (p_0, q_0) full off-diagonal bounds. This proves that Assumption 4.1.2(c) is satisfied.

Finally, for Assumption 4.1.2(d), in the proof of [AM06, Thm. 7.2 (b)] it was proved that for any ball $B(x,r)$ we have

$$\left(\fint_{B(x,r)} |G_L e^{-r^2 L} f|^{q_0} \,\mathrm{d}\mu\right)^{\frac{1}{q_0}} \lesssim \sum_{j\geq 1} d(j) \left(\fint_{2^{j+1}B(x,r)} |G_L f|^{p_0} \,\mathrm{d}\mu\right)^{\frac{1}{p_0}},$$

for some sequence of numbers $d(j) > 0$ that satisfies $\sum_{j\geq 1} d(j) \lesssim 1$. This clearly implies that

$$\left(\fint_{B(x,r)} |G_L e^{-r^2 L} f|^{q_0} \,\mathrm{d}\mu\right)^{\frac{1}{q_0}} \lesssim \inf_{y\in B(x,r)} \mathcal{M}_{p_0}(G_L f)(y), \tag{4.3.2}$$

and thus Assumption 4.1.2(d) is valid. □

Remark 4.3.5. If A is real-valued then it is known that $\mathcal{J}^0(L) = [1, \infty]$ (c.f. [AM06]). Proposition 4.3.3 will then imply that

$$\|g_L\|_{L^p(w)} \lesssim [w]_{A_p}^{\max\left(\frac{1}{p-1}, \frac{1}{2}\right)}$$

for all $w \in A_p$. If, in addition to being real-valued, A has also smooth coefficients this result was proved by Bui and Duong in [BD20a]. In this case it is known that $\mathcal{K}^0(L) = [1, \infty]$. Proposition 4.3.4 then implies that

$$\|G_L\|_{L^p(w)} \lesssim [w]_{A_p}^{\max\left(\frac{1}{p-1}, \frac{1}{2}\right)},$$

which reproduces a result in [BD20a]. Also, in the same work, the authors showed that square functions associated with $\sqrt{L}$ are dominated by the corresponding one associated with L [BD20a, Thm. 1.4]. In particular, our bounds for g_L and G_L imply the same bound for the square function $g_{\sqrt{L}}$ and $G_{\sqrt{L}}$, recovering weighted estimates for the vertical square functions in [Ler11].

Remark 4.3.6. For $A = I$ we have $L = \Delta$ and it is then known that $\mathcal{J}^0(L) = \mathcal{K}^0(L) = [1, \infty]$. We can then take $p_0 = 1$ and $q_0 = \infty$ in Propositions 4.3.3 and 4.3.4. This will produce the weighted estimates

$$\|g_\Delta\|_{L^p(w)},\ \|G_\Delta\|_{L^p(w)} \lesssim \left([w]_{A_p}\,[w]_{RH_1}\right)^{\max\left(\frac{1}{p-1},\frac{1}{2}\right)} = [w]_{A_p}^{\max\left(\frac{1}{p-1},\frac{1}{2}\right)}$$

for all $w \in A_p \cap RH_1 = A_p$. For both square functions, it is known that these estimates are optimal in the sense that they will not hold for an exponent of $[w]_{A_p}$ any smaller than the above exponent. This provides a new proof of weighted boundedness of the standard square functions associated with Δ with optimal dependence on the constant $[w]_{A_p}$.

4.3.2 Laplace–Beltrami

Let $\mathbb{X}$ be a complete, connected, non-compact Riemannian manifold. It will be assumed that the Riemannian measure μ satisfies the volume doubling property. In addition, it will also be assumed that there exists a function $\psi : (0, \infty) \to (0, \infty)$ for which

$$|B(x, r)| = \mu(B(x, r)) \eqsim \psi(r)$$

uniformly for all $x \in \mathbb{X}$ and $r > 0$. That is, the manifold is of ψ-growth. Enforcing this stronger growth condition will allow us to interchange our different notions of off-diagonal estimates (c.f. Remark 4.2.13).

Consider the Laplace–Beltrami operator Δ defined as an unbounded operator on

$L^2(\mathbb{X},\mu)$ through the integration by parts formula

$$\langle \Delta f, f\rangle = \| |\nabla f| \|_2^2$$

for $f \in C_0^\infty(\mathbb{X})$, where ∇ is the Riemannian gradient. The positivity of Δ implies that it will generate an analytic semigroup $e^{-t\Delta}$ on $L^2(\mathbb{X},\mu)$.

Recall that the heat kernel $k_t(x,y)$ of Δ is said to satisfy Gaussian upper bounds if there exists $c > 0$ such that

$$k_t(x,y) \lesssim \frac{1}{|B(x,\sqrt{t})|} e^{-c\frac{d^2(x,y)}{t}}$$

for all x, $y \in \mathbb{X}$ and $t > 0$. This is a very common assumption that is imposed when considering the boundedness of singular operators on Riemannian manifolds. For further information refer to [CD99], [ACDH04] or [AM08]. Consider the square function g_Δ defined through

$$g_\Delta f := \left(\int_0^\infty |t\Delta e^{-t\Delta} f|^2 \frac{\mathrm{d}t}{t} \right)^{1/2}.$$

The boundedness for square functions of this form on unweighted $L^p(\mathbb{X})$ with $1 < p < \infty$ is known to hold in the general symmetric Markov semigroup setting [Ste70, pg. 111]. Let us consider the weighted case on the full range of $p \in (1,\infty)$.

Proposition 4.3.7. *Suppose that the heat kernel for $\mathbb{X}$ satisfies Gaussian upper bounds. Then, for any $p \in (1,\infty)$ and $w \in A_p$,*

$$\|g_\Delta\|_{L^p(w)} \lesssim [w]_{A_p}^{\max\left(\frac{1}{2},\frac{1}{p-1}\right)}.$$

Proof. This result will follow from Corollary C provided that Assumptions 4.1.1 and 4.1.2 are verified to hold with $p_0 = 1$ and $q_0 = \infty$.

For Assumption 4.1.1, it is known that the heat kernel satisfying Gaussian upper bounds is equivalent to the semigroup $e^{-t\Delta}$ satisfying $(1,\infty)$ full off-diagonal estimates. For proof, the reader is referred to [AM07b, Prop. 2.2] and [AM07b, Prop. 3.3]. Thus

Assumption 4.1.1 will be valid.

For Assumption 4.1.2(a), the L^2-boundedness of g_Δ follows from the bounded holomorphic functional calculus of Δ on L^2. For Assumption 4.1.2(b), notice that

$$t\Delta e^{-t\Delta} = e^{-\frac{t}{3}\Delta} \cdot t\Delta e^{-\frac{t}{3}\Delta} \cdot e^{-\frac{t}{3}\Delta}.$$

Observe that since the semigroup $e^{-t\Delta}$ satisfies $(1,\infty)$ full off-diagonal estimates, $e^{-t\Delta}$ will satisfy both $(1,2)$ and $(2,\infty)$ full off-diagonal bounds. At the same time, $t\Delta e^{-t\Delta}$ is well-known to satisfy $(2,2)$ full off-diagonal bounds (c.f. [ACDH04, pg. 930] and [Dav95, Lemma 7]). It then follows from the stability of full off-diagonal bounds under composition ([AM07b, Thm. 2.3 (b)]) that $t\Delta e^{-t\Delta}$ satisfies $(1,\infty)$ full off-diagonal bounds. This proves that Assumption 4.1.2(b) is satisfied.

Assumption 4.1.2(c) follows from the expression

$$\mathcal{Q}_s(t\Delta)^N e^{-t\Delta} = \frac{st^N}{(s+t)^{N+1}} \left[(s+t)\Delta\right]^{N+1} e^{-(s+t)\Delta}$$

and the fact that the operator family $\{(r\Delta)^{N+1}e^{-r\Delta}\}_{r>0}$ satisfies $(1,\infty)$ full off-diagonal bounds by an argument similar to that of Remark 4.3.2.

Finally, the validity of Assumption 4.1.2(d) can be proved in an identical manner to the argument used to obtain (4.3.1). This argument can be found in [AM06, §7] on pages 729–730. This argument in the elliptic setting follows from a combination of the off-diagonal estimates of the constituent operators, the fact that the constituent operators are expressible in terms of the semigroup and a variation of the Marcinkiewicz–Zygmund theorem [Gra14, Thm. 5.5.1]. All three of these components will hold for our square function in this Riemannian manifold setting and thus the argument will be valid. □

Next, we will apply our sparse result to the square function

$$G_\Delta f := \left(\int_0^\infty |\sqrt{t}\nabla e^{-t\Delta} f|^2 \frac{\mathrm{d}t}{t} \right)^{1/2}.$$

Define

$$q_+ := \sup\left\{p \in (1,\infty) : \||\nabla\Delta^{-\frac{1}{2}}f|\|_p \lesssim \|f\|_p\right\}.$$

The weighted boundedness of the Riesz transforms operator $\nabla\Delta^{-\frac{1}{2}}$ on $L^p(\mathbb{X}, w\,\mathrm{d}\mu)$ was considered for $p \in (1, q_+)$ in [AM08]. Owing to the strong connection between the Riesz transforms and the square function G_Δ, the range $(1, q_+)$ will also be a natural interval over which to consider the boundedness of G_Δ. From the definition of q_+ and the L^2-boundedness of $\nabla\Delta^{-\frac{1}{2}}$, it is clear that $q_+ \geq 2$. In the below proposition we assume this inequality to be strict.

Proposition 4.3.8. *Assume that the heat kernel of $\mathbb{X}$ satisfies Gaussian upper bounds and that $q_+ > 2$. Let $2 < q_0 < q_+$ and $p \in [1, q_0)$. Then for any $w \in A_p \cap RH_{(\frac{q_0}{p})'}$,*

$$\|G_\Delta\|_{L^p(w)} \lesssim \left([w]_{A_p} \cdot [w]_{RH_{(\frac{q_0}{p})'}}\right)^{\gamma(p)}.$$

Proof. Once again, let us apply Corollary C. Assumption 4.1.1 will be true for the same reason as in Proposition 4.3.7. Assumption 4.1.2(a) is well-known and can be obtained by combining the L^2-boundedness of $\nabla\Delta^{-\frac{1}{2}}$ together with the bounded holomorphic functional calculus of Δ on L^2.

Let us show that the family of operators $\mathcal{Q}_t = \sqrt{t}\nabla e^{-t\Delta}$ satisfies $(1, q_0)$ off-diagonal estimates at scale $\sqrt{t}$ with $\rho(x) = \exp(-cx^2)$, for some $c > 0$. Fix balls B_1, $B_2 \subset \mathbb{X}$ of radius $\sqrt{t}$. From the argument in the proof of [ACDH04, Prop 1.10],

$$\left(\int_{\mathbb{X}} |\nabla_x k_t(x,y)|^{q_0} e^{c\frac{d^2(x,y)}{t}}\,\mathrm{d}\mu(x)\right)^{\frac{1}{q_0}} \lesssim \frac{1}{\sqrt{t}|B(y,\sqrt{t})|^{1-\frac{1}{q_0}}}$$

for all $t > 0$ and $y \in \mathbb{X}$, where $c > 0$ is dependent on q_0. This immediately implies that

$$\begin{aligned}\left(\fint_{B_2} |\nabla_x k_t(x,y)|^{q_0}\,\mathrm{d}\mu(x)\right)^{\frac{1}{q_0}} &\lesssim \frac{1}{\sqrt{t}} e^{-c\frac{d^2(B_1,B_2)}{t}} \frac{1}{|B(y,\sqrt{t})|^{1-\frac{1}{q_0}}|B_2|^{\frac{1}{q_0}}} \\ &\approx \frac{1}{\sqrt{t}} e^{-c\frac{d^2(B_1,B_2)}{t}} \frac{1}{\psi(\sqrt{t})},\end{aligned}$$

where the last line follows from the uniform ψ-growth condition imposed upon our manifold. For f supported in B_1, Minkowski's inequality followed by the previous estimate produces

$$\begin{aligned}
\left(\fint_{B_2} |\sqrt{t}\nabla e^{-t\Delta} f(x)|^{q_0} \, \mathrm{d}\mu(x)\right)^{\frac{1}{q_0}} &= \left(\fint_{B_2} \left|\int_{B_1} \sqrt{t}\nabla_x k_t(x,y) f(y) \, \mathrm{d}\mu(y)\right|^{q_0} \mathrm{d}\mu(x)\right)^{\frac{1}{q_0}} \\
&\leq \int_{B_1} \left(\fint_{B_2} |\sqrt{t}\nabla_x k_t(x,y)|^{q_0} \, \mathrm{d}\mu(x)\right)^{\frac{1}{q_0}} |f(y)| \, \mathrm{d}\mu(y) \\
&\lesssim \frac{1}{\psi(\sqrt{t})} e^{-c\frac{d^2(B_1,B_2)}{t}} \int_{B_1} |f(y)| \, \mathrm{d}\mu(y) \\
&\eqsim e^{-c\frac{d^2(B_1,B_2)}{t}} \fint_{B_1} |f(y)| \, \mathrm{d}\mu(y).
\end{aligned}$$

Let us now prove that Assumption 4.1.2(c) is valid. Observe that

$$\mathcal{Q}_s(t\Delta)^N e^{-t\Delta} = \frac{s^{\frac{1}{2}} t^N}{(s+t)^{N+\frac{1}{2}}} \sqrt{s+t}\nabla e^{-\frac{s+t}{2}\Delta} \left[(s+t)\Delta\right]^N e^{-\frac{s+t}{2}\Delta} =: \frac{s^{\frac{1}{2}} t^N}{(s+t)^{N+\frac{1}{2}}} \Theta^{(N)}_{s+t}.$$

Observe that the operator family $\{(r\Delta)^N e^{-r\Delta}\}_{r>0}$ satisfies $(1,\infty)$ full off-diagonal estimates. Recall that for spaces of ψ-growth the three different forms of off-diagonal estimates, Definitions 4.2.4, 4.2.11 and 4.2.12, are all equivalent. This, when combined with Hölder's inequality, implies that this operator family satisfies $(1,2)$ off-diagonal estimates at scale $\sqrt{r}$. Similarly, the family $\{\sqrt{r}\nabla e^{-r\Delta}\}_{r>0}$ satisfies $(2,q_0)$ off-diagonal estimates at scale $\sqrt{r}$. The stability of off-diagonal estimates under composition then implies that the operator family Θ_r satisfies $(1,q_0)$ at scale $\sqrt{r}$, which implies $(1,q_0)$ off-diagonal estimates at all scales. This proves Assumption 4.1.2(c).

Finally, the validity of Assumption 4.1.2 (d) can be proved in an identical manner to the argument used to obtain (4.3.2). This argument can be found in [AM06, §7] on page 732. This argument in the elliptic setting follows from a combination of the off-diagonal estimates of the constituent operators, the fact that the constituent operators are expressible in terms of the semigroup and a variation of the Marcinkiewicz–Zygmund theorem [Gra14, Thm. 5.5.1]. All three of these components will hold for our square function in this Riemannian manifold setting and thus the argument will be valid. □

4.4 Boundedness of the Maximal Function

Throughout this section, fix p_0, $q_0 \in [1, \infty]$, $N_0 \in \mathbb{N}$ and operators L and S satisfying Assumptions 4.1.1 and 4.1.2 for such a choice of p_0, q_0. For a ball B we denote by $r(B)$ its radius. Define the following maximal operator associated with the square function,

$$S^* f(x) := \sup_{\substack{B \text{ ball} \\ B \ni x}} \left(\fint_B |S^{[r(B)^2, \infty)} f|^{q_0} \, d\mu \right)^{1/q_0}$$
$$:= \sup_{\substack{B \text{ ball} \\ B \ni x}} \left(\fint_B \left(\int_{r(B)^2}^{\infty} |\mathcal{Q}_t f|^2 \frac{dt}{t} \right)^{\frac{q_0}{2}} d\mu \right)^{1/q_0}.$$

In this section, our aim is to prove the following boundedness result for S^*.

Theorem 4.4.1. *The maximal function S^* is bounded on L^2 and weak-type (p_0, p_0).*

The boundedness of this maximal function constitutes an important part of our sparse domination argument. The reliance of our argument on an associated maximal function is a well-known method for obtaining sparse bounds and finds its origins in the work of Lacey [Lac17]. It was later streamlined by Lerner [Ler16]. Quite often, the issue of proving sparse domination for a particular operator can be reduced to determining an appropriate associated maximal operator, proving its (weak) boundedness and then applying a stopping time argument that utilises this boundedness.

4.4.1 A Pointwise Estimate

In order to prove the boundedness of the operator S^* we will require a couple of preliminary lemmas. Given a ball B, we define the average of a function f over the annulus $S_k(B) := 2^{k+1}B \setminus 2^k B$ for $k \in \mathbb{N}$ as the integral over $S_k(B)$ normalised by $|2^k B|$.

Recall that A_0 is a positive number defined in Assumption 4.1.2 (c).

Lemma 4.4.2. *For any $0 < s < r^2 < t$ and $N \in \mathbb{N}$,*

$$\left(\fint_B |\mathcal{Q}_t Q_s^{(N)} f|^{q_0} \,\mathrm{d}\mu\right)^{\frac{1}{q_0}} \lesssim \frac{t^{A_0} s^N}{(t+s)^{A_0+N}} \left(\frac{\sqrt{t}}{r}\right)^{\frac{\nu}{q_0}} \sum_{j\geq 0} 2^{-j} \left(\fint_{S_j(\widetilde{B})} |f|^{p_0} \,\mathrm{d}\mu\right)^{\frac{1}{p_0}}$$

for any ball B of radius r and $\widetilde{B} := \frac{\sqrt{t}}{r} B$.

Proof. Fix B a ball of radius r. For $j \geq 0$, let $\mathcal{R}_j$ denote a collection of finite overlapping balls of radius $\sqrt{t}$ that is a cover for the set $S_j(\widetilde{B})$. Then, Assumption 4.1.2 (c) together with the triangle inequality produces

$$\begin{aligned}
\left(\fint_B |\mathcal{Q}_t Q_s^{(N)} f|^{q_0} \,\mathrm{d}\mu\right)^{\frac{1}{q_0}} &= \frac{t^{A_0} s^N}{(s+t)^{A_0+N}} \left(\fint_B |\Theta_{s+t}^{(N)} f|^{q_0} \,\mathrm{d}\mu\right)^{\frac{1}{q_0}} \\
&\leq \frac{t^{A_0} s^N}{(s+t)^{A_0+N}} \sum_{j\geq 0} \sum_{R\in\mathcal{R}_j} \left(\fint_B |\Theta_{s+t}^{(N)} (\mathbb{1}_R f)|^{q_0} \,\mathrm{d}\mu\right)^{\frac{1}{q_0}} \\
&\lesssim \frac{t^{A_0} s^N}{(s+t)^{A_0+N}} \sum_{j\geq 0} \sum_{R\in\mathcal{R}_j} \frac{|B|^{-\frac{1}{q_0}} |R|^{\frac{1}{p_0}}}{|B_{\sqrt{s+t}}|^{-\frac{1}{q_0}} |R_{\sqrt{s+t}}|^{\frac{1}{p_0}}} \left(1 + \frac{d(B,R)^2}{s+t}\right)^{-\frac{\nu+1}{2}} \left(\fint_R |f|^{p_0} \,\mathrm{d}\mu\right)^{\frac{1}{p_0}}.
\end{aligned} \tag{4.4.1}$$

On utilising the doubling property of our metric space and subsequently $s + t \eqsim t$,

$$\begin{aligned}
|B_{\sqrt{s+t}}| &\lesssim \left(\frac{\sqrt{s+t}}{r}\right)^{\nu} |B| \\
&\eqsim \left(\frac{\sqrt{t}}{r}\right)^{\nu} |B|.
\end{aligned} \tag{4.4.2}$$

This, together with the fact that $|R| \leq |R_{\sqrt{s+t}}|$ gives

$$\begin{aligned}
&\left(\fint_B |\mathcal{Q}_t Q_s^{(N)} f|^{q_0} \,\mathrm{d}\mu\right)^{\frac{1}{q_0}} \\
&\qquad \lesssim \frac{t^{A_0} s^N}{(s+t)^{A_0+N}} \left(\frac{\sqrt{t}}{r}\right)^{\frac{\nu}{q_0}} \sum_{j\geq 0} \sum_{R\in\mathcal{R}_j} \left(1 + \frac{d(B,R)^2}{s+t}\right)^{-\frac{\nu+1}{2}} \left(\fint_R |f|^{p_0} \,\mathrm{d}\mu\right)^{\frac{1}{p_0}}.
\end{aligned} \tag{4.4.3}$$

For $R \in \mathcal{R}_j$, since $d(B,R) \geq (2^j - 1)\sqrt{t} \eqsim (2^j-1)\sqrt{s+t}$ for $j \geq 1$, we have

$$\left(1 + \frac{d(B,R)^2}{s+t}\right)^{-\frac{\nu+1}{2}} \lesssim 2^{-j(\nu+1)}. \tag{4.4.4}$$

Let $\Omega = S_j(\widetilde{B})$ and $\mathcal{R}_j$ as defined above in this proof. The inclusion $\Omega \subset 2^{j+1}\widetilde{B} \subset 2^{j+2}R$ holds for any $R \in \mathcal{R}_j$ and $j \in \mathbb{N}$. Thus Lemma 4.2.8 implies that

$$\sum_{R \in \mathcal{R}_j} \left(\fint_R |f|^{p_0} \, \mathrm{d}\mu \right)^{\frac{1}{p_0}} \lesssim 2^{j\nu} \left(\fint_{S_j(\widetilde{B})} |f|^{p_0} \, \mathrm{d}\mu \right)^{\frac{1}{p_0}}.$$

Applying this estimate and (4.4.4) to (4.4.3) gives us our result. □

Using the previous lemma, the following result can then be proved using an argument identical to the first estimate of [BFP16, Lem. 4.1].

Lemma 4.4.3. *Fix $N \in \mathbb{N}$ with $N > \max(3\nu/2 + 1, N_0)$. For any ball B of radius $r(B) > 0$ and $t > r(B)^2$ we have*

$$\left(\fint_B |\mathcal{Q}_t(I - P^{(N)}_{r(B)^2}) f|^{q_0} \, \mathrm{d}\mu \right)^{\frac{1}{q_0}} \lesssim \left(\frac{r(B)^2}{t} \right)^{\frac{N}{2}} \sum_{l \geq 0} 2^{-l} \left(\fint_{2^l B} |f|^{p_0} \, \mathrm{d}\mu \right)^{\frac{1}{p_0}}. \tag{4.4.5}$$

Let $S^\#$ denote the maximal operator

$$S^\# f(x) := \sup_{\substack{B \text{ ball} \\ B \ni x}} \left(\fint_B |S P^{(N)}_{r(B)^2} f|^{q_0} \, \mathrm{d}\mu \right)^{\frac{1}{q_0}}.$$

This operator was introduced in [BFP16] and formed an important part of their sparse domination argument.

Proposition 4.4.4. *For every $x \in \mathbb{X}$,*

$$S^* f(x) \lesssim S^\# f(x) + \mathcal{M}_{p_0} f(x).$$

Proof. For $x \in \mathbb{X}$ and ball $B \subset \mathbb{X}$ containing x, the triangle inequality implies

$$\left(\fint_B \left(\int_{r(B)^2}^\infty |\mathcal{Q}_t f|^2 \frac{\mathrm{d}t}{t}\right)^{\frac{q_0}{2}} \mathrm{d}\mu\right)^{\frac{1}{q_0}} \leq \left(\fint_B \left(\int_{r(B)^2}^\infty |\mathcal{Q}_t (I - P^{(N)}_{r(B)^2}) f|^2 \frac{\mathrm{d}t}{t}\right)^{\frac{q_0}{2}} \mathrm{d}\mu\right)^{\frac{1}{q_0}} + \left(\fint_B \left(\int_{r(B)^2}^\infty |\mathcal{Q}_t P^{(N)}_{r(B)^2} f|^2 \frac{\mathrm{d}t}{t}\right)^{\frac{q_0}{2}} \mathrm{d}\mu\right)^{\frac{1}{q_0}}.$$

For the first term, apply Minkowski's inequality followed by Lemma 4.4.3 to obtain

$$\begin{aligned}
&\left(\fint_B \left(\int_{r(B)^2}^\infty |\mathcal{Q}_t (I - P^{(N)}_{r(B)^2}) f|^2 \frac{\mathrm{d}t}{t}\right)^{\frac{q_0}{2}} \mathrm{d}\mu\right)^{\frac{1}{q_0}} \\
&\quad\leq \left(\int_{r(B)^2}^\infty \left(\fint_B |\mathcal{Q}_t (I - P^{(N)}_{r(B)^2}) f|^{q_0} \mathrm{d}\mu\right)^{\frac{2}{q_0}} \frac{\mathrm{d}t}{t}\right)^{\frac{1}{2}} \\
&\quad\lesssim \left(\int_{r(B)^2}^\infty \left(\left(\frac{r(B)^2}{t}\right)^{\frac{N}{2}} \sum_{l \geq 0} 2^{-l} \left(\fint_{2^l B} |f|^{p_0} \mathrm{d}\mu\right)^{\frac{1}{p_0}}\right)^2 \frac{\mathrm{d}t}{t}\right)^{\frac{1}{2}} \\
&\quad= r(B)^N \left(\int_{r(B)^2}^\infty \frac{\mathrm{d}t}{t^{N+1}}\right)^{\frac{1}{2}} \sum_{l \geq 0} 2^{-l} \left(\fint_{2^l B} |f|^{p_0} \mathrm{d}\mu\right)^{\frac{1}{p_0}} \\
&\quad\lesssim \mathcal{M}_{p_0} f(x).
\end{aligned}$$

For the second term,

$$\begin{aligned}
\left(\fint_B \left(\int_{r(B)^2}^\infty |\mathcal{Q}_t P^{(N)}_{r(B)^2} f|^2 \frac{\mathrm{d}t}{t}\right)^{\frac{q_0}{2}} \mathrm{d}\mu\right)^{\frac{1}{q_0}} &\leq \left(\fint_B \left(\int_0^\infty |\mathcal{Q}_t P^{(N)}_{r(B)^2} f|^2 \frac{\mathrm{d}t}{t}\right)^{\frac{q_0}{2}} \mathrm{d}\mu\right)^{\frac{1}{q_0}} \\
&= \left(\fint_B |S P^{(N)}_{r(B)^2} f|^{q_0} \mathrm{d}\mu\right)^{\frac{1}{q_0}} \\
&\leq S^\# f(x).
\end{aligned}$$

We thus obtain the pointwise estimate (4.4.4). □

4.4.2 Cancellation of S with respect to L

As the operator $\mathcal{M}_{p_0}$ is L^2-bounded and weak-type (p_0, p_0), the pointwise bound of the previous section implies that in order to prove Theorem 4.4.1 it will be sufficient to show that $S^{\#}$ is L^2-bounded and weak-type (p_0, p_0). According to [BFP16, Prop. 4.6], $S^{\#}$ will be L^2-bounded and weak-type (p_0, p_0) if S satisfies the assumptions of [BFP16]. The only assumption from [BFP16] that is not included in our hypotheses is Assumption (b) of [BFP16], the cancellative property of S with respect to L. Instead, for us, the cancellation has been imposed upon the constituent operators $\mathcal{Q}_t$. In this section it will be proved that cancellation on $\mathcal{Q}_t$ with respect to L implies cancellation on S with respect to L.

Proposition 4.4.5. *There exists $\widetilde{N}_0 \geq N_0$ such that for all integers $N \geq \widetilde{N}_0$, $s > 0$ and balls B_1, B_2 of radius $\sqrt{s}$,*

$$\left(\fint_{B_2} |S Q_s^{(N)} f|^{q_0} \, \mathrm{d}\mu \right)^{\frac{1}{q_0}} \lesssim \left(1 + \frac{d(B_1, B_2)^2}{s} \right)^{-\frac{\nu+1}{2}} \left(\fint_{B_1} |f|^{p_0} \, \mathrm{d}\mu \right)^{\frac{1}{p_0}} \tag{4.4.6}$$

for all $f \in L^{p_0}(B_1)$.

Proof. For $I \subset [0, \infty)$, define the operator

$$S^I f(x) := \left(\int_I |\mathcal{Q}_t f|^2 \frac{\mathrm{d}t}{t} \right)^{1/2}.$$

In order to prove (4.4.6), it is sufficient to show that a similar estimate holds for the operators $S^{[0,s]}$ and $S^{[s,\infty)}$.

For $I \subset [0, \infty)$, Minkowski's inequality implies that

$$\begin{aligned}\left(\fint_{B_2} |S^I Q_s^{(N)} f|^{q_0} \, \mathrm{d}\mu \right)^{\frac{1}{q_0}} &= \left(\fint_{B_2} \left(\int_I |\mathcal{Q}_t Q_s^{(N)} f|^2 \frac{\mathrm{d}t}{t} \right)^{\frac{q_0}{2}} \mathrm{d}\mu \right)^{\frac{2}{q_0}\frac{1}{2}} \\ &\leq \left[\int_I \left(\fint_{B_2} |\mathcal{Q}_t Q_s^{(N)} f|^{q_0} \, \mathrm{d}\mu \right)^{\frac{2}{q_0}} \frac{\mathrm{d}t}{t} \right]^{\frac{1}{2}}.\end{aligned}$$

From Assumption 4.1.2(c) and the growth property (4.1.3), we have

$$\begin{aligned}
&\left(\fint_{B_2}|S^I Q_s^{(N)} f|^{q_0}\,\mathrm{d}\mu\right)^{\frac{1}{q_0}} \le \left[\int_I \frac{t^{2A_0}s^{2N}}{(t+s)^{2(A_0+N)}}\left(\fint_{B_2}|\Theta_{t+s}^{(N)} f|^{q_0}\,\mathrm{d}\mu\right)^{\frac{2}{q_0}}\frac{\mathrm{d}t}{t}\right]^{\frac{1}{2}}\\
&\lesssim \left[\int_I \frac{t^{2A_0}s^{2N}}{(t+s)^{2(A_0+N)}}\frac{|B_1|^{\frac{2}{p_0}}\cdot|B_2|^{-\frac{2}{q_0}}}{|B_{1,\sqrt{t+s}}|^{\frac{2}{p_0}}|B_{2,\sqrt{t+s}}|^{-\frac{2}{q_0}}}\left(1+\frac{d(B_1,B_2)^2}{t+s}\right)^{-(\nu+1)}\frac{\mathrm{d}t}{t}\right]^{1/2}\left(\fint_{B_1}|f|^{p_0}\,\mathrm{d}\mu\right)^{\frac{1}{p_0}}\\
&\eqsim \left[\int_I \frac{t^{2A_0}s^{2N}}{(t+s)^{2(A_0+N)}}\varphi\left(\frac{\sqrt{s}}{\sqrt{t+s}}\right)^{2\left(\frac{1}{p_0}-\frac{1}{q_0}\right)}\left(1+\frac{d(B_1,B_2)^2}{t+s}\right)^{-(\nu+1)}\frac{\mathrm{d}t}{t}\right]^{1/2}\left(\fint_{B_1}|f|^{p_0}\,\mathrm{d}\mu\right)^{\frac{1}{p_0}}.
\end{aligned}$$

The property that $\varphi(a)\le 1$ for $a\le 1$ then gives

$$\left(\fint_{B_2}|S^I Q_s^{(N)} f|^{q_0}\,\mathrm{d}\mu\right)^{\frac{1}{q_0}} \lesssim \left[\int_I \frac{t^{2A_0}s^{2N}}{(t+s)^{2(A_0+N)}}\left(1+\frac{d(B_1,B_2)^2}{t+s}\right)^{-(\nu+1)}\frac{\mathrm{d}t}{t}\right]^{\frac{1}{2}}\left(\fint_{B_1}|f|^{p_0}\,\mathrm{d}\mu\right)^{\frac{1}{p_0}}. \tag{4.4.7}$$

In order to prove the desired off-diagonal estimate, it is then sufficient to prove

$$\begin{aligned}
A_I &:= \int_I \frac{t^{2A_0}s^{2N}}{(t+s)^{2(A_0+N)}}\left(1+\frac{d(B_1,B_2)^2}{t+s}\right)^{-(\nu+1)}\frac{\mathrm{d}t}{t}\\
&\lesssim \left(1+\frac{d(B_1,B_2)^2}{s}\right)^{-(\nu+1)},
\end{aligned}\tag{4.4.8}$$

for both intervals $I=[0,s]$ and $I=[s,\infty)$. Consider first the interval $I=[0,s]$. For t contained in $[0,s]$ we will have $t+s\le 2s$ and therefore

$$\left(1+\frac{d(B_1,B_2)^2}{t+s}\right)^{-(\nu+1)} \lesssim \left(1+\frac{d(B_1,B_2)^2}{s}\right)^{-(\nu+1)}.$$

This gives

$$\begin{aligned}
A_I &\lesssim \left(1+\frac{d(B_1,B_2)^2}{s}\right)^{-(\nu+1)}\int_0^s \frac{t^{2A_0}s^{2N}}{(t+s)^{2(A_0+N)}}\frac{\mathrm{d}t}{t}\\
&\le \left(1+\frac{d(B_1,B_2)^2}{s}\right)^{-(\nu+1)}\frac{1}{s}\int_0^s \mathrm{d}t\\
&= \left(1+\frac{d(B_1,B_2)^2}{s}\right)^{-(\nu+1)}.
\end{aligned}$$

Applying this to (4.4.7) produces the desired off-diagonal bounds for the operator $S^{[0,s]}$.

Next, let's prove off-diagonal bounds for the operator $S^{[s,\infty)}$. Suppose first that $s > d(B_1, B_2)^2$. When this occurs, note that

$$\left(1 + \frac{d(B_1, B_2)^2}{s}\right)^{-(\nu+1)} \approx 1. \tag{4.4.9}$$

We then have,

$$\begin{aligned}
A_I &\leq \int_s^\infty \frac{t^{2A_0} s^{2N}}{(t+s)^{2(A_0+N)}} \frac{\mathrm{d}t}{t} \\
&\leq s^{2N} \int_s^\infty \frac{1}{(t+s)^{2N+1}} \,\mathrm{d}t \\
&\approx 1 \\
&\approx \left(1 + \frac{d(B_1, B_2)^2}{s}\right)^{-(\nu+1)}.
\end{aligned}$$

Applying this to (4.4.7) produces the desired off-diagonal estimates for $S^{[s,\infty)}$.

Finally, we must prove off-diagonal decay for $S^{[s,\infty)}$ for the case $s \leq d(B_1, B_2)^2$. We have,

$$\begin{aligned}
A_I &= \int_s^\infty \frac{t^{2A_0} s^{2N}}{(t+s)^{2(A_0+N)}} \left(1 + \frac{d(B_1, B_2)^2}{t+s}\right)^{-(\nu+1)} \frac{\mathrm{d}t}{t} \\
&\leq \frac{s^{2N}}{d(B_1, B_2)^{2(\nu+1)}} \int_s^\infty \frac{\mathrm{d}t}{(t+s)^{2N+1-(\nu+1)}}.
\end{aligned}$$

Select $\widetilde{N}_0 \geq N_0$ large enough so that $N \geq \widetilde{N}_0$ implies $2N > \nu + 1$. Then,

$$\begin{aligned}
A_I &\lesssim \frac{s^{\nu+1}}{d(B_1, B_2)^{2(\nu+1)}} \\
&\lesssim \left(1 + \frac{d(B_1, B_2)^2}{s}\right)^{-(\nu+1)},
\end{aligned}$$

where the last line follows from the condition $s \leq d(B_1, B_2)^2$. Applying this to (4.4.7) completes our proof. □

The below corollary, in combination with the pointwise estimate Proposition 4.4.4, completes the proof of Theorem 4.4.1.

Corollary 4.4.6 ([BFP16, Prop. 4.6])**.** *The maximal function $S^{\#}$ is bounded on L^2, and weak-type (p_0, p_0).*

4.5 Sparse Bounds

In this section we prove Theorem C. Since f has compact support, without loss of generality we can assume that its support is contained in a bounded set $E \subset \mathbb{X}$. By the Lemma 4.2.3, there exists $\alpha \geq 1$ and a partition $\mathscr{P}$ of $\mathbb{X}$ of dyadic cubes such that $\alpha Q \supseteq \operatorname{supp} f$ for every $Q \in \mathscr{P}$. Then

$$\int_{\mathbb{X}} |Sf|^2 g \, \mathrm{d}\mu = \sum_{Q \in \mathscr{P}} \int_Q |Sf|^2 g \, \mathrm{d}\mu = \sum_{Q \in \mathscr{P}} \int_Q |S(f \mathbb{1}_{\alpha Q})|^2 g \, \mathrm{d}\mu.$$

We are not concerned with the particular value of α, so we will fix $\alpha = 5$ in the following and assume that this value works for the covering lemma. Then, it is enough to show the existence of a sparse collection $\mathscr{S}_0$ inside a fixed cube Q_0 such that

$$\int_{Q_0} (Sf)^2 g \, \mathrm{d}\mu \lesssim \sum_{P \in \mathscr{S}_0} \left(\fint_{5P} |f|^{p_0} \, \mathrm{d}\mu \right)^{2/p_0} \left(\fint_{5P} |g|^{q_0^*} \, \mathrm{d}\mu \right)^{1/q_0^*} |P|.$$

We will decompose our quantity in different terms: all will be controlled by the averages of f and g but one. This last term is where f assumes a large value and it is similar to the original quantity but on a smaller scale. We can then iterate the decomposition, which terminates since the measure of the set we are decomposing shrinks geometrically at each iteration.

4.5.1 Decomposition

Denote by $\ell(P)$ the side length of the dyadic cube P. Let us consider the (localised) dyadic version of the operator introduced in §4.4,

$$\mathcal{M}^*_{Q_0,p_0} f(x) := \sup_{\substack{P\in\mathscr{D}\\ P\subseteq Q_0}} \Big(\inf_{y\in P} \mathcal{M}_{p_0} f(y) \Big) \mathbb{1}_P(x),$$

$$S^*_{Q_0} f(x) := \sup_{\substack{P\in\mathscr{D}\\ P\subseteq Q_0}} \left(\fint_P \left| \int_{\ell(P)^2}^{\infty} |\mathcal{Q}_t f|^2 \frac{\mathrm{d}t}{t} \right|^{\frac{q_0}{2}} \mathrm{d}\mu \right)^{1/q_0} \mathbb{1}_P(x).$$

For a positive η to be fixed later, consider the set

$$E(Q_0) := \left\{ x \in Q_0 \,:\, \max\left\{ \mathcal{M}^*_{Q_0,p_0} f(x), S^*_{Q_0} f(x) \right\} > \eta \left(\fint_{5Q_0} |f|^{p_0} \,\mathrm{d}\mu \right)^{1/p_0} \right\}.$$

Since the operators $\mathcal{M}^*_{Q_0,p_0}$ and $S^*_{Q_0}$ are weak-type (p_0, p_0), as shown in §4.4, there exists $\eta > 0$ such that $|E(Q_0)| \leq \frac{1}{2}|Q_0|$. Decompose our form as

$$\int_{Q_0} (Sf)^2 g \,\mathrm{d}\mu = \int_{Q_0\setminus E(Q_0)} (Sf)^2 g \,\mathrm{d}\mu + \int_{E(Q_0)} (Sf)^2 g \,\mathrm{d}\mu =: \mathrm{I} + \mathrm{II}$$

Term I is controlled by using Lebesgue differentiation theorem as in [BFP16, Lem. 4.4] since $|Sf(x)|^2 \leq |S^*_{Q_0} f(x)|^2$ for μ-almost every x. Thus, for $x \in Q_0 \setminus E(Q_0)$ we have

$$\int_{Q_0\setminus E(Q_0)} (Sf)^2 g \,\mathrm{d}\mu \lesssim \eta^2 \left(\fint_{5Q_0} |f|^{p_0} \,\mathrm{d}\mu \right)^{2/p_0} \left(\fint_{Q_0} |g|^{q_0^*} \,\mathrm{d}\mu \right)^{1/q_0^*} |Q_0|.$$

Consider term II. Let $\mathscr{E} := \{P\}_{P\in\mathscr{D}}$ be a covering of $E(Q_0)$ with maximal dyadic cubes. Then

$$\begin{aligned}
\int_{E(Q_0)} (Sf)^2 g \,\mathrm{d}\mu &= \sum_{P\in\mathscr{E}} \int_P (Sf)^2 g \,\mathrm{d}\mu \\
&= \sum_{P\in\mathscr{E}} \int_P \int_0^{\ell(P)^2} |\mathcal{Q}_t f(x)|^2 \frac{\mathrm{d}t}{t} g \,\mathrm{d}\mu + \sum_{P\in\mathscr{E}} \int_P \int_{\ell(P)^2}^{\infty} |\mathcal{Q}_t f(x)|^2 \frac{\mathrm{d}t}{t} g \,\mathrm{d}\mu \\
&=: \mathrm{II}_< + \mathrm{II}_>.
\end{aligned}$$

For each P in the covering, we write $f = f_{\mathsf{in}} + f_{\mathsf{out}}$, where $f_{\mathsf{in}} := f\mathbb{1}_{5P}$ and $f_{\mathsf{out}} := f\mathbb{1}_{(5P)^\complement}$. Then each term in $\mathrm{II}_<$ is itself decomposed into three terms

$$\int_P \int_0^{\ell(P)^2} |\mathcal{Q}_t f(x)|^2 \frac{\mathrm{d}t}{t} g \,\mathrm{d}\mu = \int_P \int_0^{\ell(P)^2} |\mathcal{Q}_t f_{\mathsf{in}}|^2 g \frac{\mathrm{d}t}{t} \,\mathrm{d}\mu \tag{$\mathrm{II}_{\mathsf{in}}$}$$

$$+ \int_P \int_0^{\ell(P)^2} |\mathcal{Q}_t f_{\mathsf{out}}|^2 g \frac{\mathrm{d}t}{t} \,\mathrm{d}\mu \tag{$\mathrm{II}_{\mathsf{out}}$}$$

$$+ 2\int_P \int_0^{\ell(P)^2} (\mathcal{Q}_t f_{\mathsf{in}})(\mathcal{Q}_t f_{\mathsf{out}}) g \frac{\mathrm{d}t}{t} \,\mathrm{d}\mu. \tag{$\mathrm{II}_{\mathsf{cross}}$}$$

Term ($\mathrm{II}_{\mathsf{in}}$) goes into the iteration. Terms ($\mathrm{II}_{\mathsf{out}}$) and ($\mathrm{II}_{\mathsf{cross}}$) are controlled by using Fubini and applying off-diagonal estimates as in the following lemma.

Lemma 4.5.1. *For a given dyadic cube P, let $S_k(P) := 2^{k+1}P \setminus 2^k P$ for $k \geq 2$. Then for any $t > 0$,*

$$\left(\fint_P |\mathcal{Q}_t f_{\mathsf{in}}|^{q_0} \,\mathrm{d}\mu \right)^{1/q_0} \lesssim \left(\frac{\ell(P)}{\sqrt{t}} \right)^{\nu} \left(\fint_{5P} |f|^{p_0} \,\mathrm{d}\mu \right)^{1/p_0} \tag{4.5.1}$$

$$\left(\fint_P |\mathcal{Q}_t f_{\mathsf{out}}|^{q_0} \,\mathrm{d}\mu \right)^{1/q_0} \lesssim \left(\frac{\ell(P)}{\sqrt{t}} \right)^{-\nu-2} \sum_{k \geq 2} 2^{-k} \left(\fint_{S_k(P)} |f|^{p_0} \,\mathrm{d}\mu \right)^{1/p_0}. \tag{4.5.2}$$

Proof of Lemma 4.5.1. The proof follows the one in [BFP16, Thm. 5.7]. For $f_{\mathsf{in}} = f\mathbb{1}_{5P}$, let $\mathcal{R}_0$ be a collection of finite overlapping balls R of radius $\sqrt{t}$ covering $5P$. By linearity of the operators, the triangle inequality, off-diagonal estimates for $\mathcal{Q}_t$ with $\rho(x) = (1 + |x|^2)^{-(\nu+1)}$ and Remark 4.2.9 we have

$$\left(\fint_P |\mathcal{Q}_t f_{\mathsf{in}}|^{q_0} \,\mathrm{d}\mu \right)^{1/q_0} \leq \sum_{R \in \mathcal{R}_0} \left(\fint_P |\mathcal{Q}_t f\mathbb{1}_R|^{q_0} \,\mathrm{d}\mu \right)^{1/q_0} \lesssim \sum_{R \in \mathcal{R}_0} \left(\fint_R |f|^{p_0} \,\mathrm{d}\mu \right)^{1/p_0}.$$

Since $5P \subseteq \frac{15\ell(P)}{\sqrt{t}} R$, Lemma 4.2.8 implies

$$\sum_{R \in \mathcal{R}_0} \left(\fint_R |f|^{p_0} \,\mathrm{d}\mu \right)^{1/p_0} \lesssim \left(\frac{5\ell(P)}{\sqrt{t}} \right)^{\nu} \left(\fint_{5P} |f|^{p_0} \,\mathrm{d}\mu \right)^{1/p_0}$$

which proves (4.5.1).

For $f_{\mathsf{out}} = f\mathbb{1}_{(5P)^{\complement}}$, decompose f on the squared annuli $S_k = S_k(P)$. Let $\mathcal{R}_k$ be the covering of S_k with finite overlapping balls R of radius $\sqrt{t}$. Linearity of the operators $\mathcal{Q}_t$, the triangle inequality and off-diagonal estimates for $\mathcal{Q}_t$ imply that

$$\begin{aligned}
\Big(\fint_P |\mathcal{Q}_t f_{\mathsf{out}}|^{q_0}\,\mathrm{d}\mu\Big)^{1/q_0} &\le \sum_{k\ge 2}\sum_{R\in\mathcal{R}_k}\Big(\fint_P |\mathcal{Q}_t f\mathbb{1}_R|^{q_0}\,\mathrm{d}\mu\Big)^{1/q_0}\\
&\lesssim \sum_{k\ge 2}\sum_{R\in\mathcal{R}_k}\rho\Big(\frac{d(P,R)}{\sqrt{t}}\Big)\Big(\fint_R |f|^{p_0}\,\mathrm{d}\mu\Big)^{1/p_0}\\
&\lesssim \sum_{k\ge 2}\rho\Big(\frac{d(P,S_k)}{\sqrt{t}}\Big)\sum_{R\in\mathcal{R}_k}\Big(\fint_R |f|^{p_0}\,\mathrm{d}\mu\Big)^{1/p_0}\\
&\lesssim \sum_{k\ge 2}\rho\Big(\frac{d(P,S_k)}{\sqrt{t}}\Big)\Big(\frac{2^{k+1}\ell(P)}{\sqrt{t}}\Big)^{\nu}\Big(\fint_{S_k} |f|^{p_0}\,\mathrm{d}\mu\Big)^{1/p_0}
\end{aligned}$$

where we used that the function ρ is monotone decreasing and $d(P,R)\ge d(P,S_k)$. The last inequality follows by applying Lemma 4.2.8, since $S_k(P)\subseteq 2^kP\subseteq \frac{2^{k+1}\ell(P)}{\sqrt{t}}R$.

Finally, we have enough decay from the remaining product, since

$$\rho\Big(\frac{d(P,S_k)}{\sqrt{t}}\Big)\Big(\frac{2^{k+1}\ell(P)}{\sqrt{t}}\Big)^{\nu}\lesssim\Big(\frac{2^{k}\ell(P)}{\sqrt{t}}\Big)^{-\nu-2}$$

This follows because $d(P,S_k) = d(P,2^{k+1}P\setminus 2^kP)$ is comparable with $2^k\ell(P)$ and the function $\rho(x) = (1+|x|^2)^{-(\nu+1)}$ decays faster than x^{ν} for $x\gg 1$. This proves estimate (4.5.2). □

We will use Lemma 4.5.1 to control the different terms left in the decomposition.

Remark 4.5.2. The geometric sum in (4.5.2) is controlled using the stopping condition: the integral over S_k is bounded by the integral over the ball $2^{k+1}P$, so

$$\begin{aligned}
\Big(\sum_{k\ge 2}2^{-k}\Big(\fint_{S_k}|f|^{p_0}\,\mathrm{d}\mu\Big)^{1/p_0}\Big)^2 &\lesssim \Big(\sup_{k\ge 2}\Big(\fint_{2^{k+1}P}|f|^{p_0}\,\mathrm{d}\mu\Big)^{1/p_0}\Big)^2\\
&\lesssim \Big(\inf_{\substack{y\in P^a\\ P^a \text{ parent of } P}}\mathcal{M}_{p_0}f(y)\Big)^2\\
&\lesssim \eta^2\Big(\fint_{5Q_0}|f|^{p_0}\,\mathrm{d}\mu\Big)^{2/p_0},
\end{aligned}$$

where we used that P is a maximal cube covering E. Similarly for the average on $5P$:

$$\left(\fint_{5P}|f|^{p_0}\,\mathrm{d}\mu\right)^{2/p_0} \lesssim \Big(\inf_{y\in P^a}\mathcal{M}_{p_0}f(y)\Big)^2 \lesssim \eta^2\left(\fint_{5Q_0}|f|^{p_0}\,\mathrm{d}\mu\right)^{2/p_0}.$$

Remark 4.5.3 (Control on the q_0^*-average of g). The sum of the q_0^*-averages of g is controlled by using Hölder's inequality in $\ell^{\frac{q_0}{2}}$. Since $\frac{2}{q_0}=1-\frac{1}{q_0^*}$, summing over all cubes P in $\mathscr{E}$ we obtain

$$\begin{aligned}\sum_P|P|\left(\fint_P|g|^{q_0^*}\,\mathrm{d}\mu\right)^{1/q_0^*} &\leq \Big(\sum_P|P|\Big)^{\frac{2}{q_0}}\left(\sum_P\int_P|g|^{q_0^*}\,\mathrm{d}\mu\right)^{1/q_0^*}\\ &\leq |Q_0|\left(\fint_{Q_0}|g|^{q_0^*}\,\mathrm{d}\mu\right)^{1/q_0^*} \lesssim |Q_0|\left(\fint_{5Q_0}|g|^{q_0^*}\,\mathrm{d}\mu\right)^{1/q_0^*}.\end{aligned} \tag{4.5.3}$$

4.5.2 Out term

Consider ($\mathrm{II}_{\mathsf{out}}$). Applying Fubini and Hölder's inequality, we have

$$\int_P\int_0^{\ell(P)^2}|\mathcal{Q}_tf_{\mathsf{out}}|^2g\frac{\mathrm{d}t}{t}\,\mathrm{d}\mu \leq \int_0^{\ell(P)^2}\left(\fint_P|\mathcal{Q}_tf_{\mathsf{out}}|^{q_0}\,\mathrm{d}\mu\right)^{2/q_0}\frac{\mathrm{d}t}{t}\left(\fint_P|g|^{q_0^*}\,\mathrm{d}\mu\right)^{1/q_0^*}|P|.$$

The average of g is controlled as in (4.5.3). Apply Lemma 4.5.1 to the first factor:

$$\begin{aligned}\int_0^{\ell(P)^2}\left(\fint_P|\mathcal{Q}_tf_{\mathsf{out}}|^{q_0}\,\mathrm{d}\mu\right)^{2/q_0}\frac{\mathrm{d}t}{t} &\lesssim \int_0^{\ell(P)^2}\left(\sum_{k\geq 2}\frac{\sqrt{t}}{2^k\ell(P)}\left(\fint_{S_k}|f|^{p_0}\,\mathrm{d}\mu\right)^{1/p_0}\right)^2\frac{\mathrm{d}t}{t}\\ &\lesssim \left(\sum_{k\geq 2}2^{-k}\left(\fint_{S_k}|f|^{p_0}\,\mathrm{d}\mu\right)^{1/p_0}\right)^2\end{aligned}$$

which is controlled as in Remark 4.5.2. This case is concluded.

4.5.3 Cross term

Consider ($\mathrm{II}_{\mathsf{cross}}$). We exchange the integrals, then an application of Hölder's and Cauchy–Schwarz inequality give

$$
\begin{aligned}
&\int_P \int_0^{\ell(P)^2} (\mathcal{Q}_t f_{\mathsf{in}})(\mathcal{Q}_t f_{\mathsf{out}}) g \frac{\mathrm{d}t}{t} \,\mathrm{d}\mu \\
&\qquad \leq \int_0^{\ell(P)^2} \left(⨍_P |(\mathcal{Q}_t f_{\mathsf{in}})(\mathcal{Q}_t f_{\mathsf{out}})|^{q_0/2} \,\mathrm{d}\mu \right)^{2/q_0} \frac{\mathrm{d}t}{t} \left(⨍_P |g|^{q_0^*} \,\mathrm{d}\mu \right)^{1/q_0^*} |P| \\
&\qquad \leq \int_0^{\ell(P)^2} \left(⨍_P |\mathcal{Q}_t f_{\mathsf{in}}|^{q_0} \,\mathrm{d}\mu \right)^{1/q_0} \left(⨍_P |\mathcal{Q}_t f_{\mathsf{out}}|^{q_0} \,\mathrm{d}\mu \right)^{1/q_0} \frac{\mathrm{d}t}{t} \left(⨍_P |g|^{q_0^*} \,\mathrm{d}\mu \right)^{1/q_0^*} |P|.
\end{aligned}
$$

The off-diagonal estimates for $\mathcal{Q}_t$ in Lemma 4.5.1 applied to f_{in} and f_{out} imply that

$$
\begin{aligned}
&\int_0^{\ell(P)^2} \left(⨍_P |\mathcal{Q}_t f_{\mathsf{in}}|^{q_0} \,\mathrm{d}\mu \right)^{1/q_0} \left(⨍_P |\mathcal{Q}_t f_{\mathsf{out}}|^{q_0} \,\mathrm{d}\mu \right)^{1/q_0} \frac{\mathrm{d}t}{t} \\
&\qquad \lesssim \int_0^{\ell(P)^2} \left(\frac{\sqrt{t}}{\ell(P)} \right)^2 \frac{\mathrm{d}t}{t} \left(⨍_{5P} |f|^{p_0} \,\mathrm{d}\mu \right)^{1/p_0} \sum_{k \geq 2} 2^{-k} \left(⨍_{S_k(P)} |f|^{p_0} \,\mathrm{d}\mu \right)^{1/p_0} \\
&\qquad \lesssim \eta^2 \left(⨍_{5Q_0} |f|^{p_0} \,\mathrm{d}\mu \right)^{2/p_0}
\end{aligned}
$$

where the last estimate follows as in Remark 4.5.2.

4.5.4 Large scales

Consider $\mathrm{II}_{>}$. Let P^a be the dyadic parent of P, so that $\ell(P^a) = 2\ell(P)$. Then

$$
\begin{aligned}
&\int_P \int_{\ell(P)^2}^{\infty} |\mathcal{Q}_t f(x)|^2 \frac{\mathrm{d}t}{t} g \,\mathrm{d}\mu \\
&= \int_P \int_{\ell(P)^2}^{\ell(P^a)^2} |\mathcal{Q}_t f(x)|^2 \frac{\mathrm{d}t}{t} g \,\mathrm{d}\mu + \int_P \int_{\ell(P^a)^2}^{\infty} |\mathcal{Q}_t f(x)|^2 \frac{\mathrm{d}t}{t} g \,\mathrm{d}\mu. \qquad (4.5.4)
\end{aligned}
$$

In the first term, we exchange the integrals and apply Hölder's inequality

$$
\int_{\ell(P)^2}^{\ell(P^a)^2} \int_P |\mathcal{Q}_t f(x)|^2 g \,\mathrm{d}\mu \frac{\mathrm{d}t}{t} \leq \int_{\ell(P)^2}^{\ell(P^a)^2} \left(⨍_P |\mathcal{Q}_t f(x)|^{q_0} \,\mathrm{d}\mu \right)^{2/q_0} \frac{\mathrm{d}t}{t} \left(⨍_P |g|^{q_0^*} \,\mathrm{d}\mu \right)^{1/q_0^*} |P|.
$$

Applying Lemma 4.5.1 and using that $\sqrt{t}$ is comparable with $\ell(P)$, we obtain

$$\begin{aligned}
&\int_{\ell(P)^2}^{\ell(P^a)^2} \left(\fint_P |\mathcal{Q}_t f|^{q_0} \,\mathrm{d}\mu \right)^{2/q_0} \frac{\mathrm{d}t}{t} \\
&\quad \lesssim \int_{\ell(P)^2}^{\ell(P^a)^2} \left(\left(\frac{\ell(P)}{\sqrt{t}} \right)^{\nu} \left(\fint_{5P} |f|^{p_0} \,\mathrm{d}\mu \right)^{1/p_0} + \sum_{k \geq 2} \frac{\sqrt{t}}{2^k \ell(P)} \left(\fint_{S_k} |f|^{p_0} \,\mathrm{d}\mu \right)^{1/p_0} \right)^2 \frac{\mathrm{d}t}{t} \\
&\quad \lesssim \left(\left(\fint_{5P} |f|^{p_0} \,\mathrm{d}\mu \right)^{1/p_0} + \sum_{k \geq 2} 2^{-k} \left(\fint_{S_k} |f|^{p_0} \,\mathrm{d}\mu \right)^{1/p_0} \right)^2,
\end{aligned}$$

which again is controlled as in Remark 4.5.2. The average of g is estimated as in (4.5.3).

The second term in (4.5.4), after applying Hölder's inequality, is controlled by the maximal truncation

$$\begin{aligned}
\int_P \int_{\ell(P^a)^2}^{\infty} |\mathcal{Q}_t f(x)|^2 \frac{\mathrm{d}t}{t} g \,\mathrm{d}\mu &\leq \left(\fint_P \left(\int_{\ell(P^a)^2}^{\infty} |\mathcal{Q}_t f(x)|^2 \frac{\mathrm{d}t}{t} \right)^{q_0/2} \mathrm{d}\mu \right)^{2/q_0} \left(\fint_P |g|^{q_0^*} \,\mathrm{d}\mu \right)^{1/q_0^*} |P| \\
&\lesssim \inf_{x \in P^a} (S^*_{Q_0} f)^2(x) \left(\fint_P |g|^{q_0^*} \,\mathrm{d}\mu \right)^{1/q_0^*} |P| \\
&\lesssim \eta^2 \left(\fint_{5Q_0} |f|^{p_0} \,\mathrm{d}\mu \right)^{2/p_0} \left(\fint_P |g|^{q_0^*} \,\mathrm{d}\mu \right)^{1/q_0^*} |P|.
\end{aligned}$$

We have shown that

$$\begin{aligned}
\int_{Q_0} \int_0^{\infty} |\mathcal{Q}_t f|^2 g \frac{\mathrm{d}t}{t} \,\mathrm{d}\mu &\lesssim \eta^2 \left(\fint_{5Q_0} |f|^{p_0} \,\mathrm{d}\mu \right)^{2/p_0} \left(\fint_{5Q_0} |g|^{q_0^*} \,\mathrm{d}\mu \right)^{1/q_0^*} |Q_0| \\
&\quad + \sum_P \int_P \int_0^{\ell(P)^2} |\mathcal{Q}_t f \mathbb{1}_{5P}|^2 \frac{\mathrm{d}t}{t} g \,\mathrm{d}\mu.
\end{aligned}$$

Let $\mathscr{S} = \{Q_0\}$. We add all P in the sum to $\mathscr{S}$ and we repeat the argument on each term in the sum. This iteration gives the desired bound: a sum of averages of f and g on cubes in the collection $\mathscr{S}$. We can choose $\eta > 0$ such that $|E(Q)| \leq \frac{1}{2}|Q|$ for each $Q \in \mathscr{S}$. Then $\mathscr{S}$ is sparse since each $Q \in \mathscr{S}$ has a subset $F_Q := Q \setminus E(Q)$ with the property that $\{F_Q\}_{Q \in \mathscr{S}}$ is a disjoint family and $|F_Q| > \frac{1}{2}|Q|$ by construction. $\square$

4.5.5 Upper bound on asymptotic behaviour

In this section we discuss the connection between sharp weighted estimates for an operator T and the asymptotic behaviour of its unweighted norm $\|T\|_{L^p\to L^p}$. In this section $\gamma(\cdot)$ will be the quantity defined below and not the power on the weighted characteristic in Proposition C, which is defined for $p\in(p_0,q_0)$. We recall the definition of $\gamma(q_0)$ from [FN19, Definition 5.1]. Let T be a bounded operator on L^p for $p\in(p_0,q_0)$.

Definition 4.5.4. For $q_0<\infty$ define

$$\gamma(q_0) := \sup\big\{\gamma\geq 0 \,|\, \forall\epsilon>0, \limsup_{p\to q_0}(q_0-p)^{\gamma-\epsilon}\|T\|_{L^p\to L^p}=\infty\big\},$$

and for $q_0=\infty$

$$\gamma(\infty) := \sup\left\{\gamma\geq 0 \,|\, \forall\epsilon>0, \limsup_{p\to\infty}\frac{\|T\|_{L^p\to L^p}}{p^{\gamma-\epsilon}}=\infty\right\}.$$

We say that an operator T admits a (p_0,q_0) quadratic sparse domination if it satisfies a bound as the one in Theorem C. We have the following upper bound on the unweighted norm of T.

Proposition 4.5.5. *Let $q^*:=(q/2)'$. If T admits a (p_0,q_0) quadratic sparse domination then for $p>2$ we have*

$$\|T\|_{L^p\to L^p}\lesssim\left[\left(\frac{p}{p_0}\right)'\right]^{\frac{1}{p_0}}\left[\left(\frac{p^*}{q_0^*}\right)'\right]^{\frac{1}{2}\frac{1}{q_0^*}}$$

and in particular

$$\gamma(q_0)\leq\frac{1}{2q_0^*}. \tag{4.5.5}$$

Proof. As in [FN19, Remark 3.4], let $\mathscr{S}$ be a η-sparse family. For $p>2$ we have

$$\sum_{P\in\mathscr{S}}\left(\fint_P|f|^{p_0}\,\mathrm{d}\mu\right)^{2/p_0}\left(\fint_P|g|^{q_0^*}\,\mathrm{d}\mu\right)^{1/q_0^*}|P|\lesssim\frac{1}{\eta}\|\mathcal{M}^{\mathscr{D}}_{\frac{p_0}{2}}(|f|^2)\|_{L^{p/2}}\|\mathcal{M}^{\mathscr{D}}_{q_0^*}g\|_{L^{(p/2)'}}$$

$$\lesssim \frac{1}{\eta}\left[\left(\frac{p}{p_0}\right)'\right]^{\frac{2}{p_0}}\left[\left(\frac{p^*}{q_0^*}\right)'\right]^{\frac{1}{q_0^*}}\|f\|_{L^p}^2\|g\|_{L^{p^*}}$$

where the last inequality follows from the bound on the L^p-norm of $\mathcal{M}^{\mathscr{D}}$ in (4.2.1), since

$$\|\mathcal{M}^{\mathscr{D}}_{\frac{p_0}{2}}(|f|^2)\|_{L^{p/2}} = \|\mathcal{M}^{\mathscr{D}}(|f|^{p_0})\|_{L^{p/p_0}}^{2/p_0}.$$

□

Remark 4.5.6. The upper bound on $\gamma(q_0)$ in (4.5.5) implies that, if $\gamma(q_0)$ equals $1/(2q_0^*)$ then the weighted estimates in Corollary C are sharp.

CHAPTER 5

TWO WEIGHT THEORY FOR THE BERGMAN PROJECTION

Apparently Bergman misunderstood the task.

G.M. D'A.

This chapter presents an ongoing work around the following question:

> What are the necessary and sufficient conditions on two weight u, ω for the boundedness of the Bergman projection $P : L^2(u) \to A^2(\omega)$?

Progress on this question has recently been obtained via sparse domination. For example, in the case of the unit disc $\mathbb{D}$, Aleman, Pott and Reguera characterised such weights in the Bergman space $A^2(\mathbb{D})$ in terms of testing conditions [APR17].

This chapter contains some sufficient conditions for the boundedness in question. These conditions are known for sparse forms [Li17], although they have not explicitly appeared in the context of Bergman spaces. In this setting the question is particularly relevant, as it is connected to a conjecture of Sarason in Operator Theory, see §1.5.1.

The sufficient condition is given in terms of a bump condition for the two weight.

Theorem D. *Let σ, ω be two weights on the unit ball $\mathbb{B}^d$ in $\mathbb{C}^d$ and let Φ, Ψ be two Young functions such that the associated maximal function is bounded on L^2. Then the Bergman projection P on $L^2(\mathbb{B}^d)$ satisfies the following bound*

$$\|P(\sigma \cdot)\|_{L^2(\sigma)\to A^2(\omega)} \leq C\,[\sigma, \omega]_{\Phi,\Psi}$$

where C is a positive constant independent of σ, ω and the quantity $[\sigma,\omega]_{\Phi,\Psi}$ is given by

$$[\sigma,\omega]_{\Phi,\Psi} := \sup_{\widehat{K}} \frac{\langle\sigma\rangle_{\widehat{K}}}{\langle\sigma^{1/2}\rangle_{\Phi,\widehat{K}}} \frac{\langle\omega\rangle_{\widehat{K}}}{\langle\omega^{1/2}\rangle_{\Psi,\widehat{K}}} \tag{5.0.1}$$

where the supremum is taken over the dyadic tents introduced in §5.1.

The condition (5.0.1) is known as bump condition, as the averages of the weights have been "bumped up" by mean of Orlicz norms.

We also refine the estimates in [RTW17] by extending the result [APR17, Theorem 5.7] to higher dimensions and to general weights σ, ω that are not dual to each other.

Theorem E. *Let σ, ω be two weights in B_∞ of the ball $\mathbb{B}^d$ such that their joint $\mathcal{B}_2$ characteristic $[\sigma, w]_{\mathcal{B}_2}$ defined in §1.5.2 is finite. The Bergman projection P on $L^2(\mathbb{B}^d)$ satisfies the following bound*

$$\|P(\sigma\,\cdot)\|_{L^2(\sigma)\to L^2(\omega)} \leq C\,[\omega,\sigma]_{\mathcal{B}_2}^{1/2}\big([\sigma]_{B_\infty}^{1/2} + [\omega]_{B_\infty}^{1/2}\big)$$

where C is a positive constant independent of σ, ω and the other quantity on the right hand side is the B_∞ characteristic defined in Definition 5.1.4.

Remark 5.0.1. The estimate in Theorem E improves on the B_2 estimates in [RTW17], since $[\sigma]_{B_\infty} \leq [\sigma]_{B_2}$. This is shown in Proposition 5.1.5.

Remark 5.0.2. A dyadic structure on convex domains of finite type can be constructed via the dyadic flow tents [GHK20]. This generalises the construction in §5.1 for the ball.

The resulting collection of dyadic flow tents is also sparse, and it produces weighted estimates for the Bergman projection of that class of domains.

Since our bump condition implies the boundedness of a sparse operator, the same condition implies the boundedness of the Bergman projection on convex domains of finite type by the pointwise control in [GHK20, Lemma 4.1].

Recall that the Bergman projection on the complex unit ball $\mathbb{B}^d \subset \mathbb{C}^d$ is the integral operator

$$Pf(z) := \int_{\mathbb{B}^d} \frac{f(\zeta)}{(1 - z\bar{\zeta})^{d+1}}\, \mathrm{d}\nu(\zeta)$$

where $\mathrm{d}\nu$ is the normalised measure on $\mathbb{B}^d$.

5.1 Dyadic structure on the complex unit ball

We borrow the dyadic structure on the ball developed by Arcozzi, Rochberg, and Sawyer [ARS02] and also used in [RTW17, §2]. This structure introduces a collection of sets called "dyadic kubes", which comes with a tree structure $\mathcal{T}$ called Bergman tree (namely a collection of partially ordered indexes $\{\alpha \in \mathcal{T}\}$. The points $\{c_\alpha\}_{\alpha\in\mathcal{T}}$ are the centres of the dyadic kubes).

We explain how the dyadic structure is constructed.

Let φ_z be the bi-holomorphic involution of the ball exchanging z and the origin:

$$\varphi_z(w) := \frac{z - \langle w, \frac{z}{|z|}\rangle \frac{z}{|z|} - \sqrt{1 - |z|^2}(w - \langle w, \frac{z}{|z|}\rangle \frac{z}{|z|})}{1 - \langle w, z\rangle}.$$

The Bergman metric on the unit ball $\mathbb{B}^d$ is defined as

$$\beta(z, w) := \frac{1}{2} \log \frac{1 + |\varphi_z(w)|}{1 - |\varphi_z(w)|}.$$

In the following, $B(z_0, r) \subset \mathbb{B}^d$ denotes the ball of centre z_0 and radius r in the Bergman metric. We also denote by $\mathbb{S}_r$ the sphere of radius r centred at the origin, so $\mathbb{S}_r = \partial B(0, r)$.

Fix $R, \delta > 0$. For $n \in \mathbb{N}$, there is a collection of points $\{z_j^n\}_{j=1}^{J_n}$ and a partition of the sphere $\mathbb{S}_{nR}$ in Borel subsets $\{\Omega_j^n\}_{j=1}^{J_n}$ such that

(i) $\mathbb{S}_{nR} = \bigsqcup_{j=1}^{J_n} \Omega_j^n$;

(ii) $\big(B(z_j, \delta) \cap \mathbb{S}_{nR}\big) \subseteq \Omega_j^n \subseteq \big(B(z_j, C\delta) \cap \mathbb{S}_{nR}\big)$ for some $C > 0$.

Let π_{nR} denote the radial projection from $\mathbb{B}^d$ onto the sphere $\mathbb{S}_{nR}$. The kubes are given by

$$K_1^0 \coloneqq B(0,R),$$
$$K_j^n \coloneqq \{\zeta \in B(0,(n+1)R) \setminus B(0,nR) : \pi_{nR}(\zeta) \in \Omega_j^n\}.$$

The centre of the kube K_j^n is $c_j^n \coloneqq \pi_{(n+\frac{1}{2})R}(z_j^n)$. We say that a point c_i^{n+1} is a child of c_k^n if $\pi_{nR}(c_i^{n+1}) \in \Omega_k^n$. Then the centres form a tree structure $\mathcal{T}$, which we will refer to as Bergman tree.

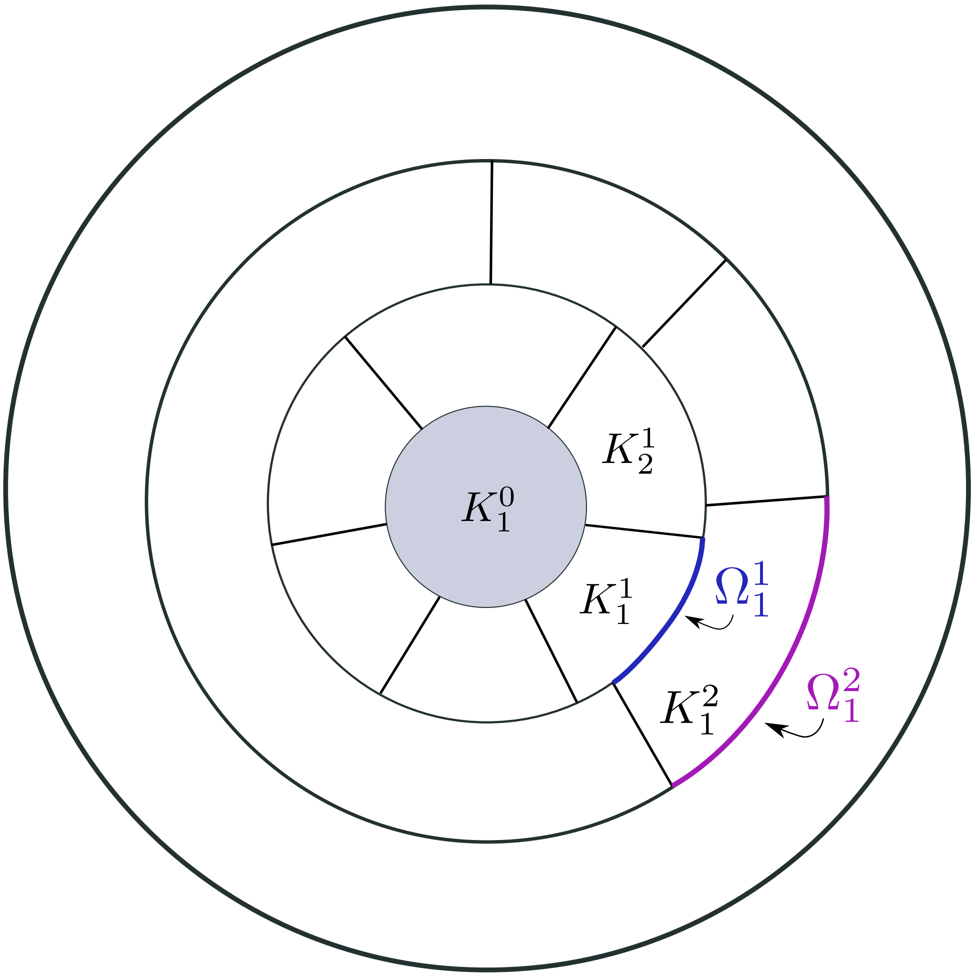

Figure 5.1: Example of the first generations of kubes and the respective Ω_k^n in the dyadic structure on $\mathbb{B}^d$.

To simplify the notation, let α be an element in $\mathcal{T}$. We denote by K_α the unique kube with centre α. If β is a descendant of α we write $\beta \geq \alpha$. Given a kube K_α, the dyadic tent $\widehat{K}_\alpha$ is the union of all kubes whose centres are descendant of α in $\mathcal{T}$, namely

$$\widehat{K}_\alpha \coloneqq \bigcup_{\beta \geq \alpha} K_\alpha.$$

The volume of K_α and $\widehat{K}_\alpha$ are comparable. This was originally proved in [ARS06, Lemma 2.8], see also [RTW17, Lemma 1].

Lemma 5.1.1 (Arcozzi, Rochberg, and Sawyer 2006)**.** *Let $\mathcal{T}$ be a Bergman tree on $\mathbb{B}^d$ with parameters R, δ. There is a universal constant $\tau > 1$, depending only on R, δ and the dimension d, such that $|\widehat{K}_\alpha| \leq \tau |K_\alpha|$ for all $\alpha \in \mathcal{T}$.*

From this lemma, and from the fact that the kubes $\{K_\alpha\}_{\alpha \in \mathcal{T}}$ are pairwise disjoint, follows immediately that the collection of dyadic tents $\mathscr{T} \coloneqq \{\widehat{K}_\alpha\}_{\alpha \in \mathcal{T}}$ is $\frac{1}{\tau}$-sparse, in the sense of Definition 1.2.8.

Definition 5.1.2. Given two weights w, σ, their joint B_p characteristic on the dyadic tents $\mathscr{T}$ is

$$[w, \sigma]_{B_p} \coloneqq \sup_{\widehat{K}_\alpha \in \mathscr{T}} \langle w \rangle_{\widehat{K}_\alpha} \langle \sigma \rangle_{\widehat{K}_\alpha}^{p-1}.$$

As before, we denote by $[w]_{B_p} \coloneqq [w, w^{1-p'}]_{B_p}$ so in particular $[w]_{B_2} \coloneqq [w, w^{-1}]_{B_2}$. We say that $w \in B_p$ if $[w]_{B_p}$ is finite.

We shall not confuse the class of weights B_p with the one of Young functions. To help the reader, we will recall which class we are referring to and we will always denote Young functions by capital Greek letters (Φ or Ψ), whilst we keep the lower case notation w, v, u, σ for weights.

Remark 5.1.3. Note that if $\sigma \in B_p$, by Hölder's inequality $\mathscr{T}$ is $(\tau^p [\sigma]_{B_p})^{-1}$-sparse with respect to the measure $\sigma \, \mathrm{d}\nu$.

Definition 5.1.4. We consider the weights in $B_\infty := \bigcup_{p>1} B_p$ and the quantity

$$[\sigma]_{B_\infty} := \sup_{\widehat{K}_\alpha \in \mathcal{T}} \frac{1}{\sigma(\widehat{K}_\alpha)} \int_{\widehat{K}_\alpha} M(\sigma \mathbb{1}_{\widehat{K}_\alpha})$$

where M is the maximal operator

$$Mf(z) := \sup_{\widehat{K} \in \mathcal{T}} \langle |f| \rangle_{\widehat{K}} \mathbb{1}_{\widehat{K}}(z).$$

The characteristic $[\sigma]_{B_\infty}$ is controlled by $[\sigma]_{B_p}$. We recall the simple proof from [APR17, Proposition 5.6].

Proposition 5.1.5 (Aleman, Pott, Reguera 2017)**.** *For $1 < p < \infty$, let w be a weight in B_p. Then we have*

$$[w]_{B_\infty} \le [w]_{B_p}.$$

Proof. Let $w \in B_p$ and let $\sigma := w^{1-p'}$ be the dual weight. By writing $1 = \sigma^{\frac{1}{p'}} \sigma^{\frac{1}{p}-1}$ and using Hölder's inequality, we have

$$\begin{aligned}
\int_{\widehat{K}} M(w\mathbb{1}_{\widehat{K}}) \sigma^{\frac{1}{p'}} \sigma^{\frac{1}{p}-1} &\le \left(\int_{\widehat{K}} M(w\mathbb{1}_{\widehat{K}})^{p'} \sigma \right)^{1/p'} \left(\int_{\widehat{K}} \sigma^{1-p} \right)^{1/p} \\
&\le \|M\|_{L^{p'}(\sigma) \to L^{p'}(\sigma)} \left(\int_{\widehat{K}} w^{p'} \sigma \right)^{1/p'} \left(\int_{\widehat{K}} \sigma^{1-p} \right)^{1/p} \\
&\lesssim_{p,d} [w]_{B_p} \int_{\widehat{K}} w
\end{aligned}$$

where we used that $w^{p'}\sigma = w = \sigma^{1-p}$ and the bound from Buckley in Theorem 1.2.5:

$$\|M\|_{L^{p'}(\sigma) \to L^{p'}(\sigma)} \lesssim_{p',d} [\sigma]_{B_{p'}}^{1/(p'-1)} = [w]_{B_p}.$$

A simple proof of the bound for the norm of M can also be found in [Ler08a]. $\square$

We will derive estimates using the above quantities and then we will compare the joint dyadic B_2 with the one on Carleson tents.

5.2 Program to deduce weighted estimates

A possible route to prove weighted estimates for P follows these steps, see also [Ler13b].

1. *(Control by a positive operator).* The modulus of the Bergman projection is controlled by the maximal Bergman projection:

$$P^+ f(z) := \int_{\mathbb{B}^d} \frac{f(\zeta)}{|1 - z\bar{\zeta}|^{d+1}} \,\mathrm{d}\nu(\zeta).$$

   Namely we have $|Pf(z)| \leq P^+|f|(z)$. Note that $P^+(|\cdot|)$ is a real-valued, positive operator. For positive weights w, v we have

$$\|P\|_{L^2(w)\to L^2(v)} \leq \|P^+\|_{L^2(w)\to L^2(v)}.$$

2. *(Equivalence with a sparse operator).* Given the dyadic structure on $\mathbb{B}^d$ and the sparse collection $\mathcal{T}$, the associated sparse operator $\Lambda_{\mathcal{T}}$ is equivalent to the maximal Bergman projection [RTW17, Lemma 5]:

$$P^+|f|(z) \eqsim_d \Lambda_{\mathcal{T}} f(z) := \sum_{\widehat{K}_\alpha \in \mathcal{T}} \langle f \rangle_{\widehat{K}_\alpha} \mathbb{1}_{\widehat{K}_\alpha}.$$

3. *(Bumps for the sparse operator).* Two-weight estimates for sparse operators are well understood. For example, they are equivalent to two-weight estimates for the maximal operator M, see [Ler13b, Theorem 1.2]. Sufficient conditions on (w, v) for the boundedness of

$$\|M\|_{L^2(w)\to L^2(v)} \quad \text{and} \quad \|\Lambda_{\mathcal{T}}\|_{L^2(w)\to L^2(v)}$$

   are known in terms of testing conditions. These are presented in §5.2.1.

These three steps are enough if we are aiming to find sufficient conditions.

If our goal is to find also necessary conditions and solving the two weight problem, another step is needed.

4. By showing the reverse inequality:

$$\|P^+\|_{L^2(w)\to L^2(v)} \lesssim \|P\|_{L^2(w)\to L^2(v)}$$

one has the equivalence

$$\|P\|_{L^2(w)\to L^2(v)} \eqsim \|\Lambda_{\mathscr{T}}\|_{L^2(w)\to L^2(v)}.$$

Then the two-weight estimates for $\Lambda_{\mathscr{T}}$ imply estimates for the Bergman projection. This has been done only for holomorphic weights $w, v \in A^2(\mathbb{D})$ in [APR17, §3].

The task of characterising weights w, v for which $\|P\|_{L^2(w)\to L^2(v)}$ is finite is still open.

5.2.1 Sawyer testing conditions

Let $\mathscr{S}$ be a sparse collection. We denote by $\Lambda_{\mathscr{S}}$ the corresponding sparse operator

$$\Lambda_{\mathscr{S}} f := \sum_{Q\in\mathscr{S}} \langle f\rangle_Q \mathbb{1}_Q.$$

The weights for which $\Lambda_{\mathscr{S}}(\sigma\,\cdot)\colon L^p(\sigma) \to L^p(w)$ holds, as well as the equivalent dual formulation $\Lambda_{\mathscr{S}}(\cdot\, w)\colon L^{p'}(w) \to L^{p'}(\sigma)$, have been characterised by Sawyer in terms of testing conditions:

$$\begin{aligned} \|\Lambda_{\mathscr{S}}(\sigma\mathbb{1}_Q)\|^p_{L^p(w)} &\le \mathfrak{T}\sigma(Q)\,, \quad \forall\, Q\in\mathscr{S} \\ \|\Lambda_{\mathscr{S}}(w\mathbb{1}_Q)\|^{p'}_{L^{p'}(\sigma)} &\le \mathfrak{T}'w(Q)\,, \quad \forall\, Q\in\mathscr{S} \end{aligned} \tag{5.2.1}$$

where the optimal testing constants are

$$\begin{aligned}\mathfrak{T} &:= \mathfrak{T}_p(w,\sigma) := \sup_Q \frac{\|\mathbb{1}_Q \Lambda_{\mathscr{S}}(\mathbb{1}_Q \sigma)\|^p_{L^p(w)}}{\sigma(Q)} \\ \mathfrak{T}' &:= \mathfrak{T}'_{p'}(w,\sigma) := \sup_Q \frac{\|\mathbb{1}_Q \Lambda_{\mathscr{S}}(\mathbb{1}_Q w)\|^{p'}_{L^{p'}(\sigma)}}{w(Q)}\end{aligned} \tag{5.2.2}$$

These conditions are named after Sawyer, who first derived them for maximal operators [Saw82] and for fractional and Poisson integrals [Saw88]. For sparse operators they have been proved in [LSU09].

Testing constants for off-diagonal estimates $\Lambda_{\mathscr{S}}(\sigma\,\cdot)\colon L^p(\sigma) \to L^q(w)$ for $q \neq p$ and more general sparse forms have also been studied, see [Li17, Theorem 1.1]. In particular we have

$$\|\Lambda_{\mathscr{S}}(\sigma\,\cdot)\|_{L^p(\sigma)\to L^p(w)} \approx \big(\mathfrak{T}^{1/p} + (\mathfrak{T}')^{1/p'}\big).$$

In the following we estimate the constants $\mathfrak{T}, \mathfrak{T}'$ from above with quantities involving the weights w, σ.

5.3 Proof of Theorem D

We derive a bump condition in L^2 for two weights w, σ in terms of Orlicz averages.

First, the maximal Bergman projection P^+ is controlled by a sparse operator Λ.

Lemma 5.3.1 ([RTW17, Lemma 5]). *There is a finite collection of Bergman trees $\{\mathcal{T}_\ell\}_{\ell=1}^N$ such that*

$$P^+|f(z)| \approx \Lambda f(z) = \sum_{\widehat{K}_\alpha \in \mathscr{T}} \langle f \rangle_{\widehat{K}_\alpha} \mathbb{1}_{\widehat{K}_\alpha}$$

where $\mathscr{T} := \cup_{\ell=1}^N \{\widehat{K}_\alpha :\ \alpha \in \mathcal{T}_\ell\}$ is a sparse collection of dyadic tents.

Then Theorem D and Theorem E follow from the respective estimates for Λ. In the rest of the chapter we give a proof of these estimates for a sparse operator Λ associated to a generic sparse collection $\mathscr{S}$. We follow the proofs of [Li17, Theorem 5.2] and [Hyt14, Theorem 6.1].

Proposition 5.3.2. *Let Λ be a sparse operator. For two weight w, σ and two Young functions $\Phi, \Psi \in B_2$, it holds*

$$\|\Lambda(\sigma \cdot)\|_{L^2(\sigma)\to L^2(w)} \lesssim [\sigma, w]_{\Phi,\Psi}.$$

We split the proof of Proposition 5.3.2 in a few simple steps, where we will use the following lemmata and the notation $\langle f\rangle_Q^\sigma = \sigma(Q)^{-1}\int_Q f\sigma$.

Lemma 5.3.3. *Let $\mathscr{S}$ be a sparse family and σ be a weight. For $1 < p < \infty$ and a function f we have*

$$\Big(\sum_{F\in\mathscr{S}} (\langle f\rangle_F^\sigma)^p \sigma(F)\Big)^{1/p} \lesssim \|f\|_{L^p(\sigma)}$$

where the implicit constant depends only on the sparse family and on the exponent p.

Proof. Assume that $\mathscr{S}$ is $\frac{1}{2}$-sparse with respect to the measure $\sigma\,\mathrm{d}x$. Then for every $F \in \mathscr{S}$ there is $E_F \subseteq F$ with $\sigma(F) \le 2\sigma(E_F)$, and the $\{E_F : F \in \mathscr{S}\}$ are disjoint. Let M^σ be the maximal function defined by

$$M^\sigma f := \sup_{F\in\mathscr{S}} \langle |f|\rangle_F^\sigma \mathbb{1}_F.$$

We bound

$$\begin{aligned}
\sum_{F\in\mathscr{S}} (\langle f\rangle_F^\sigma)^p \sigma(F) &\le 2\sum_{F\in\mathscr{S}} \big(\inf_{E_F} M^\sigma f\big)^p \sigma(E_F) \\
&\le 2\sum_F \int_{E_F} |M^\sigma f|^p \sigma\,\mathrm{d}x \\
&\le 2\|M^\sigma\|_{L^p(\sigma)\to L^p(\sigma)}^p \|f\|_{L^p(\sigma)}^p.
\end{aligned}$$

Since the norm of the dyadic maximal function $\|M^\sigma\|_{L^p(\sigma)\to L^p(\sigma)} \le p'$ and does not depend on the weight σ, the result follows. The estimate for the maximal function is classical, a proof in our case can be found in [HWW21, Lemma 3.13]. □

Lemma 5.3.4. *Let $\mathscr{S}$ be a $\frac{1}{\tau}$-sparse family, $\tau \geq 1$. For $1 < p < \infty$ let $\Psi \in B_p$ be a Young function, so that the maximal function M_Ψ is bounded on L^p. Then for any $G \in \mathscr{S}$ the following estimate holds*

$$\sum_{\substack{Q \in \mathscr{S} \\ Q \subseteq G}} \langle w^{1/p} \rangle_{\Psi,Q}^p |Q| \lesssim w(G)$$

where the implicit constant depends only on τ and $\|M_\Psi\|_{L^p \to L^p}$.

Proof. By Theorem 1.5.7, since $\Psi \in B_p$ the maximal function M_Ψ is bounded on L^p. For $Q \subseteq G$, we have $\langle w^{1/p} \rangle_{\Psi,Q} = \langle w^{1/p} \mathbb{1}_G \rangle_{\Psi,Q}$. Then $|Q| \leq \tau |E_Q|$ and

$$\begin{aligned} \sum_{\substack{Q \in \mathscr{S} \\ Q \subseteq G}} \langle w^{1/p} \rangle_{\Psi,Q}^p |Q| &\leq \tau \sum_{\substack{Q \in \mathscr{S} \\ Q \subseteq G}} \int_{E_Q} M_\Psi (w^{1/p} \mathbb{1}_G)^p \\ &\leq \tau \int_G M_\Psi (w^{1/p} \mathbb{1}_G)^p \\ &\leq \tau \|M_\Psi\|_{L^p \to L^p}^p \|w^{1/p}\|_{L^p(G)}^p. \end{aligned}$$

□

We are ready to prove Proposition 5.3.2. By symmetry, it is enough to focus on one of the two testing conditions in (5.2.1).

Reduction to dyadic form

By duality, the two-weight estimate

$$\|\Lambda(f\sigma)\|_{L^2(w)} \leq C \|f\|_{L^2(\sigma)}$$

is equivalent the supremum over $g \in L^2(w)$ of $|\langle \Lambda(f\sigma), gw \rangle|$. Then it is enough to show that for non-negative functions f and g we have

$$|\langle \Lambda(f\sigma), gw \rangle| = \sum_{Q \in \mathscr{S}} \langle f\sigma \rangle_Q \langle gw \rangle_Q |Q|$$

$$= \sum_{Q\in\mathscr{S}} \langle f\rangle_Q^\sigma \langle g\rangle_Q^w \langle\sigma\rangle_Q \langle w\rangle_Q |Q| \lesssim [\sigma, w]_{\Phi,\Psi} \|f\|_{L^2(\sigma)} \|g\|_{L^2(w)}.$$

Stopping families

We assume that both f, g are both non-negative and supported on the set Q_0. We will use a dyadic family of cubes inside Q_0 and we select special cubes for the "parallel corona" decomposition. We denote the principal cubes for (f, σ) and (g, w) by $\mathscr{F}$ and $\mathscr{G}$ respectively. These are defined as the stopping family in §3.5.1 but for the weighted averages of f and g:

$$\mathscr{A}_f^\star(Q) = \{S \in \mathscr{D}, S \subset Q \text{ maximal} : \langle f\rangle_S^\sigma > 2\langle f\rangle_Q^\sigma\},$$
$$\mathscr{A}_g^\star(Q) = \{S \in \mathscr{D}, S \subset Q \text{ maximal} : \langle g\rangle_S^w > 2\langle g\rangle_Q^w\}.$$

Then we define

$$\mathscr{F}_0 := \{Q_0\}, \qquad \mathscr{F}_{n+1} := \bigcup_{Q\in\mathscr{F}_n} \mathscr{A}_f^\star(Q), \qquad \mathscr{F} := \bigcup_{n\in\mathbb{N}} \mathscr{F}_n$$

and in a similar way for $\mathscr{G}$. As shown in Chapter 3, the families $\mathscr{F}$ and $\mathscr{G}$ constructed in this way are sparse. We denote by $\pi_{\mathscr{F}}(Q)$ the minimal cube in $\mathscr{F}$ containing Q, and similarly for $\pi_{\mathscr{G}}(Q)$. Given a pair of cubes $(F, G) \in \mathscr{F} \times \mathscr{G}$, we consider the collection of cubes such that their projection to $\mathscr{F}$ and $\mathscr{G}$ are F and G respectively. Such collection is

$$\{Q : \pi(Q) = (F, G)\}, \quad \text{where } \pi(Q) := (\pi_{\mathscr{F}}(Q), \pi_{\mathscr{G}}(Q)).$$

Using the stopping families we can write

$$\sum_{Q\in\mathscr{S}} = \sum_{F\in\mathscr{F}} \sum_{G\in\mathscr{G}} \sum_{\substack{Q\in\mathscr{S} \\ \pi(Q)=(F,G)}}.$$

Since either $F \subseteq G$ or $F \supseteq G$, by symmetry it is enough to study only one case. We focus on the latter. We have

$$\sum_{F\in\mathscr{F}}\sum_{\substack{G\in\mathscr{G}\\ G\subseteq F}}\sum_{\substack{Q\in\mathscr{S}\\ \pi(Q)=(F,G)}}\langle f\rangle_Q^\sigma\langle g\rangle_Q^w\langle\sigma\rangle_Q\langle w\rangle_Q|Q|$$
$$\leq 4\sum_{F\in\mathscr{F}}\langle f\rangle_F^\sigma\sum_{\substack{G\in\mathscr{G}\\ G\subseteq F}}\langle g\rangle_G^w\sum_{\substack{Q\in\mathscr{S}\\ \pi(Q)=(F,G)}}\langle\sigma\rangle_Q\langle w\rangle_Q|Q|. \tag{5.3.1}$$

Introduce Orlicz bumps

We focus on the last summand in (5.3.1). We see that

$$\langle\sigma\rangle_Q\langle w\rangle_Q=\left(\frac{\langle\sigma\rangle_Q\langle w\rangle_Q}{\langle\sigma^{1/2}\rangle_{\Phi,Q}\langle w^{1/2}\rangle_{\Psi,Q}}\right)\langle\sigma^{1/2}\rangle_{\Phi,Q}\langle w^{1/2}\rangle_{\Psi,Q}.$$

The supremum over all cubes $Q\in\mathscr{D}$ of the quantity in brackets is $[\sigma,w]_{\Phi,\Psi}$. Then we have

$$\sum_{\substack{Q\in\mathscr{S}\\ \pi(Q)=(F,G)}}\langle\sigma\rangle_Q\langle w\rangle_Q|Q|\leq[\sigma,w]_{\Phi,\Psi}\sum_{\substack{Q\in\mathscr{S}\\ \pi(Q)=(F,G)}}\langle\sigma^{1/2}\rangle_{\Phi,Q}\langle w^{1/2}\rangle_{\Psi,Q}|Q|.$$

Using the Cauchy–Schwarz inequality and Lemma 5.3.4 we estimate

$$\begin{aligned}&\sum_{\substack{Q\in\mathscr{S}\\ \pi(Q)=(F,G)}}\langle\sigma^{1/2}\rangle_{\Phi,Q}\langle w^{1/2}\rangle_{\Psi,Q}|Q|\\ &\qquad\leq\Big(\sum_{\substack{Q\in\mathscr{S}\\ \pi(Q)=(F,G)}}\langle\sigma^{1/2}\rangle_{\Phi,Q}^2|Q|\Big)^{1/2}\Big(\sum_{\substack{Q\in\mathscr{S}\\ \pi(Q)=(F,G)}}\langle w^{1/2}\rangle_{\Psi,Q}^2|Q|\Big)^{1/2}\\ &\qquad\lesssim\Big(\sum_{\substack{Q\in\mathscr{S}\\ \pi(Q)=(F,G)}}\langle\sigma^{1/2}\rangle_{\Phi,Q}^2|Q|\Big)^{1/2}w(G)^{1/2}.\end{aligned}$$

Putting all together, and using the Cauchy–Schwarz inequality in ℓ^2 in the third and and fifth inequality and Lemma 5.3.4 in the second and the fourth, we obtain

$$\sum_{F\in\mathscr{F}}\langle f\rangle_F^\sigma\sum_{\substack{G\in\mathscr{G}\\ G\subseteq F}}\langle g\rangle_G^w\sum_{\substack{Q\in\mathscr{S}\\ \pi(Q)=(F,G)}}\langle\sigma\rangle_Q\langle w\rangle_Q|Q|$$

$$
\begin{aligned}
&\leq [\sigma,w]_{\Phi,\Psi}\sum_{F\in\mathscr{F}}\langle f\rangle_F^\sigma\sum_{\substack{G\in\mathscr{G}\\G\subseteq F}}\langle g\rangle_G^w\sum_{\substack{Q\in\mathscr{S}\\ \pi(Q)=(F,G)}}\langle\sigma^{1/2}\rangle_{\Phi,Q}\langle w^{1/2}\rangle_{\Psi,Q}|Q|\\
&\lesssim [\sigma,w]_{\Phi,\Psi}\sum_{F\in\mathscr{F}}\langle f\rangle_F^\sigma\sum_{\substack{G\in\mathscr{G}\\G\subseteq F}}\langle g\rangle_G^w\Big(\sum_{\substack{Q\in\mathscr{S}\\ \pi(Q)=(F,G)}}\langle\sigma^{1/2}\rangle_{\Phi,Q}^2|Q|\Big)^{1/2}w(G)^{1/2}\\
&\leq [\sigma,w]_{\Phi,\Psi}\sum_{F\in\mathscr{F}}\langle f\rangle_F^\sigma\Big(\sum_{\substack{G\in\mathscr{G}\\G\subseteq F}}(\langle g\rangle_G^w)^2w(G)\Big)^{1/2}\Big(\sum_{\substack{G\in\mathscr{G}\\G\subseteq F}}\sum_{\substack{Q\in\mathscr{S}\\ \pi(Q)=(F,G)}}\langle\sigma^{1/2}\rangle_{\Phi,Q}^2|Q|\Big)^{1/2}\\
&\lesssim [\sigma,w]_{\Phi,\Psi}\sum_{F\in\mathscr{F}}\langle f\rangle_F^\sigma\Big(\sum_{\substack{G\in\mathscr{G}\\G\subseteq F}}(\langle g\rangle_G^w)^2w(G)\Big)^{1/2}\sigma(F)^{1/2}\\
&\leq [\sigma,w]_{\Phi,\Psi}\Big(\sum_{F\in\mathscr{F}}(\langle f\rangle_F^\sigma)^2\sigma(F)\Big)^{1/2}\Big(\sum_{F\in\mathscr{F}}\sum_{\substack{G\in\mathscr{G}\\G\subseteq F}}(\langle g\rangle_G^w)^2w(G)\Big)^{1/2}\\
&\lesssim [\sigma,w]_{\Phi,\Psi}\|f\|_{L^2(\sigma)}\|g\|_{L^2(w)}
\end{aligned}
$$

where the last inequality follows from Lemma 5.3.3, concluding the proof. □

5.4 Proof of Theorem E

In [APR17, §5] the authors obtained sharp one-weight estimates for the maximal Bergman projection

$$P^+f(z):=\int_{\mathbb{B}^d}\frac{f(\zeta)}{|1-z\bar{\zeta}|^{d+1}}\,\mathrm{d}\nu(\zeta)$$

in terms of the mixed B_2–B_∞ characteristics. These estimates follow from a sparse domination of P^+ and are obtained via Sawyer's testing conditions for the sparse operator presented in §5.2.1. Combining the sparse domination in [RTW17] and the estimates for sparse forms in [Li17], we derive B_2–B_∞ estimates for P^+.

We consider the sparse operator

$$\Lambda_{\mathcal{T}}f:=\sum_{\widehat{K}_\alpha\in\mathcal{T}}\langle f\rangle_{\widehat{K}_\alpha}\mathbb{1}_{\widehat{K}_\alpha}.$$

The testing conditions for the boundedness of $\|\Lambda_{\mathscr{S}}(\sigma\,\cdot)\|_{L^p(\sigma)\to L^p(w)}$ for two weight w,σ

are

$$\|\mathbb{1}_{\widehat{K}_0}\Lambda_{\mathscr{T}}(\sigma\mathbb{1}_{\widehat{K}_0})\|_{L^2(w)}^2 \lesssim [w,\sigma]_{B_2}[\sigma]_{B_\infty}\sigma(\widehat{K}_0)$$
$$\|\mathbb{1}_{\widehat{K}_0}\Lambda_{\mathscr{T}}(w\mathbb{1}_{\widehat{K}_0})\|_{L^2(\sigma)}^2 \lesssim [\sigma,w]_{B_2}[w]_{B_\infty}w(\widehat{K}_0).$$

By symmetry, it is enough to prove one of the two. We choose the first one.

Proposition 5.4.1. *Let σ, w be two weights. Then for any dyadic tent $\widehat{K}_0 \in \mathscr{T}$, we have*

$$\|\mathbb{1}_{\widehat{K}_0}\Lambda_{\mathscr{T}}\sigma\|_{L^2(w)}^2 \lesssim [w,\sigma]_{B_2}[\sigma]_{B_\infty}\sigma(\widehat{K}_0).$$

We refer the reader to [HL12, Prop. 5.2] for a version of this result for dyadic shifts. Since we deal with sparse operators, the proof we present here is simpler. It follows the approach in Hytönen's work [Hyt14, §5.A] and in [APR17, §5].

Proof of Proposition 5.4.1. For simplicity, we denote by $L_0 \in \mathscr{T}$ a fixed dyadic tent, instead of $\widehat{K}_0$. Recall that, since $\mathscr{T}$ is sparse, there is a fixed $\tau \geq 1$ such that for every $L \in \mathscr{T}$ there exists a subset $E_L \subseteq L$ with the property that $|L| \leq \tau|E_L|$ and the sets in $\{E_L : L \in \mathscr{T}\}$ are pairwise disjoint.

Then we have

$$\begin{aligned}
\|\Lambda_{\mathscr{T}}\sigma\mathbb{1}_{L_0}\|_{L^2(w)}^2 &= \int_{L_0}\Big(\sum_{L\in\mathscr{T}}\langle\sigma\rangle_L\mathbb{1}_L\Big)^2 w \\
&\leq 2\int_{L_0}\sum_{\substack{L\in\mathscr{T}\\ L\subseteq L_0}}\langle\sigma\rangle_L\sum_{\substack{L'\in\mathscr{T}\\ L'\subseteq L}}\langle\sigma\rangle_{L'}\mathbb{1}_{L'}\,w \\
&= 2\sum_{\substack{L\in\mathscr{T}\\ L\subseteq L_0}}\langle\sigma\rangle_L\sum_{\substack{L'\in\mathscr{T}\\ L'\subseteq L}}\langle\sigma\rangle_{L'}\langle w\rangle_{L'}|L'| \\
&\leq 2\sup_{L'\in\mathscr{T}}\langle\sigma\rangle_{L'}\langle w\rangle_{L'}\sum_{\substack{L\in\mathscr{T}\\ L\subseteq L_0}}\langle\sigma\rangle_L\sum_{\substack{L'\in\mathscr{T}\\ L'\subseteq L}}|L'| \\
&\leq 2\tau\sup_{L'\in\mathscr{T}}\langle\sigma\rangle_{L'}\langle w\rangle_{L'}\sum_{\substack{L\in\mathscr{T}\\ L\subseteq L_0}}\langle\sigma\rangle_L|L|
\end{aligned}$$

$$\lesssim [\sigma, w]_{B_2} \sum_{\substack{L \in \mathcal{T} \\ L \subseteq L_0}} \langle \sigma \rangle_L |L|.$$

The remaining sum is controlled by using the maximal function and the sparseness property. We have

$$\begin{aligned}
[\sigma, w]_{B_2} \sum_{\substack{L \in \mathcal{T} \\ L \subseteq L_0}} \langle \sigma \rangle_L |L| &\leq [\sigma, w]_{B_2} \sum_{\substack{L \in \mathcal{T} \\ L \subseteq L_0}} \inf_L M(\sigma \mathbb{1}_{L_0}) |L| \\
&\leq \tau [\sigma, w]_{B_2} \sum_{\substack{L \in \mathcal{T} \\ L \subseteq L_0}} \int_{E_L} M(\sigma \mathbb{1}_{L_0}) \\
&\leq \tau [\sigma, w]_{B_2} \frac{1}{\sigma(L_0)} \int_{L_0} M(\sigma \mathbb{1}_{L_0}) \, \sigma(L_0) \\
&\leq \tau [\sigma, w]_{B_2} \Big(\sup_{L_0 \in \mathcal{T}} \frac{1}{\sigma(L_0)} \int_{L_0} M(\sigma \mathbb{1}_{L_0}) \Big) \, \sigma(L_0) \\
&\lesssim [\sigma, w]_{B_2} [\sigma]_{B_\infty} \sigma(L_0).
\end{aligned}$$

This concludes the proof of the proposition. □

The proof of Theorem E follows by combining the sparse domination in Lemma 5.3.1 with the bound for sparse operator in Proposition 5.4.1. This gives the bound

$$\|P(\sigma \, \cdot)\|_{L^2(\sigma) \to L^2(\omega)} \leq C \, [\omega, \sigma]_{B_2}^{1/2} \big([\sigma]_{B_\infty}^{1/2} + [\omega]_{B_\infty}^{1/2} \big).$$

We conclude by comparing the volume of a Carleson tents T_z with the volume of a dyadic tents $\widehat{K}_\alpha$. This is the content of the following two lemmas, see [RTW17, Lemma 3] and [HW20, Lemma 2.4].

Lemma 5.4.2 (Rahm, Tchoundja, and Wick, 2017)**.** *There exists a finite collection of Bergman trees $\{\mathcal{T}_\ell\}_{\ell=1}^N$ such that for any tent T_z there is $\ell \in \{1, \ldots, N\}$ and α in $\mathcal{T}_\ell$ such that $\widehat{K}_\alpha \supseteq T_z$ and $|T_z| \approx |\widehat{K}_\alpha|$.*

Note that since a finite union of sparse families is sparse, if we denote by

$$\mathscr{T} := \bigcup_{\ell=1}^{N} \mathscr{T}_\ell \quad \text{where} \quad \mathscr{T}_\ell := \{\widehat{K}_\alpha : \alpha \in \mathcal{T}_\ell\},$$

then $\mathscr{T}$ is a sparse collection of sets in the unit ball $\mathbb{B}^d$.

Lemma 5.4.3 (Huo and Wick 2020)**.** *For any dyadic tent $\widehat{K}_\beta \in \mathscr{T}$ there exists a Carleson tent T_z such that $\widehat{K}_\beta \subseteq T_z$ and $|\widehat{K}_\beta| \approx |T_z|$.*

Then it holds that $[w,\sigma]_{B_2} \approx [w,\sigma]_{\mathcal{B}_2}$ for B_∞ weights. The proof of Theorem E is concluded.

APPENDIX A

FOURIER INVARIANCE

This is also written in Wilson's book.

F. D'E.

The Strichartz norm introduced in Chapter 2 is invariant under the action of the Fourier transform. We use the following definitions:

$$\widehat{f}(\xi) = \int_{\mathbb{R}^d} e^{-ix\cdot\xi} f(x)dx, \qquad \check{f}(y) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{iy\cdot\xi} f(\xi)d\xi.$$

With the definitions above, the Fourier transform is an isometry between $L^2(\mathbb{R}^d)$ and $L^2(\mathbb{R}^d, (2\pi)^{-d/2}\,\mathrm{d}\xi)$.

Proposition A.0.1 (Plancherel)**.** *Let $f\colon \mathbb{R}^d \to \mathbb{C}$ be a function in $L^1 \cap L^2$, then*

$$\|\check{f}\|_2 = (2\pi)^{-\frac{d}{2}}\|f\|_2.$$

We recall the Strichartz estimates for the solution $u(t,x) = e^{-it\Delta}f(x)$ of the free Schrödinger equation $i\partial_t u = \triangle u$ with initial datum f. The following estimate holds

$$\|\|e^{-it\Delta}f\|_{L^p_x(\mathbb{R}^d)}\|_{L^q_t(\mathbb{R})} \le C\|f\|_{L^2_x(\mathbb{R}^d)}, \tag{A.0.1}$$

for all admissible (d,p,q) satisfying

$$\frac{2}{q}+\frac{d}{p}=\frac{d}{2}, \quad q\geq 2, \quad (d,p,q)\neq(2,\infty,2).$$

The left hand side of (A.0.1) is Fourier invariant, namely, if one calls

$$\|f\|_{(d,p,q)}:=\|\|e^{-it\Delta}f\|_{L^p_x(\mathbb{R}^d)}\|_{L^q_t(\mathbb{R})}$$

then we have the following

Proposition A.0.2. *For all admissible (d,p,q) and every Schwartz function f it holds that*

$$\|\check{f}\|_{(d,p,q)}=(2\pi)^{-\frac{d}{2}}\|f\|_{(d,p,q)}.$$

Proof. One can write the solution of the Schrödinger equation as a convolution

$$e^{-it\Delta}f(x)=f*\frac{e^{-i\frac{|\cdot|^2}{4t}}}{(\sqrt{4\pi t})^d}(x)=\frac{1}{(\sqrt{4\pi t})^d}\int_{\mathbb{R}^d}f(y)e^{-i\frac{|x-y|^2}{4t}}dy.$$

Then we expand the square

$$\frac{|x-y|^2}{4t}=\frac{|x|^2}{4t}+\frac{1}{4t}(|y|^2-2x\cdot y).$$

Since $|e^{-\frac{i}{4t}|x|^2}|=1$, by changing variables $t=\frac{1}{s}$ and $z=sx$, one is left with

$$\|e^{-it\Delta}f\|^q_{L^q_tL^p_x}=\left(\frac{1}{4\pi}\right)^{\frac{d}{2}q}\int_{\mathbb{R}}\left(\int_{\mathbb{R}^d}\left|\int_{\mathbb{R}^d}f(y)e^{-i\frac{s}{4}|y|^2}e^{i\frac{z}{2}\cdot y}dy\right|^p dz\right)^{\frac{q}{p}}s^{q\frac{d}{2}-2}ds.$$

Write $v=\frac{z}{2}$ and $r=\frac{s}{4}$. Introducing a factor $(2\pi)^{\pm dq}$ we obtain

$$\begin{aligned}\left(\frac{1}{4\pi}\right)^{\frac{d}{2}q}\int_{\mathbb{R}}\left(\int_{\mathbb{R}^d}\left|\int_{\mathbb{R}^d}f(y)e^{-ir|y|^2}e^{iv\cdot y}\frac{dy}{(2\pi)^d}\right|^p 2^d dv\right)^{\frac{q}{p}}4(r^{q\frac{d}{2}-2-d\frac{q}{p}})dr\\ =(2\pi)^{\frac{d}{2}q}\|e^{-it\Delta}\check{f}\|^q_{L^q_tL^p_x}.\end{aligned}$$

The exponent in the Jacobian factor vanishes because of the scaling condition. Thus we have

$$\| \| e^{-it\Delta} \check{f} \|_{L^p_x(\mathbb{R}^d)} \|_{L^q_t(\mathbb{R})} = (2\pi)^{-\frac{d}{2}} \| \| e^{-it\Delta} f \|_{L^p_x(\mathbb{R}^d)} \|_{L^q_t(\mathbb{R})}.$$

$\square$

APPENDIX B

CONDITIONAL EXPECTATION AND HAAR PROJECTIONS

In this appendix we recall some known bounds for the Haar projection. These involve conditional expectation and martingales related to the Haar system, see also [Gra14, §6.4].

Let $\mathscr{S}$ be the stopping family defined in §3.5.1. Given $S \in \mathscr{S}$, let $\mathscr{A}^\star(S)$ be the maximal stopping cubes inside S. Let $\mathscr{G}_S$ be the σ-algebra generated by $\mathscr{A}^\star(S)$. A function is measurable with respect the σ-algebra $\mathscr{G}_S$ if and only if it is constant on any cube in $\mathscr{A}^\star(S)$.

B.0.1 Conditional expectation

Denote by $\mathbb{E}[\,\cdot\,|\mathscr{G}_S]$ the projection on the space of measurable functions with respect to the σ-algebra $\mathscr{G}_S$.

$$\mathbb{E}[f|\mathscr{G}_S](x) = \begin{cases} f(x) & \text{if } x \in S \setminus \mathscr{A}(S) \\ \langle f \rangle_{S'} & \text{if } x \in S' \text{ for some } S' \in \mathscr{A}^\star(S). \end{cases}$$

For more details about this operator, we refer the reader to [HvVW16, §2.6]. Let $\mathscr{S}$ be a stopping family for f. The supremum of $\mathbb{E}[f|\mathscr{G}_S]$ in S is either $f(x)$ (if $\mathscr{A}(S)$ is empty), or $\langle f \rangle_{S'}$ for some $S' \in \mathscr{A}^\star(S)$. In both cases $\|\mathbb{E}[f\mathbb{1}_S|\mathscr{G}_S]\|_{L^\infty(S)} \lesssim_d \langle f \rangle_S$, since $\langle f \rangle_{S'} \leq 2^d A \langle f \rangle_S$ by the stopping conditions.

B.0.2 Haar projection

Given $S \in \mathscr{S}$, let $\mathsf{Tree}(S) = \{Q \in \mathscr{D} : \widehat{Q} = S\}$ be the collection of cubes Q such that S is the minimal stopping cube containing Q.

The Haar projection on S is given by

$$\mathcal{P}_S f := \sum_{I \in \mathsf{Tree}(S)} \Delta_I f = \sum_{I \in \mathsf{Tree}(S)} \sum_{\epsilon \in \{0,1\}^d \setminus \{0\}^d} \langle f, h_I^\epsilon \rangle h_I^\epsilon$$

where $\{h_I^\epsilon\}_\epsilon$ are the Haar functions on I. Being a sum of Haar functions on cubes in $\mathsf{Tree}(S)$, the Haar projection $\mathcal{P}_S f$ is constant on any $S' \in \mathscr{A}^\star(S)$, so it is measurable on $\mathscr{G}_S$. It also holds that $\mathcal{P}_S f = \mathcal{P}_S \mathbb{E}[f \mathbb{1}_S | \mathscr{G}_S]$.

The Haar projection $\mathcal{P}_S f$ can be seen as a martingale transform, and so it satisfies the following

Lemma B.0.1 (L^p bound for martingale transform [Bur84])**.** *For $1 < p < \infty$ we have*

$$\|\mathcal{P}_S \mathbb{E}[f \mathbb{1}_S | \mathscr{G}_S]\|_p \leq C_p \|\mathbb{E}[f \mathbb{1}_S | \mathscr{G}_S]\|_p. \tag{B.0.1}$$

Combining (B.0.1) with the estimate for the supremum of $\mathbb{E}[f \mathbb{1}_S | \mathscr{G}_S]$ one obtains that

$$\|\mathcal{P}_S f\|_p \lesssim_p \langle f \rangle_S.$$

B.0.3 Richer σ-algebras and r-Haar projections

The same idea works with slight modifications when S is the minimal stopping cube containing the r-ancestor of Q. Let $\mathsf{Tree}_r(S)$ be the collection of cubes Q such that $\widehat{Q^{(r)}} = S$. Define the r-Haar projection on S as

$$\mathcal{P}_S^r f = \sum_{Q \in \mathsf{Tree}_r(S)} \Delta_Q f.$$

Remark B.0.2. The projection $\mathcal{P}_S^r f$ is *not* measurable on $\mathscr{G}_S$ in general, but it is measurable with respect to the richer σ-algebra generated by the r-grandchildren of $S' \in \mathscr{A}^\star(S)$, which is

$$\mathscr{G}_S^r := \sigma\Big(\{(S')_r \in \mathsf{ch}_r(S'), S' \in \mathscr{A}^\star(S)\}\Big).$$

Then $\mathcal{P}_S^r f = \mathcal{P}_S^r \mathbb{E}[f\mathbb{1}_S|\mathscr{G}_S^r]$ and we have the following

Lemma B.0.3. *Given a function f, let S be a stopping cube in $\mathscr{S}_f$ as defined in §3.5.1. Then*

$$\|\mathbb{E}[f\mathbb{1}_S|\mathscr{G}_S^r]\|_{L^\infty(S)} \lesssim_{d,r} \langle f\rangle_S.$$

Proof. Either $|f(x)| \le A\langle f\rangle_S$ for all $x \in S$, or there exists $S' \in \mathscr{A}^\star(S)$ with $x_0 \in (S')_r$ and $\mathbb{E}[f\mathbb{1}_S|\mathscr{G}_S^r](x_0) = \langle f\rangle_{(S')_r}$. Let P be the dyadic parent of $(S')_r$. Then $P \in \mathsf{Tree}_r(S)$ and we have

$$\langle f\rangle_{(S')_r} \le 2^d\langle f\rangle_P \le 2^d 2^{dr}\langle f\rangle_{P^r} \le 2^{d(r+1)}A\langle f\rangle_S$$

where we used the stopping condition in the last inequality. □

B.0.4 Haar projection on maximal cubes

For $S \in \mathscr{S}$, the r-grandchildren $\mathsf{ch}_r(S)$ are the maximal cubes in $\mathsf{Tree}_r(S)$. Then the restriction of Haar projection $\mathcal{P}_S^r$ on a $S_r \in \mathsf{ch}_r(S)$ is

$$\mathcal{P}_{S_r} f := \sum_{\substack{Q\in\mathsf{Tree}_r(S)\\ Q\subseteq S_r}} \Delta_Q f \quad \text{and satisfies} \quad \langle|\mathcal{P}_{S_r} f|\rangle_{S_r} \lesssim \langle|f|\rangle_S. \tag{3.6.1}$$

Proof of (3.6.1). The Haar projector $\mathcal{P}_{S_r} f$ is measurable with respect to the σ-algebra $\mathscr{G}_S^r$, then

$$\begin{aligned}
\int_{S_r} |\mathcal{P}_{S_r} f| = \int_{S_r} |\mathcal{P}_{S_r}\mathbb{E}[f\mathbb{1}_{S_r}|\mathscr{G}_S^r]| &\le \|\mathbb{1}_{S_r}\|_{L^{p'}}\|\mathcal{P}_{S_r}\mathbb{E}[f\mathbb{1}_{S_r}|\mathscr{G}_S^r]\|_{L^p(S_r)} \\
\text{by (B.0.1)} \quad &\lesssim_p \|\mathbb{1}_{S_r}\|_{L^{p'}}\|\mathbb{E}[f\mathbb{1}_{S_r}|\mathscr{G}_S^r]\|_{L^p(S_r)} \\
&\le |S_r|^{\frac{1}{p'}}|S_r|^{\frac{1}{p}}\|\mathbb{E}[f\mathbb{1}_{S_r}|\mathscr{G}_S^r]\|_\infty \\
\text{by Lemma B.0.3} \quad &\lesssim |S_r|\langle f\rangle_S.
\end{aligned}$$

Divide by $|S_r|$ both sides to conclude. $\square$

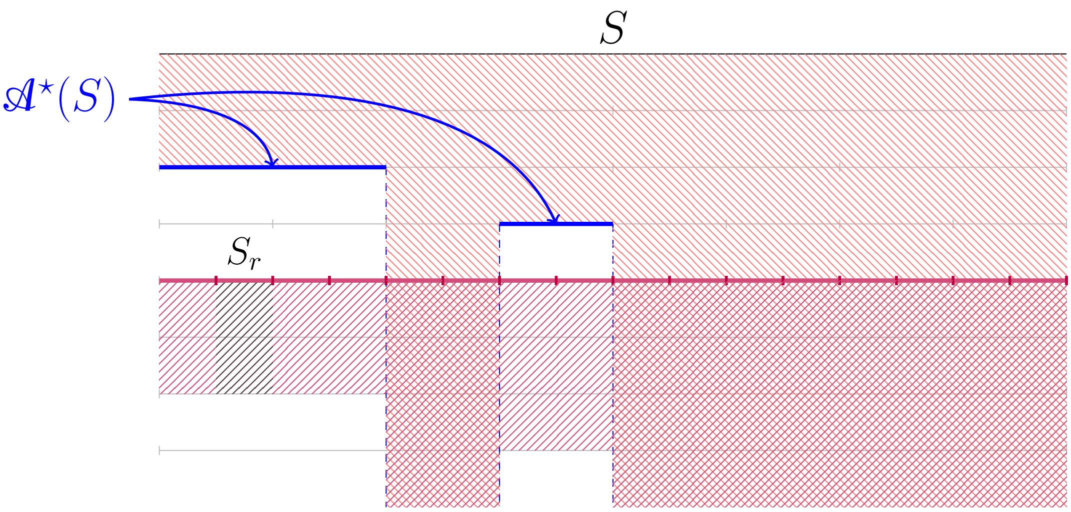

Figure B.1: An example of stopping tree $\mathsf{Tree}(S)$ and the maximal stopping cubes in $\mathscr{A}^\star(S)$. Below, shifted by r generations, there is the stopping tree $\mathsf{Tree}_r(S)$. The cubes Q in $\mathsf{Tree}_r(S)$ contained in a specific r-grandchild S_r are highlighted.

APPENDIX C

SAWYER'S DUALITY TRICK

The symmetric formulation of two weight inequalities recalled in the introduction of [TV16] is attributed to [Fef83] and [Koo80]. It goes as follow.

Proposition C.0.1. *Let u, v be two weights. Then the two weight inequality for an operator*

$$T \colon L^p(v) \to L^q(u)$$

is equivalent to

$$T(\sigma\,\cdot) \colon L^p(\sigma) \to L^q(u), \tag{C.0.1}$$

where $\sigma := v^{-p'/p} = v^{1-p'}$.

This formulation is useful since it reduces the number of different measures involved. Indeed, in (C.0.1) there are only two measures ($u\,\mathrm{d}x$ and $\sigma\,\mathrm{d}x$), instead of three: $u\,\mathrm{d}x$, $v\,\mathrm{d}x$, and the Lebesgue measure in the operator T. Moreover, it reduces the assumptions on the weights for the dual inequality.

The dual expression of $T \colon L^p(v) \to L^q(u)$ is $T^\star \colon L^{q'}(u^{1-q'}) \to L^{p'}(v^{1-p'})$. To make sense of the latter, v has to take the values 0 and ∞ only on sets of measure zero. Instead, the dual inequality of $T(\sigma\,\cdot) \colon L^p(\sigma) \to L^q(u)$ is $T^\star(\cdot\, u) \colon L^{q'}(u) \to L^{p'}(\sigma)$ and does not require to have non-negative measures, see [Cru14, §5, page 23].

Proof of Proposition C.0.1. Plugging the weight inside $|\cdot|^p$ gives

$$\Big(\int u(x)|T(fv^{1/p}v^{-1/p})(x)|^q\,\mathrm{d}x\Big)^{1/q} \le C\Big(\int |fv^{1/p}|^p\,\mathrm{d}x\Big)^{1/p}. \tag{C.0.2}$$

Let $\sigma := v^{-p'/p}$, so that $\sigma^{1/p'} = v^{-1/p}$. Also let $g = fv^{1/p}$, then (C.0.2) is equivalent to

$$\|T(g\sigma^{1/p'})\|_{L^q(u)} \le C\|g\|_{L^p}\,, \quad \forall g \in L^p.$$

Write $g = h\sigma^{1/p}$ so that $h \in L^p(\sigma)$ if and only if $g \in L^p$. We obtain

$$\|T(h\sigma^{1/p}\sigma^{1/p'})\|_{L^q(u)} \le C\|h\|_{L^p(\sigma)}\,, \quad \forall h \in L^p(\sigma)$$

which is what we wanted, since $\frac{1}{p} + \frac{1}{p'} = 1$. $\square$

En passant, note that the two weight estimate (C.0.2), as the two formulations in Proposition C.0.1, is equivalent to the *unweighted* estimate

$$u^{1/q}\,T(\cdot\,\sigma^{1/p'})\colon L^p \to L^q.$$

In [Saw82], Sawyer observed that the inequality

$$\|T(fv^{1-p'})\|_{L^q(w)} \le C\|f\|_{L^p(v^{1-p'})}$$

is equivalent to the dual inequality

$$\|T^*(gw)\|_{L^{p'}(v^{1-p'})} \le C\|g\|_{L^{q'}(w)}\,, \quad \forall g \in L^{q'}(w).$$

BIBLIOGRAPHY

[APR17] A. Aleman, S. Pott, and M. C. Reguera, "Sarason conjecture on the Bergman space," *Int. Math. Res. Not. IMRN*, no. 14, pp. 4320–4349, 2017 (cit. on pp. 3, 17, 18, 118, 119, 123, 125, 131, 132).

[And15] T. C. Anderson, "A new sufficient two-weighted bump assumption for L^p boundedness of Calderón-Zygmund operators," *Proc. Am. Math. Soc.*, vol. 143, no. 8, pp. 3573–3586, 2015 (cit. on p. 21).

[ARS06] N. Arcozzi, R. Rochberg, and E. Sawyer, "Carleson measures and interpolating sequences for Besov spaces on complex balls," *Mem. Amer. Math. Soc.*, vol. 182, no. 859, pp. vi+163, 2006 (cit. on p. 122).

[ARS02] N. Arcozzi, R. Rochberg, and E. Sawyer, "Carleson measures for analytic Besov spaces," *Rev. Mat. Iberoamericana*, vol. 18, no. 2, pp. 443–510, 2002 (cit. on p. 120).

[AIS01] K. Astala, T. Iwaniec, and E. Saksman, "Beltrami operators in the plane," *Duke Math. J.*, vol. 107, no. 1, pp. 27–56, 2001 (cit. on pp. 3, 4).

[Aus07] P. Auscher, "On necessary and sufficient conditions for L^p-estimates of Riesz transforms associated to elliptic operators on $\mathbb{R}^n$ and related estimates," *Mem. Amer. Math. Soc.*, vol. 186, no. 871, pp. xviii+75, 2007 (cit. on pp. 14, 95).

[ACDH04] P. Auscher, T. Coulhon, X. T. Duong, and S. Hofmann, "Riesz transform on manifolds and heat kernel regularity," *Ann. Sci. École Norm. Sup. (4)*, vol. 37, no. 6, pp. 911–957, 2004 (cit. on pp. 98–100).

[AM06] P. Auscher and J. M. Martell, "Weighted norm inequalities, off-diagonal estimates and elliptic operators. III: Harmonic analysis of elliptic operators," *J. Funct. Anal.*, vol. 241, no. 2, pp. 703–746, 2006 (cit. on pp. 14, 93, 95, 96, 99, 101).

[AM07a] ——, "Weighted norm inequalities, off-diagonal estimates and elliptic operators. I. General operator theory and weights," *Adv. Math.*, vol. 212, no. 1, pp. 225–276, 2007 (cit. on pp. 14, 15).

[AM07b] ——, "Weighted norm inequalities, off-diagonal estimates and elliptic operators. II: Off-diagonal estimates on spaces of homogeneous type," *J. Evol. Equ.*, vol. 7, no. 2, pp. 265–316, 2007 (cit. on pp. 14, 87, 88, 91, 93, 98, 99).

[AM08] ——, "Weighted norm inequalities, off-diagonal estimates and elliptic operators. IV. Riesz transforms on manifolds and weights," *Math. Z.*, vol. 260, no. 3, pp. 527–539, 2008 (cit. on pp. 14, 98, 100).

[AT98] P. Auscher and P. Tchamitchian, "Square root problem for divergence operators and related topics," *Astérisque*, no. 249, pp. viii+172, 1998 (cit. on p. 40).

[ACS78] S. Axler, S.-Y. A. Chang, and D. Sarason, "Products of Toeplitz operators," *Integral Equations Operator Theory*, vol. 1, no. 3, pp. 285–309, 1978 (cit. on p. 18).

[BBR20] J. Bailey, G. Brocchi, and M. C. Reguera. (2020). Quadratic sparse domination and weighted estimates for non-integral square functions. arXiv: 2007.15928 (cit. on p. 81).

[BCOR19] A. Barron, J. M. Conde-Alonso, Y. Ou, and G. Rey, "Sparse domination and the strong maximal function," *Adv. Math.*, vol. 345, pp. 1–26, 2019 (cit. on p. 8).

[BC20] D. Beltran and L. Cladek, "Sparse bounds for pseudodifferential operators," *J. Anal. Math.*, vol. 140, no. 1, pp. 89–116, 2020 (cit. on p. 8).

[BV20] D. Beltran and L. Vega, "Bilinear identities involving the k-plane transform and Fourier extension operators," *Proc. Roy. Soc. Edinburgh Sect. A*, vol. 150, no. 6, pp. 3349–3377, 2020 (cit. on p. 2).

[BB18] C. Benea and F. Bernicot, "Conservation de certaines propriétés à travers un contrôle épars d'un opérateur et applications au projecteur de Leray-Hopf," *Ann. Inst. Fourier (Grenoble)*, vol. 68, no. 6, pp. 2329–2379, 2018 (cit. on p. 9).

[BBL17] C. Benea, F. Bernicot, and T. Luque, "Sparse bilinear forms for Bochner Riesz multipliers and applications," *Trans. London Math. Soc.*, vol. 4, no. 1, pp. 110–128, 2017 (cit. on p. 8).

[BBCH09] J. Bennett, N. Bez, A. Carbery, and D. Hundertmark, "Heat-flow monotonicity of Strichartz norms," *Anal. PDE*, vol. 2, no. 2, pp. 147–158, 2009 (cit. on p. 2).

[BBF+18] J. Bennett, N. Bez, T. C. Flock, S. Gutiérrez, and M. Iliopoulou, "A sharp k-plane Strichartz inequality for the Schrödinger equation," *Trans. Amer. Math. Soc.*, vol. 370, no. 8, pp. 5617–5633, 2018 (cit. on pp. 2, 3, 25, 32).

[BBJP17] J. Bennett, N. Bez, C. Jeavons, and N. Pattakos, "On sharp bilinear Strichartz estimates of Ozawa–Tsutsumi type," *Journal of the Mathematical Society of Japan*, vol. 69, no. 2, pp. 459–476, 2017 (cit. on p. 32).

[BN21] J. Bennett and S. Nakamura, "Tomography bounds for the fourier extension operator and applications," *Mathematische Annalen*, pp. 1–41, 2021 (cit. on p. 2).

[BFP16] F. Bernicot, D. Frey, and S. Petermichl, "Sharp weighted norm estimates beyond Calderón-Zygmund theory," *Anal. PDE*, vol. 9, no. 5, pp. 1079–1113, 2016 (cit. on pp. i, 8, 14, 15, 84, 85, 104, 106, 109–111).

[BK03] S. Blunck and P. C. Kunstmann, "Calderón-Zygmund theory for non-integral operators and the H^∞ functional calculus," *Rev. Mat. Iberoamericana*, vol. 19, no. 3, pp. 919–942, 2003 (cit. on p. 14).

[Bro20] G. Brocchi, "A sparse quadratic $T(1)$ theorem," *New York J. Math.*, vol. 26, pp. 1232–1272, 2020 (cit. on p. 11).

[BOQ20] G. Brocchi, D. Oliveira e Silva, and R. Quilodrán, "Sharp Strichartz inequalities for fractional and higher-order Schrödinger equations," *Anal. PDE*, vol. 13, no. 2, pp. 477–526, 2020.

[Buc93] S. M. Buckley, "Estimates for operator norms on weighted spaces and reverse Jensen inequalities," *Trans. Amer. Math. Soc.*, vol. 340, no. 1, pp. 253–272, 1993 (cit. on pp. 6, 7, 12).

[BCDH17] T. A. Bui, J. M. Conde-Alonso, X. T. Duong, and M. Hormozi, "A note on weighted bounds for singular operators with nonsmooth kernels," *Studia Math.*, vol. 236, no. 3, pp. 245–269, 2017 (cit. on p. 14).

[BD20a] T. A. Bui and X. T. Duong, "Sharp weighted estimates for square functions associated to operators on spaces of homogeneous type," *J. Geom. Anal.*, vol. 30, no. 1, pp. 874–900, 2020 (cit. on pp. 14, 80, 92, 96, 97).

[BD20b] ——, "Sharp weighted norm inequalities for singular integrals with non-smooth kernels," *Math. Z.*, vol. 295, no. 3-4, pp. 1733–1750, 2020 (cit. on p. 14).

[Bur84] D. L. Burkholder, "Boundary value problems and sharp inequalities for martingale transforms," *Ann. Probab.*, vol. 12, no. 3, pp. 647–702, 1984 (cit. on p. 139).

[CD17] M. J. Carro and C. Domingo-Salazar, "Stein's square function G_α and sparse operators," *J. Geom. Anal.*, vol. 27, no. 2, pp. 1624–1635, 2017 (cit. on p. 8).

[CJ87] M. Christ and J.-L. Journé, "Polynomial growth estimates for multilinear singular integral operators," *Acta Math.*, vol. 159, no. 1-2, pp. 51–80, 1987 (cit. on p. 10).

[CW71] R. R. Coifman and G. Weiss, *Analyse harmonique non-commutative sur certains espaces homogènes*, ser. Lecture Notes in Mathematics, Vol. 242. Springer-Verlag, Berlin-New York, 1971, pp. v+160, Étude de certaines intégrales singulières (cit. on p. 87).

[CCDO17] J. M. Conde-Alonso, A. Culiuc, F. Di Plinio, and Y. Ou, "A sparse domination principle for rough singular integrals," *Anal. PDE*, vol. 10, no. 5, pp. 1255–1284, 2017 (cit. on pp. 8, 9).

[CR16] J. M. Conde-Alonso and G. Rey, "A pointwise estimate for positive dyadic shifts and some applications," *Math. Ann.*, vol. 365, no. 3-4, pp. 1111–1135, 2016 (cit. on p. 8).

[CD99] T. Coulhon and X. T. Duong, "Riesz transforms for $1 \leq p \leq 2$," *Trans. Am. Math. Soc.*, vol. 351, no. 3, pp. 1151–1169, 1999 (cit. on p. 98).

[Cru94] D. Cruz-Uribe, "The invertibility of the product of unbounded Toeplitz operators," *Integral Equations Operator Theory*, vol. 20, no. 2, pp. 231–237, 1994 (cit. on p. 18).

[Cru14] ——, (2014). Two weight norm inequalities for fractional integral operators and commutators. arXiv: `1412.4157` (cit. on p. 142).

[CMP12] D. Cruz-Uribe, J. M. Martell, and C. Pérez, "Sharp weighted estimates for classical operators." *Adv. Math.*, vol. 229, no. 1, pp. 408–441, 2012 (cit. on p. 13).

[CDO18a] A. Culiuc, F. Di Plinio, and Y. Ou, "Domination of multilinear singular integrals by positive sparse forms," *J. Lond. Math. Soc. (2)*, vol. 98, no. 2, pp. 369–392, 2018 (cit. on pp. 8, 14).

[CDO18b] ——, "Uniform sparse domination of singular integrals via dyadic shifts," *Math. Res. Lett.*, vol. 25, no. 1, pp. 21–42, 2018 (cit. on p. 8).

[DJ84] G. David and J.-L. Journé, "A boundedness criterion for generalized Calderón-Zygmund operators," *Ann. of Math. (2)*, vol. 120, no. 2, pp. 371–397, 1984 (cit. on p. 9).

[Dav95] E. B. Davies, "Uniformly elliptic operators with measurable coefficients," *J. Funct. Anal.*, vol. 132, no. 1, pp. 141–169, 1995 (cit. on p. 99).

[DDU18] F. Di Plinio, Y. Q. Do, and G. N. Uraltsev, "Positive sparse domination of variational Carleson operators," *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, vol. 18, no. 4, pp. 1443–1458, 2018 (cit. on p. 8).

[DWW20] F. Di Plinio, B. D. Wick, and T. Williams. (2020). Wavelet representation of singular integral operators. arXiv: `2009.01212` (cit. on p. 13).

[DLR16] C. Domingo-Salazar, M. Lacey, and G. Rey, "Borderline weak-type estimates for singular integrals and square functions," *Bull. Lond. Math. Soc.*, vol. 48, no. 1, pp. 63–73, 2016 (cit. on p. 13).

[DGPP05] O. Dragičević, L. Grafakos, M. C. Pereyra, and S. Petermichl, "Extrapolation and sharp norm estimates for classical operators on weighted Lebesgue spaces," *Publ. Mat.*, vol. 49, no. 1, pp. 73–91, 2005 (cit. on p. 13).

[Dro14] A. Drouot, "Sharp constant for a k-plane transform inequality," *Anal. PDE*, vol. 7, no. 6, pp. 1237–1252, 2014 (cit. on p. 3).

[Fef83] C. L. Fefferman, "The uncertainty principle," *Bull. Amer. Math. Soc. (N.S.)*, vol. 9, no. 2, pp. 129–206, 1983 (cit. on p. 142).

[FKP91] R. A. Fefferman, C. E. Kenig, and J. Pipher, "The theory of weights and the Dirichlet problem for elliptic equations," *Ann. of Math. (2)*, vol. 134, no. 1, pp. 65–124, 1991 (cit. on p. 3).

[FN19] D. Frey and Z. Nieraeth, "Weak and strong type $A_1 - A_\infty$ estimates for sparsely dominated operators," *J. Geom. Anal.*, vol. 29, no. 1, pp. 247–282, 2019 (cit. on pp. 10, 80, 116).

[GHK20] C. Gan, B. Hu, and I. Khan. (2020). Dyadic decomposition of convex domains of finite type and applications. arXiv: `2001.07200` (cit. on p. 119).

[Gra14] L. Grafakos, *Classical Fourier analysis*, Third, ser. Graduate Texts in Mathematics. Springer, New York, 2014, vol. 249, pp. xviii+638 (cit. on pp. 6, 7, 13, 14, 39, 99, 101, 138).

[Haa06] M. Haase, *The functional calculus for sectorial operators*. Basel: Birkhäuser, 2006, vol. 169, pp. xiv + 392 (cit. on p. 83).

[Hän18] T. S. Hänninen, "Equivalence of sparse and Carleson coefficients for general sets," *Ark. Mat.*, vol. 56, no. 2, pp. 333–339, 2018 (cit. on p. 8).

[Hof08] S. Hofmann, "A local Tb theorem for square functions." In *Perspectives in partial differential equations, harmonic analysis and applications. A volume in honor of Vladimir G. Maz'ya's 70th birthday*, Providence, RI: American Mathematical Society (AMS), 2008, pp. 175–185 (cit. on p. 11).

[Hof10] ——, "Local $T(b)$ theorems and applications in PDE," in *Harmonic analysis and partial differential equations*, ser. Contemp. Math. Vol. 505, Amer. Math. Soc., Providence, RI, 2010, pp. 29–52 (cit. on p. 11).

[HLM+11] S. Hofmann, G. Lu, D. Mitrea, M. Mitrea, and L. Yan, "Hardy spaces associated to non-negative self-adjoint operators satisfying Davies-Gaffney estimates," *Mem. Amer. Math. Soc.*, vol. 214, no. 1007, pp. vi+78, 2011 (cit. on p. 90).

[HM03] S. Hofmann and J. M. Martell, "L^p bounds for Riesz transforms and square roots associated to second order elliptic operators," *Publ. Mat.*, vol. 47, no. 2, pp. 497–515, 2003 (cit. on p. 14).

[HTV00] S. Hukovic, S. Treil, and A. Volberg, "The Bellman functions and sharp weighted inequalities for square functions," in *Complex analysis, operators, and related topics*, ser. Oper. Theory Adv. Appl. Vol. 113, Birkhäuser, Basel, 2000, pp. 97–113 (cit. on p. 12).

[HMW73] R. Hunt, B. Muckenhoupt, and R. Wheeden, "Weighted norm inequalities for the conjugate function and Hilbert transform," *Trans. Amer. Math. Soc.*, vol. 176, pp. 227–251, 1973 (cit. on p. 3).

[HWW21] Z. Huo, N. A. Wagner, and B. D. Wick, "Bekollé-Bonami estimates on some pseudoconvex domains," *Bull. Sci. Math.*, vol. 170, p. 102 993, 2021 (cit. on p. 127).

[HW20] Z. Huo and B. D. Wick, "Weighted estimates for the Bergman projection on the Hartogs triangle," *J. Funct. Anal.*, vol. 279, no. 9, pp. 108727, 34, 2020 (cit. on p. 133).

[HK12] T. Hytönen and A. Kairema, "Systems of dyadic cubes in a doubling metric space," *Colloq. Math.*, vol. 126, no. 1, pp. 1–33, 2012 (cit. on pp. 85, 86).

[HPTV14] T. Hytönen, C. Pérez, S. Treil, and A. Volberg, "Sharp weighted estimates for dyadic shifts and the A_2 conjecture," *J. Reine Angew. Math.*, vol. 687, pp. 43–86, 2014, arXiv: `1010.0755 [math.CA]` (cit. on p. 71).

[HvVW16] T. Hytönen, J. van Neerven, M. Veraar, and L. Weis, *Analysis in Banach spaces. Volume I. Martingales and Littlewood-Paley theory.* Cham: Springer, 2016, vol. 63, pp. xvii + 614 (cit. on p. 138).

[Hyt12] T. P. Hytönen, "The sharp weighted bound for general Calderón-Zygmund operators." *Ann. Math. (2)*, vol. 175, no. 3, pp. 1473–1506, 2012 (cit. on pp. 6, 38).

[Hyt14] ——, "The A_2 theorem: Remarks and complements," in *Harmonic analysis and partial differential equations*, ser. Contemp. Math. Vol. 612, Amer. Math. Soc., Providence, RI, 2014, pp. 91–106 (cit. on pp. 7, 126, 132).

[Hyt17] ——, "Representation of singular integrals by dyadic operators, and the A_2 theorem." *Expo. Math.*, vol. 35, no. 2, pp. 166–205, 2017, arXiv: `1108.5119 [math.CA]` (cit. on pp. 38, 47).

[HL12] T. P. Hytönen and M. T. Lacey, "The A_p-A_∞ inequality for general Calderón-Zygmund operators," *Indiana Univ. Math. J.*, vol. 61, no. 6, pp. 2041–2092, 2012 (cit. on p. 132).

[HLP13] T. P. Hytönen, M. T. Lacey, and C. Pérez, "Sharp weighted bounds for the q-variation of singular integrals." *Bull. Lond. Math. Soc.*, vol. 45, no. 3, pp. 529–540, 2013 (cit. on p. 7).

[HL18] T. P. Hytönen and K. Li, "Weak and strong A_p-A_∞ estimates for square functions and related operators," *Proc. Amer. Math. Soc.*, vol. 146, no. 6, pp. 2497–2507, 2018 (cit. on p. 13).

[HRT17] T. P. Hytönen, L. Roncal, and O. Tapiola, "Quantitative weighted estimates for rough homogeneous singular integrals," *Israel J. Math.*, vol. 218, no. 1, pp. 133–164, 2017 (cit. on p. 7).

[HLM+12] T. P. Hytönen *et al.*, "Weak and strong type estimates for maximal truncations of Calderón-Zygmund operators on A_p weighted spaces," *J. Anal. Math.*, vol. 118, no. 1, pp. 177–220, 2012 (cit. on p. 7).

[Koo80] P. Koosis, "Moyennes quadratiques pondérées de fonctions périodiques et de leurs conjuguées harmoniques," *CR Acad. Sci. Paris Sér. AB*, vol. 291, no. 4, 1980 (cit. on p. 142).

[Kra01] S. G. Krantz, *Function theory of several complex variables. Reprint of the 1992 2nd ed. with corrections*, Reprint of the 1992 2nd ed. with corrections. Providence, RI: American Mathematical Society (AMS), AMS Chelsea Publishing, 2001, pp. xvi + 564 (cit. on p. 17).

[LS12] M. T. Lacey and J. Scurry. (2012). Weighted weak type estimates for square functions. arXiv: `1211.4219 [math.CA]` (cit. on p. 13).

[Lac17] M. T. Lacey, "An elementary proof of the A_2 bound," *Israel J. Math.*, vol. 217, no. 1, pp. 181–195, 2017 (cit. on pp. 7, 102).

[LL16] M. T. Lacey and K. Li, "On A_p–A_∞ type estimates for square functions," *Math. Z.*, vol. 284, no. 3-4, pp. 1211–1222, 2016 (cit. on p. 13).

[LM17a] M. T. Lacey and H. Martikainen, "Local Tb theorem with L^2 testing conditions and general measures: Square functions," *J. Anal. Math.*, vol. 133, pp. 71–89, 2017 (cit. on pp. 11, 46, 54).

[LMR19] M. T. Lacey, D. Mena, and M. C. Reguera, "Sparse bounds for Bochner-Riesz multipliers," *J. Fourier Anal. Appl.*, vol. 25, no. 2, pp. 523–537, 2019 (cit. on p. 8).

[LM17b] M. T. Lacey and D. A. Mena, "The sparse T1 theorem." *Houston J. Math.*, vol. 43, no. 1, pp. 111–127, 2017, arXiv: `1610.01531 [math.CA]` (cit. on pp. 9, 71, 73).

[LPR10] M. T. Lacey, S. Petermichl, and M. C. Reguera, "Sharp A_2 inequality for Haar shift operators," *Math. Ann.*, vol. 348, no. 1, pp. 127–141, 2010 (cit. on p. 6).

[LSU09] M. T. Lacey, E. T. Sawyer, and I. Uriarte-Tuero. (2009). Two weight inequalities for discrete positive operators. arXiv: `0911.3437 [math.CA]`, [Online]. Available: `https://ddd.uab.cat/record/60902` (cit. on p. 126).

[LS17] M. T. Lacey and S. Spencer, "Sparse bounds for oscillatory and random singular integrals." *New York J. Math.*, vol. 23, pp. 119–131, 2017 (cit. on p. 8).

[Ler06] A. K. Lerner, "On some sharp weighted norm inequalities," *J. Funct. Anal.*, vol. 232, no. 2, pp. 477–494, 2006 (cit. on pp. 12, 80).

[Ler08a] ——, "An elementary approach to several results on the Hardy-Littlewood maximal operator," *Proc. Amer. Math. Soc.*, vol. 136, no. 8, pp. 2829–2833, 2008 (cit. on p. 123).

[Ler08b] ——, "On some weighted norm inequalities for littlewood–paley operators," *Illinois Journal of Mathematics*, vol. 52, no. 2, pp. 653–666, 2008 (cit. on p. 12).

[Ler11] ——, "Sharp weighted norm inequalities for Littlewood-Paley operators and singular integrals," *Adv. Math.*, vol. 226, no. 5, pp. 3912–3926, 2011, arXiv: `1005.1422 [math.CA]` (cit. on pp. 13, 80, 97).

[Ler13a] ——, "A simple proof of the A_2 conjecture," *Int. Math. Res. Not. IMRN*, no. 14, pp. 3159–3170, 2013 (cit. on p. 7).

[Ler13b] ——, "On an estimate of Calderón-Zygmund operators by dyadic positive operators," *J. Anal. Math.*, vol. 121, pp. 141–161, 2013 (cit. on p. 124).

[Ler16] ——, "On pointwise estimates involving sparse operators," *New York J. Math.*, vol. 22, pp. 341–349, 2016, arXiv: `1512.07247 [math.CA]` (cit. on pp. 7, 8, 102).

[LN19] A. K. Lerner and F. Nazarov, "Intuitive dyadic calculus: The basics," *Expo. Math.*, vol. 37, no. 3, pp. 225–265, 2019, arXiv: `1508.05639 [math.CA]`.

[Li17] K. Li, "Two weight inequalities for bilinear forms," *Collect. Math.*, vol. 68, no. 1, pp. 129–144, 2017 (cit. on pp. 22, 118, 126, 131).

[Lor21] E. Lorist, "On Pointwise ℓ^r-Sparse Domination in a Space of Homogeneous Type," *J. Geom. Anal.*, vol. 31, no. 9, pp. 9366–9405, 2021, arXiv: `1907.00690 [math.CA]` (cit. on p. 86).

[LN20] E. Lorist and Z. Nieraeth. (2020). Sparse domination implies vector-valued sparse domination. arXiv: `2003.02233 [math.CA]` (cit. on pp. 10, 81).

[MM14] H. Martikainen and M. Mourgoglou, "Square functions with general measures." *Proc. Am. Math. Soc.*, vol. 142, no. 11, pp. 3923–3931, 2014 (cit. on pp. 12, 44, 46, 54).

[MM15] ——, "Boundedness of non-homogeneous square functions and L^q type testing conditions with $q \in (1, 2)$," *Math. Res. Lett.*, vol. 22, no. 5, pp. 1417–1457, 2015, arXiv: `1401.5457 [math.CA]` (cit. on p. 44).

[MMV19] H. Martikainen, M. Mourgoglou, and J. Vuorinen, "A new approach to non-homogeneous local *tb* theorems," *Journal d'Analyse Mathematique*, 2019, arXiv: `1511.00528v2 [math.CA]` (cit. on p. 34).

[NTV03] F. Nazarov, S. Treil, and A. Volberg, "The *Tb*-theorem on non-homogeneous spaces," *Acta Math.*, vol. 190, no. 2, pp. 151–239, 2003 (cit. on p. 36).

[Nie19] Z. Nieraeth, "Quantitative estimates and extrapolation for multilinear weight classes," *Math. Ann.*, vol. 375, no. 1-2, pp. 453–507, 2019 (cit. on pp. 80, 81).

[Obe19] R. Oberlin, "Sparse bounds for a prototypical singular Radon transform," *Canad. Math. Bull.*, vol. 62, no. 2, pp. 405–415, 2019 (cit. on p. 8).

[Per19] M. C. Pereyra, "Dyadic harmonic analysis and weighted inequalities: The sparse revolution," in *New Trends in Applied Harmonic Analysis, Volume 2*, Springer, 2019, pp. 159–239, arXiv: `1812.00850 [math.CA]` (cit. on pp. 38, 47).

[Pér95] C. Pérez, "On sufficient conditions for the boundedness of the Hardy-Littlewood maximal operator between weighted L^p-spaces with different weights," *Proc. London Math. Soc. (3)*, vol. 71, no. 1, pp. 135–157, 1995 (cit. on p. 21).

[Pet07] S. Petermichl, "The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical A_p characteristic," *Amer. J. Math.*, vol. 129, no. 5, pp. 1355–1375, 2007 (cit. on p. 6).

[Pet08] S. Petermichl, "The sharp weighted bound for the Riesz transforms," *Proc. Amer. Math. Soc.*, vol. 136, no. 4, pp. 1237–1249, 2008 (cit. on p. 6).

[PV02] S. Petermichl and A. Volberg, "Heating of the Ahlfors-Beurling operator: Weakly quasiregular maps on the plane are quasiregular," *Duke Math. J.*, vol. 112, no. 2, pp. 281–305, 2002 (cit. on pp. 3, 4).

[PR13] S. Pott and M. C. Reguera, "Sharp Békollé estimates for the Bergman projection," *J. Funct. Anal.*, vol. 265, no. 12, pp. 3233–3244, 2013 (cit. on p. 17).

[RTW17] R. Rahm, E. Tchoundja, and B. D. Wick, "Weighted estimates for the Berezin transform and Bergman projection on the unit ball," *Math. Z.*, vol. 286, no. 3-4, pp. 1465–1478, 2017 (cit. on pp. 17, 22, 119, 120, 122, 124, 126, 131, 133).

[Rub18] B. Rubin, "A note on the Blaschke-Petkantschin formula, Riesz distributions, and Drury's identity," *Fract. Calc. Appl. Anal.*, vol. 21, no. 6, pp. 1641–1650, 2018 (cit. on pp. 28, 29).

[Saw82] E. T. Sawyer, "A characterization of a two-weight norm inequality for maximal operators," *Studia Math.*, vol. 75, no. 1, pp. 1–11, 1982 (cit. on pp. 126, 143).

[Saw88] ——, "A characterization of two weight norm inequalities for fractional and Poisson integrals," *Trans. Amer. Math. Soc.*, vol. 308, no. 2, pp. 533–545, 1988 (cit. on p. 126).

[Sem90] S. Semmes, "Square function estimates and the $T(b)$ theorem," *Proc. Amer. Math. Soc.*, vol. 110, no. 3, pp. 721–726, 1990.

[Ste70] E. M. Stein, *Topics in harmonic analysis related to the Littlewood-Paley theory*, ser. Annals of Mathematics Studies, No. 63. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1970, pp. viii+146 (cit. on p. 98).

[Str77] R. S. Strichartz, "Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations," *Duke Mathematical Journal*, vol. 44, no. 3, pp. 705–714, 1977 (cit. on p. 24).

[TV16] S. Treil and A. Volberg, "Entropy conditions in two weight inequalities for singular integral operators," *Adv. Math.*, vol. 301, pp. 499–548, 2016 (cit. on p. 142).

[VZ18] A. Volberg and P. Zorin-Kranich, "Sparse domination on non-homogeneous spaces with an application to A_p weights," *Rev. Mat. Iberoam.*, vol. 34, no. 3, pp. 1401–1414, 2018 (cit. on p. 8).

[Wil07] M. Wilson, "The intrinsic square function," *Rev. Mat. Iberoam.*, vol. 23, no. 3, pp. 771–791, 2007 (cit. on p. 13).

[Wit02] J. Wittwer, "A sharp estimate on the norm of the continuous square function," *Proc. Amer. Math. Soc.*, vol. 130, no. 8, pp. 2335–2342, 2002 (cit. on p. 12).

[Zor16] P. Zorin-Kranich. (2016). Intrinsic square functions with arbitrary aperture. arXiv: `1605.02936` (cit. on p. 13).