HARMONIC ANALYSIS AND SQUARE FUNCTIONS, LECTURE NOTES, UNIVERSITY OF HELSINKI, FALL 2015

HENRI MARTIKAINEN

ABSTRACT. In these lecture notes we study non-homogeneous analysis in $\mathbb{R}^n$. This means that we do harmonic analysis with rather general Borel measures μ. Usually we assume that our measure μ is of order m i.e. $\mu(B(x,r)) \le Cr^m$ for some $m \in (0,n]$. No control from below, or doubling properties, are assumed. We study the boundedness properties of vertical square functions and prove various different Tb theorems.

These lecture notes are mainly based on the article [1] by Martikainen, Mourgoglou and Vuorinen, and the book [4] by X. Tolsa.

CONTENTS

1. SQUARE FUNCTIONS

We work in $\mathbb{R}^n$ using m-dimensional objects, where $m \in (0,n]$ does not need to be an integer.

H.M. is supported by the Academy of Finland through the grant Multiparameter dyadic harmonic analysis and probabilistic methods, and is a member of the Finnish Centre of Excellence in Analysis and Dynamics Research.

Littlewood–Paley family of kernels. We say that a family of functions

$$s_t\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C},\ t > 0,$$

is an m-dimensional Littlewood–Paley family (m–LP–family), if for some $\alpha > 0$ and $C < \infty$ we have the size condition

$$|s_t(x,y)| \le C \frac{t^\alpha}{(t + |x-y|)^{m+\alpha}} \tag{1.1}$$

and the y-Hölder condition

$$|s_t(x,y) - s_t(x,z)| \le C \frac{|y-z|^\alpha}{(t + |x-y|)^{m+\alpha}} \tag{1.2}$$

whenever $|y - z| < t/2$.

We say that $(s_t)_{t>0}$ is an x-continuous m–LP–family if, in addition, we have the x-Hölder condition

$$|s_t(x,y) - s_t(z,y)| \le C \frac{|x-z|^\alpha}{(t + |x-y|)^{m+\alpha}} \tag{1.3}$$

whenever $|x - z| < t/2$.

Measures of order m. We say that a Borel measure μ in $\mathbb{R}^n$ is of order m if for some $C_\mu < \infty$ we have

$$\mu(B(x,r)) \le C_\mu r^m, \qquad x \in \mathbb{R}^n,\ r > 0.$$

We do not assume any control from below (in the much easier ADR situation one would also have $\mu(B(x,r)) \ge c_\mu r^m$). Also, we do not assume any doubling (i.e. $\mu(B(x,2r)) \le C\mu(B(x,r))$ property of the measure.

Integral operators θ_t and the vertical square function V. Let us be given an m–LP–family $(s_t)_{t>0}$ and a measure μ which is either finite or of order m. For $f \in \bigcup_{p\in[1,\infty]} L^p(\mu)$ and $x \in \mathbb{R}^n$ we set

$$\theta_t^\mu f(x) = \int s_t(x,y) f(y)\, d\mu(y).$$

This is an absolutely convergent integral (exercise). Our main object of study is the vertical square function operator

$$V_\mu f(x) = \Big(\int_0^\infty |\theta_t^\mu f(x)|^2 \,\frac{dt}{t}\Big)^{1/2}.$$

Why square functions? One of the main application areas is the theory of partial differential equations. A prototypical example of an n–LP–family of convolution form is

$$s_t(x-y) = t\frac{\partial}{\partial t}p_t(x-y),$$

where

$$p_t(x) = \frac{2t}{\sigma_n(t^2+|x|^2)^{(n+1)/2}}$$

is the classical Poisson kernel for the Laplacian in the upper half-space $\mathbb{R}^{n+1} = \{(x,t) : x \in \mathbb{R}^n, t > 0\}$, with σ_n denoting the volume of the unit n-sphere in $\mathbb{R}^{n+1}$. Let m_n denote the Lebesgue measure in $\mathbb{R}^n$. Then

$$\theta_t^{m_n} f = s_t * f$$

and the corresponding square function plays an important role in the theory of harmonic functions in $\mathbb{R}^{n+1}_+$.

An important example of non-convolution kernels arises in the theory of divergence form elliptic operators. Generalising what was done above we can define s_t to be the Poisson kernel p_t^L associated to some second order divergence operator L. Under appropriate assumptions also such kernels are of Littlewood–Paley type by the De Giorgi/Nash regularity theory

In these lecture notes we are interested in the L^2 (or L^p) boundedness of square functions. In the Lebesgue measure case the convolution kernel situation is classical. It reduces to Fourier transform considerations and using Planceherel identity. We are interested in situations where such easy solutions do not exist. This will lead us to the so called Tb theorems, which characterise the L^2 boundedness for general LP-families which need not be of convolution form. Another significant difficulty is that we will consider quite irregular measures μ (this is so called non-homogeneous theory since our measure is not doubling).

Before getting to the various Tb theorems for square functions we will spend some time on general non-homogeneous analysis and other boundedness criteria for square functions.

2. Preliminaries from non-homogeneous analysis

We denote by C a general big constant, which can change from line to line. If we write C_1, C_2 etc. we refer to specific fixed constants. If we want to highlight the dependence on some parameter, say α, we may write C_α.

2.1. **Complex measures.** A complex measure ν is a complex-valued and countably additive function (i.e. $\nu(\bigcup A_i) = \sum_i \nu(A_i)$ for disjoint measurable A_i) defined in some σ-algebra. To every such measure we associate its variation measure

$$|\nu|(A) = \sup \sum_i |\nu(A_i)|,$$

where the supremum is taken over all the measurable sequences of disjoint sets (A_i) satisfying $A = \bigcup A_i$. This is a finite positive measure. There also exists a measurable function h so that $|h(x)| = 1$ for all $x \in \mathbb{R}^n$, and

$$d\nu = h\, d|\nu| \qquad \text{i.e. } \nu(A) = \int_A h\, d|\nu|.$$

Moreover, if μ is a measure, $f \in L^1(\mu)$ and we define

$$\nu = f\, d\mu$$

then

$$|\nu| = |f|\, d\mu.$$

We define the total variation

$$\|\nu\| = |\nu|(\mathbb{R}^n).$$

Let $M(\mathbb{R}^n)$ denote the vector-space of all complex Borel measures in $\mathbb{R}^n$. This becomes a Banach space when equipped with the norm $\|\cdot\|$. In the exercises we will consider some details. For more details about complex measures you can also consult Chapter 6 of Rudin's book [3].

It is convenient to extend the above definition of vertical square functions as follows. Given an LP-family $(s_t)_{t>0}$ define

$$\theta_t \nu(x) = \int s_t(x, y)\, d\nu(y), \qquad \nu \in M(\mathbb{R}^n),\ x \in \mathbb{R}^n,$$

and

$$V\nu(x) = \Big(\int_0^\infty |\theta_t \nu(x)|^2 \frac{dt}{t} \Big)^{1/2}, \qquad \nu \in M(\mathbb{R}^n),\ x \in \mathbb{R}^n.$$

Given a measure μ we say that V maps $M(\mathbb{R}^n) \to L^{1,\infty}(\mu)$ if for some $C < \infty$ we have

$$\mu(\{x\colon\ V\nu(x) > \lambda\}) \le C \frac{\|\nu\|}{\lambda}$$

for every $\nu \in M(\mathbb{R}^n)$ and $\lambda > 0$.

2.2. **Maximal functions.** Let μ be an arbitrary Radon (i.e. locally finite Borel) measure in $\mathbb{R}^n$.

2.1. *Remark.* Our definition of a Radon measure μ is that μ is a positive Borel measure in $\mathbb{R}^n$ such that $\mu(K) < \infty$ for every compact set K. It is then a non-trivial fact that μ automatically enjoys other regularity properties (see Theorem 2.18 in [3]):

$$\mu(E) = \sup\{\mu(K)\colon\ K \subset E,\ K \text{ compact}\} = \inf\{\mu(V)\colon\ E \subset V,\ V \text{ open}\}$$

for every Borel set E. We could take this as the definition of a Radon measure, but it is, like said, automatic from the local finiteness.

We define the centred maximal function M_μ acting on a complex measure ν by

$$M_\mu \nu(x) = \sup_{r>0} \frac{|\nu|(B(x,r))}{\mu(B(x,r))}.$$

For $f \in L^1_{\mathrm{loc}}(\mu)$ we set $M_\mu f := M_\mu(f d\mu)$ i.e.

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

2.2. **Proposition.** *Let μ be a Radon measure in $\mathbb{R}^n$. The centred maximal operator M_μ maps $L^p(\mu) \to L^p(\mu)$ for $1 < p \le \infty$ and $M(\mathbb{R}^n) \to L^{1,\infty}(\mu)$.*

Proof. The case $p = \infty$ is trivial. The $L^p(\mu)$, $1 < p < \infty$, boundedness then follows from the Marcinkiewicz interpolation theorem if one shows the $L^1(\mu) \to L^{1,\infty}(\mu)$ boundedness. But this follows from the more general statement that $M_\mu \colon M(\mathbb{R}^n) \to L^{1,\infty}(\mu)$ boundedly. This is what we will prove now.

Fix $\nu \in M(\mathbb{R}^n)$ and $\lambda > 0$. We need to show that that

$$\mu(\{x \colon\ M_\mu \nu(x) > \lambda\}) \le C \frac{\|\nu\|}{\lambda}$$

for some $C < \infty$. Given $R > 0$ define

$$\Omega_R = \{x \in B(0,R) \cap \operatorname{spt} \mu \colon\ M_\mu \nu(x) > \lambda\}).$$

The localisation to a ball is made just to make this set bounded, since the covering theorem by Besicovitch requires this.

For each $x \in \Omega_R$ let B_x be a ball centred at x such that

$$\frac{|\nu|(B_x)}{\mu(B_x)} > \lambda.$$

By the Besicovitch covering theorem we can choose a subfamily $\{B_i\}_i \subset \{B_x\}_{x \in \Omega_R}$ such that

$$\Omega_R \subset \bigcup_i B_i \qquad \text{and} \qquad \sum_i 1_{B_i} \le C_n,$$

where $C_n < \infty$ is some dimensional constant. We get

$$\mu(\Omega_R) \le \sum_i \mu(B_i) \le \frac{1}{\lambda} \sum_i |\nu|(B_i) \le C_n \frac{\|\nu\|}{\lambda}.$$

This is a uniform bound in R so the claim follows by letting $R \to \infty$. □

Let $M_\mu^{\mathcal{Q}}$ be the centred maximal function, where we take supremum over cubes $Q(x,r)$ centred at x instead of balls $B(x,r)$. The above theorem holds also for $M_\mu^{\mathcal{Q}}$, since the Besicovitch covering theorem holds also for cubes.

An important difficulty in the non-homogeneous analysis is that the non-centred maximal function

$$M_\mu^{\mathrm{nc}} f(x) = \sup \Big\{ \frac{1}{\mu(B)} \int_B |f| \, d\mu \colon\ B \text{ closed ball},\ x \in B \Big\}$$

may fail to be of weak type $(1,1)$. One needs to use the centred maximal function M_μ or some other variant. Sometimes useful is the following variant

$$\widetilde{M}_\mu f(x) := \sup \Big\{ \frac{1}{\mu(5B)} \int_B |f| \, d\mu \colon \ B \text{ closed ball, } x \in B \Big\},$$

which satisfies the same boundedness properties as M_μ. This is easy to see using the standard $5r$-covering theorem.

2.3. **Local dyadic grids.** Let $Q_0 \subset \mathbb{R}^n$ be a half-open cube in $\mathbb{R}^n$ i.e.

$$Q_0 = [x_1, x_1 + \ell) \times \cdots \times [x_n, x_n + \ell).$$

We let $\mathcal{D}_0(Q_0) = \{Q_0\}$. A dyadic child of Q_0 is any of the 2^n cubes obtained by partitioning Q_0 by n median hyperplanes (these are the hyperplanes parallel to the faces of Q_0 and dividing each edge into 2 equal parts). The collection of the dyadic children of Q_0 is denoted ch(Q_0). We also define $\mathcal{D}_1(Q_0) = \text{ch}(Q_0)$. In general, given $k > 0$ we define inductively

$$\mathcal{D}_k(Q_0) = \bigcup_{Q \in \mathcal{D}_{k-1}(Q_0)} \text{ch}(Q).$$

For each $k \geq 0$ we have that

$$Q_0 = \bigcup_{Q \in \mathcal{D}_k(Q_0)} Q,$$

where the union is disjoint. We define the local dyadic grid $\mathcal{D}(Q_0)$ by setting

$$\mathcal{D}(Q_0) = \bigcup_{k=0}^{\infty} \mathcal{D}_k(Q_0).$$

If $Q \in \mathcal{D}_k(Q_0)$ we call Q a dyadic subcube of Q_0 of generation k, and denote its side length by $\ell(Q) = 2^{-k}\ell(Q_0)$.

2.4. **The dyadic maximal function.** Given a half-open cube Q_0 and a Radon measure μ we define the (local) dyadic maximal function

$$M_{\mathcal{D}(Q_0),\mu}\nu(x) = \sup_{Q \in \mathcal{D}(Q_0)} 1_Q(x) \frac{|\nu|(Q)}{\mu(Q)},$$

and as usual denote $M_{\mathcal{D}(Q_0),\mu} f := M_{\mathcal{D}(Q_0),\mu}(f \, d\mu)$.

2.3. **Proposition.** *Let μ be a Radon measure in $\mathbb{R}^n$. Given any half-open cube Q_0 the dyadic maximal operator $M_{\mathcal{D}(Q_0),\mu}$ maps $L^p(\mu) \to L^p(\mu)$ for $1 < p \leq \infty$ and $M(\mathbb{R}^n) \to L^{1,\infty}(\mu)$.*

Proof. As usual, it suffices to prove the boundedness $M(\mathbb{R}^n) \to L^{1,\infty}(\mu)$. So fix $\nu \in M(\mathbb{R}^n)$ and $\lambda > 0$. Choose maximal $Q \in \mathcal{D}(Q_0)$ so that

$$\frac{|\nu|(Q)}{\mu(Q)} > \lambda,$$

and denote their collection by $\mathcal{F}(Q_0)$. We have

$$\{x\colon\ M_{\mathcal{D}(Q_0),\mu}\nu(x) > \lambda\} = \bigcup_{Q\in\mathcal{F}(Q_0)} Q,$$

where the union is disjoint by maximality. Therefore, we have that

$$\mu(\{x\colon\ M_{\mathcal{D}(Q_0),\mu}\nu(x) > \lambda\}) = \sum_{Q\in\mathcal{F}(Q_0)} \mu(Q) \le \frac{1}{\lambda}\sum_{Q\in\mathcal{F}(Q_0)} |\nu|(Q) \le \frac{\|\nu\|}{\lambda}.$$

This ends the proof (notice that the above weak type bound does not even depend on n). $\square$

2.5. **Doubling cubes.** A cube $Q \subset \mathbb{R}^d$ is said μ-(α,β)-doubling (or just (α,β)-doubling if the measure μ is clear from the context) if

$$\mu(\alpha Q) \le \beta\mu(Q),$$

where αQ is the cube concentric with Q with diameter $\alpha\operatorname{diam}(Q)$. We record the following result:

2.4. **Lemma.** *Let μ be a measure of order m and $\beta > \alpha^m$. For every $x \in \operatorname{spt}\mu$ and $c > 0$ there exist some (α,β)-doubling cube Q centred at x with $\ell(Q) \ge c$.*

Proof. Exercise. $\square$

2.5. **Lemma.** *Let μ be a Radon measure in $\mathbb{R}^n$ and $\beta > \alpha^n$. Then for μ-a.e. $x \in \mathbb{R}^n$ there exists a sequence of (α,β)-doubling cubes $(Q_k)_k$ centred at x with $\ell(Q_k) \to 0$ when $k \to \infty$.*

Proof. For $j \ge 0$ define

$$F_j = \{x\colon\ \text{there is no } (\alpha,\beta)\text{-doubling cubes } Q \text{ centred at } x \text{ with } \ell(Q) \le 2^{-j}\}.$$

It suffices to fix j and prove $\mu(F_j) = 0$. For this it suffices to fix an arbitrary cube Q_0 with $\ell(Q_0) = 2^{-j}$ and show that $\mu(F_j \cap Q_0) = 0$. Fix now an integer $k \ge 1$. For each $y \in F_j \cap Q_0$ let Q_y be some cube centred at y with side length $\alpha^{-k}\ell(Q_0)$. Notice that the cubes $\alpha^s Q_y$ are not (α,β)-doubling for all $s = 0,\dots,k-1$ and that $\alpha^k Q_y \subset 2Q_0$. This implies that

$$\mu(Q_y) \le \beta^{-1}\mu(\alpha Q_y) \le \cdots \le \beta^{-k}\mu(\alpha^k Q_y) \le \beta^{-k}\mu(2Q_0). \tag{2.6}$$

Using Besicovitch covering theorem we choose $y_m \in Q_0 \cap F_j$ so that

$$Q_0 \cap F_j \subset \bigcup_m Q_{y_m} \qquad \text{and} \qquad \sum_m 1_{Q_{y_m}} \le C_n. \tag{2.7}$$

Let $N = \#\{y_m\}$. We will show $N < \infty$ and derive a useful quantitative estimate. Let $|A|$ denote the Lebesgue measure of a set $A \subset \mathbb{R}^n$. Notice that

$$N(\alpha^{-k}\ell(Q_0))^n = \sum_{m=1}^N |Q_{y_m}| \le C_n|2Q_0| = 2^n C_n \ell(Q_0)^n$$

so that

$$N \leq 2^n C_n \alpha^{kn}.$$

Using (2.6) and (2.7) we derive

$$\mu(Q_0 \cap F_j) \leq \sum_{m=1}^{N} \mu(Q_{y_m}) \leq N\beta^{-k}\mu(2Q_0) \leq 2^n C_n \Big(\frac{\alpha^n}{\beta}\Big)^k \mu(2Q_0) \to 0,$$

when $k \to \infty$ (since $\beta > \alpha^n$). Therefore $\mu(Q_0 \cap F_j) = 0$ and we are done. $\square$

2.6. **Cubes with small boundaries.** Given $t > 0$ we say that a cube $Q \subset \mathbb{R}^n$ has t-small boundary with respect to the measure μ if

$$\mu(\{x \in 2Q \colon \; \operatorname{dist}(x, \partial Q) \leq \lambda \ell(Q)\}) \leq t\lambda\mu(2Q)$$

for every $\lambda > 0$.

2.8. **Lemma.** *Let μ be a Radon measure in $\mathbb{R}^n$. Let $t > 0$ be some constant big enough (depending only on n). Then, given a cube $Q \subset \mathbb{R}^n$, there exists a concentric cube Q' so that $Q \subset Q' \subset 1.1Q$ which has t-small boundary with respect to μ.*

Proof. We may assume for simplicity that $c_Q = 0$ (where c_Q is the centre of Q). Given $a \in \mathbb{R}$ define the hyperplanes

$$H_j(a) = \{x \in \mathbb{R}^n \colon \; x_j = a\}, \qquad j = 1, \ldots, n,$$

and the ϵ-neighbourhoods

$$H_{j,\epsilon}(a) = \{x \in \mathbb{R}^n \colon \; \operatorname{dist}(x, H_j(a)) \leq \epsilon\}, \qquad \epsilon > 0.$$

Define the measure $\sigma = \mu\lfloor 2Q$ so that $\|\sigma\| = \mu(2Q)$. It is enough to find $a \in [\ell(Q), 1.05\ell(Q)]$ so that for *all* $j = 1, \ldots, n$ and $\lambda > 0$ we have

$$\frac{\sigma(H_{j,\lambda\ell(Q)}(a))}{\lambda\ell(Q)} \leq t\frac{\|\sigma\|}{\ell(Q)} \qquad \text{and} \qquad \frac{\sigma(H_{j,\lambda\ell(Q)}(-a))}{\lambda\ell(Q)} \leq t\frac{\|\sigma\|}{\ell(Q)}. \tag{2.9}$$

The existence of a is shown as follows. Define the projections $\pi_j(x) = x_j$ and $\tilde{\pi}_j(x) = -x_j$. We define the image measures $\nu_j = (\pi_j)_{\#}\sigma$ and $\tilde{\nu}_j = (\tilde{\pi}_j)_{\#}\sigma$ (here e.g. $\nu_j(A) = \sigma(\pi_j^{-1}A)$ for $A \subset \mathbb{R}$). Notice that $H_{j,\epsilon}(a) = \pi_j^{-1}I(a,\epsilon)$, where $I(a,\epsilon) = [a-\epsilon, a+\epsilon]$. Therefore, we have that $\nu_j(I(a,\lambda\ell(Q))) = \sigma(\pi_j^{-1}I(a,\lambda\ell(Q))) = \sigma(H_{j,\lambda\ell(Q)}(a))$. This means that (2.9) is equivalent to

$$\frac{\nu_j(I(a,\lambda\ell(Q)))}{\lambda\ell(Q)} \leq t\frac{\|\sigma\|}{\ell(Q)} \qquad \text{and} \qquad \frac{\tilde{\nu}_j(I(a,\lambda\ell(Q)))}{\lambda\ell(Q)} \leq t\frac{\|\sigma\|}{\ell(Q)}$$

This follows if

$$M\nu_j(a) \leq t\frac{\|\sigma\|}{2\ell(Q)} \qquad \text{and} \qquad M\tilde{\nu}_j(a) \leq t\frac{\|\sigma\|}{2\ell(Q)},$$

where M is the centred maximal function defined using the Lebesgue measure in $\mathbb{R}$.

Define the measure $\nu = \sum_{j=1}^n (\nu_j + \tilde{\nu}_j)$. Since $\|\nu\| = \|\tilde{\nu}\| = \|\sigma\|$ for all j, we have $\|\nu\| = 2n\|\sigma\|$. Therefore, it suffices that

$$M\nu(a) \le t\frac{\|\sigma\|}{2\ell(Q)} = t\frac{\|\nu\|}{4n\ell(Q)}.$$

Since M is bounded from $M(\mathbb{R}) \to L^{1,\infty}(m_1)$ we have for some $C < \infty$ that

$$m_1\Big(\Big\{a \in \mathbb{R}\colon\ M\nu(a) > t\frac{\|\nu\|}{4n\ell(Q)}\Big\}\Big) \le C\frac{4n\ell(Q)}{t\|\nu\|}\|\nu\| = \frac{4Cn}{t}\ell(Q) \le \frac{\ell(Q)}{100}$$

for all t large enough. We conclude that for all large enough t there must exist $a \in [\ell(Q), 1.05\ell(Q)]$ so that

$$M\nu(a) \le t\frac{\|\nu\|}{4n\ell(Q)}.$$

This completes the proof. $\square$

3. Calderón–Zygmund decomposition and weak $(1,1)$ boundedness of square functions

3.1. **Proposition** (Non-homogeneous Calderón–Zygmund decomposition). *Let μ be a Radon measure in $\mathbb{R}^n$. For every $\nu \in M(\mathbb{R}^n)$ with compact support and every $\lambda > 2^{n+1}\|\nu\|/\|\mu\|$, we have:*

(1) There exists a family of cubes $(Q_i)_i$ so that $\sum_i 1_{Q_i} \le C_n$ and a function $f \in L^1(\mu)$ such that

$$|\nu|(Q_i) > \frac{\lambda}{2^{n+1}}\mu(2Q_i), \tag{3.2}$$

$$|\nu|(\eta Q_i) \le \frac{\lambda}{2^{n+1}}\mu(2\eta Q_i) \text{ for } \eta > 2, \tag{3.3}$$

$$\nu = f\,d\mu \text{ in } \mathbb{R}^n \setminus \bigcup_i Q_i, \text{ with } |f| \le \lambda\,\mu\text{-a.e.} \tag{3.4}$$

(2) Suppose that for each i we are given a $(6,\beta_0)$-μ-doubling cube R_i such that it is concentric with Q_i and $Q_i \subset R_i$. For each i set

$$w_i = \frac{1_{Q_i}}{\sum_k 1_{Q_k}}.$$

Then there exists a family of functions (φ_i) (of the form $\varphi_i = \alpha_i h_i$ for some constant $\alpha_i \in \mathbb{C}$ and non-negative function $h_i \ge 0$) such that

$$\operatorname{spt}\varphi_i \subset R_i, \tag{3.5}$$

$$\int \varphi_i\,d\mu = \int w_i\,d\nu, \tag{3.6}$$

$$\sum_i |\varphi_i| \le B\lambda \ \ (B \text{ depends only on } \beta_0, n), \tag{3.7}$$

$$\|\varphi_i\|_{L^\infty(\mu)}\mu(R_i) \le 2|\nu|(Q_i). \tag{3.8}$$

Proof. Fix $\nu \in M(\mathbb{R}^n)$ and $\lambda > 2^{n+1}\|\nu\|/\|\mu\|$.

(1): Let

$$H = \Big\{x \in \operatorname{spt}\nu\colon \text{ there exists cube } Q \text{ centred at } x \text{ so that } |\nu|(Q) > \frac{\lambda}{2^{n+1}}\mu(2Q)\Big\}.$$

For $x \in H$ let

$$\ell(x) = \sup\Big\{\ell(Q)\colon\ Q \text{ is a cube centred at } x \text{ so that } |\nu|(Q) > \frac{\lambda}{2^{n+1}}\mu(2Q)\Big\}.$$

Let $\Lambda > 1$ be so that $\lambda \geq \Lambda 2^{n+1}\|\nu\|/\|\mu\|$. Given x there exists $C_x < \infty$ so that for all cubes Q centred at x for which $\ell(Q) \geq C_x$ we have

$$\frac{\mu(2Q)}{\|\mu\|} \geq \frac{1}{\Lambda}.$$

This means that for all such Q we have

$$|\nu|(Q) \leq \|\nu\| \leq \frac{\lambda}{2^{n+1}}\frac{1}{\Lambda}\|\mu\| \leq \frac{\lambda}{2^{n+1}}\mu(2Q).$$

This proves that for $x \in H$ we have $0 < \ell(x) \leq C_x < \infty$. For every $x \in H$ let Q_x be some cube centred at x so that $\ell(Q_x) \geq \ell(x)/2$ and

$$|\nu|(Q_x) > \frac{\lambda}{2^{n+1}}\mu(2Q_x).$$

Then for every cube Q centred at x for which $\ell(Q) > 2\ell(Q_x)$ we have $\ell(Q) > \ell(x)$ and so

$$|\nu|(Q) \leq \frac{\lambda}{2^{n+1}}\mu(2Q).$$

Notice that H is a bounded set, since $\operatorname{spt}\nu$ is compact. Therefore, we can use Besicovitch covering theorem to choose $\{Q_i\}_i \subset \{Q_x\}_{x\in H}$ so that

$$H \subset \bigcup_i Q_i \qquad \text{and} \qquad \sum_i 1_{Q_1} \leq C_n.$$

The cubes Q_i satisfy (3.2) and (3.3) by construction.

Next, we will prove (3.4). Let F consist of those $y \in \operatorname{spt}\nu$ for which there does not exist a sequence of $(2, 2^{n+1})$-$|\nu|$-doubling cubes centred at y with side length tending to 0. By Lemma 2.5 we have $|\nu|(F) = 0$. Suppose $x \in \operatorname{spt}\nu \setminus (H \cup F)$. Then choose a sequence P_k of $(2, 2^{n+1})$-$|\nu|$-doubling cubes centred at x so that $\ell(P_k) \to 0$ when $k \to \infty$. Since $x \notin H$ we have

$$|\nu|(P_k) \leq \frac{\lambda}{2^{n+1}}\mu(2P_k)$$

and so

$$|\nu|(2P_k) \leq 2^{n+1}|\nu|(P_k) \leq \lambda\mu(2P_k).$$

This yields

$$\liminf_{r\to 0}\frac{|\nu|(Q(x,r))}{\mu(Q(x,r))} \leq \lambda, \qquad x \in \operatorname{spt}\nu \setminus (H \cup F).$$

This implies that for every $A \subset (H \cup F)^c$ we have $|\nu|(A) \leq \lambda \mu(A)$. Since $|\nu|(F) = 0$, we have that $|\nu|\lfloor H^c \ll \mu$.

The Radon–Nikodym theorem guarantees the existence of $g \in L^1(\mu)$, $g \geq 0$, so that $|\nu|\lfloor H^c = g\, d\mu$. Let us see that $g(x) \leq \lambda$ for μ-a.e. x. Notice first that $g(x) = 0$ for μ-a.e. $x \in H \cup F$, since $0 = |\nu|\lfloor H^c(H \cup F) = \int_{H \cup F} g\, d\mu$. Define

$$A_j = \{x \in (H \cup F)^c \colon\ g(x) \geq \lambda + 1/j\}$$

and notice that

$$(\lambda + 1/j)\mu(A_j) \leq \int_{A_j} g\, d\mu = |\nu|(A_j) \leq \lambda \mu(A_j)$$

implying that $\mu(A_j) = 0$. It follows that $g(x) \leq \lambda$ for μ-a.e. x.

Let us then write $d\nu = h\, d|\nu|$ for a function h satisfying $|h| = 1$. Now $\nu\lfloor H^c = hg\, d\mu =: f\, d\mu$, and $|f| = |hg| = g \leq \lambda$ μ-a.e. Since $H \subset \bigcup_i Q_i$ this completes the proof of (3.4).

(2): Suppose R_i and w_i are like in the assumptions. Let us first assume that the family $(Q_i)_i$ is finite. Then we may assume it is enumerated so that $\ell(R_{i+1}) \geq \ell(R_i)$ for all i. The functions φ_i will be now inductively constructed, and they will have the form $\varphi_i = \alpha_i 1_{A_i}$ for some $\alpha_i \in \mathbb{C}$ and $A_i \subset R_i$.

We begin by setting $A_1 = R_1$ and $\varphi_1 = \alpha_1 1_{R_1}$, where α_1 is chosen so that

$$\int_{Q_1} w_1\, d\nu = \int \varphi_1\, d\mu \text{ i.e. } \alpha_1 = \frac{1}{\mu(R_1)} \int_{Q_1} w_1\, d\nu.$$

Notice that

$$|\varphi_1| \leq |\alpha_1| \leq \frac{|\nu|(Q_1)}{\mu(R_1)}$$

which readily yields

$$\|\varphi_1\|_{L^\infty(\mu)} \mu(R_1) \leq |\nu|(Q_1).$$

Moreover, since

$$|\nu|(Q_1) \leq |\nu|(3Q_1) \leq \frac{\lambda}{2^{n+1}} \mu(6Q_1) \leq \frac{\lambda}{2^{n+1}} \mu(6R_1) \leq \frac{\beta_0}{2^{n+1}} \mu(R_1) \cdot \lambda$$

so that

$$|\varphi_1| \leq C_1 \lambda.$$

We are done with the initial step of the induction. Suppose then that $\varphi_1, \ldots, \varphi_{k-1}$ have been constructed so that they are of the form $\varphi_i = \alpha_i 1_{A_i}$, $A_i \subset R_i$, satisfy (3.6), (3.8) and

$$\sum_{i=1}^{k-1} |\varphi_i| \leq B\lambda$$

for some B which will be fixed below.

Let $R_{s_1}, \ldots, R_{s_p}$ be the subfamily of $R_1, \ldots, R_{k-1}$ such that $R_{s_j} \cap R_k \neq \emptyset$. As $\ell(R_{s_j}) \leq \ell(R_k)$ we have $R_{s_j} \subset 3R_k$. We know for $i = 1, \ldots, k-1$ that

$$\int |\varphi_i| \, d\mu = \Big| \int \varphi_i \, d\mu \Big| = \Big| \int_{Q_i} w_i \, d\nu \Big| \leq |\nu|(Q_i).$$

Using this we get

$$\begin{aligned}\sum_j \int |\varphi_{s_j}| \, d\mu &\leq \sum_j |\nu|(Q_{s_j}) \\ &= \int_{3R_k} \sum_j 1_{Q_{s_j}} \, d|\nu| \\ &\leq C_n |\nu|(3R_k) \leq \frac{C_n}{2^{n+1}} \lambda \mu(6R_k) \leq \frac{C_n \beta_0}{2^{n+1}} \lambda \mu(R_k) =: C_2 \lambda \mu(R_k).\end{aligned}$$

Therefore, we see that

$$\mu\Big(\Big\{ \sum_j |\varphi_{s_j}| > 2C_2\lambda \Big\}\Big) \leq \frac{1}{2C_2\lambda} \sum_j \int |\varphi_{s_j}| \, d\mu \leq \frac{\mu(R_k)}{2}.$$

We now set

$$A_k = R_k \cap \Big\{ \sum_j |\varphi_{s_j}| \leq 2C_2\lambda \Big\}$$

so that $\mu(A_k) \geq \mu(R_k)/2$.

We choose the constant α_k so that

$$\int \varphi_k \, d\mu = \int_{Q_k} w_k \, d\nu \text{ i.e. } \alpha_k = \frac{1}{\mu(A_k)} \int_{Q_k} w_k \, d\nu.$$

We have

$$|\alpha_k| \leq 2 \frac{|\nu|(Q_k)}{\mu(R_k)} \leq 2C_1\lambda.$$

This gives that

$$|\varphi_k| + \sum_j |\varphi_{s_j}| \leq 2(C_1 + C_2)\lambda \qquad \text{in } A_k.$$

Letting $B = 2(C_1 + C_2)$ yields (3.7) by induction.

That (3.8) holds is easy:

$$\|\varphi_k\|_{L^\infty(\mu)} \mu(R_k) \leq 2|\alpha_k| \mu(A_k) = 2\Big| \int_{Q_k} w_k \, d\nu \Big| \leq 2|\nu|(Q_k).$$

We are done with the proof in the case that $(Q_i)_i$ is finite.

The general case (where $(Q_i)_i$ is not necessarily finite) follows from this by rather standard convergence arguments (but it does require some knowledge of topology). $\square$

3.9. **Example.** The previous Proposition is used as follows. Let μ be a measure of order m in $\mathbb{R}^n$, $\nu \in M(\mathbb{R}^n)$ with compact support and $\lambda > 2^{n+1}\|\nu\|/\|\mu\|$. Let the cubes $(Q_i)_i$ and the function $f \in L^1(\mu)$ be like in (1) of Proposition 3. For each i let R_i be the smallest $(6, 6^{m+1})$-μ-doubling cube of the form $6^k Q_i$, $k \geq 0$ (the existence of such a cube follows from the simple arguments used to prove Lemma 2.4). Then let w_i and φ_i be like in (2) of Proposition 3.

We begin by writing

$$\begin{aligned}
\nu &= \nu\lfloor\Big(\mathbb{R}^n \setminus \bigcup_i Q_i\Big) + \nu\lfloor\Big(\bigcup_i Q_i\Big) \\
&= f\,d\mu + \sum_i w_i\,d\nu \\
&= f\,d\mu + \sum_i \varphi_i\,d\mu + \sum_i (w_i\,d\nu - \varphi_i\,d\mu) \\
&=: g\,d\mu + \sum_i \beta_i,
\end{aligned}$$

where the function g is defined by

$$g = f + \sum_i \varphi_i$$

and the complex measure β_i is defined by

$$\beta_i = w_i\,d\nu - \varphi_i\,d\mu.$$

We now go through the important properties of this Calderón–Zygmund decomposition of ν with respect to the non-homogeneous measure μ.

The property of fundamental importance is the the following measure estimate:

$$\mu\Big(\bigcup_i 2Q_i\Big) \leq \sum_i \mu(2Q_i) \leq \frac{2^{n+1}}{\lambda}\sum_i |\nu(Q_i)| \leq \frac{2^{n+1}C_n}{\lambda}\|\nu\| = \frac{C_3}{\lambda}\|\nu\|. \tag{3.10}$$

We continue with the properties of the good function g. It holds that

$$\|g\|_{L^\infty(\mu)} \leq (1+B)\lambda$$

and

$$\begin{aligned}
\|g\|_{L^1(\mu)} \leq \int |f|\,d\mu + \sum_i \int |\varphi_i|\,d\mu &= |\nu|\Big(\mathbb{R}^n \setminus \bigcup_i Q_i\Big) + \sum_i \Big|\int \varphi_i\,d\mu\Big| \\
&\leq \|\nu\| + \sum_i \Big|\int_{Q_i} w_i\,d\nu\Big| \\
&\leq \|\nu\| + \sum_i |\nu|(Q_i) \leq (1+C_n)\|\nu\|.
\end{aligned}$$

Combining these bounds we get the following important $L^2(\mu)$ bound

$$\|g\|_{L^2(\mu)}^2 \le \|g\|_{L^\infty(\mu)}\|g\|_{L^1(\mu)} \le C_4\lambda\|\nu\|. \tag{3.11}$$

Regarding the complex measures β_i we have the following:

(a) $\operatorname{spt}\beta_i \subset R_i$,
(b) $\beta_i(R_i) = \int w_i\,d\nu - \int \varphi_i\,d\mu = 0$,
(c)

$$\|\beta_i\| \le \int |w_i|\,d\nu + \int |\varphi_i|\,d\mu = \int |w_i|\,d\nu + \Big|\int w_i\,d\nu\Big| \le 2|\nu|(Q_i).$$

The point that R_i is the *smallest* $(6, 6^{m+1})$-μ-doubling cube of the form 6^kQ_i, $k \ge 0$, is utilised as follows. It implies that

$$\int_{R_i\setminus Q_i} \frac{d\mu(x)}{|x - c_{Q_i}|^m} \le C_5. \tag{3.12}$$

This is an exercise.

Using the Calderón–Zygmund decomposition from the previous example we now show the following important result.

3.13. **Theorem** (Weak $(1,1)$ boundedness of square functions)**.** *Let μ be a measure of order m in $\mathbb{R}^n$, $(s_t)_{t>0}$ be an m-LP-family, and V be the corresponding vertical square function. Suppose that $V_\mu\colon L^2(\mu) \to L^2(\mu)$ boundedly. Then $V\colon M(\mathbb{R}^n) \to L^{1,\infty}(\mu)$ boundedly.*

Proof. We need to show that for some $C < \infty$ we have for every $\nu \in M(\mathbb{R}^n)$ and $\lambda > 0$ that

$$\mu(\{x\colon\ V\nu(x) > \lambda\}) \le C\frac{\|\nu\|}{\lambda}.$$

Let us first assume that ν has compact support. If $\lambda \le 2^{n+1}\|\nu\|/\|\mu\|$ we have the trivial estimate

$$\mu(\{x\colon\ V\nu(x) > \lambda\}) \le \|\mu\| \le 2^{n+1}\frac{\|\nu\|}{\lambda}.$$

So we may suppose that $\lambda > 2^{n+1}\|\nu\|/\|\mu\|$. Then we are in the position to perform the Calderón–Zygmund decomposition of ν with respect to μ like in Example (3.9):

$$\nu = g\,d\mu + \sum_i \beta_i$$

with all the properties (and Q_i, R_i, w_i, φ_i etc) exactly like above.

Using the subadditivity of V we see that

$$V\nu(x) \le V_\mu g(x) + \sum_i V\beta_i(x).$$

Therefore, it suffices to bound the terms

$$\mu(\{x\colon\ V_\mu g(x) > \lambda/2\}) \qquad \text{and} \qquad \mu\Big(\Big\{x\colon\ \sum_i V\beta_i(x) > \lambda/2\Big\}\Big)$$

Using that $V_\mu \colon L^2(\mu) \to L^2(\mu)$ boundedly and (3.11) we see that

$$\begin{aligned}\mu(\{x\colon\ V_\mu g(x) > \lambda/2\}) &\le \frac{4}{\lambda^2}\|V_\mu g\|_{L^2(\mu)}^2 \\ &\le \frac{4\|V_\mu\|_{L^2(\mu)\to L^2(\mu)}^2}{\lambda^2}\|g\|_{L^2(\mu)}^2 \le 4C_4\|V_\mu\|_{L^2(\mu)\to L^2(\mu)}^2\frac{\|\nu\|}{\lambda}.\end{aligned}$$

Because of (3.10) it only remains to prove that

$$\mu\Big(\Big\{x \in \mathbb{R}^n \setminus \bigcup_j 2Q_j\colon\ \sum_i V\beta_i(x) > \lambda/2\Big\}\Big) \le C\frac{\|\nu\|}{\lambda}.$$

We bound

$$\begin{aligned}\mu\Big(\Big\{x \in \mathbb{R}^n \setminus \bigcup_j 2Q_j\colon\ \sum_i V\beta_i(x) > \lambda/2\Big\}\Big) &\le \frac{2}{\lambda}\int_{\mathbb{R}^n\setminus\bigcup_j 2Q_j}\sum_i V\beta_i\,d\mu \\ &= \frac{2}{\lambda}\sum_i\int_{\mathbb{R}^n\setminus\bigcup_j 2Q_j} V\beta_i\,d\mu \\ &\le \frac{2}{\lambda}\sum_i\int_{\mathbb{R}^n\setminus 2Q_i} V\beta_i\,d\mu.\end{aligned}$$

Let us fix i for the moment. We will prove that

$$\int_{\mathbb{R}^n\setminus 2Q_i} V\beta_i\,d\mu \le C|\nu|(Q_i).$$

This will end the proof, since

$$\sum_i |\nu|(Q_i) \le C_n\|\nu\|.$$

Let us write

$$\int_{\mathbb{R}^n\setminus 2Q_i} V\beta_i\,d\mu = \int_{\mathbb{R}^n\setminus 4R_i} V\beta_i\,d\mu + \int_{4R_i\setminus 2Q_i} V\beta_i\,d\mu = I + II.$$

We estimate

$$II \le \int_{4R_i\setminus 2Q_i} V_\nu w_i\,d\mu + \int_{4R_i} V_\mu\varphi_i\,d\mu = II_a + II_b.$$

Using the $L^2(\mu)$ boundedness of V_μ we see that

$$\begin{aligned}II_b &\le \mu(4R_i)^{1/2}\Big(\int |V_\mu\varphi_i|^2\,d\mu\Big)^{1/2} \\ &\le 6^{(m+1)/2}\|V_\mu\|_{L^2(\mu)\to L^2(\mu)}\mu(R_i)^{1/2}\Big(\int_{R_i}|\varphi_i|^2\,d\mu\Big)^{1/2} \\ &\le 6^{(m+1)/2}\|V_\mu\|_{L^2(\mu)\to L^2(\mu)}\|\varphi_i\|_{L^\infty(\mu)}\mu(R_i) \\ &\le 2\cdot 6^{(m+1)/2}\|V_\mu\|_{L^2(\mu)\to L^2(\mu)}|\nu|(Q_i),\end{aligned}$$

where the last inequality used (3.8).

Let us then estimate II_a. It is an exercise to show that for every $x \in \mathbb{R}^n \setminus 2Q_i$ we have

$$V_\nu w_i(x) \le C \frac{|\nu|(Q_i)}{|x - c_{Q_i}|^m}.$$

This gives that

$$II_a \le C|\nu|(Q_i) \int_{4R_i \setminus Q_i} \frac{d\mu(x)}{|x - c_{Q_i}|^m}.$$

But notice that

$$\int_{4R_i \setminus Q_i} \frac{d\mu(x)}{|x - c_{Q_i}|^m} = \int_{4R_i \setminus R_i} \frac{d\mu(x)}{|x - c_{Q_i}|^m} + \int_{R_i \setminus Q_i} \frac{d\mu(x)}{|x - c_{Q_i}|^m}.$$

It is recorded in (3.12) that the second term is bounded by C_5. The first term is clearly bounded by a constant, since $\mu(4R_i) \le C\ell(R)^m$ and $|x - c_{Q_i}|^m = |x - c_{R_i}|^m \ge c\ell(R)^m$ for $x \notin R_i$. This yields that

$$II_a \le C|\nu|(Q_i).$$

Combining what was done above we have shown that

$$II \le C|\nu|(Q_i).$$

It remains to show

$$I = \int_{\mathbb{R}^n \setminus 4R_i} V\beta_i \, d\mu \le C|\nu|(Q_i).$$

For all $x \in \mathbb{R}^n \setminus 4R_i$ we will show that

$$V\beta_i(x) \le C\Big(\frac{\ell(R_i)^\alpha}{|x - c_{R_i}|^{m+\alpha}} + \frac{\ell(R_i)^{\alpha/2}}{|x - c_{R_i}|^{m+\alpha/2}}\Big)|\nu|(Q_i). \tag{3.14}$$

This is enough, since

$$\int_{\mathbb{R}^n \setminus B(x_0,r)} \frac{d\mu(x)}{|x - x_0|^{m+\epsilon}} \le C_\epsilon r^{-\epsilon}, \qquad x_0 \in \mathbb{R}^n, r > 0, \epsilon > 0.$$

So we fix $x \in \mathbb{R}^n \setminus (4R_i)$, and estimate

$$\begin{aligned} V\beta_i(x) &\le \Big(\int_0^{2\,\mathrm{diam}(R_i)} |\theta_t\beta_i(x)|^2 \frac{dt}{t}\Big)^{1/2} \\ &\quad + \Big(\int_{2\,\mathrm{diam}(R_i)}^{|x-c_{R_i}|} |\theta_t\beta_i(x)|^2 \frac{dt}{t}\Big)^{1/2} \\ &\quad + \Big(\int_{|x-c_{R_i}|}^{\infty} |\theta_t\beta_i(x)|^2 \frac{dt}{t}\Big)^{1/2} = A_1 + A_2 + A_3. \end{aligned}$$

In A_3 we use all of the properties of β_i listed in Example 3.9. First, since $\mathrm{spt}\,\beta_i \subset R_i$ and $\beta_i(R_i) = 0$ we may estimate

$$|\theta_t\beta_i(x)| = \Big|\int_{R_i} [s_t(x,y) - s_t(x, c_{R_i})]\, d\beta_i(y)\Big| \le \int_{R_i} |s_t(x,y) - s_t(x,c_{R_i})|\, d|\beta_i|(y).$$

Recall that in A_3 we have $t \geq |x - c_{R_i}| > 2\,\mathrm{diam}(R_i)$ (since $x \notin 4R_i$). Therefore, we have for $y \in R_i$ that

$$|y - c_{R_i}| \leq \mathrm{diam}(R_i) < t/2.$$

Using (1.2) we then get

$$|\theta_t \beta_i(x)| \leq C \frac{\ell(R_i)^\alpha}{t^{m+\alpha}} \|\beta_i\|.$$

Recalling from Example 3.9 that $\|\beta_i\| \leq 2|\nu|(Q_i)$ we now get

$$A_3 \leq C|\nu|(Q_i)\ell(R_i)^\alpha \Big(\int_{|x-c_{R_i}|}^\infty t^{-2m-2\alpha-1}\,dt \Big)^{1/2} \leq C \frac{\ell(R_i)^\alpha}{|x - c_{R_i}|^{m+\alpha}} |\nu|(Q_i).$$

In A_2 we use again the y-Hölder-continuity argument from above. This time we estimate

$$|s_t(x,y) - s_t(x, c_{R_i})| \leq C \frac{\ell(R_i)^\alpha}{|x - c_{R_i}|^{m+\alpha}} \leq C \frac{\ell(R_i)^{\alpha/2} t^{\alpha/2}}{|x - c_{R_i}|^{m+\alpha}}$$

yielding

$$|\theta_t \beta_i(x)| \leq C \frac{\ell(R_i)^{\alpha/2} t^{\alpha/2}}{|x - c_{R_i}|^{m+\alpha}} |\nu|(Q_i).$$

From this we get

$$A_2 \leq C \frac{\ell(R_i)^{\alpha/2}}{|x - c_{R_i}|^{m+\alpha}} |\nu|(Q_i) \Big(\int_0^{|x-c_{R_i}|} t^{\alpha-1}\,dt \Big)^{1/2} \leq C \frac{\ell(R_i)^{\alpha/2}}{|x - c_{R_i}|^{m+\alpha/2}} |\nu|(Q_i).$$

Finally, we estimate A_1. Here we use (1.1) to the effect that

$$|\theta_t \beta_i(x)| \leq C t^\alpha \int_{R_i} \frac{d|\beta_i|(y)}{|x-y|^{m+\alpha}} \leq C \frac{t^\alpha}{|x - c_{R_i}|^{m+\alpha}} |\nu|(Q_i).$$

Using this we see

$$A_1 \leq C \frac{1}{|x - c_{R_i}|^{m+\alpha}} |\nu|(Q_i) \Big(\int_0^{2\,\mathrm{diam}(R_i)} t^{2\alpha-1}\,dt \Big)^{1/2} \leq C \frac{\ell(R_i)^\alpha}{|x - c_{R_i}|^{m+\alpha}} |\nu|(Q_i).$$

Combining the bounds for A_i, $i = 1, 2, 3$, we have proved (3.14). This completes the proof in the case that ν has compact support.

Let us now prove the result in full generality. So assume $\nu \in M(\mathbb{R}^n)$ (not necessarily of compact support) and $\lambda > 0$. Assume first that μ has compact support, and choose N_0 so that $\mathrm{spt}\,\mu \subset B(0, N_0)$. Define the compactly supported measure $\nu_N = \nu\lfloor B(0,N)$, $N > N_0$. Using (1.1) we have for every $x \in \mathrm{spt}\,\mu \subset$

$B(0, N_0)$ that

$$
\begin{aligned}
V(\nu - \nu_N)(x) &\leq \Big(\int_0^{N-N_0} |\theta_t(\nu - \nu_N)(x)|^2 \frac{dt}{t} \Big)^{1/2} + \Big(\int_{N-N_0}^{\infty} |\theta_t(\nu - \nu_N)(x)|^2 \frac{dt}{t} \Big)^{1/2} \\
&\leq \frac{C\|\nu\|}{(N-N_0)^{m+\alpha}} \Big(\int_0^{N-N_0} t^{2\alpha - 1}\, dt \Big)^{1/2} + C\|\nu\| \Big(\int_{N-N_0}^{\infty} t^{-2m-1}\, dt \Big)^{1/2} \\
&\leq \frac{C\|\nu\|}{(N-N_0)^m} \leq \frac{\lambda}{2}
\end{aligned}
$$

fixing N large enough. So if $x \in \operatorname{spt} \mu$ is such that $V\nu(x) > \lambda$ then $V\nu_N(x) > \lambda/2$. Therefore, we have

$$
\mu(\{V\nu > \lambda\}) \leq \mu(\{V\nu_N > \lambda/2\}) \leq C\frac{\|\nu_N\|}{\lambda} \leq C\frac{\|\nu\|}{\lambda}
$$

using the fact that ν_N has compact support.

Disposing of the assumption that μ has compact support is trivial. Define $\mu_N = \mu\lfloor B(0, N)$, $N > 0$. This is a compactly supported measure of order m (uniformly in N). Since V_μ is bounded in $L^2(\mu)$, we see that V_{μ_N} is bounded in $L^2(\mu_N)$ uniformly in N. But then we can conclude that

$$
\mu_N(\{V\nu > \lambda\}) \leq C\frac{\|\nu\|}{\lambda}
$$

uniformly in N. Letting $N \to \infty$ gives the claim. We have proved the result in full generality. □

4. THE NON-HOMOGENEOUS GOOD LAMBDA METHOD

4.1. **Theorem** (The non-homogeneous good lambda method)**.** *Let μ be a measure of order m in $\mathbb{R}^n$, $(s_t)_{t>0}$ be an x-continuous m-LP-family, and V be the corresponding vertical square function. Let $\beta > 0$ and $C_1 > 0$ be big enough numbers, depending only on the dimension n, and assume $\theta \in (0,1)$. Suppose for each $(2,\beta)$-doubling cube Q with C_1-small boundary there exists a subset $G_Q \subset Q$ such that $\mu(G_Q) \geq \theta\mu(Q)$ and $V\colon M(\mathbb{R}^n) \to L^{1,\infty}(\mu\lfloor G_Q)$ is bounded with a uniform constant independent of Q. Then V_μ is bounded in $L^p(\mu)$ for all $1 < p < \infty$ with a constant depending on p and on the preceding constants.*

4.2. *Remark.* One can also assume that $V_{\mu\lfloor G_Q}\colon L^2(\mu\lfloor G_Q) \to L^2(\mu\lfloor G_Q)$ with norm bounded uniformly on Q, since then V is bounded from $M(\mathbb{R}^n)$ to $L^{1,\infty}(\mu\lfloor G_Q)$ by Theorem 3.13. This will be important for us later.

To prove Theorem 4.1 we will use a Whitney's decomposition of some open set. In the next lemma we show the precise version of the required decomposition.

4.3. **Lemma.** *If $\Omega \subset \mathbb{R}^n$ is open, $\Omega \neq \mathbb{R}^n$, then Ω can be decomposed as*

$$
\Omega = \bigcup_{i \in I} Q_i,
$$

where Q_i, $i \in I$, are closed dyadic cubes with disjoint interiors such that for some constants $R > 20$ and $D_0 \geq 1$ depending only on n the following holds:

(i) $10Q_i \subset \Omega$ *for each* $i \in I$.
(ii) $RQ_i \cap \Omega^c \neq \varnothing$ *for each* $i \in I$.
(iii) *For each cube Q_i, there are at most D_0 cubes Q_j such that $10Q_i \cap 10Q_j \neq \varnothing$. Further, for such cubes Q_i, Q_j, we have $\ell(Q_i) \approx \ell(Q_j)$.*

Moreover, if μ is a positive Radon measure on $\mathbb{R}^n$ and $\mu(\Omega) < \infty$, there is a family of cubes $\{\widetilde{Q}_j\}_{j\in S}$, with $S \subset I$, so that $Q_j \subset \widetilde{Q}_j \subset 1.1Q_j$, satisfying the following:

(a) *Each cube $\widetilde{Q}_j$, $j \in S$, is $(9, 2D_0)$-doubling and has C_1-small boundary.*
(b) *The cubes $\widetilde{Q}_j$, $j \in S$, are pairwise disjoint.*
(c)

$$\mu\Big(\bigcup_{j\in S} \widetilde{Q}_j\Big) \geq \frac{1}{8D_0}\,\mu(\Omega). \tag{4.4}$$

Proof. A Whitney decomposition using dyadic cubes satisfying (i), (ii) and (iii) is easy. To this end, let $\mathcal{D}_0$ be the collection of standard dyadic cubes in $\mathbb{R}^n$ i.e.

$$\mathcal{D}_0 = \bigcup_{k\in\mathbb{Z}} \mathcal{D}_0^k, \qquad \mathcal{D}_0^k = \{2^{-k}([0,1)^n + m) \colon\ m \in \mathbb{Z}^n\}.$$

The main properties of dyadic cubes are that each $\mathcal{D}_0^k$ is a partition (a pairwise disjoint cover) of $\mathbb{R}^n$ and each $\mathcal{D}_0^{k+1}$ is a refinement of the previous $\mathcal{D}_0^k$ (cubes from $\mathcal{D}_0^k$ can be written as the disjoint union of their 2^n dyadic children from $\mathcal{D}_0^{k+1}$). Moreover, maximal dyadic cubes with respect to some property are always disjoint. Now, choose maximal dyadic cubes $Q \in \mathcal{D}_0$ so that

$$10\overline{Q} \subset \Omega.$$

It is almost immediate that the cubes $\overline{Q}$ can be taken to be the cubes Q_i, $i \in I$. Details are left as an exercise.

To prove the existence of the family $\{\widetilde{Q}_j\}_{j\in S}$, we denote by $I_{db} \subset I$ the subfamily of the indices such that the cubes from $\{Q_i\}_{i\in I_{db}}$ are $(10, 2D_0)$-doubling. Then notice that

$$\mu(Q_j) \leq \frac{1}{2D_0}\mu(10Q_j) \qquad \text{if } j \in I \setminus I_{db}.$$

Since

$$\sum_{j\in I} 1_{10Q_j} \leq D_0 1_\Omega,$$

we deduce that

$$\sum_{j\in I\setminus I_{db}} \mu(Q_j) \leq \frac{1}{2D_0}\sum_{j\in I}\mu(10Q_j) \leq \frac{1}{2}\mu(\Omega).$$

Thus,

$$\mu\Big(\bigcup_{j\in I_{db}} Q_j\Big) \geq \mu(\Omega) - \sum_{j\in I\setminus I_{db}} \mu(Q_j) \geq \frac{1}{2}\,\mu(\Omega). \tag{4.5}$$

Choose a finite subset $I_{db}^1 \subset I_{db}$ so that

$$\mu\Big(\bigcup_{j\in I_{db}} Q_j\Big) \leq 2\mu\Big(\bigcup_{j\in I_{db}^1} Q_j\Big)$$

Since I_{db}^1 is finite there are no problems with using the $5r$-covering theorem: there exists $S \subset I_{db}^1$ so that

$$\bigcup_{j\in I_{db}^1} Q_j \subset \bigcup_{j\in I_{db}^1} 2Q_j \subset \bigcup_{j\in S} 10Q_j$$

and so that the cubes $2Q_j$, $j \in S$, are pairwise disjoint. For each $j \in S$, we consider a cube $\widetilde{Q}_j$ with C_1-small boundary so that $Q_j \subset \widetilde{Q}_j \subset 1.1Q_j$. The existence of these cubes is guaranteed by Lemma 2.8. It is clear that the cubes $\widetilde{Q}_j$, $j \in S$, are pairwise disjoint since the cubes $2Q_j$, $j \in S$, are pairwise disjoint.

Next, notice that

$$\mu(9\widetilde{Q}_j) \leq \mu(10Q_j) \leq 2D_0\,\mu(Q_j) \leq 2D_0\,\mu(\widetilde{Q}_j),$$

since $S \subset I_{db}$. Therefore, each cube $\widetilde{Q}_j$, $j \in S$, is $(9, 2D_0)$-doubling and has C_1-small boundary. It remains to prove the measure estimate (c):

$$\begin{aligned}\mu(\Omega) &\leq 2\,\mu\Big(\bigcup_{j\in I_{db}} Q_j\Big) \leq 4\mu\Big(\bigcup_{j\in I_{db}^1} Q_j\Big) \leq 4\mu\Big(\bigcup_{j\in S} 10Q_j\Big)\\ &\leq 4\sum_{j\in S}\mu(10Q_j) \leq 8D_0\sum_{j\in S}\mu(Q_j) \leq 8D_0\sum_{j\in S}\mu(\widetilde{Q}_j) = 8D_0\mu\Big(\bigcup_{j\in S}\widetilde{Q}_j\Big).\end{aligned}$$

□

Proof of Theorem 4.1. For technical reasons we will consider the t_0-truncated, $t_0 > 0$, operators

$$V_{\mu,t_0}f(x) := \Big(\int_{t_0}^{\infty} |\theta_t^\mu f(x)|^2\frac{dt}{t}\Big)^{1/2}.$$

It is clearly enough to prove that V_{μ,t_0} is bounded in $L^p(\mu)$ uniformly in $t_0 > 0$. Fix $p \in (1,\infty)$ and consider $f \in L^p(\mu)$.

The first useful fact is that $x \mapsto V_{\mu,t_0}f(x)$ is continuous. Let us prove this. Notice that

$$V_{\mu,t_0}f(x) = \|t \mapsto \theta_t^\mu f(x)\|_{L^2((t_0,\infty),dt/t)}$$

so that using the triangle inequality we have

$$|V_{\mu,t_0}f(x) - V_{\mu,t_0}f(y)| \leq \Big(\int_{t_0}^{\infty} |\theta_t^\mu f(x) - \theta_t^\mu f(y)|^2\frac{dt}{t}\Big)^{1/2}.$$

If $|x-y|<t_0/2$ we can use (1.3) to conclude for every $t \ge t_0$ that

$$\begin{aligned}
|\theta_t^\mu f(x) - \theta_t^\mu f(y)| &\le \int_{\mathbb{R}^n} |s_t(x,z) - s_t(y,z)|\,|f(z)|\,d\mu(z) \\
&\le C\int_{\mathbb{R}^n} \frac{|x-y|^\alpha}{(t+|x-z|)^{m+\alpha}}\,|f(z)|\,d\mu(z) \\
&\le C\frac{|x-y|^\alpha}{t^{\alpha/2}} \int_{\mathbb{R}^n} \frac{|f(z)|}{(t_0+|x-z|)^{m+\alpha/2}}\,d\mu(z) \\
&\le C\frac{|x-y|^\alpha}{t^{\alpha/2}} \|f\|_{L^p(\mu)} \Big(\int_{\mathbb{R}^n} \frac{d\mu(z)}{(t_0+|x-z|)^{(m+\alpha/2)p'}} \Big)^{1/p'} \\
&\le C(t_0)\frac{|x-y|^\alpha}{t^{\alpha/2}} \|f\|_{L^p(\mu)}.
\end{aligned}$$

From this it follows that for $|x-y|<t_0/2$ we have

$$|V_{\mu,t_0} f(x) - V_{\mu,t_0} f(y)| \le C(t_0)|x-y|^\alpha \|f\|_{L^p(\mu)},$$

since $t^{-\alpha-1}$ can be integrated over (t_0,∞) to get another big dependence on t_0. Therefore, continuity follows. This is used purely qualitatively to conclude that the sets

$$\Omega_\lambda = \Omega_{\lambda,t_0} := \{V_{\mu,t_0} f > \lambda\}, \qquad \lambda > 0,$$

are open.

For our argument we will also need the a priori information $\Omega_\lambda \ne \mathbb{R}^n$ and $\|V_{\mu,t_0} f\|_{L^p(\mu)} < \infty$ (then also $\mu(\Omega_\lambda) < \infty$). We can achieve this by assuming that f is compactly supported and bounded (such functions are dense). Let $R>0$ be so that $\operatorname{spt} f \subset B(0,R)$. Notice that for $x \in \mathbb{R}^n$ and $t \ge t_0$ we have

$$\begin{aligned}
|\theta_t^\mu f(x)| &\le C(f)\int_{B(0,R)} \frac{t^\alpha}{(t+|x-y|)^{m+\alpha}}\,d\mu(y) \\
&\le C(f)\int_{B(0,R)} \frac{1}{(t+|x-y|)^{m}}\,d\mu(y) \\
&\le C(f)\frac{R^m}{(t+\operatorname{dist}(x,B(0,R)))^m} \\
&\le C(f)R^m \frac{1}{(t_0+\operatorname{dist}(x,B(0,R)))^{m-\gamma}}\frac{1}{t^\gamma},
\end{aligned}$$

where we used some very small auxiliary parameter $\gamma > 0$. For the previous calculation it only matters that $m-\gamma>0$, but we will take it so small that $p(m-\gamma) > m$. Integrating $t^{-2\gamma-1}$ over (t_0,∞) it follows that

$$V_{\mu,t_0} f(x) \le C(f,t_0)R^m \frac{1}{(t_0+\operatorname{dist}(x,B(0,R)))^{m-\gamma}}.$$

From here we can read that $V_{\mu,t_0} f(x) \to 0$, when $|x| \to \infty$. In particular, Ω_λ is contained in some large ball and therefore not equal to $\mathbb{R}^n$, and $\mu(\Omega_\lambda) < \infty$. Also

$$\int_{\mathbb{R}^n} \frac{d\mu(x)}{(t_0 + \mathrm{dist}(x, B(0,R)))^{p(m-\gamma)}} < \infty$$

since $p(m-\gamma) > m$. This gives that $\|V_{\mu,t_0} f\|_{L^p(\mu)} < \infty$.

We are done with the technical preliminaries. Since $\Omega_\lambda \neq \mathbb{R}^n$ is open and satisfies $\mu(\Omega_\lambda) < \infty$ we can use Lemma 4.3 to write

$$\Omega_\lambda = \bigcup_{i \in I} Q_i,$$

and also to extract $\{\widetilde{Q}_j\}_{j \in S}$, where $S \subset I$, so that all the properties of the lemma hold. For $j \in S$ let us write $P_j = \widetilde{Q}_j$. The cubes P_j have C_1-small boundary and are $(9, 2D_0)$-doubling, in particular $(2, 2D_0)$-doubling. So assuming that the parameter β from the assumptions is larger than $2D_0$, we have by assumption that there exists $G_{P_j} \subset P_j$ so that $\mu(G_{P_j}) \geq \theta\mu(P_j)$ and $V\colon\ M(\mathbb{R}^n) \to L^{1,\infty}(\mu\lfloor G_{P_j})$ boundedly with a constant A which is uniform in $j \in S$. For $j \in S$ denote $G_j = G_{P_j}$.

The idea is to prove using the previous cubes that given $\epsilon, \lambda > 0$ there exists $\delta = \delta(\epsilon, \theta, A) = \delta(\epsilon) > 0$ (θ and A are fixed constants from the assumptions) so that

$$\mu(\{x \in \mathbb{R}^n\colon\ V_{\mu,t_0} f(x) > (1+\epsilon)\lambda,\ M_\mu^{\mathcal{Q}} f(x) \leq \delta\lambda\}) \leq \Big(1 - \frac{\theta}{16D_0}\Big)\mu(\Omega_\lambda). \tag{4.6}$$

Let us show why this implies that V_{μ,t_0} maps $L^p(\mu)$ to $L^p(\mu)$ boundedly. So assume we have proved (4.6). Begin by fixing $\epsilon = \epsilon(\theta, p) > 0$ so that

$$(1+\epsilon)^p\Big(1 - \frac{\theta}{16D_0}\Big) = 1 - \frac{\theta}{32D_0}.$$

Then let $\delta = \delta(\epsilon) > 0$ be so that (4.6) holds. We have

$$\begin{aligned}
&\mu(\{x\colon\ V_{\mu,t_0} f(x) > (1+\epsilon)\lambda\}) \\
&\leq \mu(\{x\colon\ V_{\mu,t_0} f(x) > (1+\epsilon)\lambda,\ M_\mu^{\mathcal{Q}} f(x) \leq \delta\lambda\}) + \mu(\{x\colon\ M_\mu^{\mathcal{Q}} f(x) > \delta\lambda\}) \\
&\leq \Big(1 - \frac{\theta}{16D_0}\Big)\mu(\{x\colon\ V_{\mu,t_0} f(x) > \lambda\}) + \mu(\{x\colon\ M_\mu^{\mathcal{Q}} f(x) > \delta\lambda\}).
\end{aligned}$$

Recall for $g \geq 0$ the formula

$$\int g^p\, d\mu = p\int_0^\infty \lambda^{p-1}\mu(\{x\colon\ g(x) > \lambda\})\, d\lambda$$

and notice that for $s > 0$ we have

$$\int_0^\infty \lambda^{p-1}\mu(\{x\colon\ g(x) > s\lambda\})\, d\lambda = \frac{1}{s^p}\int_0^\infty \lambda^{p-1}\mu(\{x\colon\ g(x) > \lambda\})\, d\lambda.$$

Therefore, we get by multiplying by $p\lambda^{p-1}$ and integrating over $\lambda \in (0,\infty)$ that
$$\frac{1}{(1+\epsilon)^p}\|V_{\mu,t_0}f\|^p_{L^p(\mu)} \le \Big(1-\frac{\theta}{16D_0}\Big)\|V_{\mu,t_0}f\|^p_{L^p(\mu)} + \frac{1}{\delta^p}\|M^{\mathcal{Q}}_\mu f\|^p_{L^p(\mu)}.$$
which implies that
$$\|V_{\mu,t_0}f\|^p_{L^p(\mu)} \le \Big(1-\frac{\theta}{32D_0}\Big)\|V_{\mu,t_0}f\|^p_{L^p(\mu)} + C\|f\|^p_{L^p(\mu)}.$$
Since $\|V_{\mu,t_0}f\|_{L^p(\mu)} < \infty$ this implies the desired quantitative bound.

So it only remains to prove the good lambda inequality (4.6). Notice that
$$\{V_{\mu,t_0}f > (1+\epsilon)\lambda,\, M^{\mathcal{Q}}_\mu f \le \delta\lambda\} \subset \Omega_\lambda = \Big[\Omega_\lambda \setminus \bigcup_{j\in S} P_j\Big] \cup \bigcup_{j\in S} P_j.$$
Therefore, we have
$$\begin{aligned}
&\mu(\{x\in\mathbb{R}^n\colon\, V_{\mu,t_0}f(x) > (1+\epsilon)\lambda,\, M^{\mathcal{Q}}_\mu f(x) \le \delta\lambda\}) \\
&\le \mu\Big(\Omega_\lambda \setminus \bigcup_{j\in S} P_j\Big) + \sum_{j\in S}\mu(\{x\in P_j\colon\, V_{\mu,t_0}f(x) > (1+\epsilon)\lambda,\, M^{\mathcal{Q}}_\mu f(x) \le \delta\lambda\}) \\
&\le \mu\Big(\Omega_\lambda \setminus \bigcup_{j\in S} P_j\Big) + \sum_{j\in S}\mu(P_j\setminus G_j) \\
&\qquad + \sum_{j\in S}\mu(\{x\in G_j\colon\, V_{\mu,t_0}f(x) > (1+\epsilon)\lambda,\, M^{\mathcal{Q}}_\mu f(x) \le \delta\lambda\}).
\end{aligned}$$
Notice that
$$\sum_{j\in S}\mu(P_j\setminus G_j) \le (1-\theta)\sum_{j\in S}\mu(P_j) = (1-\theta)\mu\Big(\bigcup_{j\in S}P_j\Big)$$
recalling that the cubes P_j are disjoint. Therefore, by also using that
$$\mu\Big(\bigcup_{j\in S}P_j\Big) \ge \frac{1}{8D_0}\mu(\Omega_\lambda)$$
we have
$$\begin{aligned}
\mu\Big(\Omega_\lambda \setminus \bigcup_{j\in S} P_j\Big) + \sum_{j\in S}\mu(P_j\setminus G_j) &\le \mu(\Omega_\lambda) - \theta\mu\Big(\bigcup_{j\in S}P_j\Big) \\
&\le \Big(1-\frac{\theta}{8D_0}\Big)\mu(\Omega_\lambda).
\end{aligned}$$
We are reduced to proving that
$$\sum_{j\in S}\mu(\{x\in G_j\colon\, V_{\mu,t_0}f(x) > (1+\epsilon)\lambda,\, M^{\mathcal{Q}}_\mu f(x) \le \delta\lambda\}) \le \frac{\theta}{16D_0}\mu(\Omega_\lambda) \tag{4.7}$$
if $\delta = \delta(\epsilon) > 0$ is small enough.

We will do this by showing that if $x\in P_j$ is such that
$$V_{\mu,t_0}f(x) > (1+\epsilon)\lambda \qquad \text{and} \qquad M^{\mathcal{Q}}_\mu f(x) \le \delta\lambda$$

then

$$V_{\mu,t_0}(f1_{2P_j})(x) > \frac{\epsilon}{2}\lambda,$$

in particular $V_\mu(f1_{2P_j})(x) > \frac{\epsilon}{2}\lambda$. Let us show how the claim follows by assuming this for the moment. Then it holds that

$$\begin{aligned}
\mu(\{x \in G_j\colon\ &V_{\mu,t_0}f(x) > (1+\epsilon)\lambda,\ M_\mu^{\mathcal{Q}}f(x) \le \delta\lambda\}) \\
&\le \mu\Big(\Big\{x \in G_j\colon\ V_\mu(f1_{2P_j})(x) > \frac{\epsilon}{2}\lambda\Big\}\Big) \\
&= (\mu\lfloor G_j)\Big(\Big\{x\colon\ V(f1_{2P_j}\,d\mu)(x) > \frac{\epsilon}{2}\lambda\Big\}\Big) \\
&\le A\frac{2}{\epsilon\lambda}\|f1_{2P_j}\|_{L^1(\mu)}.
\end{aligned}$$

Let us estimate this further. For this we can clearly assume that there exists $x_0 \in P_j$ such that $M_\mu^{\mathcal{Q}}f(x_0) \le \delta\lambda$. Notice that $2P_j \subset Q_{x_0} \subset 10Q_j$, where Q_{x_0} is a cube centred at x_0 with diameter $4d(P_j)$. Using this we have

$$\int_{2P_j} |f|\,d\mu \le \int_{Q_{x_0}} |f|\,d\mu \le \mu(Q_{x_0})M_\mu^{\mathcal{Q}}f(x_0) \le \mu(10Q_j)\delta\lambda.$$

This means that

$$\mu(\{x \in G_j\colon\ V_{\mu,t_0}f(x) > (1+\epsilon)\lambda,\ M_\mu^{\mathcal{Q}}f(x) \le \delta\lambda\}) \le \frac{2A\delta}{\epsilon}\mu(10Q_j),$$

and so

$$\begin{aligned}
\sum_{j\in S} \mu(\{x \in G_j\colon\ &V_{\mu,t_0}f(x) > (1+\epsilon)\lambda,\ M_\mu^{\mathcal{Q}}f(x) \le \delta\lambda\}) \\
&\le \frac{2A\delta}{\epsilon}\sum_{j\in I}\mu(10Q_j) \le \frac{2D_0A\delta}{\epsilon}\mu(\Omega_\lambda) \le \frac{\theta}{16D_0}\mu(\Omega_\lambda)
\end{aligned}$$

by assuming that

$$\delta \le \frac{\theta\epsilon}{32D_0^2A}.$$

We have shown that (4.7) follows.

Let now $x \in P_j$ be such that $V_{\mu,t_0}f(x) > (1+\epsilon)\lambda$ and $M_\mu^{\mathcal{Q}}f(x) \le \delta\lambda$. We have

$$(1+\epsilon)\lambda < V_{\mu,t_0}f(x) \le V_{\mu,t_0}(f1_{2P_j})(x) + V_{\mu,t_0}(f1_{(2P_j)^c})(x)$$

and so

$$V_{\mu,t_0}(f1_{2P_j})(x) > (1+\epsilon)\lambda - V_{\mu,t_0}(f1_{(2P_j)^c})(x).$$

It is enough to show that

$$V_{\mu,t_0}(f1_{(2P_j)^c})(x) \le \Big(1+\frac{\epsilon}{2}\Big)\lambda.$$

For convenience we assume that $t_0 < 2Rd(P_j)$ – the opposite case is the same calculation except easier, since there we do not need to split the integration in the t variable. We have

$$V_{\mu,t_0}(f1_{(2P_j)^c})(x) \le \Big(\int_{t_0}^{2Rd(P_j)} |\theta_t^\mu (f1_{(2P_j)^c})(x)|^2 \, \frac{dt}{t} \Big)^{1/2} + \Big(\int_{2Rd(P_j)}^{\infty} |\theta_t^\mu (f1_{(2P_j)^c})(x)|^2 \, \frac{dt}{t} \Big)^{1/2} = I + II.$$

Notice that for all $t > 0$ there holds that

$$\begin{aligned} |\theta_t^\mu (f1_{(2P_j)^c})(x)| &\le Ct^\alpha \int_{(2P_j)^c} \frac{|f(y)|}{|x-y|^{m+\alpha}} \, d\mu(y) \\ &\le Ct^\alpha \int_{\mathbb{R}^n \setminus B(x, c\ell(P_j))} \frac{|f(y)|}{|x-y|^{m+\alpha}} \, d\mu(y) \\ &= Ct^\alpha \sum_{j=0}^{\infty} \int_{2^j c\ell(P_j) \le |x-y| < 2^{j+1} c\ell(P_j)} \frac{|f(y)|}{|x-y|^{m+\alpha}} \, d\mu(y) \\ &\le Ct^\alpha \sum_{j=0}^{\infty} (2^j \ell(P_j))^{-m-\alpha} \int_{Q(x, C2^{j+1}\ell(P_j))} |f(y)| \, d\mu(y) \\ &\le Ct^\alpha M_\mu^{\mathcal{Q}} f(x) \sum_{j=0}^{\infty} (2^j \ell(P_j))^{-\alpha} \le C \frac{t^\alpha}{\ell(P_j)^\alpha} \delta\lambda. \end{aligned}$$

This gives that

$$I \le \frac{C\delta\lambda}{\ell(P_j)^\alpha} \Big(\int_0^{2Rd(P_j)} t^{2\alpha - 1} \, dt \Big)^{1/2} \le C\delta\lambda.$$

Recall that $RQ_j \cap \Omega_\lambda^c \ne \emptyset$ and so $RP_j \cap \Omega_\lambda^c \ne \emptyset$. Using this we fix $z \in RP_j$ so that $V_{\mu,t_0} f(z) \le \lambda$. We now estimate

$$II \le \Big(\int_{2Rd(P_j)}^{\infty} |\theta_t^\mu (f1_{(2P_j)^c})(x) - \theta_t^\mu (f1_{(2P_j)^c})(z)|^2 \, \frac{dt}{t} \Big)^{1/2} + \Big(\int_{2Rd(P_j)}^{\infty} |\theta_t^\mu (f1_{(2P_j)^c})(z)|^2 \, \frac{dt}{t} \Big)^{1/2} = III + IV.$$

It holds that

$$\begin{aligned} IV &\le \Big(\int_{t_0}^{\infty} |\theta_t^\mu f(z)|^2 \, \frac{dt}{t} \Big)^{1/2} + \Big(\int_{2Rd(P_j)}^{\infty} |\theta_t^\mu (f1_{2P_j})(z)|^2 \, \frac{dt}{t} \Big)^{1/2} \\ &\le \lambda + \Big(\int_{2Rd(P_j)}^{\infty} |\theta_t^\mu (f1_{2P_j})(z)|^2 \, \frac{dt}{t} \Big)^{1/2}. \end{aligned}$$

Notice that

$$\begin{aligned} |\theta_t^\mu(f1_{2P_j})(z)| \le Ct^{-m} \int_{2P_j} |f|\,d\mu &\le Ct^{-m} \int_{Q(x,C\ell(P_j))} |f|\,d\mu \\ &\le Ct^{-m}\ell(P_j)^m M_\mu^{\mathcal{Q}} f(x) \\ &\le Ct^{-m}\ell(P_j)^m \delta\lambda. \end{aligned}$$

This gives that

$$\Big(\int_{2Rd(P_j)}^\infty |\theta_t^\mu(f1_{2P_j})(z)|^2 \,\frac{dt}{t}\Big)^{1/2} \le C\ell(P_j)^m\delta\lambda \Big(\int_{2Rd(P_j)}^\infty t^{-2m-1}\,dt\Big)^{1/2} \le C\delta\lambda,$$

and so

$$IV \le \lambda + C\delta\lambda.$$

We now estimate III. Notice that

$$|\theta_t^\mu(f1_{(2P_j)^c})(x) - \theta_t^\mu(f1_{(2P_j)^c})(z)| \le \int_{\mathbb{R}^n} |s_t(x,y) - s_t(z,y)||f(y)|\,d\mu(y).$$

Since here $|x-z| \le Rd(P_j) < t/2$, we have

$$|s_t(x,y) - s_t(z,y)| \le C\frac{\ell(P_j)^\alpha}{(t+|x-y|)^{m+\alpha}},$$

and so

$$\begin{aligned} |\theta_t^\mu(f1_{(2P_j)^c})(x) - \theta_t^\mu(f1_{(2P_j)^c})(z)| &\le C\frac{\ell(P_j)^\alpha}{t^\alpha} \cdot t^\alpha \int_{\mathbb{R}^n} \frac{|f(y)|}{(t+|x-y|)^{m+\alpha}}\,d\mu(y) \\ &\le C\frac{\ell(P_j)^\alpha}{t^\alpha} M_\mu^{\mathcal{Q}} f(x) \le C\frac{\ell(P_j)^\alpha}{t^\alpha}\delta\lambda. \end{aligned}$$

This implies that

$$III \le C\ell(P_j)^\alpha\delta\lambda\Big(\int_{2Rd(P_j)}^\infty t^{-2\alpha-1}\,dt\Big)^{1/2} \le C\delta\lambda.$$

We have shown that

$$V_{\mu,t_0}(f1_{(2P_j)^c})(x) \le (1+C\delta)\lambda \le \Big(1+\frac{\epsilon}{2}\Big)\lambda$$

for all small enough $\delta = \delta(\epsilon)$. This completes the proof. □

5. Big pieces global Tb

In this section we prove the "big pieces" global Tb theorem for square functions. It will be highly useful in combination with the above presented good lambda method (Theorem 4.1).

5.1. **Definition.** Given a cube $Q \subset \mathbb{R}^n$ we consider the following random dyadic grid. For small notational convenience assume that $c_Q = 0$ (that is, Q is centred at the origin). Let $N \in \mathbb{Z}$ be defined by the requirement $2^{N-3} \le \ell(Q) < 2^{N-2}$. Consider the random square $Q^* = Q^*(w) = w + [-2^N, 2^N)^n$, where $w \in$

$[-2^{N-1}, 2^{N-1})^n =: \Omega_N = \Omega$. The set Ω is equipped with the normalised Lebesgue measure $\mathbb{P}_N = \mathbb{P}$. We define the grid $\mathcal{D}(w) := \mathcal{D}(Q^*(w))$. Notice that $Q \subset \alpha Q^*(w)$ for some $\alpha < 1$, and $\ell(Q) \sim \ell(Q^*(w))$.

Given a cube Q let us also denote the square function restricted to $(0, \ell(Q))$ by V_Q i.e.

$$V_Q\nu(x) = \Big(\int_0^{\ell(Q)} |\theta_t \nu(x)|^2 \frac{dt}{t} \Big)^{1/2}.$$

5.2. **Theorem.** *Suppose $(s_t)_{t>0}$ is an m-LP-family and V is the corresponding vertical square function. Let $Q \subset \mathbb{R}^n$ be a cube. Let σ be a finite Borel measure in $\mathbb{R}^n$ so that $\operatorname{spt}\sigma \subset Q$. Suppose b is a function satisfying that $\|b\|_{L^\infty(\sigma)} \le C_b$. For every w let T_w be the union of the maximal dyadic cubes $R \in \mathcal{D}(w)$ for which*

$$\Big| \int_R b \, d\sigma \Big| < c_{\text{acc}} \sigma(R).$$

We are also given a measurable set $H \subset \mathbb{R}^n$ satisfying the following properties.

- *There is $\delta_0 < 1$ so that $\sigma(H \cup T_w) \le \delta_0 \sigma(Q)$ for every w.*
- *Every ball B_r of radius r satisfying $\sigma(B_r) > C_0 r^m$ satisfies $B_r \subset H$.*
- *We have for some $s > 0$ the estimate*

$$\sup_{\lambda > 0} \lambda^s \sigma(\{x \in Q \setminus H \colon\ V_{\sigma,Q} b(x) > \lambda\}) \le C_1 \sigma(Q).$$

Then there is a measurable set G_Q satisfying $G_Q \subset Q \setminus H$ and the following properties:
(a) $\sigma(G_Q) \ge c\sigma(Q)$.
(b) $\|1_{G_Q} V_{\sigma,Q} f\|_{L^2(\sigma)} \le C\|f\|_{L^2(\sigma)}$ for every $f \in L^2(\sigma)$.
The constants c and C depend only on the preceding constants.

Proof. In this proof we write $A \lesssim B$, if there is a constant (absolute or depending on fixed constants) $C > 0$ so that $A \le CB$. We may also write $A \sim B$ if $B \lesssim A \lesssim B$.

Suppressing the operator. We begin by suppressing our operator appropriately. Set

$$S_0 = \{x \in Q \colon\ V_{\sigma,Q} b(x) > \lambda_0\},$$

where $0 < \lambda_0 \lesssim 1$ is large enough. Now simply define

$$\tilde{s}_t(x,y) = s_t(x,y) 1_{\mathbb{R}^n \setminus S_0}(x).$$

Notice that $(\tilde{s}_t)_{t>0}$ is a measurable family of kernels satisfying (1.1) and (1.2), which is all we shall need in what follows. Now, $\widetilde{V}_{\sigma,Q}$ (and similar objects) are defined in the natural way using the kernels $\tilde{s}_t$. Then for any f we have

$$\widetilde{V}_{\sigma,Q} f(x) = V_{\sigma,Q} f(x) 1_{\mathbb{R}^n \setminus S_0}(x) = V_{\sigma,Q} f(x) 1_{\mathbb{R}^n \setminus (Q \cap \{V_{\sigma,Q} b > \lambda_0\})}(x), \tag{5.3}$$

and from here we can easily read two key things about these suppressed operators. The first is that for any f we have

$$\widetilde{V}_{\sigma,Q} f(x) = V_{\sigma,Q} f(x) \text{ for } x \in \mathbb{R}^n \setminus S_0, \tag{5.4}$$

and the second is that

$$\widetilde{V}_{\sigma,Q}b(x) \le \lambda_0 \text{ for every } x \in Q. \tag{5.5}$$

Finally, with a large enough choice of λ_0 we have (for every w) that $\sigma(H \cup T_w \cup S_0) \le \delta_1 \sigma(Q)$ for some $\delta_1 < 1$. Indeed,

$$\sigma(S_0 \setminus H) \le \sigma(\{x \in Q \setminus H \colon V_{\sigma,Q}b(x) > \lambda_0\}) \le \frac{C_1}{\lambda_0^s}\sigma(Q).$$

At this point $\lambda_0 \lesssim 1$ can be fixed by demanding that it satisfies

$$\lambda_0^s > \frac{2C_1}{1-\delta_0},$$

whence we conclude that

$$\sigma(H \cup T_w \cup S_0) \le \sigma(H \cup T_w) + \sigma(S_0 \setminus H) \le \frac{1+\delta_0}{2}\sigma(Q) =: \delta_1 \sigma(Q),\ \delta_1 < 1. \tag{5.6}$$

We are now done with suppressing the operator.

Definition of the set G_Q. We will next define the set G_Q. This is done by setting

$$p_0(x) = \mathbb{P}(\{w \in \Omega \colon x \in Q \setminus [H \cup T_w \cup S_0]\}),$$

and then defining

$$G_Q = \Big\{x \in Q \colon p_0(x) > \frac{1-\delta_1}{2} =: \tau\Big\} \subset Q \setminus H.$$

We will show that $\sigma(G_Q) \gtrsim \sigma(Q)$. Notice first that by (5.6) we have that

$$\int_Q p_0(x)\,d\sigma(x) = \int_\Omega \sigma(Q \setminus [H \cup T_w \cup S_0])\,d\mathbb{P}(w) \ge (1-\delta_1)\sigma(Q).$$

Since $1 - p_0 \ge 0$ everywhere, and $1 - p_0 \ge 1 - \tau = (1+\delta_1)/2$ on $Q \setminus G_Q$, we have

$$\int_Q (1-p_0(x))\,d\sigma(x) \ge \int_{Q\setminus G_Q} (1-p_0(x))\,d\sigma(x) \ge \frac{1+\delta_1}{2}\sigma(Q \setminus G_Q).$$

We conclude that

$$\sigma(Q \setminus G_Q) \le \frac{2}{1+\delta_1}\Big(\sigma(Q) - \int_Q p_0(x)\,d\sigma(x)\Big) \le \frac{2\delta_1}{1+\delta_1}\sigma(Q),$$

and so

$$\sigma(G_Q) \ge \Big(1 - \frac{2\delta_1}{1+\delta_1}\Big)\sigma(Q) = \frac{1-\delta_1}{1+\delta_1}\sigma(Q).$$

Beginning of the proof of the L^2 bound. It remains to prove the L^2 estimate $\|1_{G_Q} V_{\sigma,Q} f\|_{L^2(\sigma)} \lesssim \|f\|_{L^2(\sigma)}$ for every $f \in L^2(\sigma)$. The key property of G_Q is as follows. Suppose $h \geq 0$ is any positive function. Then we have that

$$\int_{G_Q} h(x)\, d\sigma(x) \leq \tau^{-1} \int_{G_Q} p_0(x) h(x)\, d\sigma(x) = \tau^{-1} E_w \int_{G_Q \setminus [H \cup T_w \cup S_0]} h(x)\, d\sigma(x).$$

We apply this in the following way:

$$\begin{aligned} \|1_{G_Q} V_{\sigma,Q} f\|_{L^2(\sigma)}^2 &= \int_{G_Q} \int_0^{\ell(Q)} |\theta_t^\sigma f(x)|^2 \frac{dt}{t}\, d\sigma(x) \\ &\leq \tau^{-1} E_w \int_{G_Q \setminus [H \cup T_w \cup S_0]} \int_0^{\ell(Q)} |\theta_t^\sigma f(x)|^2 \frac{dt}{t}\, d\sigma(x) \\ &= \tau^{-1} E_w \sum_{R \in \mathcal{D}_0} \int_{[R \cap G_Q] \setminus [H \cup T_w \cup S_0]} \int_{\ell(R)/2}^{\min(\ell(R), \ell(Q))} |\theta_t^\sigma f(x)|^2 \frac{dt}{t}\, d\sigma(x), \end{aligned}$$

where $\mathcal{D}_0 = \mathcal{D}(0)$.

Given w we then write

$$\sum_{R \in \mathcal{D}_0} = \sum_{\substack{R \in \mathcal{D}_0 \\ R \text{ is } \mathcal{D}(w)\text{-good}}} + \sum_{\substack{R \in \mathcal{D}_0 \\ R \text{ is } \mathcal{D}(w)\text{-bad}}},$$

where $R \in \mathcal{D}_0$ is said to be $\mathcal{D}(w)$-good if $d(R, \partial P) > \ell(R)^\gamma \ell(P)^{1-\gamma}$ for every $P \in \mathcal{D}(w)$ satisfying $\ell(P) \geq 2^r \ell(R)$. Here $r \lesssim 1$ is a fixed large enough parameter, and $\gamma := \alpha/(2m + 2\alpha)$. It is a standard fact by Nazarov–Treil–Volberg that given $R \in \mathcal{D}_0$ we have that

$$\mathbb{P}(\{w \in \Omega\colon\ R \text{ is } \mathcal{D}(w)\text{-bad}\}) \leq \tau/2 \tag{5.7}$$

for a large enough fixed r. This is an exercise in elementary probability, which we skip here.

Using (5.7) we estimate

$$\begin{aligned} &E_w \sum_{\substack{R \in \mathcal{D}_0 \\ R \text{ is } \mathcal{D}(w)\text{-bad}}} \int_{[R \cap G_Q] \setminus [H \cup T_w \cup S_0]} \int_{\ell(R)/2}^{\min(\ell(R), \ell(Q))} |\theta_t^\sigma f(x)|^2 \frac{dt}{t}\, d\sigma(x) \\ &\leq \sum_{R \in \mathcal{D}_0} \mathbb{P}(\{w \in \Omega\colon\ R \text{ is } \mathcal{D}(w)\text{-bad}\}) \int_{R \cap G_Q} \int_{\ell(R)/2}^{\min(\ell(R), \ell(Q))} |\theta_t^\sigma f(x)|^2 \frac{dt}{t}\, d\sigma(x) \\ &\leq \frac{\tau}{2} \int_{G_Q} \int_0^{\ell(Q)} |\theta_t^\sigma f(x)|^2 \frac{dt}{t}\, d\sigma(x). \end{aligned}$$

To be precise, for the following we would need the a priori finiteness of this term. However, this is easy to arrange in a multiple of ways (e.g. do these calculations first by replacing $\int_0^{\ell(Q)}$ with $\int_\epsilon^{\ell(Q)}$, and let $\epsilon \to 0$ in the end), so we skip this technicality. We may now conclude (using also that $\theta_t^\sigma f(x) = \widetilde{\theta}_t^\sigma f(x)$ for every $x \in Q \setminus S_0$ by (5.4)) that

$$\begin{aligned}
&\|1_{G_Q} V_{\sigma,Q} f\|_{L^2(\sigma)}^2 \\
&\quad \le 2\tau^{-1} E_w \sum_{\substack{R \in \mathcal{D}_0 \\ R \text{ is } \mathcal{D}(w)\text{-good}}} \int_{[R \cap G_Q] \setminus [H \cup T_w \cup S_0]} \int_{\ell(R)/2}^{\min(\ell(R),\ell(Q))} |\widetilde{\theta}_t^\sigma f(x)|^2 \, \frac{dt}{t} \, d\sigma(x) \\
&\quad \lesssim E_w \sum_{\substack{R \in \mathcal{D}_0 \\ R \text{ is } \mathcal{D}(w)\text{-good} \\ R \not\subset H \cup T_w}} \int_R \int_{\ell(R)/2}^{\min(\ell(R),\ell(Q))} |\widetilde{\theta}_t^\sigma f(x)|^2 \, \frac{dt}{t} \, d\sigma(x).
\end{aligned}$$

We will now fix w, write $\mathcal{D} = \mathcal{D}(w)$ and $T = T_w$, and prove that

$$\sum_{\substack{R \in \mathcal{D}_0 \\ R \text{ is } \mathcal{D}\text{-good} \\ R \not\subset H \cup T}} \int_R \int_{\ell(R)/2}^{\min(\ell(R),\ell(Q))} |\widetilde{\theta}_t^\sigma f(x)|^2 \, \frac{dt}{t} \, d\sigma(x) \lesssim \|f\|_{L^2(\sigma)}^2. \tag{5.8}$$

This will then end the proof.

The important property of the set T is that if $R \in \mathcal{D}$ and $R \not\subset T$ then

$$\Big| \int_R b \, d\sigma \Big| \gtrsim \sigma(R),$$

while the important property of the set H is that if $L \subset \mathbb{R}^n$ is an arbitrary cube satisfying $L \not\subset H$ then $\sigma(\lambda L) \lesssim \lambda^m \ell(L)^m$ for all $\lambda \ge 1$. It is useful to say that $R \in \mathcal{D}_0^{tr}$ (tr stands for *transit*) if $R \in \mathcal{D}_0$, $\sigma(R) \neq 0$ and $R \not\subset H \cup T$, and $P \in \mathcal{D}^{tr}$ if $P \in \mathcal{D}$, $\sigma(P) \neq 0$ and $P \not\subset H \cup T$. Note that $\mathcal{D}_0^{tr}$ really means w-transit cubes from $\mathcal{D}_0$ (and one should really write $\mathcal{D}_0^{tr}(w)$), but w is fixed and so T is fixed and we do not need to insist on this.

Martingale decomposition of f. It is time to expand the function f in the grid $\mathcal{D}$ using b-adapted martingales only in the transit cubes $P \in \mathcal{D}^{tr}$. Denote $\langle f \rangle_A = \langle f \rangle_A^\sigma = \sigma(A)^{-1} \int_A f \, d\sigma$, if $\sigma(A) \neq 0$. Let $P_0 = Q^*(w)$ (see the Definition 5.1) so that all $P \in \mathcal{D}$ satisfy $P \subset P_0$. Without loss of generality we can assume that $\operatorname{spt} b \subset Q$ and $\operatorname{spt} f \subset Q$. Define

$$E_{P_0} f = \frac{\langle f \rangle_{P_0}}{\langle b \rangle_{P_0}} b.$$

(This is actually independent of w since it just equals $E_Q f$, because $\operatorname{spt} \sigma \subset Q \subset P_0$). For any cube $P \in \mathcal{D}^{tr}$ define the function $\Delta_P f$ as follows:

$$\Delta_P f = \sum_{P' \in \operatorname{ch}(P)} A_{P'}(f) 1_{P'},$$

where

$$A_{P'}(f) = \begin{cases} \left(\frac{\langle f\rangle_{P'}}{\langle b\rangle_{P'}} - \frac{\langle f\rangle_P}{\langle b\rangle_P}\right)b & \text{if } P' \in \mathcal{D}^{tr}, \\ f - \frac{\langle f\rangle_P}{\langle b\rangle_P} b & \text{if } P' \notin \mathcal{D}^{tr}. \end{cases}$$

It is easy to see that $\int_P \Delta_P f \, d\sigma = 0$, $P \in \mathcal{D}^{tr}$.

Notice that $P_0 \in \mathcal{D}^{tr}$, since $\sigma(P_0) = \sigma(Q)$ and every non-transit cube P has to satisfy $\sigma(P) \le \sigma(H \cup T) \le \delta_0 \sigma(Q)$. It holds that

$$f = \sum_{P \in \mathcal{D}^{tr}} \Delta_P f + E_{P_0} f$$

σ-a.e. and in $L^2(\sigma)$, and that

$$\sum_{P \in \mathcal{D}^{tr}} \|\Delta_P f\|^2_{L^2(\sigma)} + \|E_{P_0} f\|_{L^2(\sigma)} \lesssim \|f\|^2_{L^2(\sigma)}. \tag{5.9}$$

Let us prove the previous claims now. Given $x \in P_0$ and $k \ge 0$ let P_k^x denote the unique cube $P \in \mathcal{D}$ for which $\ell(P) = 2^{-k}\ell(P)$ and $x \in P$ (so that in particular $P_0 = P_0^x$). For $k_0 > 0$ define

$$S_{k_0} f(x) = \sum_{\substack{P \in \mathcal{D}^{tr} \\ \ell(P) > 2^{-k_0}\ell(P_0)}} \Delta_P f(x) + E_{P_0} f(x).$$

Suppose first that x is such that $P_x^k \in \mathcal{D}^{tr}$ for all $k \ge 0$. Then we have that

$$S_{k_0} f(x) = \frac{\langle f\rangle_{P^x_{k_0}}}{\langle b\rangle_{P^x_{k_0}}} b(x),$$

since the sum telescopes. Suppose then that x is such that $P_k^x \notin \mathcal{D}^{tr}$ for some k. Let $s(x) \ge 1$ be an integer so that $P^x_{s(x)-1} \in \mathcal{D}^{tr}$ but $P^x_{s(x)} \notin \mathcal{D}^{tr}$. Then for $k_0 < s(x)$ we have that

$$S_{k_0} f(x) = \frac{\langle f\rangle_{P^x_{k_0}}}{\langle b\rangle_{P^x_{k_0}}} b(x),$$

and for $k_0 \ge s(x)$ that

$$S_{k_0} f(x) = f(x).$$

We infer that for σ-a.e. $x \in Q$ we have

$$\lim_{k_0 \to \infty} S_{k_0} f(x) = f(x)$$

and

$$|S_{k_0} f(x)| \lesssim M_{\mathcal{D},\sigma} f(x).$$

By dominated convergence it then also follows that in $L^2(\sigma)$ we have

$$\lim_{k_0 \to \infty} S_{k_0} f = f$$

We have shown that

$$f = \sum_{P \in \mathcal{D}^{tr}} \Delta_P f + E_{P_0} f$$

σ-a.e. and in $L^2(\sigma)$.

We still need to prove the estimate

$$\sum_{P\in\mathcal{D}^{tr}} \|\Delta_P f\|_{L^2(\sigma)}^2 + \|E_{P_0} f\|_{L^2(\sigma)} \lesssim \|f\|_{L^2(\sigma)}^2.$$

It is obvious that $\|E_{P_0} f\|_{L^2(\sigma)} \lesssim \|f\|_{L^2(\sigma)}^2$. Let us then split

$$\begin{aligned}\sum_{P\in\mathcal{D}^{tr}} \|\Delta_P f\|_{L^2(\sigma)}^2 &= \sum_{P\in\mathcal{D}^{tr}} \sum_{\substack{P'\in\mathrm{ch}(P)\\ P'\in\mathcal{D}^{tr}}} \|1_{P'}\Delta_P f\|_{L^2(\sigma)}^2 + \sum_{P\in\mathcal{D}^{tr}} \sum_{\substack{P'\in\mathrm{ch}(P)\\ P'\notin\mathcal{D}^{tr}}} \|1_{P'}\Delta_P f\|_{L^2(\sigma)}^2 \\ &= I + II.\end{aligned}$$

Define also the standard martingales

$$D_P f(x) = \sum_{P'\in\mathrm{ch}(P)} [\langle f\rangle_{P'} - \langle f\rangle_P] 1_{P'}(x), \qquad P\in\mathcal{D}.$$

It is clear by orthogonality that

$$\sum_{P\in\mathcal{D}} \|D_P f\|_{L^2(\sigma)}^2 \le \|f\|_{L^2(\sigma)}^2.$$

We now estimate the term I above. Given $P\in\mathcal{D}^{tr}$ and $P'\in\mathrm{ch}(P)$ for which $P'\in\mathcal{D}^{tr}$ notice that

$$\begin{aligned}\Delta_P f|_{P'} &= \Big(\frac{\langle f\rangle_{P'}}{\langle b\rangle_{P'}} - \frac{\langle f\rangle_P}{\langle b\rangle_P}\Big) b \\ &= \Big(\frac{\langle f\rangle_{P'}}{\langle b\rangle_{P'}} - \frac{\langle f\rangle_{P'}}{\langle b\rangle_P}\Big) b + \Big(\frac{\langle f\rangle_{P'}}{\langle b\rangle_P} - \frac{\langle f\rangle_P}{\langle b\rangle_P}\Big) b \\ &= \frac{\langle b\rangle_P - \langle b\rangle_{P'}}{\langle b\rangle_P\langle b\rangle_{P'}}\langle f\rangle_{P'} b + (\langle f\rangle_{P'} - \langle f\rangle_P)\frac{b}{\langle b\rangle_P}.\end{aligned}$$

This implies that

$$|\Delta_P f|_{P'}| \lesssim |D_P b|_{P'}||\langle f\rangle_{P'}| + |D_P f|_{P'}|,$$

so that

$$\begin{aligned}I &\lesssim \sum_{P\in\mathcal{D}} \|D_P f\|_{L^2(\sigma)}^2 + \sum_{P\in\mathcal{D}}\sum_{P'\in\mathrm{ch}(P)} \|1_{P'} D_P b\|_{L^2(\sigma)}^2 |\langle f\rangle_{P'}|^2 \\ &\le \|f\|_{L^2(\sigma)}^2 + \sum_{\substack{P\in\mathcal{D}\\ P\subsetneq P_0}} \|1_P D_{P^{(1)}} b\|_{L^2(\sigma)}^2 |\langle f\rangle_P|^2 \lesssim \|f\|_{L^2(\sigma)}^2,\end{aligned}$$

where the last estimate used the fact that the sequence $a_P := \|1_P D_{P^{(1)}} b\|_{L^2(\sigma)}^2$ satisfies the Carleson estimate

$$\sum_{P\subset R} a_P \lesssim \sigma(R), \qquad R\in\mathcal{D}.$$

This is easy to see:

$$\sum_{P\subset R} a_P = \|1_R D_{R^{(1)}} b\|^2_{L^2(\sigma)} + \sum_{P\subsetneq R} \|1_P D_{P^{(1)}} b\|^2_{L^2(\sigma)}$$
$$\lesssim \sigma(R) + \sum_{P\in\mathcal{D}} \|D_P(b1_R)\|^2_{L^2(\sigma)} \le \sigma(R) + \|b1_R\|^2_{L^2(\sigma)} \lesssim \sigma(R).$$

We now deal with the term II. Given $P \in \mathcal{D}^{tr}$ and $P' \in \mathrm{ch}(P)$ for which $P' \notin \mathcal{D}^{tr}$ and $\sigma(P') \neq 0$, notice that

$$\Delta_P f|_{P'} = f - \frac{\langle f\rangle_P}{\langle b\rangle_P} b = \Big(f - \frac{\langle f\rangle_{P'}}{\langle b\rangle_P} b\Big) + \Big(\frac{\langle f\rangle_{P'}}{\langle b\rangle_P} b - \frac{\langle f\rangle_P}{\langle b\rangle_P} b\Big).$$

This implies that

$$|\Delta_P f|_{P'}| \lesssim |f| + |\langle f\rangle_{P'}| + |D_P f|_{P'}|,$$

and so

$$II \lesssim \sum_{P\in\mathcal{D}^{tr}} \sum_{\substack{P'\in \mathrm{ch}(P)\\ P'\notin\mathcal{D}^{tr}}} \int_{P'} |f|^2\,d\sigma + \sum_{P\in\mathcal{D}} \|\Delta_P f\|^2_{L^2(\sigma)} \lesssim \|f\|^2_{L^2(\sigma)}.$$

Here we used that the maximal non-transit cubes are disjoint. This completes our proof of the L^2 estimate for the martingales.

In what follows it will be convenient to exploit notation by redefining on the largest level P_0 the operator $\Delta_{P_0} f$ to be $\Delta_{P_0} f + E_{P_0} f$. This means that we may write

$$f = \sum_{P\in\mathcal{D}^{tr}} \Delta_P f,$$

where $\int_P \Delta_P f\,d\sigma = 0$ unless $P = P_0$.

Continuation of the proof of the L^2 bound. Going back to (5.8) we see that we need to prove that

$$\sum_{\substack{R\in\mathcal{D}_0^{tr}\\ R \text{ is } \mathcal{D}\text{-good}}} \int_R \int_{\ell(R)/2}^{\min(\ell(R),\ell(Q))} \Big|\sum_{P\in\mathcal{D}^{tr}} \widetilde{\theta}_t^\sigma \Delta_P f(x)\Big|^2 \frac{dt}{t}\,d\sigma(x) \lesssim \|f\|^2_{L^2(\sigma)}.$$

Given $R \in \mathcal{D}_0^{tr}$ which is $\mathcal{D}$-good, the $P \in \mathcal{D}^{tr}$ summation is split in to the following four pieces:

(1) P: $\ell(P) < \ell(R)$;
(2) P: $\ell(P) \ge \ell(R)$ and $d(P,R) > \ell(R)^\gamma \ell(P)^{1-\gamma}$;
(3) P: $\ell(R) \le \ell(P) \le 2^r\ell(R)$ and $d(P,R) \le \ell(R)^\gamma \ell(P)^{1-\gamma}$;
(4) P: $\ell(P) > 2^r\ell(R)$ and $d(P,R) \le \ell(R)^\gamma \ell(P)^{1-\gamma}$.

The matrix A_{PR}. Define

$$\begin{aligned} A_{PR} &:= \frac{\ell(P)^{\alpha/2}\ell(R)^{\alpha/2}}{D(P,R)^{m+\alpha}}\sigma(P)^{1/2}\sigma(R)^{1/2}; \\ D(P,R) &:= \ell(P)+\ell(R)+d(P,R). \end{aligned}$$

We will prove the following extremely useful estimate

$$\sum_{\substack{P\in\mathcal{D}^{tr}\\ R\in\mathcal{D}_0^{tr}}} A_{PR}x_Py_R \lesssim \Big(\sum_{P\in\mathcal{D}^{tr}} x_P^2\Big)^{1/2}\Big(\sum_{R\in\mathcal{D}_0^{tr}} y_R^2\Big)^{1/2}$$

for every $x_P, y_R \ge 0$. Therefore, we have

$$\Big(\sum_{R\in\mathcal{D}_0^{tr}}\Big[\sum_{P\in\mathcal{D}^{tr}} A_{PR}x_P\Big]^2\Big)^{1/2} \lesssim \Big(\sum_{P\in\mathcal{D}^{tr}} x_P^2\Big)^{1/2}, \tag{5.10}$$

which is an estimate we shall have frequent use for.

By symmetry it is enough to prove that

$$I := \sum_{\substack{P\in\mathcal{D}^{tr}\\ R\in\mathcal{D}_0^{tr}\\ \ell(P)\le\ell(R)}} A_{PR}x_Py_R \lesssim \Big(\sum_{P\in\mathcal{D}^{tr}} x_P^2\Big)^{1/2}\Big(\sum_{R\in\mathcal{D}_0^{tr}} y_R^2\Big)^{1/2}.$$

Let us first note that this claim follows if for every $R\in\mathcal{D}_0^{tr}$ we have

$$I_R := \sum_{\substack{P\in\mathcal{D}^{tr}\\ \ell(P)\le\ell(R)}} A_{PR}\sigma(P)^{1/2} \lesssim \sigma(R)^{1/2}$$

and for every $P\in\mathcal{D}^{tr}$ we have

$$I_P := \sum_{\substack{R\in\mathcal{D}_0^{tr}\\ \ell(P)\le\ell(R)}} A_{PR}\sigma(R)^{1/2} \lesssim \sigma(P)^{1/2}.$$

Indeed, assuming these inequalities we have for given $R\in\mathcal{D}_0^{tr}$ that

$$\begin{aligned} \sum_{\substack{P\in\mathcal{D}^{tr}\\ \ell(P)\le\ell(R)}} A_{PR}x_P &= \sum_{\substack{P\in\mathcal{D}^{tr}\\ \ell(P)\le\ell(R)}} A_{PR}^{1/2}\sigma(P)^{1/4}\cdot A_{PR}^{1/2}\sigma(P)^{-1/4}x_P \\ &\le I_R^{1/2}\Big(\sum_{\substack{P\in\mathcal{D}^{tr}\\ \ell(P)\le\ell(R)}} A_{PR}\sigma(P)^{-1/2}x_P^2\Big)^{1/2} \\ &\lesssim \sigma(R)^{1/4}\Big(\sum_{\substack{P\in\mathcal{D}^{tr}\\ \ell(P)\le\ell(R)}} A_{PR}\sigma(P)^{-1/2}x_P^2\Big)^{1/2}, \end{aligned}$$

and so

$$I \lesssim \sum_{R\in\mathcal{D}_0^{tr}} y_R\sigma(R)^{1/4}\Big(\sum_{\substack{P\in\mathcal{D}^{tr}\\ \ell(P)\leq\ell(R)}} A_{PR}\sigma(P)^{-1/2}x_P^2\Big)^{1/2}$$

$$\leq \Big(\sum_{R\in\mathcal{D}_0^{tr}} y_R^2\Big)^{1/2}\Big(\sum_{P\in\mathcal{D}^{tr}} x_P^2\sigma(P)^{-1/2}I_P\Big)^{1/2} \lesssim \Big(\sum_{P\in\mathcal{D}^{tr}} x_P^2\Big)^{1/2}\Big(\sum_{R\in\mathcal{D}_0^{tr}} y_R^2\Big)^{1/2}.$$

So will just prove these inequalities for I_R and I_P.

Let us first fix $R \in \mathcal{D}_0^{tr}$ and write I_R as follows

$$\begin{aligned} I_R &= \sigma(R)^{1/2} \sum_{\substack{P\in\mathcal{D}^{tr}\\ \ell(P)\leq\ell(R)}} \Big(\frac{\ell(P)}{\ell(R)}\Big)^{\alpha/2}\frac{\ell(R)^\alpha}{D(P,R)^{m+\alpha}}\sigma(P) \\ &= \sigma(R)^{1/2}\sum_{k\geq 0} 2^{-\alpha k/2} \sum_{\substack{P\in\mathcal{D}^{tr}\\ \ell(P)=2^{-k}\ell(R)}} \frac{\ell(R)^\alpha}{D(P,R)^{m+\alpha}}\sigma(P). \end{aligned}$$

We fix $k \geq 0$ and then show that the last sum is $\lesssim 1$, which yields the desired bound for I_R. This can be seen as follows

$$\begin{aligned} &\sum_{\substack{P\in\mathcal{D}^{tr}\\ \ell(P)=2^{-k}\ell(R)}} \frac{\ell(R)^\alpha}{D(P,R)^{m+\alpha}}\sigma(P) \\ &\leq \frac{1}{\ell(R)^m}\sum_{\substack{P\in\mathcal{D}^{tr}\\ \ell(P)=2^{-k}\ell(R)\\ d(P,R)\leq\ell(R)}} \sigma(P) + \sum_{j\geq 0} 2^{-\alpha j}\frac{1}{(2^j\ell(R))^m}\sum_{\substack{P\in\mathcal{D}^{tr}\\ \ell(P)=2^{-k}\ell(R)\\ 2^j\ell(R)\leq d(P,R)\leq 2^{j+1}\ell(R)}} \sigma(P) \\ &\leq \frac{\sigma(10R)}{\ell(R)^m} + \sum_{j\geq 0} 2^{-\alpha j}\frac{\sigma(2^{j+5}R)}{(2^j\ell(R))^m} \lesssim 1, \end{aligned}$$

where we used that $R \not\subset H$ by transitivity, and so $\sigma(\lambda R) \lesssim \lambda^m\ell(R)^m$ for all $\lambda \geq 1$.

Next, we fix $P \in \mathcal{D}^{tr}$ and write I_P as follows

$$I_P = \sigma(P)^{1/2}\sum_{k\geq 0} 2^{-\alpha k/2} \sum_{\substack{R\in\mathcal{D}_0^{tr}\\ \ell(R)=2^k\ell(P)}} \frac{\ell(R)^\alpha}{D(P,R)^{m+\alpha}}\sigma(R).$$

Again, for a fixed $k \geq 0$ we will show that the last sum is $\lesssim 1$. Let P_k be a cube concentric with P and of side length $\ell(P_k) = 2^k \ell(P) = \ell(R)$. We have that

$$\begin{aligned}
&\sum_{\substack{R \in \mathcal{D}_0^{tr} \\ \ell(R) = 2^k \ell(P)}} \frac{\ell(R)^\alpha}{D(P,R)^{m+\alpha}} \sigma(R) \\
&\leq \frac{1}{\ell(P_k)^m} \sum_{\substack{R \in \mathcal{D}_0^{tr} \\ \ell(R) = \ell(P_k) \\ d(P,R) \leq \ell(P_k)}} \sigma(R) + \sum_{j \geq 0} 2^{-\alpha j} \frac{1}{(2^j \ell(P_k))^m} \sum_{\substack{R \in \mathcal{D}_0^{tr} \\ \ell(R) = \ell(P_k) \\ 2^j \ell(P_k) \leq d(P,R) \leq 2^{j+1} \ell(P_k)}} \sigma(R) \\
&\leq \frac{\sigma(10P_k)}{\ell(P_k)^m} + \sum_{j \geq 0} 2^{-\alpha j} \frac{\sigma(2^{j+5} P_k)}{(2^j \ell(P_k))^m} \lesssim 1,
\end{aligned}$$

where we used that $P_k \not\subset H$ (since $P \not\subset H$ by transitivity), and so $\sigma(\lambda P_k) \lesssim \lambda^m \ell(P_k)^m$ for all $\lambda \geq 1$.

Summations (1) and (2). Notice that in (1) we have $\ell(P) < \ell(R) \leq \ell(P_0)$ so that $\int \Delta_P f \, d\sigma = 0$. Therefore, using the y-Hölder for $\tilde{s}_t$ we get for $(x,t) \in W_R := R \times [\ell(R)/2, \ell(R))$ that

$$\begin{aligned}
|\widetilde{\theta}_t^\sigma \Delta_P f(x)| &= \Big| \int_P [\tilde{s}_t(x,y) - \tilde{s}_t(x,c_P)] \Delta_P f(y) \, d\sigma(y) \Big| \\
&\lesssim \frac{\ell(P)^\alpha}{(\ell(R) + d(P,R))^{m+\alpha}} \int_P |\Delta_P f(y)| \, d\sigma(y).
\end{aligned}$$

The Hölder estimate can be used, since here

$$|y - c_P| \lesssim \ell(P) \leq \ell(R) \lesssim t.$$

This implies that

$$|\widetilde{\theta}_t^\sigma \Delta_P f(x)| \lesssim A_{PR} \sigma(R)^{-1/2} \|\Delta_P f\|_{L^2(\sigma)}, \qquad (x,t) \in W_R. \tag{5.11}$$

In the case (2), the size estimate for $\tilde{s}_t$ yields

$$|\widetilde{\theta}_t^\sigma \Delta_P f(x)| \lesssim \frac{\ell(R)^\alpha}{d(P,R)^{m+\alpha}} \sigma(P)^{1/2} \|\Delta_P f\|_{L^2(\sigma)}, \qquad (x,t) \in W_R.$$

But this yields the same bound as in (5.11), since here

$$\frac{\ell(R)^\alpha}{d(P,R)^{m+\alpha}} \sigma(P)^{1/2} \lesssim A_{PR} \sigma(R)^{-1/2}.$$

To see this, notice that it is obvious if $d(P,R) \geq \ell(P)$. In the opposite case note that $d(P,R)^{m+\alpha} \gtrsim D(P,R)^{m+\alpha} \ell(P)^{-\alpha/2} \ell(R)^{\alpha/2}$. This is seen by combining the facts that $d(P,R) > \ell(R)^\gamma \ell(P)^{1-\gamma}$, $\gamma m + \gamma \alpha = \alpha/2$ and $D(P,R) \lesssim \ell(P)$:

$$d(P,R)^{m+\alpha} > \ell(R)^{\gamma(m+\alpha)} \ell(P)^{m+\alpha} \ell(P)^{-\gamma(m+\alpha)} \gtrsim \ell(R)^{\alpha/2} \ell(P)^{-\alpha/2} D(P,R)^{m+\alpha}.$$

Thus, also in the case (2) the estimate (5.11) holds. The cases (1) and (2) are therefore under control via the estimate

$$\sum_{R\in\mathcal{D}_0^{tr}}\Big[\sum_{P\in\mathcal{D}^{tr}} A_{PR}\|\Delta_P f\|_{L^2(\sigma)}\Big]^2 \lesssim \sum_{P\in\mathcal{D}^{tr}}\|\Delta_P f\|^2_{L^2(\sigma)} \lesssim \|f\|^2_{L^2(\sigma)}.$$

Here we used (5.10) and (5.9).

Summation (3). The summation (3) is even easier. Using that P and R are both transit, $t\sim\ell(R)\sim\ell(P)$ and the size estimate for $\tilde{s}_t$ we see that

$$|\widetilde{\theta}_t^\sigma\Delta_P f(x)| \lesssim t^{-m}\sigma(P)^{1/2}\|\Delta_P f\|_{L^2(\sigma)} \lesssim \sigma(R)^{-1/2}\|\Delta_P f\|_{L^2(\sigma)}, \qquad (x,t)\in W_R.$$

This can then easily be summed, since given R there are only finitely many P such that $\ell(P)\sim\ell(R)$ and $d(P,R)\lesssim\min(\ell(P),\ell(R))$. Denote this by $P\sim R$, and simply bound

$$\sum_{R\in\mathcal{D}_0^{tr}}\Big[\sum_{\substack{P\in\mathcal{D}^{tr}\\ P\sim R}}\|\Delta_P f\|_{L^2(\sigma)}\Big]^2 \lesssim \sum_{P\in\mathcal{D}^{tr}}\|\Delta_P f\|^2_{L^2(\sigma)}\sum_{\substack{R\in\mathcal{D}_0^{tr}\\ R\sim P}} 1 \lesssim \|f\|^2_{L^2(\sigma)}.$$

Summation (4). In this summation the $\mathcal{D}$-goodness of R forces that $R\subset P$. For each $R\in\mathcal{D}_0^{tr}$ satisfying that R is $\mathcal{D}$-good, $R\subset P_0$ and $\ell(R)<2^{-r}\ell(P_0)$ we let $P_{R,k}\in\mathcal{D}$, $k\in\{r,r+1,\ldots,\log_2[\ell(P_0)/\ell(R)]\}$, be the unique $\mathcal{D}$-cube satisfying that $\ell(P_{R,k})=2^k\ell(R)$ and $R\subset P_{R,k}$. Such a cube exists since R is $\mathcal{D}$-good. Moreover, since $R\not\subset H\cup T$ then also $P_{R,k}\not\subset H\cup T$ i.e. $P_{R,k}\in\mathcal{D}^{tr}$. We see that we only need to prove that

$$\sum_{\substack{R\in\mathcal{D}_0^{tr}:\, R\subset P_0\\ R \text{ is } \mathcal{D}\text{-good}\\ \ell(R)<2^{-r}\ell(P_0)}}\int_R\int_{\ell(R)/2}^{\min(\ell(R),\ell(Q))}\Big|\sum_{k=r+1}^{\log_2[\ell(P_0)/\ell(R)]}\widetilde{\theta}_t^\sigma\Delta_{P_{R,k}}f(x)\Big|^2\frac{dt}{t}\,d\sigma(x)\lesssim\|f\|^2_{L^2(\sigma)}.$$

Define

$$B_{P_{R,k-1}} = \langle\Delta_{P_{R,k}}f/b\rangle_{P_{R,k-1}} = \begin{cases}\frac{\langle f\rangle_{P_{R,k-1}}}{\langle b\rangle_{P_{R,k-1}}}-\frac{\langle f\rangle_{P_{R,k}}}{\langle b\rangle_{P_{R,k}}}, & \text{if } r+1\le k<\log_2\frac{\ell(P_0)}{\ell(R)},\\ \frac{\langle f\rangle_{P_{R,k-1}}}{\langle b\rangle_{P_{R,k-1}}}, & k=\log_2\frac{\ell(P_0)}{\ell(R)}.\end{cases}$$

Notice that the fact that $P_{R,k}\in\mathcal{D}^{tr}$ for all $k\in\{r,r+1,\ldots,\log_2[\ell(P_0)/\ell(R)]\}$ was used here. We write

$$\Delta_{P_{R,k}}f = 1_{P_{R,k}\setminus P_{R,k-1}}\Delta_{P_{R,k}}f + 1_{P_{R,k-1}}\Delta_{P_{R,k}}f,$$

where

$$1_{P_{R,k-1}}\Delta_{P_{R,k}}f = 1_{P_{R,k-1}}B_{P_{R,k-1}}b = B_{P_{R,k-1}}b - 1_{\mathbb{R}^n\setminus P_{R,k-1}}B_{P_{R,k-1}}b.$$

We can now see that $\sum_{k=r+1}^{\log_2[\ell(P_0)/\ell(R)]} \widetilde{\theta}_t^\sigma \Delta_{P_{R,k}} f$ equals

$$-\sum_{k=r+1}^{\log_2[\ell(P_0)/\ell(R)]} B_{P_{R,k-1}} \widetilde{\theta}_t^\sigma (1_{\mathbb{R}^n \setminus P_{R,k-1}} b) \\ + \sum_{k=r+1}^{\log_2[\ell(P_0)/\ell(R)]} \widetilde{\theta}_t^\sigma (1_{P_{R,k} \setminus P_{R,k-1}} \Delta_{P_{R,k}} f) + \frac{\langle f \rangle_{P_{R,r}}}{\langle b \rangle_{P_{R,r}}} \widetilde{\theta}_t^\sigma b,$$

where we used that

$$\sum_{k=r+1}^{\log_2[\ell(P_0)/\ell(R)]} B_{P_{R,k-1}} = \frac{\langle f \rangle_{P_{R,r}}}{\langle b \rangle_{P_{R,r}}}.$$

Let us start deciphering this by proving that the term

$$\Pi := \sum_{\substack{R \in \mathcal{D}_0^{tr}:\, R \subset P_0 \\ R \text{ is } \mathcal{D}\text{-good} \\ \ell(R) < 2^{-r}\ell(P_0)}} \Big| \frac{\langle f \rangle_{P_{R,r}}}{\langle b \rangle_{P_{R,r}}} \Big|^2 \int_R \int_{\ell(R)/2}^{\min(\ell(R),\ell(Q))} |\widetilde{\theta}_t^\sigma b(x)|^2 \frac{dt}{t}\, d\sigma(x)$$

is under control. We simply estimate

$$\Pi \lesssim \sum_{P \in \mathcal{D}^{tr}} |\langle f \rangle_P|^2 a_P, \qquad a_P := \sum_{\substack{R \in \mathcal{D}_0^{tr}:\, R \subset P_0 \\ R \text{ is } \mathcal{D}\text{-good} \\ \ell(R) < 2^{-r}\ell(P_0) \\ P_{R,r} = P}} \int_R \int_{\ell(R)/2}^{\min(\ell(R),\ell(Q))} |\widetilde{\theta}_t^\sigma b(x)|^2 \frac{dt}{t}\, d\sigma(x).$$

To have $\Pi \lesssim \|f\|_{L^2(\sigma)}^2$ it is enough to verify the Carleson property of $(a_P)_{P \in \mathcal{D}}$. To this end, let $S \in \mathcal{D}$ be arbitrary. We have that

$$\begin{aligned} \sum_{\substack{P \in \mathcal{D} \\ P \subset S}} a_P &\le \sum_{\substack{R \in \mathcal{D}_0^{tr} \\ R \subset S}} \iint_{[S \times (0,\ell(Q))] \cap W_R} |\widetilde{\theta}_t^\sigma b(x)|^2 \frac{dt}{t}\, d\sigma(x) \\ &\le \iint_{S \times (0,\ell(Q))} |\widetilde{\theta}_t^\sigma b(x)|^2 \frac{dt}{t}\, d\sigma(x) = \int_S [\widetilde{V}_{\sigma,Q} b(x)]^2\, d\sigma(x) \lesssim \sigma(S), \end{aligned}$$

since $\widetilde{V}_{\sigma,Q} b(x) \lesssim 1$ for every $x \in \operatorname{spt}\sigma$ by (5.5).

We now deal with the rest of the terms. Let us control $|B_{P_{R,k-1}} \widetilde{\theta}_t^\sigma (1_{\mathbb{R}^n \setminus P_{R,k-1}} b)(x)|$ for $(x,t) \in W_R$. Notice that $R \subset B(x, d(R, \partial P_{R,k-1})/2)$, since $d(R, \partial P_{R,k-1}) \ge 2^{r(1-\gamma)}\ell(R) \ge C_d \ell(R)$ by having r large enough to begin with. The point of this observation is that $B(x, d(R, \partial P_{R,k-1})/2) \not\subset H$. Moreover, we clearly have that

$B(x, d(R, \partial P_{R,k-1})/2) \subset P_{R,k-1}$. Using these facts we get

$$\begin{aligned} |\widetilde{\theta}_t^\sigma(1_{\mathbb{R}^n \setminus P_{R,k-1}} b)(x)| &\lesssim \int_{\mathbb{R}^n \setminus B(x, d(R,\partial P_{R,k-1})/2)} \frac{\ell(R)^\alpha}{|x-y|^{m+\alpha}}\, d\sigma(y) \\ &\lesssim \ell(R)^\alpha d(R, \partial P_{R,k-1})^{-\alpha} \lesssim \Big(\frac{\ell(R)}{\ell(P_{R,k-1})} \Big)^{\alpha/2} \sim 2^{-\alpha k/2}, \end{aligned}$$

where we also used that $d(R, \partial P_{R,k-1}) \ge \ell(R)^{1/2} \ell(P_{R,k-1})^{1/2}$ (which follows since R is $\mathcal{D}$-good). Since $P_{R,k-1} \not\subset T$ we have

$$\begin{aligned} |B_{P_{R,k-1}}| \sigma(P_{R,k-1}) &\lesssim \Big| \int_{P_{R,k-1}} B_{P_{R,k-1}} b\, d\sigma \Big| \\ &= \Big| \int_{P_{R,k-1}} \Delta_{P_{R,k}} f\, d\sigma \Big| \le \sigma(P_{R,k-1})^{1/2} \|\Delta_{P_{R,k}} f\|_{L^2(\sigma)}. \end{aligned}$$

Combining these estimates we get for $(x,t) \in W_R$ that

$$|B_{P_{R,k-1}} \widetilde{\theta}_t^\sigma(1_{\mathbb{R}^n \setminus P_{R,k-1}} b)(x)| \lesssim 2^{-\alpha k/2} \sigma(P_{R,k-1})^{-1/2} \|\Delta_{P_{R,k}} f\|_{L^2(\sigma)}. \tag{5.12}$$

Let us still estimate $|\widetilde{\theta}_t^\sigma(1_{P_{R,k} \setminus P_{R,k-1}} \Delta_{P_{R,k}} f)(x)|$ for $(x,t) \in W_R$. Let $S \in \mathrm{ch}(P_{R,k})$, $S \ne P_{R,k-1}$. We do not know whether this cube is transitive or not, but it shall not matter. Indeed, we just estimate

$$\begin{aligned} |\widetilde{\theta}_t^\sigma(1_S \Delta_{P_{R,k}} f)(x)| &\lesssim \frac{\ell(R)^\alpha}{d(R,S)^{m+\alpha}} \int_{P_{R,k}} |\Delta_{P_{R,k}} f(y)|\, d\sigma(y) \\ &\lesssim \Big(\frac{\ell(R)}{\ell(P_{R,k-1})} \Big)^{\alpha/2} \frac{\sigma(P_{R,k})^{1/2}}{\ell(P_{R,k-1})^m} \|\Delta_{P_{R,k}} f\|_{L^2(\sigma)} \\ &\lesssim 2^{-\alpha k/2} \sigma(P_{R,k-1})^{-1/2} \|\Delta_{P_{R,k}} f\|_{L^2(\sigma)}, \end{aligned}$$

where we used that $\ell(S) = \ell(P_{R,k-1})$, $d(R,S)^{m+\alpha} \ge \ell(R)^{\alpha/2} \ell(S)^{\alpha/2} \ell(S)^m$ and the transitivity of $P_{R,k-1}, P_{R,k}$. So $|\widetilde{\theta}_t^\sigma(1_{P_{R,k} \setminus P_{R,k-1}} \Delta_{P_{R,k}} f)(x)|$ satisfies the same estimate as in (5.12).

We are now ready to complete the whole proof. The following estimate is all that remains:

$$\begin{aligned}
&\sum_{\substack{R\in\mathcal{D}_0^{tr}:\, R\subset P_0\\ R \text{ is } \mathcal{D}\text{-good}\\ \ell(R)<2^{-r}\ell(P_0)}} \sigma(R)\Big[\sum_{k=r+1}^{\log_2[\ell(P_0)/\ell(R)]} 2^{-\alpha k/2}\sigma(P_{R,k-1})^{-1/2}\|\Delta_{P_{R,k}}f\|_{L^2(\sigma)}\Big]^2\\
&\lesssim \sum_{\substack{R\in\mathcal{D}_0^{tr}:\, R\subset P_0\\ R \text{ is } \mathcal{D}\text{-good}\\ \ell(R)<2^{-r}\ell(P_0)}} \sigma(R) \sum_{k=r+1}^{\log_2[\ell(P_0)/\ell(R)]} 2^{-\alpha k/2}\sigma(P_{R,k-1})^{-1}\|\Delta_{P_{R,k}}f\|^2_{L^2(\sigma)}\\
&= \sum_{u=r+1}^{\infty} \sum_{\substack{R\in\mathcal{D}_0^{tr}:\, R\subset P_0\\ R \text{ is } \mathcal{D}\text{-good}\\ \ell(R)=2^{-u}\ell(P_0)}} \sigma(R) \sum_{k=r+1}^{u} 2^{-\alpha k/2}\sigma(P_{R,k-1})^{-1}\|\Delta_{P_{R,k}}f\|^2_{L^2(\sigma)}\\
&= \sum_{k=r+1}^{\infty} 2^{-\alpha k/2}\sum_{u=k}^{\infty} \sum_{\substack{R\in\mathcal{D}_0^{tr}:\, R\subset P_0\\ R \text{ is } \mathcal{D}\text{-good}\\ \ell(R)=2^{-u}\ell(P_0)}} \sigma(R)\sigma(P_{R,k-1})^{-1}\|\Delta_{P_{R,k}}f\|^2_{L^2(\sigma)}\\
&= \sum_{k=r+1}^{\infty} 2^{-\alpha k/2}\sum_{u=k}^{\infty} \sum_{\substack{P\in\mathcal{D}^{tr}\\ \ell(P)=2^{k-1-u}\ell(P_0)}} \|\Delta_{P^{(1)}}f\|^2_{L^2(\sigma)}\frac{1}{\sigma(P)} \sum_{\substack{R\in\mathcal{D}_0^{tr}:\, R\subset P_0\\ R \text{ is } \mathcal{D}\text{-good}\\ P_{R,k-1}=P}} \sigma(R)\\
&\le \sum_{k=r+1}^{\infty} 2^{-\alpha k/2}\sum_{u=k}^{\infty} \sum_{\substack{P\in\mathcal{D}^{tr}\\ \ell(P)=2^{k-1-u}\ell(P_0)}} \|\Delta_{P^{(1)}}f\|^2_{L^2(\sigma)}\\
&\lesssim \sum_{k=r+1}^{\infty} 2^{-\alpha k/2}\sum_{u=k}^{\infty} \sum_{\substack{P\in\mathcal{D}^{tr}\\ \ell(P)=2^{k-u}\ell(P_0)}} \|\Delta_{P}f\|^2_{L^2(\sigma)}\\
&= \Big(\sum_{k=r+1}^{\infty} 2^{-\alpha k/2}\Big)\Big(\sum_{P\in\mathcal{D}^{tr}} \|\Delta_{P}f\|^2_{L^2(\sigma)}\Big) \lesssim \|f\|^2_{L^2(\sigma)}.
\end{aligned}$$

We have proved the estimate $\|1_{G_Q}V_{\sigma,Q}f\|_{L^2(\sigma)} \lesssim \|f\|_{L^2(\sigma)}$ for every $f\in L^2(\sigma)$, and so the proof of the big pieces Tb is now complete.

□

6. LOCAL Tb THEOREM

The following local Tb theorem is the main result of [1].

6.1. **Theorem.** *Let μ be a measure of order m in $\mathbb{R}^n$ and $B_1, B_2 < \infty$, $\epsilon_0 \in (0,1)$ be given constants. Let also $(s_t)_{t>0}$ be an x-continuous m-LP-family, and V be the corresponding vertical square function. Let $\beta > 0$ and C_1 be large enough (depending only on n). Suppose that for every $(2,\beta)$-doubling cube $Q \subset \mathbb{R}^n$ with C_1-small boundary there exists a complex measure ν_Q so that*

(1) *spt $\nu_Q \subset Q$;*
(2) *$\mu(Q) = \nu_Q(Q)$;*
(3) *$\|\nu_Q\| \le B_1\mu(Q)$;*
(4) *For all Borel sets $A \subset Q$ satisfying $\mu(A) \le \epsilon_0\mu(Q)$ we have*
$$|\nu_Q|(A) \le \frac{\|\nu_Q\|}{32B_1}.$$

Suppose there exist $s > 0$ and for all Q as above a Borel set $U_Q \subset \mathbb{R}^n$ such that $|\nu_Q|(U_Q) \le \frac{\|\nu_Q\|}{16B_1}$ and
$$\sup_{\lambda>0} \lambda^s\mu(\{x \in Q \setminus U_Q \colon V_Q\nu_Q(x) > \lambda\}) \le B_2\|\nu_Q\|.$$
Then $V_\mu \colon L^p(\mu) \to L^p(\mu)$ for every $p \in (1,\infty)$.

6.2. *Remark.* If $V_\mu \colon L^2(\mu) \to L^2(\mu)$ boundedly, then $V \colon M(\mathbb{R}^n) \to L^{1,\infty}(\mu)$ boundedly. In this case, given ν_Q like above one has to have
$$\sup_{\lambda>0} \lambda\mu(\{x \in Q \colon V_Q\nu_Q(x) > \lambda\}) \le \sup_{\lambda>0} \lambda\mu(\{x \colon V\nu_Q(x) > \lambda\}) \le C\|\nu_Q\|.$$
This makes the assumptions necessary.

6.3. *Remark.* Set $\nu_Q = b_Q\,d\mu$ for some function b_Q supported in Q satisfying that $\mu(Q) = \int_Q b_Q\,d\mu$ and $\int_Q |b_Q|^q\,d\mu \lesssim \mu(Q)$. If $q > 1$ we automatically have (4) (and of course (3)) using Hölder's inequality. But one can have $q = 1$ if one has (4) by some other virtue. The testing condition on the operator side is extremely weak, e.g.
$$\sup_{\lambda>0} \lambda\mu(\{x \in Q \colon V_{\mu,Q}b_Q(x) > \lambda\}) \le B_2\mu(Q)$$
suffices. In the previously known best results one needed an L^q norm also on the operator side if $b_Q \in L^q$, $q > 1$, like in the above discussion.

We also allow to work with measures, allow a small exceptional set U_Q, and require the existence of ν_Q only in very regular cubes Q.

We record the following easy lemma.

6.4. **Lemma.** *Let a cube $Q \subset \mathbb{R}^n$ be given and $G \subset Q$. Suppose also that $\nu(Q) \lesssim \ell(Q)^m$. If $\|1_G V_{\nu,Q} f\|_{L^2(\nu)} \lesssim \|f\|_{L^2(\nu)}$ for every $f \in L^2(\nu)$ satisfying spt $f \subset G$, then also $\|1_G V_\nu f\|_{L^2(\nu)} \lesssim \|f\|_{L^2(\nu)}$ for every $f \in L^2(\nu)$ satisfying spt $f \subset G$.*

Proof. This follows from the Exercise 3 in Set 3. □

Next, we prove the main Proposition.

6.5. **Proposition.** *Let μ be a measure of order m and $B_1, B_2 < \infty$, $\epsilon_0 \in (0,1)$ be given constants. Let also $(s_t)_{t>0}$ be an m-LP-family, and V be the corresponding vertical square function. Let $Q \subset \mathbb{R}^n$ be a fixed cube. Assume that there exists a complex measure $\nu = \nu_Q$ such that*

(1) spt $\nu \subset Q$;
(2) $\mu(Q) = \nu(Q)$;
(3) $\|\nu\| \le B_1 \mu(Q)$;
(4) For all Borel sets $A \subset Q$ satisfying $\mu(A) \le \epsilon_0 \mu(Q)$ we have

$$|\nu|(A) \le \frac{\|\nu\|}{32 B_1}.$$

Suppose there exist $s > 0$ and a Borel set $U_Q \subset \mathbb{R}^n$ for which $|\nu|(U_Q) \le \frac{\|\nu\|}{16B_1}$ so that

$$\sup_{\lambda > 0} \lambda^s \mu(\{x \in Q \setminus U_Q \colon\, V_Q\nu(x) > \lambda\}) \le B_2 \|\nu\|.$$

Then, there is some subset $G_Q \subset Q \setminus U_Q$ such that $\mu(G_Q) \gtrsim \mu(Q)$ and

$$\|1_{G_Q} V_\mu f\|_{L^2(\mu)} \lesssim \|f\|_{L^2(\mu)}$$

for every $f \in L^2(\mu)$ satisfying that spt $f \subset G_Q$.

Proof. We can assume that spt $\mu \subset Q$. Indeed, if we have proved the theorem for such measures, we can then apply it to $\mu\lfloor Q$. Let us denote $\sigma = |\nu|$, where $|\nu|$ is the variation measure of ν. Also, let us write the polar decomposition of the complex measure ν as $\nu = b\, d\sigma$, where b is a function so that $|b(x)| = 1$ always.

The idea is to apply the big pieces global Tb theorem from Section 5 (Theorem 5.2). It will be applied to the measure σ and the bounded function b. Using stopping times we need to construct some exceptional sets so that the assumptions of that theorem are verified. Moreover, we need to be able to come back to the μ measure – this requires encompassing additional stopping times to the construction.

We fix w, and write $\mathcal{D}(w) = \mathcal{D}$. We also write $\mathcal{D}_0 = \mathcal{D}(0)$. Let $\mathcal{A} = \mathcal{A}_w$ consist of the maximal dyadic cubes $R \in \mathcal{D}$ for which

$$\Big| \int_R b\, d\sigma \Big| < \eta \sigma(R),$$

where $\eta := \frac{1}{2} B_1^{-1}$. We set

$$T = T_w = \bigcup_{R \in \mathcal{A}} R \subset \mathbb{R}^n.$$

Notice that

$$\sigma(Q) = \|\nu\| \le B_1 \mu(Q) = B_1 \nu(Q) = B_1 \int_Q b\, d\sigma.$$

Then estimate

$$\int_Q b\, d\sigma = \Big| \int_Q b\, d\sigma \Big| = \Big| \int_{Q \setminus T} b\, d\sigma + \sum_{R \in \mathcal{A}} \int_R b\, d\sigma \Big| \le \sigma(Q \setminus T) + \eta \sigma(Q).$$

Since $\eta B_1 = 1/2$ we conclude that

$$\sigma(Q) \leq B_1 \sigma(Q \setminus T) + \frac{1}{2}\sigma(Q),$$

and so

$$\sigma(Q) \leq 2B_1[\sigma(Q) - \sigma(T)].$$

From here we can read that

$$\sigma(T) \leq (1-\eta)\sigma(Q).$$

Next, let $\mathcal{F}$ consist of the maximal dyadic cubes $R \in \mathcal{D}_0$ for which

$$\sigma(R) > \frac{B_1}{\epsilon_0}\mu(R)$$

or

$$\sigma(R) < \delta\mu(R),$$

where $\delta := \eta/16 = \frac{1}{32}B_1^{-1}$. Let $\mathcal{F}_1$ be the collection of maximal cubes $R \in \mathcal{D}_0$ satisfying the first condition, and define $\mathcal{F}_2$ analogously. Note that

$$\mu\Big(\bigcup_{R\in\mathcal{F}_1} R\Big) \leq \epsilon_0\mu(Q),$$

so that we have by assumption (4) that

$$\sigma\Big(\bigcup_{R\in\mathcal{F}_1} R\Big) \leq \frac{1}{32B_1}\sigma(Q) = \delta\sigma(Q).$$

Finally, we record that

$$\sigma\Big(\bigcup_{R\in\mathcal{F}_2} R\Big) = \sum_{R\in\mathcal{F}_2}\sigma(R) \leq \delta\sum_{R\in\mathcal{F}_2}\mu(R) = \delta\mu\Big(\bigcup_{R\in\mathcal{F}_2} R\Big) \leq \delta\mu(Q) \leq \delta\sigma(Q).$$

We may conclude that the set

$$H_1 = \bigcup_{R\in\mathcal{F}} R$$

satisfies $\sigma(H_1) \leq 2\delta\sigma(Q) = \frac{\eta}{8}\sigma(Q)$.

We now record the important property of the exceptional set H_1. Let $x \in Q\setminus H_1$. For any $R \in \mathcal{D}_0$ satisfying that $x \in R$ we have that

$$\frac{1}{32B_1} = \delta \leq \frac{\sigma(R)}{\mu(R)} \leq \frac{B_1}{\epsilon_0}.$$

From this we can conclude (using a dyadic variant of Lemma 2.13 of [2] or Lemma A.5; see also Exercise 5 in Set 2) that for all Borel sets $A \subset \mathbb{R}^n$ there holds that

$$\delta\mu(A\cap(Q\setminus H_1)) \leq \sigma(A\cap(Q\setminus H_1)) \leq \frac{B_1}{\epsilon_0}\mu(A\cap(Q\setminus H_1)).$$

In particular, we have that $\sigma\lfloor(Q \setminus H_1) \ll \mu\lfloor(Q \setminus H_1)$. Using Radon–Nikodym theorem we let $\varphi \geq 0$ be a function so that

$$\sigma(A) = \int_A \varphi \, d\mu$$

for all Borel sets $A \subset Q \setminus H_1$. We obviously have that $\varphi \sim 1$ for μ-a.e. $x \in Q \setminus H_1$.

We need another exceptional set H_2. To this end, let

$$p(x) = \sup_{r>0} \frac{\sigma(B(x,r))}{r^m} =: M_{R,m}\nu(x).$$

For $p_0 > 0$ let $E_{p_0} = \{p \geq p_0\}$. Using that $M_{R,m} \colon M(\mathbb{R}^n) \to L^{1,\infty}(\mu)$ boundedly we see that

$$\mu(E_{p_0}) = \mu(\{M_{R,m}\nu \geq p_0\}) \leq \frac{C}{p_0}\|\nu\| \leq \frac{CB_1}{p_0}\mu(Q).$$

We fix $p_0 \lesssim 1$ so large that $\mu(E_{p_0/2^m}) \leq \epsilon_0\mu(Q)$, so that in particular $\sigma(E_{p_0/2^m}) \leq \frac{\eta}{8}\sigma(Q)$. For $x \in \{p > p_0\}$ define

$$r(x) = \sup\{r > 0 \colon \sigma(B(x,r)) > p_0 r^m\},$$

and then set

$$H_2 := \bigcup_{x \in \{p > p_0\}} B(x, r(x)).$$

It is clear that every ball B_r with $\sigma(B_r) > p_0 r^m$ satisfies $B_r \subset H_2$. Notice that if $y \in H_2$, then there is $x \in \{p > p_0\}$ so that $y \in B(x, r(x))$, and so $\sigma(B(y, 2r(x)) \geq \sigma(B(x, r(x)) \geq p_0 r(x)^m = p_0 2^{-m}[2r(x)]^m$. We conclude that $H_2 \subset E_{p_0/2^m}$, and so $\sigma(H_2) \leq \frac{\eta}{8}\sigma(Q)$.

The assumption about the set U_Q reads $\sigma(U_Q) \leq \frac{\eta}{8}\sigma(Q)$. Define now $H = H_1 \cup H_2 \cup U_Q$. The properties of H are as follows:

(1) We have $\sigma(H) \leq \frac{\eta}{2}\sigma(Q)$, and so $\sigma(H \cup T_w) \leq (1 - \eta/2)\sigma(Q) = \tau_1\sigma(Q)$, $\tau_1 < 1$.
(2) If $\sigma(B_r) > p_0 r^m$, then $B_r \subset H$.
(3) We have a function φ so that

$$\sigma(A) = \int_A \varphi \, d\mu$$

for all Borel sets $A \subset Q \setminus H$, and $\varphi \sim 1$ for μ-a.e. $x \in Q \setminus H$.

We also have for every $\lambda > 0$ that

$$\begin{aligned}
&\lambda^s\sigma(\{x \in Q \setminus H \colon V_{\sigma,Q}b(x) > \lambda\}) \\
&\quad = \lambda^s\sigma(\{x \in Q \setminus H \colon V_Q\nu(x) > \lambda\}) \\
&\quad \lesssim \lambda^s\mu(\{x \in Q \setminus U_Q \colon V_Q\nu(x) > \lambda\}) \leq B_2\|\nu\| = B_2\sigma(Q).
\end{aligned}$$

Appealing to Theorem 5.2 with the measure σ and the L^∞ function b we find $G_Q \subset Q \setminus H \subset Q \setminus U_Q$ so that $\sigma(G_Q) \gtrsim \sigma(Q)$ and

$$\|1_{G_Q} V_{\sigma,Q} f\|_{L^2(\sigma)} \lesssim \|f\|_{L^2(\sigma)} \tag{6.6}$$

for every $f \in L^2(\sigma)$.

Suppose now that $g \in L^2(\mu)$ and $\operatorname{spt} g \subset G_Q$. We apply Equation (6.6) with $f = g/\varphi$ (since $G_Q \subset Q \setminus H$ we have $\varphi \sim 1$ μ-a.e. on the support of g). Notice that

$$\|1_{G_Q} V_{\sigma,Q}(g/\varphi)\|_{L^2(\sigma)} = \|1_{G_Q} V_{\mu,Q} g\|_{L^2(\sigma)} \gtrsim \|1_{G_Q} V_{\mu,Q} g\|_{L^2(\mu)}$$

so that

$$\|1_{G_Q} V_{\mu,Q} g\|_{L^2(\mu)} \lesssim \|g/\varphi\|_{L^2(\sigma)} \lesssim \|g\|_{L^2(\mu)}.$$

Applying Lemma 6.4 we conclude that

$$\|1_{G_Q} V_\mu f\|_{L^2(\mu)} \lesssim \|f\|_{L^2(\mu)}$$

for every $f \in L^2(\mu)$ satisfying that $\operatorname{spt} f \subset G_Q$. Moreover, we have that

$$\mu(Q) \le \sigma(Q) \lesssim \sigma(G_Q) = \int_{G_Q} \varphi \, d\mu \lesssim \mu(G_Q).$$

We are done. □

We are ready to prove the local Tb theorem.

Proof of Theorem 6.1. Proposition 6.5 gives for every $(2,\beta)$-doubling cube $Q \subset \mathbb{R}^n$ with C_1-small boundary a subset $G_Q \subset Q$ such that $\mu(G_Q) \gtrsim \mu(Q)$ and

$$\|1_{G_Q} V_\mu f\|_{L^2(\mu)} \lesssim \|f\|_{L^2(\mu)}$$

for every $f \in L^2(\mu)$ with $\operatorname{spt} f \subset G_Q$. Applying the non-homogeneous good lambda method i.e. Theorem 4.1 and Remark 4.2 gives the result. □

Appendix A. Some standard results from geometric and harmonic analysis

A.1. Covering theorems.

A.1. Theorem ($5r$-covering theorem). *Let $\mathcal{B}$ be a family of either closed or open balls (or cubes) in $\mathbb{R}^n$ such that*

$$\sup_{B \in \mathcal{B}} \operatorname{diam}(B) < \infty.$$

Then there exists $B_1, B_2, \ldots \in \mathcal{B}$ so that $B_i \cap B_j = \emptyset$ for $i \neq j$, and

$$\bigcup_{B \in \mathcal{B}} B \subset \bigcup_i 5B_i.$$

A.2. Theorem (Besicovitch covering theorem). *Suppose $A \subset \mathbb{R}^n$ is bounded and that for every $x \in A$ we are given some closed ball (or cube) B_x centred at x. Then there are $\{B_i\}_i \subset \{B_x\}_{x \in A}$ so that*

$$A \subset \bigcup_i B_i \qquad \text{and} \qquad \sum_i 1_{B_i} \le C_n,$$

where $C_n < \infty$ is a purely dimensional constant.

A.2. **Absolute continuity, derivation of measures, Radon–Nikodym.** Consider two Radon measures μ and σ in $\mathbb{R}^n$. We say that σ is absolutely continuous with respect to μ if $\mu(A) = 0$ implies $\sigma(A) = 0$. This is denoted $\sigma \ll \mu$.

A.3. **Example.** Suppose μ is given and $f \geq 0$ is locally integrable. Define

$$\sigma = f\, d\mu \qquad i.e. \qquad \sigma(A) = \int_A f\, d\mu.$$

Then clearly $\sigma \ll \mu$.

An extremely useful result, the Radon–Nikodym theorem, says that all absolutely continuous measures arise in this way.

A.4. **Theorem** (Radon–Nikodym)**.** *Consider two Radon measures μ and σ in $\mathbb{R}^n$. Suppose that $\sigma \ll \mu$. Then*

$$\sigma(A) = \int_A f\, d\mu$$

for all Borel sets $A \subset \mathbb{R}^n$, where f can be defined for μ-a.e. x by

$$f = D(\sigma, \mu, x) := \lim_{r \to 0} \frac{\sigma(B(x,r))}{\mu(B(x,r))}.$$

The most natural way to check whether $\sigma \ll \mu$ is as follows.

A.5. **Lemma.** *Define*

$$\underline{D}(\sigma, \mu, x) := \liminf_{r \to 0} \frac{\sigma(B(x,r))}{\mu(B(x,r))}.$$

Suppose that $A \subset \mathbb{R}^n$ is a Borel set so that for some constant $\lambda > 0$ we have

$$\sup_{x \in A} \underline{D}(\sigma, \mu, x) \leq \lambda.$$

Then $\sigma(A) \leq \lambda \mu(A)$. In particular, if $B \subset A$ and $\mu(B) = 0$ then $\sigma(B) = 0$, i.e. $\sigma\lfloor A \ll \mu\lfloor A$.

For the proofs of these results, see Section 2 of [2]. These results (and more about absolute continuity) are also proved in the course Real Analysis II. We will also consider some details in the exercises.

References

[1] H. Martikainen, M. Mourgoglou, E. Vuorinen, A new approach to non-homogeneous local Tb theorems: Square functions and weak $(1,1)$ testing with measures, preprint, arXiv:1511.00528, 2015.

[2] P. Mattila, Geometry of sets and measures in Euclidean spaces: Fractals and rectifiability, Cambridge Studies in Advanced Mathematics, Vol. 44, Cambridge University Press, Cambridge, 1995.

[3] W. Rudin, Real and Complex Analysis, McGraw-Hill, New York, 1966.

[4] X. Tolsa, Analytic capacity, the Cauchy transform, and non-homogeneous Calderón–Zygmund theory, Progress in Mathematics, Vol. 307, Birkhäuser Verlag, Basel, 2014.

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

E-mail address: henri.martikainen@helsinki.fi