BOUNDEDNESS OF NON-HOMOGENEOUS SQUARE FUNCTIONS AND L^q TYPE TESTING CONDITIONS WITH $q \in (1,2)$

HENRI MARTIKAINEN AND MIHALIS MOURGOGLOU

ABSTRACT. We continue the study of local Tb theorems for square functions defined in the upper half-space $(\mathbb{R}^{n+1}_+, \mu \times dt/t)$. Here μ is allowed to be a non-homogeneous measure in $\mathbb{R}^n$. In this paper we prove a boundedness result assuming local L^q type testing conditions in the difficult range $q \in (1,2)$. Our theorem is a non-homogeneous version of a result of S. Hofmann valid for the Lebesgue measure. It is also an extension of the recent results of M. Lacey and the first named author where non-homogeneous local L^2 testing conditions have been considered.

1. INTRODUCTION

We study the boundedness of the vertical square function

$$Vf(x) = \left(\int_0^\infty |\theta_t f(x)|^2 \, \frac{dt}{t} \right)^{1/2}.$$

Here the linear operators θ_t, $t > 0$, have the form

$$\theta_t f(x) = \int_{\mathbb{R}^n} s_t(x,y) f(y) \, d\mu(y). \tag{1.1}$$

The appearing measure μ is a Borel measure in $\mathbb{R}^n$ which is only assumed to satisfy, for some m, the upper bound

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

Moreover, for some $\alpha > 0$, the kernels s_t satisfy the size and continuity conditions

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

and

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

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

2010 *Mathematics Subject Classification.* 42B20.

Key words and phrases. Square function, non-homogeneous analysis, local Tb, L^q test functions.

Research of H.M. is supported by the Academy of Finland through the grant Multiparameter dyadic harmonic analysis and probabilistic methods. H.M. would like to thank Michael Lacey and the School of Mathematics of Georgia Tech, where part of this research was carried, for their hospitality.

Research of M.M. is supported by the Institut des Hautes Études Scientifiques and was partially supported by the ANR project "Harmonic analysis at its boundaries" ANR-12-BS01-0013-01. The hospitality of T. Hytönen and the Department of Mathematics and Statistics, University of Helsinki, where part of this research was carried, is gratefully acknowledged by M.M.

The following is our main theorem.

1.4. **Theorem.** *Let $q \in (1,2)$ be a fixed number. Assume that to every open cube $Q \subset \mathbb{R}^n$ with $\mu(\partial Q) = 0$ there is associated a test function b_Q satisfying that*

(1) *spt $b_Q \subset Q$;*
(2) *$\int_Q b_Q \, d\mu = \mu(Q)$;*
(3) *$\|b_Q\|_{L^q(\mu)}^q \lesssim \mu(Q)$;*
(4)
$$\int_Q \Big(\int_0^{\ell(Q)} |\theta_t b_Q(x)|^2 \frac{dt}{t} \Big)^{q/2} d\mu(x) \lesssim \mu(Q).$$

Then for every $p \in (1,q]$ we have that

$$\|V\|_{L^p(\mu) \to L^p(\mu)} \lesssim 1 + V_{\mathrm{loc},q},$$

where

$$V_{\mathrm{loc},q} = \Big[\sup_Q \frac{1}{\mu(Q)} \int_Q \Big(\int_0^{\ell(Q)} |\theta_t b_Q(x)|^2 \frac{dt}{t} \Big)^{q/2} d\mu(x) \Big]^{1/q}$$

and the implicit constant depends only on n, m, q, p, the kernel constants and the constant in the testing condition (3).

Suppose we, in addition, have the x-continuity of the kernels s_t, i.e.,

$$|s_t(x,y) - s_t(z,y)| \lesssim \frac{|x-z|^\alpha}{(t+|x-y|)^{m+\alpha}} \tag{1.5}$$

whenever $|x-z| < t/2$. Then for every $p \in (1,\infty)$ we have that

$$\|V\|_{L^p(\mu) \to L^p(\mu)} \lesssim 1 + V_{\mathrm{loc},q}.$$

After some reductions the proof boils down to proving the L^p estimate for $p = q$. However, the conclusion of the theorem for $p = 2$ under x-continuity can be considered to be the main point. As such it is an extremely general form of a local Tb theorem with non-homogeneous measures and non-scale-invariant $L^{1+\epsilon}$ type testing conditions. It should also be noted that an example from [23] shows that when dealing with the vertical square function (as we are here) one cannot derive the $L^2(\mu)$ estimate from the $L^q(\mu)$ estimate without x-continuity. This fails even in the case that μ is the Lebesgue measure.

Hofmann [11] proved the L^2 boundedness of the square function under these local L^q testing conditions (and x-continuity) in the case that μ is the Lebesgue measure. Our proof in this general context is completely different, and we find it necessary to first work directly in $L^q(\mu)$ rather than in $L^2(\mu)$. To this end, we need to begin by establishing that

$$\|V\|_{L^q(\mu) \to L^q(\mu)} \lesssim 1 + \sup_Q \Big[\frac{1}{\mu(10Q)} \int_Q \Big(\int_0^{\ell(Q)} |\theta_t 1_Q(x)|^2 \frac{dt}{t} \Big)^{q/2} d\mu(x) \Big]^{1/q}.$$

We will then bound this Carleson norm of L^q type by $V_{\mathrm{loc},q}$. Hofmann, however, uses a $T1$ in L^2 and bounds, by a clever argument, a Carleson norm of L^2 type by a Carleson norm of L^q, $q < 2$, type. We find his argument to be specific to the doubling situation. Our strategy has the additional bonus of giving the L^q result without x-continuity.

We also rely on and extend the very recent non-homogeneous twisted martingale bounds by Lacey and one of us [18]. On the other hand, we develop an L^q analog of

the good Whitney averaging trick first used by us [22]. Lastly, we make extensive use of the L^p, $p \neq 2$, techniques, including Stein's inequality. The complete outline of the proof is given in Subsection 1.1.

Very recently Lacey and the first named author [19] managed to prove the L^2 boundedness in the non-homogeneous case but only with local L^2 testing conditions. Our main theorem is an extension of these two state of the art results [11], [19]. Indeed, we consider general measures and general exponents simultaneously. The aforementioned two references are the most obvious predecessors of our main theorem, but the whole story up to this point is rather long.

One can consider Tb theorems at least for square functions and Calderón–Zygmund operators. Then they can be global or local. And if they are local, they can be with the easier L^∞/BMO/$T^{2,\infty}$ type testing assumptions, or with the more general L^s, $s < \infty$, type assumptions. Moreover, in the latter case the range of the exponents (in the Calderón–Zygmund world more than one set of testing functions appear) one can use is a very significant problem. Lastly, the fact that whether one considers the homogeneous or non-homogeneous theory is a major factor. All of these big story arcs are relevant for the context of the current paper. We now try to give at least some of the key references of local Tb theorems.

The first local Tb theorem, with L^∞ control of the test functions and their images, is by Christ [7]. Nazarov, Treil and Volberg [24] proved a non-homogeneous version of this theorem. The point compared to the global Tb theorems is as follows. The accretivity of a given test function b_Q is only assumed on its supporting cube Q, i.e., $|\int_Q b_Q \, d\mu| \gtrsim \mu(Q)$. While in a global Tb one needs a function which is simultaneously accretive on all scales. But the remaining conditions are still completely scale invariant: $b_Q \in L^\infty(\mu)$ and $Tb_Q \in L^\infty(\mu)$. This scale invariance of the testing conditions is the main thing one wants to get rid of.

The non-scale-invariant L^s type testing conditions were introduced by Auscher, Hofmann, Muscalu, Tao and Thiele [3]. Their theorem is for perfect dyadic singular integral operators and the assumptions are of the form $\int_Q |b_Q^1|^p \lesssim |Q|$, $\int_Q |b_Q^2|^q \lesssim |Q|$, $\int_Q |Tb_Q^1|^{q'} \lesssim |Q|$ and $\int_Q |T^* b_Q^2|^{p'} \lesssim |Q|$, $1 < p, q \leq \infty$. Extending the result to general Calderón–Zygmund operators is complicated (it is almost done by now – but not completely). Hofmann [10] established the result for general operators but only assuming the existence of $L^{2+\epsilon}$ test functions mapping to L^2. Auscher and Yang [5] removed the ϵ by proving the theorem in the sub-dual case $1/p + 1/q \leq 1$. Auscher and Routin [4] considered the general case under some additional assumptions. The full super-dual case $1/p + 1/q > 1$ is by Hytönen and Nazarov [17], but even then with the additional buffer assumption $\int_{2Q} |Tb_Q^1|^{q'} \lesssim |Q|$ and $\int_{2Q} |T^* b_Q^2|^{p'} \lesssim |Q|$.

This was the main story for the Calderón–Zygmund operators for doubling measures. For square functions the situation is a bit more clear with the need for only one exponent q. The case $q = 2$ is implicit in the Kato square root papers [2, 12, 13] and explicitly stated and proved in [1] and [9]. The case $q > 2$ is weaker than this. The hardest case $q \in (1, 2)$ is due to Hofmann [11] as already mentioned. Some key applications really need the fact that one can push the integrability of the test functions to $1 + \epsilon$ (see again [11]).

The non-homogeneous world is yet another story. The whole usage of these non-scale-invariant testing conditions is a huge source of problem in this context. One reason lies in the fact that even if we have performed a stopping time argument which gives us that

a fixed test function b_F behaves nicely on a cube Q, for example that $\int_Q |b_F|^2 \, d\mu \lesssim \mu(Q)$, we cannot say much what happens in the stopping children of Q. That is, in a stopping child Q' of Q we cannot use the simple argument

$$\int_{Q'} |b_F|^2 \, d\mu \leq \int_Q |b_F|^2 \, d\mu \lesssim \mu(Q) \lesssim \mu(Q')$$

which would only be available if μ were doubling. The non-homogeneous case $q = 2$ for square functions is the very recent work of Lacey and the first named author [19]. The case $p = q = 2$ for Calderón–Zygmund operators is by the same authors [18]. For relevant dyadic techniques see also the Lacey–Vähäkangas papers [20, 21] and Hytönen–Martikainen [16]. To recap the context, in this paper we consider non-homogeneous square functions and push q to the range $q \in (1, 2)$.

We still mention that the study of the boundedness of non-homogeneous square functions was initiated by the recent authors in [22]. This was a global Tb. The key technique was the usage of good (in a probabilistic sense) Whitney regions. A scale invariant local Tb is by the current authors together with T. Orponen [23]. In that paper we also study the end point theory, L^p theory, and various counter-examples (e.g. the failure of the change of aperture with general measures and the difference between conical and vertical square functions).

1.1. **Outline of the proof.** Let us give a brief point-by-point outline of the proof:

(1) We prove a $T1$ in $L^q(\mu)$ which states that

$$\|V\|_{L^q(\mu) \to L^q(\mu)} \lesssim 1 + \mathrm{Car}_V(q, 10),$$

where

$$\mathrm{Car}_V(q, 10) = \sup_Q \Big[\frac{1}{\mu(10Q)} \int_Q \Big(\int_0^{\ell(Q)} |\theta_t 1_Q(x)|^2 \frac{dt}{t} \Big)^{q/2} d\mu(x) \Big]^{1/q}.$$

(2) Using the above we reduce to bounded functions $|f| \leq 1$ in various martingale estimates. This is key for us. Indeed, we need to be able to control norms of the type

$$\Big\| \Big(\sum_k |\Delta_k f|^2 \Big)^{1/2} \Big\|_{L^q(\mu)},$$

where the (twisted) martingales Δ_k are constructed (in a dyadic grid) via stopping times from the test function b_Q. This is very delicate with general measures and such general test function. Indeed, estimates of this type have only very recently been established by Lacey and one of us [18] in the $q = 2$ situation.

(3) Sufficient control of such martingales for bounded functions is then achieved by a non-homogeneous John–Nirenberg type argument. This generalises the arguments in [18] from the L^2 case to the L^q case.

(4) The estimation of $\mathrm{Car}_V(q, 10)$ by $V_{\mathrm{loc},q}$ starts by a probabilistic good Whitney averaging argument in L^q. This requires some rethinking compared to the $q = 2$ case of [22]. Essentially, we fix an arbitrary cube $Q_0 \subset \mathbb{R}^n$ with side-length $\ell(Q_0) \sim 2^s$ and an arbitrary function f with $|f| \leq 1_{Q_0}$. The good Whitney means that we

reduce to bounding

$$\Big[\int_{Q_0}\Big(\sum_{\substack{R\in\mathcal{D}_{\mathrm{good}}\\ \ell(R)\le 2^s}} 1_R(x)\int_{\ell(R)/2}^{\ell(R)}|\theta_t f(x)|^2\,\frac{dt}{t}\Big)^{q/2}d\mu(x)\Big]^{1/q}$$

by $C(1+V_{\mathrm{loc},q})\mu(10Q_0)^{1/q}$, where $\mathcal{D}$ is an arbitrary fixed dyadic grid.

(5) We then expand $f=\sum_{Q\in\mathcal{D}}\Delta_Q f$ using the constructed twisted martingales. This sum is not restricted to good cubes. Still, the existence of common dyadic parents of Q and R of controlled size is needed. This is because the L^q, $q\neq 2$, theory requires the usage of Stein's inequality, and the needed martingale structure is constructed via these common ancestors. But only the goodness of R is used for this (and for multiple other bounds). This is special to the square function case and is a massive technical simplification: we avoid showing that twisted martingale decompositions restricted to bad cubes would be, on average, small in the L^q norm, and our paraproducts readily collapse (see Remark 4.1 of [16] and Remark 2.14 of [21]).

(6) After completing the estimation of various series we conclude the $p=q$ case of our theorem. The general case is achieved by the non-homogeneous vector-valued theory of Calderón–Zygmund operators by García-Cuerva and Martell [8].

We conclude the introduction by additional notes, which are collected to the following remark. This can be safely skipped in the first reading. We then indicate, as an extremely brief example, an interesting connection of square functions to geometry. After that we set up some notation.

1.6. *Remark.* The estimate $\|V\|_{L^q(\mu)\to L^q(\mu)}\lesssim 1+V_{\mathrm{loc},q}$ can be proved assuming only that $\mu(B(x,r))\le\lambda(x,r)$ for some $\lambda\colon\mathbb{R}^n\times(0,\infty)\to(0,\infty)$ satisfying that $r\mapsto\lambda(x,r)$ is non-decreasing and $\lambda(x,2r)\le C_\lambda\lambda(x,r)$ for all $x\in\mathbb{R}^n$ and $r>0$. In this case one only needs to replace the kernel estimates by

$$|s_t(x,y)|\lesssim\frac{t^\alpha}{t^\alpha\lambda(x,t)+|x-y|^\alpha\lambda(x,|x-y|)}$$

and

$$|s_t(x,y)-s_t(x,z)|\lesssim\frac{|y-z|^\alpha}{t^\alpha\lambda(x,t)+|x-y|^\alpha\lambda(x,|x-y|)}$$

whenever $|y-z|<t/2$. This is done in the global situation in [22]. Here we skip the required modifications. Such formalism lets one capture the doubling theory as a by-product, and allows some more general upper bounds than r^m.

We also point out that the assumption (1), i.e. $\operatorname{spt} b_Q\subset Q$, is not necessary. But then one does need to understand that the assumption (3) reads $\int_{\mathbb{R}^n}|b_Q|^q\,d\mu\lesssim\mu(Q)$. A careful reader can see that everything goes through as written except that Section 8 requires some small technical modifications. One can assume that the dyadic grid $\mathcal{D}$ one is working with has the following "no quadrants" property: Whenever $I_k\in\mathcal{D}$, $k\in\mathbb{N}$, is a strictly increasing sequence (meaning that $I_k\subsetneq I_{k+1}$) of dyadic cubes, then this sequence exhausts all of $\mathbb{R}^n$ (meaning that $\bigcup_{k\in\mathbb{N}}I_k=\mathbb{R}^n$). One can assume this since almost every random dyadic grid is such. Then it is easy to see that Lemma 8.2 remains true. This is almost everything one needs to change in the proof, we leave the details as an exercise.

Let us also make the standard remark that the assumption (2) can be replaced by $|\int_Q b_Q\,d\mu| \gtrsim \mu(Q)$. One then just considers the scaled test functions

$$\tilde{b}_Q = \frac{\mu(Q)}{\int_Q b_Q\,d\mu} b_Q.$$

1.7. **Example.** We briefly point out an interesting connection of the boundedness of square functions (our topic) and geometry. Let $E \subset \mathbb{R}^n$ be a closed set which is m-ADR for some integer $0 < m < n$, i.e., $\mu := \mathcal{H}^m|_E$ satisfies $\mu(B(x,r)) \approx r^m$ for all $x \in E$ and $r \in (0, \operatorname{diam}(E))$. Instead of going through the extensive literature of uniformly rectifiable sets we content by just citing the very recent result of Chousionis, Garnett, Le and Tolsa [6]. Let $s_t(x,y) = t\partial_t[t^{-m}\phi([x-y]/t)]$, where $\phi(x) = (1+|x|^2)^{-(m+1)/2}$. This kernel satisfies (1.2), (1.3) and (1.5). Let V be the corresponding square function. One of the results of [6] says that E is uniformly m-rectifiable (for the definition see [6]) if and only if V is $L^2(\mu)$ bounded.

Notation. We write $A \lesssim B$, if there is a constant $C > 0$ so that $A \leq CB$. We may also write $A \approx B$ if $B \lesssim A \lesssim B$. For a number a we write $a \sim 2^k$ if $2^k \leq a < 2^{k+1}$.

We then set some dyadic notation. Consider a dyadic grid $\mathcal{D}$ in $\mathbb{R}^n$. For $Q, R \in \mathcal{D}$ we use the following notation:

- $\ell(Q)$ is the side-length of Q;
- $d(Q,R)$ denotes the distance between the cubes Q and R;
- $D(Q,R) := d(Q,R) + \ell(Q) + \ell(R)$ is the long distance;
- $W_Q = Q \times (\ell(Q)/2, \ell(Q))$ is the Whitney region associated with Q;
- $\operatorname{ch}(Q) = \{Q' \in \mathcal{D} : Q' \subset Q, \ell(Q') = \ell(Q)/2\}$;
- $\operatorname{gen}(Q)$ is determined by $\ell(Q) = 2^{\operatorname{gen}(Q)}$;
- $Q^{(k)} \in \mathcal{D}$ is the unique cube for which $\ell(Q^{(k)}) = 2^k \ell(Q)$ and $Q \subset Q^{(k)}$;
- $\langle f \rangle_Q = \mu(Q)^{-1} \int_Q f\,d\mu$.

2. Reductions and start of the proof

2.1. **Reduction to the case $p = q$.** Suppose we have proved the bound $\|V\|_{L^q(\mu)\to L^q(\mu)} \lesssim 1 + V_{\mathrm{loc},q}$ already. Define $\mathbb{A} = \mathbb{C}$ and $\mathbb{B} = L^2((0,\infty), dt/t)$. Let $K\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{B} = \mathcal{L}(\mathbb{A}, \mathbb{B})$ be defined by

$$K(x,y) = [t \mapsto s_t(x,y)],$$

and let $T\colon L^q(\mu) = L^q_{\mathbb{A}}(\mu) \to L^q_{\mathbb{B}}(\mu)$ be defined by

$$Tf(x) = \int_{\mathbb{R}^n} K(x,y) f(y)\,d\mu(y) = [t \mapsto \theta_t f(x)].$$

It is not hard to see that

$$\|K(x,y)\|_{\mathbb{B}} \lesssim \frac{1}{|x-y|^m}$$

and

$$\|K(x,y) - K(x,z)\|_{\mathbb{B}} \lesssim \frac{|y-z|^{\alpha/2}}{|x-y|^{m+\alpha/2}}, \qquad 2|y-z| \leq |x-y|.$$

The non-homogeneous vector-valued theory of Calderón–Zygmund operators by García-Cuerva and Martell [8] yields that $\|T\|_{L^p(\mu)\to L^p_{\mathbb{B}}(\mu)} \lesssim 1 + V_{\mathrm{loc},q}$ for $p \in (1,q]$. If we have

the x-continuity of the kernels s_t, then we also have that

$$\|K(x,y)-K(z,y)\|_{\mathbb{B}} \lesssim \frac{|x-z|^{\alpha/2}}{|x-y|^{m+\alpha/2}}, \qquad 2|x-z| \le |x-y|.$$

The bound $\|T\|_{L^p(\mu)\to L^p_{\mathbb{B}}(\mu)} \lesssim 1+V_{\mathrm{loc},q}$ holds for all $p\in(1,\infty)$ in this case. Since we have that $\|Tf\|_{L^p_{\mathbb{B}}(\mu)} = \|Vf\|_{L^p(\mu)}$, the L^p theory of V follows from the $p=q$ case.

2.2. **Reduction to a priori bounded operators V.** In this subsection we say the following. Suppose we have proved the $L^q(\mu)$ bound of Theorem 1.4, i.e., the quantitative bound $\|V\|_{L^q(\mu)\to L^q(\mu)} \lesssim 1+V_{\mathrm{loc},q}$, under the additional a priori finiteness assumption $\|V\|_{L^q(\mu)\to L^q(\mu)}<\infty$. Then the $L^q(\mu)$ bound of Theorem 1.4 automatically follows without the a priori assumption.

To this end, define $s^i_t(x,y)=s_t(x,y)$ if $1/i\le t\le i$, and $s^i_t(x,y)=0$ otherwise. These kernels are clearly in our original class – they satisfy (1.2) and (1.3) with kernel constants bounded by those of V. Define

$$V_i f(x) := \left(\int_{1/i}^{i} |\theta_t f(x)|^2 \,\frac{dt}{t}\right)^{1/2} = \left(\int_0^{\infty} |\theta^i_t f(x)|^2 \,\frac{dt}{t}\right)^{1/2},$$

where

$$\theta^i_t f(x) = \int_{\mathbb{R}^n} s^i_t(x,y) f(y)\, d\mu(y).$$

Let us note that the V_i are bounded operators on $L^q(\mu)$. Let

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

This centred maximal function is a bounded operator on $L^p(\mu)$ for every $p\in(1,\infty)$. Notice that $|\theta_t f(x)| \lesssim M_\mu f(x)$ for every $t>0$ and $x\in\mathbb{R}^n$. Using this we see that $\|V_i\|_{L^q(\mu)\to L^q(\mu)} \le [2\log i]^{1/2}\|M_\mu\|_{L^q(\mu)\to L^q(\mu)}<\infty$.

By monotone convergence we have that

$$\begin{aligned}
\|Vf\|_{L^q(\mu)} &= \lim_{i\to\infty} \|V_i f\|_{L^q(\mu)} \\
&\le \limsup_{i\to\infty} \|V_i\|_{L^q(\mu)\to L^q(\mu)} \|f\|_{L^q(\mu)} \\
&\lesssim \limsup_{i\to\infty} (1+V^i_{\mathrm{loc},q}) \|f\|_{L^q(\mu)} \\
&\le (1+V_{\mathrm{loc},q})\|f\|_{L^q(\mu)}.
\end{aligned}$$

So it suffices to prove Theorem 1.4 under the assumption $\|V\|_{L^q(\mu)\to L^q(\mu)}<\infty$ – a piece of information that will be used purely in a qualitative way.

2.3. **Reduction to a q-Carleson estimate.** We begin by stating a $T1$ in $L^q(\mu)$ (the case $q=2$ is in [22]). The proof of this $T1$ is indicated in Appendix A. Define, say for $\lambda\ge 3$,

$$\mathrm{Car}_V(q,\lambda) := \sup_Q \Big[\frac{1}{\mu(\lambda Q)}\int_Q \Big(\int_0^{\ell(Q)} |\theta_t 1_Q(x)|^2 \frac{dt}{t}\Big)^{q/2} d\mu(x)\Big]^{1/q}$$

and

$$\widetilde{\mathrm{Car}}_V(q,\lambda) := \sup_Q \Big[\frac{1}{\mu(\lambda Q)}\int_Q \Big(\int_0^{\ell(Q)} |\theta_t 1(x)|^2 \frac{dt}{t}\Big)^{q/2} d\mu(x)\Big]^{1/q}.$$

Here the supremums run over all the open cubes $Q \subset \mathbb{R}^n$. Then, for $q \in (1, 2]$, we have that there holds that

$$\|V\|_{L^q(\mu)\to L^q(\mu)} \le C_1(1 + \widetilde{\mathrm{Car}}_V(q, 30)) \le C_2(1 + \mathrm{Car}_V(q, 10)). \tag{2.1}$$

Assuming the existence of the L^q test functions as in Theorem 1.4 we then prove that

$$\mathrm{Car}_V(q, 10) \le C_3(1 + V_{\mathrm{loc},q}) + C_2^{-1}\|V\|_{L^q(\mu)\to L^q(\mu)}/2. \tag{2.2}$$

We call this the key inequality. Combining (2.1) and (2.2) gives that

$$\|V\|_{L^q(\mu)\to L^q(\mu)} \le C(1 + V_{\mathrm{loc},q}) + \|V\|_{L^q(\mu)\to L^q(\mu)}/2$$

ending the proof.

We will now start the proof of the key inequality (2.2). This task is completed in Section 8. In Appendix A we indicate the proof of the $T1$ theorem in $L^q(\mu)$, i.e., the first estimate of (2.1).

3. Random and stopping cubes/ Martingale difference operators

3.1. **Random dyadic grids.** At this point we need to set up the basic notation for random dyadic grids (these facts are essentially presented in this way by Hytönen [14]).

Let $\mathcal{D}_0$ denote the standard dyadic grid, consisting of all the cubes of the form $2^k(\ell + [0, 1)^n)$, where $k \in \mathbb{Z}$ and $\ell \in \mathbb{Z}^n$. We also denote $\mathcal{D}_{0,k} = \{Q \in \mathcal{D}_0 : \ell(Q) = 2^k\}$. A generic dyadic grid, parametrized by $w \in (\{0, 1\}^n)^{\mathbb{Z}}$, is of the form $\mathcal{D}(w) = \cup_{k\in\mathbb{Z}}\mathcal{D}_k(w)$, where $\mathcal{D}_k(w) = \{Q_0 + x_k(w) : Q_0 \in \mathcal{D}_{0,k}\}$ and $x_k(w) = \sum_{j<k} w_j 2^j$. The notation $Q_0 + w := Q_0 + \sum_{j<k} w_j 2^j$, $Q_0 \in \mathcal{D}_{0,k}$, is convenient. We get random dyadic grids by placing the natural product probability measure P_w on $(\{0, 1\}^n)^{\mathbb{Z}}$ (thus the coordinate functions w_j are independent and $P_w(w_j = \eta) = 2^{-n}$ if $\eta \in \{0, 1\}^n$).

We fix the constant $\gamma \in (0, 1)$ to be so small that

$$\gamma \le \alpha/(2m + 2\alpha) \quad \text{and} \quad m\gamma/(1 - \gamma) \le \alpha/4,$$

where $\alpha > 0$ appears in the kernel estimates and m appears in $\mu(B(x, r)) \lesssim r^m$. A cube $R \in \mathcal{D}$ is called $\mathcal{D}$-bad if there exists another cube $Q \in \mathcal{D}$ so that $\ell(Q) \ge 2^r\ell(R)$ and $d(R, \partial Q) \le \ell(R)^\gamma\ell(Q)^{1-\gamma}$. Otherwise it is good. We denote the collections of good and bad cubes by $\mathcal{D}_{\mathrm{good}}$ and $\mathcal{D}_{\mathrm{bad}}$ respectively. The following properties are known (see e.g. [14]).

- For a fixed $Q_0 \in \mathcal{D}_0$ the set $Q_0 + w$ depends on w_j with $2^j < \ell(Q_0)$, while the goodness (or badness) of $Q_0 + w$ depends on w_j with $2^j \ge \ell(Q_0)$. In particular, these notions are independent (meaning that for any fixed $Q_0 \in \mathcal{D}_0$ the random variable $w \mapsto 1_{\mathrm{good}}(Q_0 + w)$ and any random variable that depends only on the cube $Q_0 + w$ as a set, like $w \mapsto \int_{Q_0+w} f\, d\mu$, are independent).
- The probability $\pi_{\mathrm{good}} := P_w(Q_0 + w \text{ is good})$ is independent of $Q_0 \in \mathcal{D}_0$.
- $\pi_{\mathrm{bad}} := 1 - \pi_{\mathrm{good}} \lesssim 2^{-r\gamma}$, with the implicit constant independent of r.

The parameter $r \lesssim 1$ is a fixed constant which is at least so large that $2^{r(1-\gamma)} \ge 100$.

The following lemma is stated without proof since the first part was proved on page 25 of [15] and the second is lemma 2.10 of [20].

3.1. **Lemma.** *Let $Q \in \mathcal{D}$ and $R \in \mathcal{D}_{\mathrm{good}}$, and set $\theta(u) := \left\lceil \frac{\gamma u + r}{1-\gamma} \right\rceil$, $u \in \mathbb{N}$.*

(1) *Assume* $\ell(Q) < \ell(R)$. *Let* $\ell(R)/\ell(Q) = 2^{\ell}$ *and* $D(Q,R)/\ell(R) \sim 2^{j}$ *for* $\ell \geq 1$ *and* $j \geq 0$. *Then, there holds that*

$$R \subset Q^{(\ell+j+\theta(j))}.$$

(2) *Assume* $\ell(R) \leq \ell(Q)$. *Let* $\ell(Q)/\ell(R) = 2^{\ell}$ *and* $D(Q,R)/\ell(Q) \sim 2^{j}$ *for* $\ell, j \geq 0$. *Then there holds that*

$$R \subset Q^{(j+\theta(j+\ell))}.$$

3.2. **Collections of stopping cubes.** Let $\mathcal{D}$ be a dyadic grid in $\mathbb{R}^n$ for which $\mu(\partial Q) = 0$ for all $Q \in \mathcal{D}$. It will be enough to work with such dyadic grids since for almost every $w \in (\{0,1\}^n)^{\mathbb{Z}}$ there holds that $\mathcal{D} = \mathcal{D}(w)$ satisfies this property. For $Q \in \mathcal{D}$ we may now set $b_Q := b_{\mathrm{int}(Q)}$, where the latter function exists by assumption. Now also b_Q satisfies the assumptions (1)-(4) listed in Theorem 1.4.

Let $Q^* \in \mathcal{D}$ be a fixed dyadic cube with $\ell(Q^*) = 2^s$. Set $\mathcal{F}^0_{Q^*} = \{Q^*\}$ and let $\mathcal{F}^1_{Q^*}$ be the (disjoint) collection of the maximal cubes $Q \in \mathcal{D}$, $Q \subset Q^*$, for which at least one of the following two conditions holds:

(1) $|\langle b_{Q^*}\rangle_Q| < 1/2$;
(2) $\langle |b_{Q^*}|^q\rangle_Q > 2^{q'+1}A^{q'}$.

Here A is a constant such that $\|b_R\|^q_{L^q(\mu)} \leq A\mu(R)$ for every R.

Next, we repeat the previous procedure by replacing Q^* with a fixed $Q \in \mathcal{F}^1_{Q^*}$. The combined collection of stopping cubes resulting from this is called $\mathcal{F}^2_{Q^*}$. This is continued and we set $\mathcal{F}_{Q^*} = \bigcup_{j=0}^{\infty} \mathcal{F}^j_{Q^*}$. Finally, for every $Q \in \mathcal{D}$, $Q \subset Q^*$, we let $Q^{\alpha} \in \mathcal{F}_{Q^*}$ be the minimal cube $R \in \mathcal{F}_{Q^*}$ for which $Q \subset R$.

3.2. **Lemma.** *If* $F \in \mathcal{F}^j_{Q^*}$ *for some* $j \geq 0$, *then there holds that*

$$\sum_{\substack{S \in \mathcal{F}^{j+1}_{Q^*} \\ S \subset F}} \mu(S) \leq \tau\mu(F), \ \tau := 1 - \frac{1}{2}\frac{A}{(2A)^{q'}} \in (0,1). \tag{3.3}$$

Proof. Let $F \in \mathcal{F}_{Q^*}$. Consider a disjoint collection $\{Q^1_i\}_i \subset \mathcal{D}$ for which $Q^1_i \subset F$ and $|\langle b_F\rangle_{Q^1_i}| < 1/2$. We have that

$$\begin{aligned}
\mu(F) = \int_F b_F\,d\mu &= \int_{F\setminus\bigcup_i Q^1_i} b_F\,d\mu + \sum_i \int_{Q^1_i} b_F\,d\mu \\
&\leq \mu\Big(F\setminus\bigcup_i Q^1_i\Big)^{1/q'}\Big(\int_F |b_F|^q\,d\mu\Big)^{1/q} + \frac{1}{2}\sum_i \mu(Q^1_i) \\
&\leq A^{1/q}\mu\Big(F\setminus\bigcup_i Q^1_i\Big)^{1/q'}\mu(F)^{1/q} + \frac{1}{2}\mu(F),
\end{aligned}$$

which implies that

$$\mu(F) \leq \big(2A^{1/q}\big)^{q'}\cdot\mu\Big(F\setminus\bigcup_i Q^1_i\Big) = \frac{(2A)^{q'}}{A}\Big[\mu(F) - \Big(\bigcup_i Q^1_i\Big)\Big].$$

Therefore, we obtain

$$\mu\Big(\bigcup_i Q_i^1\Big) \leq \Big(1 - \frac{A}{(2A)^{q'}}\Big)\mu(F).$$

Next, we consider a disjoint collection $\{Q_i^2\}_i \subset \mathcal{D}$ for which $Q_i^2 \subset F$ and $\langle |b_F|^q \rangle_{Q_i^2} > 2^{q'+1}A^{q'}$. Then, one may notice that

$$\mu\Big(\bigcup_i Q_i^2\Big) \leq 2^{-q'-1}A^{-q'}\int_F |b_F|^q \, d\mu \leq \frac{1}{2}\frac{A}{(2A)^{q'}}\mu(F).$$

Combining the analysis we conclude that (3.3) holds. □

The next lemma follows.

3.4. **Lemma.** *The following is a Carleson sequence: $\alpha_Q = 0$ if Q is not from $\bigcup_j \mathcal{F}_{Q^*}^j$, and it equals $\mu(Q)$ otherwise. This means that $\sum_{Q\subset R} a_Q \lesssim \mu(R)$ for every dyadic R.*

We now state the classical Carleson embedding theorem.

3.5. **Proposition.** *Given a Carleson sequence $(A_Q)_{Q\in\mathcal{D}}$ we have for every $f \in L^p(\mu)$, $1 < p < \infty$, that*

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

3.6. *Remark.* Note that q is always reserved to be the fixed index $q \in (1,2)$ appearing in the testing conditions.

The next proposition is a Carleson embedding on $L^p(\mu)$, where the Carleson condition itself depends on p. This kind of Carleson is also well-known, of course, but we state and prove this general version here for the convenience of the reader.

3.7. **Proposition.** *Let $\mathcal{D}$ be a dyadic grid in $\mathbb{R}^n$ and $p \in (1,2]$ be a fixed number. Suppose that for each $Q \in \mathcal{D}$ we have a function A_Q satisfying that spt $A_Q \subset Q$ and*

$$\mathrm{Car}_p((A_Q)_{Q\in\mathcal{D}}) := \Big(\sup_{R\in\mathcal{D}} \frac{1}{\mu(R)}\int_R \Big[\sum_{\substack{Q\in\mathcal{D}\\ Q\subset R}} |A_Q(x)|^2\Big]^{p/2} d\mu(x)\Big)^{1/p} < \infty. \tag{3.8}$$

Then we have that

$$\int \Big[\sum_{Q\in\mathcal{D}} |\langle f\rangle_Q|^2 |A_Q(x)|^2\Big]^{p/2} d\mu(x) \lesssim \mathrm{Car}_p((A_Q)_{Q\in\mathcal{D}})^p \|f\|_{L^p(\mu)}^p. \tag{3.9}$$

Proof. For each fixed $j \in \mathbb{Z}$ let $(R_j^i)_i$ denote the maximal $R \in \mathcal{D}$ for which $|\langle f\rangle_R| > 2^j$. We have that

$$\begin{aligned}
\int \Big[\sum_{Q\in\mathcal{D}} |\langle f\rangle_Q|^2 |A_Q(x)|^2\Big]^{p/2} d\mu(x) &= \int \Big[\sum_{j\in\mathbb{Z}} \sum_{\substack{Q\in\mathcal{D}\\ |\langle f\rangle_Q|\sim 2^j}} |\langle f\rangle_Q|^2 |A_Q(x)|^2\Big]^{p/2} d\mu(x)\\
&\lesssim \int \Big[\sum_{j\in\mathbb{Z}} 2^{2j} \sum_i \sum_{\substack{Q\in\mathcal{D}\\ Q\subset R_j^i}} |A_Q(x)|^2\Big]^{p/2} d\mu(x)\\
&\le \sum_{j\in\mathbb{Z}} 2^{pj} \sum_i \int_{R_j^i} \Big[\sum_{\substack{Q\in\mathcal{D}\\ Q\subset R_j^i}} |A_Q(x)|^2\Big]^{p/2} d\mu(x)\\
&\le \mathrm{Car}_p((A_Q)_{Q\in\mathcal{D}})^p \sum_{j\in\mathbb{Z}} 2^{pj} \mu\Big(\bigcup_i R_j^i\Big)\\
&\le \mathrm{Car}_p((A_Q)_{Q\in\mathcal{D}})^p \sum_{j\in\mathbb{Z}} 2^{pj} \mu(\{M_\mu^{\mathcal{D}} f > 2^j\})\\
&\approx \mathrm{Car}_p((A_Q)_{Q\in\mathcal{D}})^p \|M_\mu^{\mathcal{D}} f\|_{L^p(\mu)}^p\\
&\lesssim \mathrm{Car}_p((A_Q)_{Q\in\mathcal{D}})^p \|f\|_{L^p(\mu)}^p,
\end{aligned}$$

where the last estimate follows form the $L^p(\mu) \to L^p(\mu)$ boundedness of the dyadic Hardy-Littlewood maximal operator $M_\mu^{\mathcal{D}}$. In the proof we also used the assumption $p \in (1,2]$ simply via the fact that $(a+b)^\gamma \le a^\gamma + b^\gamma$ for $a, b \ge 0$ and $\gamma \in (0,1]$. □

3.3. **Twisted martingale difference operators and square function estimates.** If $Q \in \mathcal{D}$, $Q \subset Q^*$, and $f \in L^1_{\mathrm{loc}}(\mu)$, we define the twisted martingale difference operators

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

Note that on the largest Q^* level we agree (by abuse of notation) that $\Delta_{Q^*} = E_{Q^*}^b + \Delta_{Q^*}$, where $E_{Q^*}^b f = \langle f\rangle_{Q^*} b_{Q^*}$. Therefore, we have that $\int \Delta_Q f \, d\mu = 0$ if $Q \subsetneq Q^*$. We also define

$$\Delta_k f = \Delta_k^{Q^*} f := \sum_{Q\in\mathcal{D}_k : Q\subset Q^*} \Delta_Q f.$$

Notice that if $\ell(Q^*) = 2^s$, then $k \le s$, that is, only cubes inside the fixed Q^* are considered.

We now state some lemmata which contain the square function estimates we need in our proof. The first one was proved by Stein on page 103 of [25]:

3.10. **Lemma.** *Let (M,ν) be a σ-finite measure space and let $\mathfrak{M}$ denote the family of measurable subsets of M. Suppose that $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots$ is an infinite increasing sequence of (σ-finite) σ-subalgebras of $\mathfrak{M}$. Let $E_k = E(\cdot|\mathcal{F}_k)$ denote the conditional expectation operator with respect to $\mathcal{F}_k$. Assume that $\{f_k\}_k$ is any sequence of functions on (M,ν), where f_k is not assumed to be $\mathcal{F}_k$-measurable, and let $(n_k)_k$ be any sequence of positive integers. Then there holds that*

$$\Big\|\Big(\sum_{k\ge 1} |E_{n_k} f_k|^2\Big)^{1/2}\Big\|_{L^p(\nu)} \le A_p \Big\|\Big(\sum_{k\ge 1} |f_k|^2\Big)^{1/2}\Big\|_{L^p(\nu)}, \quad 1 < p < \infty, \tag{3.11}$$

where A_p depends only on p.

The proof of the next lemma is quite hard. It was proved by Lacey and the first named author [18] (but only in $L^2(\mu)$).

3.12. **Lemma.** *Suppose $F \in \mathcal{F}_{Q^*}$ and $f \in L^q(\mu)$. Suppose also that we have constants ϵ_Q, $Q \in \mathcal{D}$, which satisfy $|\epsilon_Q| \leq 1$. Then there holds that*

$$\Big\| \sum_{\substack{Q \in \mathcal{D} \\ Q^\alpha = F}} \epsilon_Q \Delta_Q f \Big\|_{L^q(\mu)}^q \lesssim \|f\|_{L^q(\mu)}^q.$$

We don't need the full strength of this. Therefore, instead of using it we will give a somewhat simpler proof in the $|f| \leq 1$ case, which is the only thing we will need. This also contains the modifications needed in the $q \neq 2$ case.

3.13. **Lemma.** *Suppose $F \in \mathcal{F}_{Q^*}$ and $|f| \leq 1$. Suppose also that we have constants ϵ_Q, $Q \in \mathcal{D}$, which satisfy $|\epsilon_Q| \leq 1$. Then there holds that*

$$\Big\| \sum_{\substack{Q \in \mathcal{D} \\ Q^\alpha = F}} \epsilon_Q \Delta_Q f \Big\|_{L^q(\mu)}^q \lesssim \mu(F). \tag{3.14}$$

Proof. For the fixed $F \in \mathcal{F}_{Q^*}$, we let $j \in \mathbb{N}$ be such that $F \in \mathcal{F}_{Q^*}^j$ and define $\mathcal{H} = \mathcal{H}_F = \{H \in \mathcal{F}_{Q^*}^{j+1} : H \subset F\}$. For a cube $Q \in \mathcal{D}$ for which $Q^a = F$ we set

$$D_Q f := \sum_{Q' \in \mathrm{ch}(Q) \setminus \mathcal{H}} \Big[\frac{\langle f \rangle_{Q'}}{\langle b_F \rangle_{Q'}} - \frac{\langle f \rangle_Q}{\langle b_F \rangle_Q} \Big] 1_{Q'}.$$

The initial step is that

$$\Big\| \sum_{\substack{Q \in \mathcal{D} \\ Q^\alpha = F}} \epsilon_Q \Delta_Q f \Big\|_{L^q(\mu)}^q \lesssim \|f 1_F\|_{L^q(\mu)}^q + \Big\| \sup_{\epsilon > 0} \Big| \sum_{\substack{Q \in \mathcal{D} \\ Q^a = F \\ \ell(Q) > \epsilon}} \epsilon_Q D_Q f \Big| \Big\|_{L^q(\mu)}^q.$$

This works exactly as in [18], proof of Proposition 2.4.

The second step is to show that

$$\Big\| \sup_{\epsilon > 0} \Big| \sum_{\substack{Q \in \mathcal{D} \\ Q^a = F \\ \ell(Q) > \epsilon}} \epsilon_Q D_Q f \Big| \Big\|_{L^p(\mu)}^p \lesssim \mu(F), \qquad |f| \leq 1,\ p \in (0, \infty). \tag{3.15}$$

The argument we will next give shows that for (3.15) it is enough to show that for a fixed $s \in (0, \infty)$ but for all $P \in \mathcal{D}$ there holds that

$$\Big\| \sup_{\epsilon > 0} \Big| \sum_{\substack{Q \in \mathcal{D} :\, Q \subset P \\ Q^a = F \\ \ell(Q) > \epsilon}} \epsilon_Q D_Q f \Big| \Big\|_{L^s(\mu)}^s \leq C_1 \mu(P). \tag{3.16}$$

Consider a fixed function f for which $|f| \leq 1$. Let us define the function $\varphi_Q = C_2^{-1} \epsilon_Q D_Q f$, if $Q^a = F$, and $\varphi_Q = 0$ otherwise. Notice that $\|\varphi_Q\|_{L^\infty(\mu)} \leq 1$ if $C_2 \geq 4$.

Notice also that φ_Q is supported on Q and constant on the children $Q' \in \mathrm{ch}(Q)$. For $P \in \mathcal{D}$ we define

$$\Phi_P := \sup_{\epsilon>0} \Big| \sum_{\substack{Q\in\mathcal{D}\\ Q\subset P\\ \ell(Q)>\epsilon}} \varphi_Q \Big| = C_2^{-1} \sup_{\epsilon>0} \Big| \sum_{\substack{Q\in\mathcal{D}:\, Q\subset P\\ Q^a=F\\ \ell(Q)>\epsilon}} \epsilon_Q D_Q f \Big|.$$

Suppose we have (3.16) with some s and for all P. Then for all $P \in \mathcal{D}$ we have that

$$\begin{aligned}\mu(\{x \in P:\, \Phi_P(x) > 1\}) \le \int_P \Phi_P^s \, d\mu &= C_2^{-s} \Big\| \sup_{\epsilon>0} \Big| \sum_{\substack{Q\in\mathcal{D}:\, Q\subset P\\ Q^a=F\\ \ell(Q)>\epsilon}} \epsilon_Q D_Q f \Big| \Big\|_{L^s(\mu)}^s \\ &\le C_2^{-s} C_1 \mu(P) \\ &\le \mu(P)/2,\end{aligned}$$

if $C_2 \ge C_1^{1/s} 2^{1/s}$. So let us fix C_2 large enough.

The non-homogeneous John–Nirenberg principle (see e.g. Lemma 2.8 of [18]) now tells us that for every $P \in \mathcal{D}$ and $t > 1$ there holds that

$$\mu(\{x \in P : \Phi_P(x) > t\}) \le 2^{-(t-1)/2} \mu(P).$$

But then we have for every $p \in (0,\infty)$ and $P \in D$ that

$$\Big\| \sup_{\epsilon>0} \Big| \sum_{\substack{Q\in\mathcal{D}:\, Q\subset P\\ Q^a=F\\ \ell(Q)>\epsilon}} \epsilon_Q D_Q f \Big| \Big\|_{L^p(\mu)}^p \lesssim \int_P \Phi_P^p \, d\mu \lesssim \mu(P). \tag{3.17}$$

With the choice $P = F$ we have (3.15).

So we have reduced to showing (3.16) with some exponent $s \in (0,\infty)$ and for all dyadic cubes $P \in \mathcal{D}$. We will first do this with $f = 1$ and $s = 1/2$, i.e., we will prove that for every $P \in \mathcal{D}$ there holds that

$$\int_P \Big[\sup_{\epsilon>0} \Big| \sum_{\substack{Q\in\mathcal{D}\\ Q^a=F,\, Q\subset P\\ \ell(Q)>\epsilon}} \epsilon_Q D_Q 1 \Big| \Big]^{1/2} d\mu \lesssim \mu(P).$$

Let us write

$$\frac{1}{\langle b_F\rangle_{Q'}} - \frac{1}{\langle b_F\rangle_Q} = \frac{\langle b_F\rangle_Q - \langle b_F\rangle_{Q'}}{\langle b_F\rangle_Q^2} + \frac{[\langle b_F\rangle_Q - \langle b_F\rangle_{Q'}]^2}{\langle b_F\rangle_Q{}^2 \langle b_F\rangle_{Q'}}.$$

Define $\tilde{\epsilon}_Q := \epsilon_Q/\langle b_F\rangle_Q^2$, $Q^a = F$. Note that $|\tilde{\epsilon}_Q| \lesssim 1$, and then that

$$\begin{aligned}
&\int_P \Big[\sup_{\epsilon>0}\Big|\sum_{\substack{Q\in\mathcal{D}\\ Q^a=F,\, Q\subset P\\ \ell(Q)>\epsilon}} \tilde{\epsilon}_Q \sum_{Q'\in \mathrm{ch}(Q)\setminus\mathcal{H}} [\langle b_F\rangle_{Q'} - \langle b_F\rangle_Q]1_{Q'}\Big| \, d\mu\Big]^{1/2}\\
&\le \mu(P)^{1-1/(2q)}\Big\|\sup_{\epsilon>0}\Big|\sum_{\substack{Q\in\mathcal{D}\\ Q^a=F,\, Q\subset P\\ \ell(Q)>\epsilon}} \tilde{\epsilon}_Q \sum_{Q'\in \mathrm{ch}(Q)\setminus\mathcal{H}} [\langle b_F\rangle_{Q'} - \langle b_F\rangle_Q]1_{Q'}\Big|\Big\|_{L^q(\mu)}^{1/2}\\
&\le \mu(P)^{1-1/(2q)}\|1_P b_F\|_{L^q(\mu)}^{1/2} \lesssim \mu(P).
\end{aligned}$$

The penultimate estimate follows from Corollary 2.10 of [18] (with $p = q$). For the last inequality we have the following explanation. It is trivial if $F \cap P = \emptyset$ or $F \subset P$. Otherwise, we may assume that there is a Q for which $Q^a = F$ and $Q \subset P \subset F$. But then $P^a = F$.

The exponent $s = 1/2$ is more useful now when we are dealing with the second term:

$$\begin{aligned}
&\int_P \Big[\sup_{\epsilon>0}\Big|\sum_{\substack{Q\in\mathcal{D}\\ Q^a=F,\, Q\subset P\\ \ell(Q)>\epsilon}} \epsilon_Q \sum_{Q'\in \mathrm{ch}(Q)\setminus\mathcal{H}} \frac{[\langle b_F\rangle_Q - \langle b_F\rangle_{Q'}]^2}{\langle b_F\rangle_Q{}^2\langle b_F\rangle_{Q'}} 1_{Q'}\Big|\Big]^{1/2} d\mu\\
&\lesssim \int_P \Big[\sum_{Q\in\mathcal{D}} |\Delta_Q^c(1_P b_F)|^2\Big]^{1/2} d\mu\\
&\le \mu(P)^{1-1/q}\Big\|\Big[\sum_{Q\in\mathcal{D}} |\Delta_Q^c(1_P b_F)|^2\Big]^{1/2}\Big\|_{L^q(\mu)}\\
&\lesssim \mu(P)^{1-1/q}\|1_P b_F\|_{L^q(\mu)} \lesssim \mu(P).
\end{aligned}$$

Here

$$\Delta_Q^c f = \sum_{Q'\in\mathrm{ch}(Q)} [\langle f\rangle_{Q'} - \langle f\rangle_Q]1_{Q'}$$

is the classical martingale difference. So we have proved (3.16) with $s = 1/2$ and $f = 1$ for every $P \in \mathcal{D}$. That means that for $f = 1$ we have (3.17) with every $p \in (0,\infty)$ and $P \in \mathcal{D}$.

Consider now a function f for which $|f| \le 1$. Using the above special case we will now prove (3.16) for every $P \in \mathcal{D}$ with $s = 1$. Let us write

$$\begin{aligned}
\frac{\langle f\rangle_{Q'}}{\langle b_F\rangle_{Q'}} - \frac{\langle f\rangle_Q}{\langle b_F\rangle_Q} &= \Big\{\frac{\langle f\rangle_{Q'}}{\langle b_F\rangle_Q} - \frac{\langle f\rangle_Q}{\langle b_F\rangle_Q}\Big\} + \Big\{\frac{\langle f\rangle_{Q'}}{\langle b_F\rangle_{Q'}} - \frac{\langle f\rangle_{Q'}}{\langle b_F\rangle_Q}\Big\}\\
&= \frac{1}{\langle b_F\rangle_Q}\big\{\langle f\rangle_{Q'} - \langle f\rangle_Q\big\} \qquad (3.18)\\
&\quad + \big\{\langle f\rangle_{Q'} - \langle f\rangle_Q\big\}\Big\{\frac{1}{\langle b_F\rangle_{Q'}} - \frac{1}{\langle b_F\rangle_Q}\Big\} \qquad (3.19)\\
&\quad + \langle f\rangle_Q\Big\{\frac{1}{\langle b_F\rangle_{Q'}} - \frac{1}{\langle b_F\rangle_Q}\Big\}. \qquad (3.20)
\end{aligned}$$

The terms (3.18)-(3.20) give us the corresponding decomposition

$$\epsilon_Q D_Q f = \epsilon_Q^1 \dot{\Delta}_Q^c f + \epsilon_Q \Delta_Q^c f \cdot D_Q 1 + \epsilon_Q^2 D_Q 1,$$

where Δ_Q^c is the classical martingale defined above, $\dot{\Delta}_Q^c$ is the stopped classical martingale

$$\dot{\Delta}_Q^c f = \sum_{Q' \in \mathrm{ch}(Q) \setminus \mathcal{H}} [\langle f \rangle_{Q'} - \langle f \rangle_Q] 1_{Q'},$$

and the bounded constants ϵ_Q^1 and ϵ_Q^2 are defined by

$$\epsilon_Q^1 = \frac{\epsilon_Q}{\langle b_F \rangle_Q}, \qquad \epsilon_Q^2 = \epsilon_Q \langle f \rangle_Q.$$

The first term can be bounded by Hölder (say with $p = 2$) and using Corollary 2.10 of [18] (with $p = 2$). The rest exploit the special case $f = 1$. The second term can be bounded by bringing the absolute values in, using Hölder to the sums with $p = 2$, and then using Hölder in the integral with $p = 2$. Here one needs (3.17) with $f = 1$ and $p = 2$. The last term is just (3.17) with $f = 1$ and $p = 1$. We are done. □

In the $|f| \leq 1$ case we can get rid of the assumption $Q^a = F$ as follows:

3.21. **Proposition.** *Let $|f| \leq 1$. Then there holds that*

$$\Big\| \Big(\sum_k |\Delta_k f|^2 \Big)^{1/2} \Big\|_{L^q(\mu)}^q \lesssim \mu(Q^*). \tag{3.22}$$

Proof. By Khinchine's inequality there holds that

$$\Big\| \Big(\sum_k |\Delta_k f|^2 \Big)^{1/2} \Big\|_{L^q(\mu)} \lesssim \Big\| \sum_k \varepsilon_k \Delta_k f \Big\|_{L^q(\mu \times \mathbb{P})},$$

where $(\varepsilon_k)_{k \in \mathbb{Z}}$ is a random sequence of Rademacher functions, i.e., a sequence of independent random variables attaining values ± 1 with an equal probability $\mathbb{P}(\varepsilon_k = 1) = \mathbb{P}(\varepsilon_k = -1) = 1/2$. If we set $\epsilon_Q = \varepsilon_k$, when $Q \in \mathcal{D}_k$, we have that

$$\begin{aligned}
&\Big\| \sum_k \varepsilon_k \sum_{Q \in \mathcal{D}_k : Q \subset Q^*} \Delta_Q f \Big\|_{L^q(\mu \times \mathbb{P})} = \Big\| \sum_{Q \in \mathcal{D} : Q \subset Q^*} \epsilon_Q \Delta_Q f \Big\|_{L^q(\mu \times \mathbb{P})} \\
&\leq \sum_{j \geq 0} \Big(\sum_{F \in \mathcal{F}_{Q^*}^j} \Big\| \sum_{\substack{Q \in \mathcal{D} : Q \subset Q^* \\ Q^\alpha = F}} \epsilon_Q \Delta_Q f \Big\|_{L^q(\mu \times \mathbb{P})}^q \Big)^{1/q} \\
&\lesssim \sum_{j \geq 0} \Big(\sum_{F \in \mathcal{F}_{Q^*}^j} \mu(F) \Big)^{1/q} \lesssim \mu(Q^*)^{1/q},
\end{aligned}$$

where the second-to-last inequality follows from (3.14) and $\int d\mathbb{P} = 1$, and the last one from (3.3). □

4. Reductions towards the proof of the key inequality

We will estimate the quantity

$$\Big[\int_{Q_0}\Big(\int_0^{\ell(Q_0)}|\theta_t f(x)|^2\frac{dt}{t}\Big)^{q/2}d\mu(x)\Big]^{1/q}$$

for an arbitrary fixed cube $Q_0\subset\mathbb{R}^n$ and for an arbitrary fixed function f satisfying that $|f|\le 1_{Q_0}$ (the choice $f=1_{Q_0}$ would suffice). Let s be defined by $2^{s-1}\le \ell(Q_0)<2^s$.

4.1. **Reduction to a dyadic setting of good geometric data.** For a fixed $w\in(\{0,1\}^n)^{\mathbb{Z}}$ and $x\in Q_0$ we have that

$$\int_0^{\ell(Q_0)}|\theta_t f(x)|^2\,\frac{dt}{t}\le\sum_{\substack{R\in\mathcal{D}(w)\\ \ell(R)\le 2^s}}1_R(x)\int_{\ell(R)/2}^{\ell(R)}|\theta_t f(x)|^2\,\frac{dt}{t}.$$

Recall the constants from (2.1). To prove (2.2) we note that by above it is enough to prove that

$$E_w\Big[\int_{\mathbb{R}^n}1_{Q_0}(x)\Big(\sum_{\substack{R\in\mathcal{D}(w)\\ \ell(R)\le 2^s}}1_R(x)\int_{\ell(R)/2}^{\ell(R)}|\theta_t f(x)|^2\,\frac{dt}{t}\Big)^{q/2}d\mu(x)\Big]^{1/q}\tag{4.1}$$

can be bounded by

$$[C_3(1+V_{\mathrm{loc},q})+C_2^{-1}\|V\|_{L^q(\mu)\to L^q(\mu)}/2]\mu(10Q_0)^{1/q}.$$

And here we may take the average over only those w for which there holds that $\mathcal{D}(w)$ satisfies that $\mu(\partial R)=0$ for every $R\in\mathcal{D}(w)$. This is because almost every w is such.

We can estimate the quantity in (4.1) by

$$E_w\Big[\int_{\mathbb{R}^n}1_{Q_0}(x)\Big(\sum_{\substack{R\in\mathcal{D}(w)_{\mathrm{good}}\\ \ell(R)\le 2^s}}1_R(x)\int_{\ell(R)/2}^{\ell(R)}|\theta_t f(x)|^2\,\frac{dt}{t}\Big)^{q/2}d\mu(x)\Big]^{1/q}$$
$$+E_w\Big[\int_{\mathbb{R}^n}1_{Q_0}(x)\Big(\sum_{\substack{R\in\mathcal{D}(w)_{\mathrm{bad}}\\ \ell(R)\le 2^s}}1_R(x)\int_{\ell(R)/2}^{\ell(R)}|\theta_t f(x)|^2\,\frac{dt}{t}\Big)^{q/2}d\mu(x)\Big]^{1/q}.$$

It is actually only now that these quantities inside the average really start to depend on w! Using $Eg^\alpha\le(Eg)^\alpha$ for $\alpha\in(0,1]$, we see (with $\alpha=1/q$ and $\alpha=q/2$) that

$$E_w\Big[\int_{\mathbb{R}^n}1_{Q_0}(x)\Big(\sum_{\substack{R\in\mathcal{D}(w)_{\mathrm{bad}}\\ \ell(R)\le 2^s}}1_R(x)\int_{\ell(R)/2}^{\ell(R)}|\theta_t f(x)|^2\,\frac{dt}{t}\Big)^{q/2}d\mu(x)\Big]^{1/q}$$
$$\le\Big[\int_{\mathbb{R}^n}1_{Q_0}(x)\Big(E_w\sum_{\substack{R\in\mathcal{D}(w)_{\mathrm{bad}}\\ \ell(R)\le 2^s}}1_R(x)\int_{\ell(R)/2}^{\ell(R)}|\theta_t f(x)|^2\,\frac{dt}{t}\Big)^{q/2}d\mu(x)\Big]^{1/q}.$$

Using the fact that $w \mapsto 1_{\mathrm{bad}}(R_0+w)$ is independent of $w \mapsto 1_{R_0+w}(x)$ for every $R_0 \in \mathcal{D}_0$, and that $E_w 1_{\mathrm{bad}}(R_0+w) \le c(r) \to 0$ when $r \to \infty$, we have

$$E_w\Big[\int_{\mathbb{R}^n} 1_{Q_0}(x)\Big(\sum_{\substack{R\in\mathcal{D}(w)_{\mathrm{bad}}\\ \ell(R)\le 2^s}} 1_R(x)\int_{\ell(R)/2}^{\ell(R)} |\theta_t f(x)|^2\,\frac{dt}{t}\Big)^{q/2} d\mu(x)\Big]^{1/q}$$
$$\le c(r)^{1/2}\|Vf\|_{L^q(\mu)} \le (2C_2)^{-1}\|V\|_{L^q(\mu)\to L^q(\mu)}\mu(10Q_0)^{1/q}$$

fixing $r \lesssim 1$ large enough (note that $c(r) = C(n,\alpha,m)2^{-r\gamma}$).

We have reduced to showing that uniformly on $w \in (\{0,1\}^n)^{\mathbb{Z}}$ the quantity

$$\Big[\int_{\mathbb{R}^n} 1_{Q_0}(x)\Big(\sum_{\substack{R\in\mathcal{D}(w)_{\mathrm{good}}\\ \ell(R)\le 2^s}} 1_R(x)\int_{\ell(R)/2}^{\ell(R)} |\theta_t f(x)|^2\,\frac{dt}{t}\Big)^{q/2} d\mu(x)\Big]^{1/q}$$

can be dominated by $C_3(1+V_{\mathrm{loc},q})\mu(10Q_0)^{1/q}$. We fix one w and write $\mathcal{D} = \mathcal{D}(w)$. More specifically, w can also be taken such that $\mu(\partial R) = 0$ for every $R \in \mathcal{D}$.

4.2. **Decomposition of f.** Since $f \in L^q(\mu)$ is supported in Q_0 we may expand

$$f = \sum_{\substack{Q^*\in\mathcal{D}\\ \ell(Q^*)=2^s\\ Q_0\cap Q^*\neq\emptyset}} \sum_{\substack{Q\in\mathcal{D}\\ Q\subset Q^*}} \Delta_Q f. \tag{4.2}$$

Notice that there are only finitely many such Q^* and always $Q^* \subset 10Q_0$. Define

$$A_\kappa f(x) := \Big(\sum_{\substack{R\in\mathcal{D}_{\mathrm{good}}\\ 2^{-\kappa}<\ell(R)\le 2^s}} 1_R(x)\int_{\ell(R)/2}^{\ell(R)} |\theta_t f(x)|^2\,\frac{dt}{t}\Big)^{1/2}$$

and

$$Af(x) := \Big(\sum_{\substack{R\in\mathcal{D}_{\mathrm{good}}\\ \ell(R)\le 2^s}} 1_R(x)\int_{\ell(R)/2}^{\ell(R)} |\theta_t f(x)|^2\,\frac{dt}{t}\Big)^{1/2}.$$

Notice that for $x \in Q_0$ there holds that

$$\begin{aligned}
&\Big|Af(x) - A_\kappa\Big(\sum_{Q^*}\sum_{\substack{Q\subset Q^*\\ \ell(Q)>2^{-\kappa}}} \Delta_Q f\Big)(x)\Big| \\
&\le |Af(x) - A_\kappa f(x)| + \Big|A_\kappa f(x) - A_\kappa\Big(\sum_{Q^*}\sum_{\substack{Q\subset Q^*\\ \ell(Q)>2^{-\kappa}}} \Delta_Q f\Big)(x)\Big| \\
&\le \Big(\sum_{\substack{R\in\mathcal{D}_{\text{good}}\\ \ell(R)\le 2^{-\kappa}}} 1_R(x)\int_{\ell(R)/2}^{\ell(R)} |\theta_t f(x)|^2\,\frac{dt}{t}\Big)^{1/2} + A_\kappa\Big(f - \sum_{Q^*}\sum_{\substack{Q\subset Q^*\\ \ell(Q)>2^{-\kappa}}} \Delta_Q f\Big)(x) \\
&\le \Big(\int_0^{2^{-\kappa}} |\theta_t f(x)|^2\,\frac{dt}{t}\Big)^{1/2} + V\Big(f - \sum_{Q^*}\sum_{\substack{Q\subset Q^*\\ \ell(Q)>2^{-\kappa}}} \Delta_Q f\Big)(x).
\end{aligned}$$

It follows by dominated convergence and the fact that V is bounded on $L^q(\mu)$ that

$$\lim_{\kappa\to\infty}\Big\|1_{Q_0}\Big(Af - A_\kappa\Big(\sum_{Q^*}\sum_{\substack{Q\subset Q^*\\ \ell(Q)>2^{-\kappa}}} \Delta_Q f\Big)\Big)\Big\|_{L^q(\mu)} = 0.$$

We have reduced to showing that

$$\Big\|1_{Q_0}(x)\Big(\sum_{\substack{R\in\mathcal{D}_{\text{good}}\\ 2^{-\kappa}<\ell(R)\le 2^s}} 1_R(x)\int_{\ell(R)/2}^{\ell(R)}\Big|\sum_{\substack{Q\in\mathcal{D}\\ Q\subset Q^*\\ \ell(Q)>2^{-\kappa}}} \theta_t\Delta_Q f(x)\Big|^2\,\frac{dt}{t}\Big)^{1/2}\Big\|_{L^q(\mu)} \tag{4.3}$$

can be dominated by $C_3(1+V_{\text{loc},q})\mu(Q^*)^{1/q}$ for every fixed κ and for every fixed Q^*. We used the fact that

$$\theta_t\Big(\sum_{Q^*}\sum_{\substack{Q\subset Q^*\\ \ell(Q)>2^{-\kappa}}} \Delta_Q f\Big) = \sum_{Q^*}\sum_{\substack{Q\subset Q^*\\ \ell(Q)>2^{-\kappa}}} \theta_t\Delta_Q f,$$

since the sum is finite for every κ. To fix only one $Q^* \subset 10Q_0$ we used the fact that $\#\{Q^*\in\mathcal{D}:\ \ell(Q^*)=2^s \text{ and } Q^*\cap Q_0\neq\emptyset\}\lesssim 1$.

4.3. **Splitting the summation.** We will split the sum (4.3) in to the following four pieces:

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

We call the second sum the separated sum, the third sum the diagonal sum and the last sum the nested sum. Thus, (4.3) is bounded by

$$I_{\ell(Q)<\ell(R)} + I_{\text{sep}} + I_{\text{diag}} + I_{\text{nested}}.$$

We bound these four pieces in the four subsequent chapters.

4.4. *Remark.* The κ and the s are fixed and sometimes such implicit conditions on the generations of the cubes are not written down.

5. The case $\ell(Q) < \ell(R)$

We start by proving the following lemma.

5.1. **Lemma.** *Let $Q, R \in \mathcal{D}$ be such that $\ell(R)/\ell(Q) = 2^\ell$ and $D(Q,R)/\ell(R) \sim 2^j$ for $\ell \geq 1$ and $j \geq 0$. Then, if $S_0 = Q^{(\ell+j+\theta(j))}$, $x \in R$ and $y \in Q$, there holds that*

$$|s_t(x,y) - s_t(x,c_Q)| \lesssim 2^{-\alpha\ell} 2^{-3\alpha j/4} \ell(S_0)^{-m}, \quad t \in (\ell(R)/2, \ell(R)). \tag{5.2}$$

Here c_Q denotes the centre of Q.

Proof. First, notice that for every $y \in Q$ we have that $|y - c_Q| \leq \ell(Q)/2 \leq \ell(R)/4 < t/2$. Therefore, we may use (1.3) to obtain

$$|s_t(x,y) - s_t(x,c_Q)| \lesssim \frac{\ell(Q)^\alpha}{(\ell(R) + d(Q,R))^{m+\alpha}} \lesssim \frac{\ell(Q)^\alpha}{D(Q,R)^{m+\alpha}},$$

where we used that obviously $D(Q,R) \lesssim \ell(R) + d(Q,R)$ in our situation. Next, observe that

$$\frac{\ell(Q)^\alpha}{D(Q,R)^{m+\alpha}} \approx 2^{-\alpha\ell} 2^{-(m+\alpha)j} \ell(R)^{-m}.$$

Using the estimate $m\gamma/(1-\gamma) < \alpha/4$ and the definition of S_0 we see that

$$\ell(S_0)^{-m} \gtrsim 2^{-mj - \alpha j/4} \ell(R)^{-m}.$$

Combining we get (5.2). □

Let $Q \in \mathcal{D}$ and $R \in \mathcal{D}_{\text{good}}$ be such that $\ell(R)/\ell(Q) = 2^\ell$ and $D(Q,R)/\ell(R) \sim 2^j$ for $\ell \geq 1$ and $j \geq 0$. Assume also that $(x,t) \in W_R$. Since $\ell(Q) < \ell(R) \leq 2^s$, we have $\int \Delta_Q f \, d\mu = 0$. Using this we write

$$|\theta_t \Delta_Q f(x)| = \Big| \int_Q [s_t(x,y) - s_t(x,c_Q)] \Delta_Q f(y) \, d\mu(y) \Big|.$$

Using the estimate (5.2) we now see that

$$|\theta_t \Delta_Q f(x)| \lesssim 2^{-\alpha\ell} 2^{-3\alpha j/4} \ell(S_0)^{-m} \int_Q |\Delta_Q f(y)| d\mu(y),$$

where $Q, R \subset S_0 := Q^{(\ell+j+\theta(j))}$ (by (1) of Lemma 3.1).

We can now see that $I_{\ell(Q)<\ell(R)}$ can be dominated by

$$\sum_{j,\ell} b_{j,\ell} \Big\| \Big(\sum_{k \leq s} \sum_{R \in \mathcal{D}_{k,\text{good}}} 1_R \Big(\sum_{\substack{Q \in \mathcal{D}_{k-\ell} : Q \subset Q^* \\ D(Q,R)/\ell(R) \sim 2^j}} \ell(S_0)^{-m} \int |\Delta_Q f| d\mu \Big)^2 \Big)^{1/2} \Big\|_{L^q(\mu)},$$

where $b_{j,\ell} := 2^{-\alpha(\ell+\frac{3}{4}j)}$. Let us fix j, ℓ, k. Set $\tau_j(k) := j + \theta(j) + k = \mathrm{gen}(S_0)$. We have by disjointness considerations and the fact that $Q, R \subset S_0$ that

$$\begin{aligned}
&\sum_{R\in\mathcal{D}_{k,\mathrm{good}}} 1_R\Big(\sum_{\substack{Q\in\mathcal{D}_{k-\ell}\colon Q\subset Q^*\\ D(Q,R)/\ell(R)\sim 2^j}} \ell(S_0)^{-m}\int |\Delta_Q f|\,d\mu\Big)^2\\
&\quad=\Big(\sum_{R\in\mathcal{D}_{k,\mathrm{good}}} 1_R \sum_{\substack{Q\in\mathcal{D}_{k-\ell}\colon Q\subset Q^*\\ D(Q,R)/\ell(R)\sim 2^j}} 2^{-m\tau_j(k)}\int |\Delta_Q f|\,d\mu\Big)^2\\
&\quad=\Big(\sum_{S\in\mathcal{D}_{\tau_j(k)}}\sum_{\substack{R\in\mathcal{D}_{k,\mathrm{good}}\\ R\subset S}} 1_R \sum_{\substack{Q\in\mathcal{D}_{k-\ell}\colon Q\subset Q^*\\ D(Q,R)/\ell(R)\sim 2^j}} 2^{-m\tau_j(k)}\int |\Delta_Q f|\,d\mu\Big)^2\\
&\quad\lesssim\Big(\sum_{S\in\mathcal{D}_{\tau_j(k)}}\frac{1_S}{\mu(S)}\int_S |\Delta_{k-\ell}f|d\mu\Big)^2\\
&\quad=[E_{\tau_j(k)}(|\Delta_{k-\ell}f|)]^2.
\end{aligned}$$

Note that for fixed j, ℓ there holds by Stein's inequality (Lemma 3.10) and estimate (3.22) that

$$\Big\|\Big(\sum_{k\le s}[E_{\tau_j(k)}(|\Delta_{k-\ell}f|)]^2\Big)^{1/2}\Big\|_{L^q(\mu)}\lesssim\Big\|\Big(\sum_{k\le s}|\Delta_k f|^2\Big)^{1/2}\Big\|_{L^q(\mu)}\lesssim\mu(Q^*)^{1/q}.$$

We may now conclude that $I_{\ell(Q)<\ell(R)}\lesssim\mu(Q^*)^{1/q}$.

6. The separated sum

We first prove the following lemma.

6.1. **Lemma.** *Let $Q, R\in\mathcal{D}$ be so that $d(Q,R) > \ell(R)^\gamma\ell(Q)^{1-\gamma}$, $\ell(Q)/\ell(R) = 2^\ell$ and $D(Q,R)/\ell(Q)\sim 2^j$ for $\ell, j\ge 0$. Then, if $S_0 = Q^{(j+\theta(j+\ell))}$, $x\in R$ and $y\in Q$, there holds that*

$$|s_t(x,y)|\lesssim 2^{-\alpha\ell/4}2^{-3\alpha j/4}\ell(S_0)^{-m},\quad t\in(\ell(R)/2,\ell(R)). \tag{6.2}$$

Proof. We begin by noting that

$$|s_t(x,y)|\lesssim\frac{\ell(R)^\alpha}{d(Q,R)^{m+\alpha}}\lesssim\frac{\ell(Q)^{\alpha/2}\ell(R)^{\alpha/2}}{D(Q,R)^{m+\alpha}}=2^{-\alpha\ell/2}2^{-(m+\alpha)j}\ell(Q)^{-m}.$$

The second estimate is a standard fact and follows since $(m+\alpha)\gamma\le\alpha/2$, $\ell(R)\le\ell(Q)$ and $d(Q,R) > \ell(R)^\gamma\ell(Q)^{1-\gamma}$.

On the other hand it is easy to see that

$$\ell(S_0)^{-m}\gtrsim 2^{-mj-(\ell+j)\alpha/4}\ell(Q)^{-m}.$$

This uses just the definition of S_0 and the bound $m\gamma/(1-\gamma) < \alpha/4$. Combining the estimates we have (6.2). □

Let $Q\in\mathcal{D}$, $R\in\mathcal{D}_{\mathrm{good}}$ be such that $d(Q,R) > \ell(R)^\gamma\ell(Q)^{1-\gamma}$, $\ell(Q)/\ell(R) = 2^\ell$ and $D(Q,R)/\ell(Q)\sim 2^j$ for $\ell, j\ge 0$. If $(x,t)\in W_R$ we have by (6.2) that

$$|\theta_t\Delta_Q f(x)|\lesssim 2^{-\alpha(\ell+j)/4}\ell(S_0)^{-m}\int|\Delta_Q f(y)|\,d\mu(y),$$

where $Q, R \subset S_0 := Q^{(j+\theta(j+\ell))}$ (by (2) of Lemma 3.1).

If we denote $\tilde{b}_{j,\ell} := 2^{-\alpha(\ell+j)/4}$, we may deduce that I_{sep} can be dominated by

$$\sum_{j,\ell} \tilde{b}_{j,\ell} \Big\| \Big(\sum_{k\leq s} \sum_{R\in\mathcal{D}_{k,\text{good}}} 1_R \Big(\sum_{\substack{Q\in\mathcal{D}_{k+\ell}: Q\subset Q^* \\ D(Q,R)\sim 2^j \ell(Q)}} \ell(S_0)^{-m} \int |\Delta_Q f|\, d\mu \Big)^2 \Big)^{1/2} \Big\|_{L^q(\mu)}.$$

A completely analogous estimate to that of the previous section shows that $I_{\text{sep}} \lesssim \mu(Q^*)^{1/q}$.

7. The diagonal sum

Let $Q \in \mathcal{D}$, $R \in \mathcal{D}_{\text{good}}$ be such that $\ell(Q)/\ell(R) = 2^\ell$ and $D(Q,R)/\ell(Q) \sim 2^j$. Since we are in the diagonal summation I_{diag} we have that $\ell, j \lesssim 1$. If $(x,t) \in W_R$ we have that

$$|s_t(x,y)| \lesssim t^{-m} \approx \ell(R)^{-m} \approx \ell(S_0)^{-m},$$

where $Q, R \subset S_0 := Q^{(j+\theta(j+\ell))}$ (by (2) of Lemma 3.1). It is now clear by the previous arguments that $I_{\text{diag}} \lesssim \mu(Q^*)^{1/q}$.

8. The nested sum

In this case one uses the goodness of R to conclude that one must actually have that $R \subset Q$. Therefore, things reduce to proving that

$$\Big\| 1_{Q_0} \Big(\sum_{\substack{R\in\mathcal{D}_{\text{good}}:\, R\subset Q^* \\ 2^{-\kappa}<\ell(R)<2^{s-r}}} 1_R \int_{\ell(R)/2}^{\ell(R)} \Big| \sum_{\ell=r+1}^{s-\text{gen(R)}} \theta_t \Delta_{R^{(\ell)}} f \Big|^2 \frac{dt}{t} \Big)^{1/2} \Big\|_{L^q(\mu)} \lesssim \mu(Q^*)^{1/q}.$$

We bound the left-hand side of the above inequality by $I_{\text{nested},1} + I_{\text{nested},2}$, where

$$I_{\text{nested},1} :=$$
$$\Big\| 1_{Q_0} \Big(\sum_{\substack{R\in\mathcal{D}_{\text{good}}:\, R\subset Q^* \\ 2^{-\kappa}<\ell(R)<2^{s-r}}} 1_R \int_{\ell(R)/2}^{\ell(R)} \Big| \sum_{\ell=r+1}^{s-\text{gen(R)}} \theta_t (1_{R^{(\ell)}\setminus R^{(\ell-1)}} \Delta_{R^{(\ell)}} f) \Big|^2 \frac{dt}{t} \Big)^{1/2} \Big\|_{L^q(\mu)}$$

and

$$I_{\text{nested},2} :=$$
$$\Big\| 1_{Q_0} \Big(\sum_{\substack{R\in\mathcal{D}_{\text{good}}:\, R\subset Q^* \\ 2^{-\kappa}<\ell(R)<2^{s-r}}} 1_R \int_{\ell(R)/2}^{\ell(R)} \Big| \sum_{\ell=r+1}^{s-\text{gen(R)}} \theta_t (1_{R^{(\ell-1)}} \Delta_{R^{(\ell)}} f) \Big|^2 \frac{dt}{t} \Big)^{1/2} \Big\|_{L^q(\mu)}.$$

8.1. **The sum $I_{\text{nested},1}$.** The following lemma is the key to handling this sum.

8.1. **Lemma.** *For $\ell \geq r+1$ and $R \in \mathcal{D}_{k,\text{good}}$ we have for $(x,t) \in W_R$ that there holds that*

$$|\theta_t(1_{R^{(\ell)}\setminus R^{(\ell-1)}} \Delta_{R^{(\ell)}} f)(x)| \lesssim 2^{-\alpha\ell/2} 2^{-(k+\ell)m} \int |\Delta_{R^{(\ell)}} f(y)|\, d\mu(y).$$

Proof. If $S \in \mathrm{ch}(R^{(\ell)})$, $S \neq R^{(\ell-1)}$, and $(x,t) \in W_R$, we have by the size estimate (1.2) that

$$\begin{aligned}|\theta_t(1_S \Delta_{R^{(\ell)}} f)(x)| &\lesssim \int_S \frac{\ell(R)^\alpha}{d(S,R)^{m+\alpha}} |\Delta_{R^{(\ell)}} f(y)|\, d\mu(y) \\ &\lesssim \int_S \Big(\frac{\ell(R)}{\ell(S)}\Big)^{\alpha/2} \frac{1}{\ell(S)^m} |\Delta_{R^{(\ell)}} f(y)|\, d\mu(y)\end{aligned}$$

Here we used that by goodness $d(R,S) \geq \ell(R)^\gamma \ell(S)^{1-\gamma}$, and that we have $\gamma \leq \alpha/(2m+2\alpha)$. □

Let us denote $\hat{b}_\ell := 2^{-\alpha\ell/2}$. We now see using this lemma that $I_{\mathsf{nested},1}$ can be dominated by

$$\begin{aligned}&\sum_{\ell \geq r+1} \hat{b}_\ell \Big\| 1_{Q_0} \Big(\sum_{k \leq s-\ell} \Big(\sum_{\substack{R \in \mathcal{D}_{k,\mathrm{good}} \\ R \subset Q^*}} 1_R 2^{-(k+\ell)m} \int |\Delta_{R^{(\ell)}} f(y)|\, d\mu(y) \Big)^2 \Big)^{1/2} \Big\|_{L^q(\mu)} \\ &\lesssim \sum_{\ell \geq r+1} \hat{b}_\ell \Big\| 1_{Q_0} \Big(\sum_{k \leq s-\ell} \Big(\sum_{\substack{S \in \mathcal{D}_{k+\ell} \\ S \subset Q^*}} \frac{1_S}{\mu(S)} \int |\Delta_S f(y)|\, d\mu(y) \Big)^2 \Big)^{1/2} \Big\|_{L^q(\mu)} \\ &= \sum_{\ell \geq r+1} \hat{b}_\ell \Big\| 1_{Q_0} \Big(\sum_{k \leq s-\ell} \Big(\sum_{S \in \mathcal{D}_{k+\ell}} \frac{1_S}{\mu(S)} \int_S |\Delta_{k+\ell} f(y)|\, d\mu(y) \Big)^2 \Big)^{1/2} \Big\|_{L^q(\mu)} \\ &\lesssim \Big\| \Big(\sum_{k \leq s} [E_k(|\Delta_k f|)]^2 \Big)^{1/2} \Big\|_{L^q(\mu)} \lesssim \mu(Q^*)^{1/q}.\end{aligned}$$

The last inequality follows from Stein's inequality (3.11) and (3.22).

8.2. **The sum $I_{\mathsf{nested},2}$.** We begin by recording the following bound:

8.2. **Lemma.** *For $\ell \geq r+1$ and $R \in \mathcal{D}_{\mathrm{good}}$ we have for $(x,t) \in W_R$ that there holds that*

$$|\theta_t(1_{(R^{(\ell-1)})^c} b_{(R^{(\ell)})^a})(x)| \lesssim 2^{-\alpha\ell/2}.$$

Proof. Choose N_0 so that $(R^{(\ell)})^a = R^{(\ell+N_0)}$. Notice that since R is good there holds that

$$d\big(R, \partial R^{(\ell+j-1)}\big)^{m+\alpha} \gtrsim 2^{\alpha\ell/2} 2^{\alpha j/2} \ell(R)^\alpha \mu(R^{(\ell+j)}).$$

Here we used that $\gamma(\alpha+m) \leq \alpha/2$.

Therefore, for $(x,t) \in W_R$, the above estimate, the size bound (1.2) and the stopping conditions show that

$$\begin{aligned}|\theta_t(1_{(R^{(\ell-1)})^c} b_{(R^{(\ell)})^a}(x)| &\lesssim \sum_{j=0}^{N_0} \int_{R^{(\ell+j)} \setminus (R^{(\ell+j-1)}} \frac{\ell(R)^\alpha}{|x-y|^{m+\alpha}} |b_{(R^{(\ell)})^a}(y)|\, d\mu(y) \\ &\lesssim \sum_{j=0}^{N_0} \frac{\ell(R)^\alpha \mu(R^{(\ell+j)})}{2^{\alpha\ell/2} 2^{\alpha j/2} \ell(R)^\alpha \mu(R^{(\ell+j)})} \\ &\lesssim 2^{-\alpha\ell/2}.\end{aligned}$$

□

We now have to do a case study.

8.3. **The case** $(R^{(\ell-1)})^a = (R^{(\ell)})^a$**.** In this case we may write

$$1_{R^{(\ell-1)}}\Delta_{R^{(\ell)}}f = -1_{(R^{(\ell-1)})^c}B_{R^{(\ell-1)}}b_{(R^{(\ell)})^a} + B_{R^{(\ell-1)}}b_{(R^{(\ell)})^a}, \tag{8.3}$$

where

$$B_{R^{(\ell-1)}} = \frac{\langle f\rangle_{R^{(\ell-1)}}}{\langle b_{(R^{(\ell-1)})^a}\rangle_{R^{(\ell-1)}}} - \frac{\langle f\rangle_{R^{(\ell)}}}{\langle b_{(R^{(\ell)})^a}\rangle_{R^{(\ell)}}}$$

with the minus term missing if $\ell(R^{(\ell)}) = 2^s$.

Accretivity condition gives that

$$|B_{R^{(\ell-1)}}|\mu(R^{(\ell-1)}) \lesssim \Big|\int_{R^{(\ell-1)}} B_{R^{(\ell-1)}}b_{(R^{(\ell)})^a}\,d\mu\Big| = \Big|\int_{R^{(\ell-1)}} \Delta_{R^{(\ell)}}f\,d\mu\Big|.$$

Combining with Lemma 8.2 we see that for $(x,t)\in W_R$ there holds that

$$|\theta_t(1_{(R^{(\ell-1)})^c}B_{R^{(\ell-1)}}b_{(R^{(\ell)})^a})(x)| \lesssim 2^{-\alpha\ell/2}\frac{1}{\mu(R^{(\ell-1)})}\int_{R^{(\ell-1)}}|\Delta_{R^{(\ell)}}f|\,d\mu.$$

So to control the sum with the first term of (8.3) it is enough to note that for a fixed $\ell \geq r+1$ there holds that

$$\begin{aligned}
&\Big\|1_{Q_0}\Big(\sum_{k\leq s-\ell}\Big(\sum_{R\in\mathcal{D}_k:\,R\subset Q^*}\frac{1_R}{\mu(R^{(\ell-1)})}\int_{R^{(\ell-1)}}|\Delta_{R^{(\ell)}}f(y)|\,d\mu(y)\Big)^2\Big)^{1/2}\Big\|_{L^q(\mu)}\\
&\leq \Big\|1_{Q_0}\Big(\sum_{k\leq s-\ell}\Big(\sum_{S\in\mathcal{D}_{k+\ell-1}}\frac{1_S}{\mu(S)}\int_S|\Delta_{k+\ell}f(y)|\,d\mu(y)\Big)^2\Big)^{1/2}\Big\|_{L^q(\mu)}\\
&= \Big\|\Big(\sum_{k\leq s}[E_{k-1}(|\Delta_k f|)]^2\Big)^{1/2}\Big\|_{L^q(\mu)} \lesssim \mu(Q^*)^{1/q}.
\end{aligned}$$

In the last step we again used Stein's inequality (3.11) and (3.22). We will not touch the second term of (8.3) yet – it will become part of the paraproduct.

8.4. **The case** $(R^{(\ell-1)})^a = R^{(\ell-1)}$**.** We decompose

$$\begin{aligned}
1_{R^{(\ell-1)}}\Delta_{R^{(\ell)}}f = \Big(\frac{\langle f\rangle_{R^{(\ell-1)}}}{\langle b_{R^{(\ell-1)}}\rangle_{R^{(\ell-1)}}}b_{R^{(\ell-1)}} - \frac{\langle f\rangle_{R^{(\ell)}}}{\langle b_{(R^{(\ell)})^a}\rangle_{R^{(\ell)}}}b_{(R^{(\ell)})^a}\Big)\\
+ 1_{(R^{(\ell-1)})^c}\frac{\langle f\rangle_{R^{(\ell)}}}{\langle b_{(R^{(\ell)})^a}\rangle_{R^{(\ell)}}}b_{(R^{(\ell)})^a}.
\end{aligned}$$

The term in the parenthesis will become part of the paraproduct, and we do not touch it further in this subsection.

For the second term, using the construction of the stopping time and Lemma (8.2), we have for $(x,t)\in W_R$ that

$$|\Theta_{R,\ell}(x,t)| := \Big|\theta_t\Big(1_{(R^{(\ell-1)})^c}\frac{\langle f\rangle_{R^{(\ell)}}}{\langle b_{(R^{(\ell)})^a}\rangle_{R^{(\ell)}}}b_{(R^{(\ell)})^a}\Big)(x)\Big| \lesssim 2^{-\alpha\ell/2}|\langle f\rangle_{R^{(\ell)}}|.$$

We say that $R \in \mathcal{S}_\ell$, if $(R^{(\ell-1)})^a = R^{(\ell-1)}$. To control the corresponding sum we note that

$$\begin{aligned}
&\Big\| 1_{Q_0} \Big(\sum_{\substack{R\in\mathcal{D}_{\text{good}}\colon R\subset Q^* \\ 2^{-\kappa}<\ell(R)<2^{s-r}}} 1_R \int_{\ell(R)/2}^{\ell(R)} \Big| \sum_{\substack{\ell=r+1 \\ R\in\mathcal{S}_\ell}}^{s-\text{gen(R)}} \Theta_{R,\ell}(\cdot,t) \Big|^2 \frac{dt}{t} \Big)^{1/2} \Big\|_{L^q(\mu)} \\
&\quad \lesssim \Big\| 1_{Q_0} \Big(\sum_{\substack{R\in\mathcal{D}_{\text{good}}\colon R\subset Q^* \\ 2^{-\kappa}<\ell(R)<2^{s-r}}} 1_R \sum_{\substack{\ell=r+1 \\ R\in\mathcal{S}_\ell}}^{s-\text{gen(R)}} 2^{-\alpha\ell/2} |\langle f\rangle_{R^{(\ell)}}|^2 \Big)^{1/2} \Big\|_{L^q(\mu)} \\
&\quad \le \Big\| 1_{Q_0} \Big(\sum_{\ell\ge r+1} 2^{-\alpha\ell/2} \sum_{S\in\mathcal{D}\colon S\subset Q^*} |\langle f\rangle_S|^2 A_S \Big)^{1/2} \Big\|_{L^q(\mu)} \\
&\quad \lesssim \Big\| 1_{Q_0} \Big(\sum_{S\in\mathcal{D}\colon S\subset Q^*} |\langle f\rangle_S|^2 A_S \Big)^{1/2} \Big\|_{L^q(\mu)} \lesssim \mu(Q^*)^{1/q},
\end{aligned}$$

where we denoted

$$A_S(x) := \sum_{\substack{S'\in \text{ch}(S) \\ (S')^a = S'}} 1_{S'}(x).$$

For the final estimate one can use the fact that $|f| \le 1$ to throw away the averages, and then use Hölder with exponent $p := 2/q > 1$ together with Lemma 3.4:

$$\begin{aligned}
\Big\| \Big(\sum_{S\in\mathcal{D}\colon S\subset Q^*} A_S \Big)^{1/2} \Big\|_{L^q(\mu)} &\le \Big[\mu(Q^*)^{1-1/p} \Big(\sum_{F\in\mathcal{F}_{Q^*}} \mu(F) \Big)^{1/p} \Big]^{1/q} \\
&\lesssim (\mu(Q^*)^{1-1/p} \mu(Q^*)^{1/p})^{1/q} = \mu(Q^*)^{1/q}.
\end{aligned}$$

8.5. **The Carleson estimate for the paraproduct.** Combining the above two cases and collapsing the remaining telescoping summation we are left with:

$$\begin{aligned}
&\Big\|1_{Q_0}\Big(\sum_{\substack{R\in\mathcal{D}_{\text{good}}:\,R\subset Q^*\\ 2^{-\kappa}<\ell(R)<2^{s-r}}} 1_R\int_{\ell(R)/2}^{\ell(R)}\Big|\frac{\langle f\rangle_{R^{(r)}}}{\langle b_{(R^{(r)})^a}\rangle_{R^{(r)}}}\theta_t b_{(R^{(r)})^a}\Big|^2\frac{dt}{t}\Big)^{1/2}\Big\|_{L^q(\mu)}\\
&\quad\lesssim \Big\|1_{Q_0}\Big(\sum_{\substack{S\in\mathcal{D}\\ S\subset Q^*}}\sum_{\substack{R\in\mathcal{D}\\ R^{(r)}=S}} 1_R\int_{\ell(R)/2}^{\ell(R)}\big|\theta_t b_{S^a}\big|^2\frac{dt}{t}\Big)^{1/2}\Big\|_{L^q(\mu)}\\
&\quad= \Big\|1_{Q_0}\Big(\sum_{F\in\mathcal{F}_{Q^*}}\sum_{S:S^\alpha=F}\sum_{\substack{R\in\mathcal{D}\\ R^{(r)}=S}} 1_R\int_{\ell(R)/2}^{\ell(R)}\big|\theta_t b_F\big|^2\frac{dt}{t}\Big)^{1/2}\Big\|_{L^q(\mu)}\\
&\quad\le \Big\|1_{Q_0}\Big(\sum_{F\in\mathcal{F}_{Q^*}}\sum_{R:R\subset F} 1_R\int_{\ell(R)/2}^{\ell(R)}\big|\theta_t b_F\big|^2\frac{dt}{t}\Big)^{1/2}\Big\|_{L^q(\mu)}\\
&\quad\le \Big\|1_{Q_0}\Big(\sum_{F\in\mathcal{F}_{Q^*}} 1_F\int_0^{\ell(F)}\big|\theta_t b_F\big|^2\frac{dt}{t}\Big)^{1/2}\Big\|_{L^q(\mu)}\\
&\quad\le \sum_{j\ge 0}\Big(\sum_{F\in\mathcal{F}_{Q^*}^j}\Big\|\Big(1_F\int_0^{\ell(F)}\big|\theta_t b_F\big|^2\frac{dt}{t}\Big)^{1/2}\Big\|_{L^q(\mu)}^q\Big)^{1/q}\\
&\quad\le V_{\text{loc},q}\sum_{j\ge 0}\Big(\sum_{F\in\mathcal{F}_{Q^*}^j}\mu(F)\Big)^{1/q}\lesssim V_{\text{loc},q}\mu(Q^*)^{1/q}.
\end{aligned}$$

In the first inequality we used the stopping time conditions and the fact that $|f|\le 1$, while the penultimate inequality follows from assumption (4) of Theorem 1.4.

APPENDIX A. $T1$ THEOREM IN $L^q(\mu)$

Let us recall the definition of our Carleson constant:

$$\widetilde{\mathrm{Car}}_V(q,\lambda) := \sup_{\substack{Q_0\subset\mathbb{R}^n\\ \text{cube}}} \Big[\frac{1}{\mu(\lambda Q_0)}\int_{Q_0}\Big(\int_0^{\ell(Q_0)} |\theta_t 1(x)|^2\frac{dt}{t}\Big)^{q/2} d\mu(x)\Big]^{1/q}.$$

Recall also that $q\in(1,2]$. We are interested in proving the following $T1$ theorem.

A.1. **Theorem.** *We have the quantitative bound*

$$\|V\|_{L^q(\mu)\to L^q(\mu)} \lesssim 1+\widetilde{\mathrm{Car}}_V(q,30). \tag{A.2}$$

We now indicate the proof of this theorem. We can again, without loss of generality, assume that $\|V\|_{L^q(\mu)\to L^q(\mu)} < \infty$.

A.1. **Reduction to a dyadic setting of good geometric data.** Since we are not so well localised yet this part of the argument has a few more steps than that of the main theorem. We write

$$\begin{aligned}\int_0^\infty |\theta_t f(x)|^2\,\frac{dt}{t} &= \sum_{R\in\mathcal{D}(w)} 1_R(x)\int_{\ell(R)/2}^{\ell(R)} |\theta_t f(x)|^2\,\frac{dt}{t}\\ &= \lim_{s\to\infty}\sum_{\substack{R\in\mathcal{D}(w)\\ \ell(R)\le 2^s}} 1_R(x)\int_{\ell(R)/2}^{\ell(R)} |\theta_t f(x)|^2\,\frac{dt}{t}.\end{aligned}$$

By monotone convergence we have that

$$\|Vf\|_{L^q(\mu)} = \lim_{s\to\infty}\Big[\int_{\mathbb{R}^n}\Big(\sum_{\substack{R\in\mathcal{D}(w)\\ \ell(R)\le 2^s}} 1_R(x)\int_{\ell(R)/2}^{\ell(R)} |\theta_t f(x)|^2\,\frac{dt}{t}\Big)^{q/2} d\mu(x)\Big]^{1/q}.$$

We take the expectation E_w of this identity. Notice that there holds that

$$\Big[\int_{\mathbb{R}^n}\Big(\sum_{\substack{R\in\mathcal{D}(w)\\ \ell(R)\le 2^s}} 1_R(x)\int_{\ell(R)/2}^{\ell(R)} |\theta_t f(x)|^2\,\frac{dt}{t}\Big)^{q/2} d\mu(x)\Big]^{1/q} \le \|Vf\|_{L^q(\mu)} \in L^1((\{0,1\}^n)^{\mathbb{Z}}).$$

Indeed, $\|Vf\|_{L^q(\mu)} < \infty$ and $E_w 1 = 1$. Therefore we have by dominated convergence that

$$\|Vf\|_{L^q(\mu)} = \lim_{s\to\infty} E_w\Big[\int_{\mathbb{R}^n}\Big(\sum_{\substack{R\in\mathcal{D}(w)\\ \ell(R)\le 2^s}} 1_R(x)\int_{\ell(R)/2}^{\ell(R)} |\theta_t f(x)|^2\,\frac{dt}{t}\Big)^{q/2} d\mu(x)\Big]^{1/q}.$$

We now write

$$\|V\|_{L^q(\mu)\to L^q(\mu)} = \sup_{\substack{f \text{ compactly supported}\\ \|f\|_{L^q(\mu)}\le 1}} \|Vf\|_{L^q(\mu)}.$$

Fix such f, and then fix N so that $\operatorname{spt} f \subset B(0,2^N)$. It is enough to prove that for every $s \geq N$ there holds that

$$E_w\Big[\int_{\mathbb{R}^n}\Big(\sum_{\substack{R\in\mathcal{D}(w)\\ \ell(R)\leq 2^s}} 1_R(x)\int_{\ell(R)/2}^{\ell(R)}|\theta_t f(x)|^2\,\frac{dt}{t}\Big)^{q/2}d\mu(x)\Big]^{1/q}$$
$$\leq C(1+\widetilde{\operatorname{Car}}_V(q,30))+\|V\|_{L^q(\mu)\to L^q(\mu)}/2.$$

Now also fix $s \geq N$. One may argue as in Subsection 4.1 and reduce to showing that uniformly on $w\in(\{0,1\}^n)^{\mathbb{Z}}$ there holds that

$$\Big[\int_{\mathbb{R}^n}\Big(\sum_{\substack{R\in\mathcal{D}(w)_{\text{good}}\\ \ell(R)\leq 2^s}} 1_R(x)\int_{\ell(R)/2}^{\ell(R)}|\theta_t f(x)|^2\,\frac{dt}{t}\Big)^{q/2}d\mu(x)\Big]^{1/q}\lesssim (1+\widetilde{\operatorname{Car}}_V(q,30)).$$

We fix w and write $\mathcal{D}=\mathcal{D}(w)$.

A.2. **Expanding f and splitting the summation.** We now expand the fixed f in $L^q(\mu)$ as follows:

$$f=\lim_{\kappa\to\infty}\sum_{\substack{Q^*\in\mathcal{D}\\ \ell(Q^*)=2^s\\ Q^*\cap B(0,2^N)\neq\emptyset}}\ \sum_{\substack{Q\in\mathcal{D}\\ Q\subset Q^*\\ \ell(Q)>2^{-\kappa}}}\Delta_Q f. \tag{A.3}$$

This time the martingales are simple: $\Delta_Q f=\sum_{Q'\in\operatorname{ch}(Q)}[\langle f\rangle_{Q'}-\langle f\rangle_Q]1_{Q'}$ with the understanding that $\Delta_{Q^*}f=\sum_{Q'\in\operatorname{ch}(Q^*)}\langle f\rangle_{Q'}1_{Q'}$, $\ell(Q^*)=2^s$. The argument of Subsection 4.2 shows that it is enough to be able to bound the quantity

$$\Big[\int_{\mathbb{R}^n}\Big(\sum_{\substack{R\in\mathcal{D}_{\text{good}}\\ 2^{-\kappa}<\ell(R)\leq 2^s}} 1_R(x)\int_{\ell(R)/2}^{\ell(R)}\Big|\sum_{\substack{Q\in\mathcal{D}\\ Q\subset Q^*\\ \ell(Q)>2^{-\kappa}}}\theta_t\Delta_Q f(x)\Big|^2\,\frac{dt}{t}\Big)^{q/2}d\mu(x)\Big]^{1/q} \tag{A.4}$$

with $C(1+\widetilde{\operatorname{Car}}_V(q,30))$ for every fixed κ and for every fixed Q^*.

The splitting of the summation is the same as in the proof of the main theorem: the quantity in (A.4) is dominated by $I_{\ell(Q)<\ell(R)}+I_{\text{sep}}+I_{\text{diag}}+I_{\text{nested}}$. The first three terms are treated using similar arguments to the corresponding ones found in Sections 5, 6 and 7, and allows us to obtain

$$I_{\ell(Q)<\ell(R)}+I_{\text{sep}}+I_{\text{diag}}\lesssim 1.$$

Indeed, notice that in these sections things boil down to the martingale estimate

$$\Big\|\Big(\sum_{Q\subset Q^*}|\Delta_Q f|^2\Big)^{1/2}\Big\|_{L^q(\mu)}\lesssim\|f\|_{L^q(\mu)}=1, \tag{A.5}$$

which is easy for the classical martingales. These sections don't depend on the finer structure of the martingales.

The only difference lies in the treatment of the nested sum. Mostly it is much easier because of the simple martingales. But the thing that is more complicated is that now only $f\in L^q(\mu)$ (and not bounded). The moral of the story: only the paraproduct requires a different argument.

A.3. **The paraproduct in $T1$.** We need to show that

$$\Big\| \Big(\sum_{\substack{S\in\mathcal{D}\\ S\subset Q^*}} |\langle f\rangle_S|^2 A_S^2 \Big)^{1/2} \Big\|_{L^q(\mu)} \lesssim \widetilde{\mathrm{Car}}_V(q,30),$$

where

$$A_S(x)^2 := \sum_{\substack{R\in\mathcal{D}_{\mathrm{good}}\\ R^{(r)}=S}} 1_R(x) \int_{\ell(R)/2}^{\ell(R)} \big|\theta_t 1(x)\big|^2 \frac{dt}{t}.$$

By Proposition 3.7 it is enough to show the next lemma.

A.6. **Lemma.** *There holds that*

$$\mathrm{Car}_q((A_S)_S) \lesssim \widetilde{\mathrm{Car}}_V(q,30).$$

Proof. Let $Q\in\mathcal{D}$, $Q\subset Q^*$. We have that

$$\begin{aligned}\int_Q \Big[\sum_{\substack{S\in\mathcal{D}\\ S\subset Q}} A_S(x)^2\Big]^{q/2} d\mu(x) &= \int_Q \Big[\sum_{\substack{S\in\mathcal{D}\\ S\subset Q}} \sum_{\substack{R\in\mathcal{D}_{\mathrm{good}}\\ R^{(r)}=S}} 1_R(x) \int_{\ell(R)/2}^{\ell(R)} \big|\theta_t 1(x)\big|^2 \frac{dt}{t}\Big]^{q/2} d\mu(x)\\ &\le \int_Q \Big[\sum_{\substack{R\in\mathcal{D}\\ d(R,Q^c)\ge 100\ell(R)}} 1_R(x) \int_{\ell(R)/2}^{\ell(R)} \big|\theta_t 1(x)\big|^2 \frac{dt}{t}\Big]^{q/2} d\mu(x).\end{aligned}$$

Here we used that each appearing $R\in\mathcal{D}_{\mathrm{good}}$ satisfies that $R\subset Q$ and $\ell(R)\le 2^{-r}\ell(Q)$. Therefore, we have that $d(R,Q^c)\ge \ell(R)^\gamma \ell(Q)^{1-\gamma}\ge 2^{r(1-\gamma)}\ell(R)\ge 100\ell(R)$. Let $\mathcal{R}(Q)$ denote the maximal $R\in\mathcal{D}$ for which $d(R,Q^c)\ge 100\ell(R)$.

We have reduced to bounding

$$\begin{aligned}&\int_Q \Big[\sum_{R\in\mathcal{R}(Q)} \sum_{\substack{H\in D\\ H\subset R}} 1_H(x) \int_{\ell(H)/2}^{\ell(H)} \big|\theta_t 1(x)\big|^2 \frac{dt}{t}\Big]^{q/2} d\mu(x)\\ &\le \int_Q \Big[\sum_{R\in\mathcal{R}(Q)} 1_R(x) \int_0^{\ell(R)} \big|\theta_t 1(x)\big|^2 \frac{dt}{t}\Big]^{q/2} d\mu(x)\\ &= \sum_{R\in\mathcal{R}(Q)} \int_R \Big[\int_0^{\ell(R)} \big|\theta_t 1(x)\big|^2 \frac{dt}{t}\Big]^{q/2} d\mu(x)\\ &\le \widetilde{\mathrm{Car}}_V(q,30)^q \sum_{R\in\mathcal{R}(Q)} \mu(100R)\\ &\lesssim \widetilde{\mathrm{Car}}_V(q,30)^q \mu(Q).\end{aligned}$$

Recall that the supremum in the definition of $\widetilde{\mathrm{Car}}_V(q,30)$ runs over all the open cubes. Therefore, we used that $R\subset 2\mathrm{int}(R)$ and then simply that $60\mathrm{int}(R)\subset 100R$. We also used the disjointness of the cubes in $\mathcal{R}(Q)$ and the bounded overlap property $\sum_{R\in\mathcal{R}(Q)} 1_{100R}\lesssim 1_Q$. We are done. □

This completes our proof of Theorem A.1.

REFERENCES

[1] P. Auscher, *Lectures on the Kato square root problem*, Surveys in analysis and operator theory (Canberra, 2001), Proc. Centre Math. Appl. Austral. Nat. Univ., vol. 40, Austral. Nat. Univ., Canberra, 2002, pp. 1–18.

[2] P. Auscher, S. Hofmann, M. Lacey, A. McIntosh, and Ph. Tchamitchian, *The solution of the Kato square root problem for second order elliptic operators on* $\mathbb{R}^n$, Ann. of Math. (2) **156** (2002), no. 2, 633–654.

[3] P. Auscher, S. Hofmann, C. Muscalu, T. Tao, and C. Thiele, *Carleson measures, trees, extrapolation, and* $T(b)$ *theorems*, Publ. Mat. **46** (2002), no. 2, 257–325.

[4] P. Auscher and E. Routin, *Local* Tb *theorems and Hardy inequalities*, J. Geom. Anal. **23** (2013), no. 1, 303–374.

[5] P. Auscher and Q. X. Yang, BCR *algorithm and the* $T(b)$ *theorem*, Publ. Mat. **53** (2009), no. 1, 179–196.

[6] V. Chousionis, J. Garnett, T. Le, and X. Tolsa, *Square functions and uniform rectifiability* (2014), available at `http://arxiv.org/abs/1401.3382`.

[7] M. Christ, *A* $T(b)$ *theorem with remarks on analytic capacity and the Cauchy integral*, Colloquium Mathematicum **50/51** (1990), 601–628.

[8] J. García-Cuerva and J. M. Martell, *Weighted inequalities and vector-valued Calderón–Zygmund operators on non-homogeneous spaces*, Publ. Mat. **44** (2000), no. 2, 613–640.

[9] S. Hofmann, *Local* Tb *theorems and applications in PDE*, International Congress of Mathematicians. Vol. II, Eur. Math. Soc., Zürich, 2006, pp. 1375–1392.

[10] ———, *A proof of the local* Tb *theorem for standard Calderón-Zygmund operators* (2007), available at `http://arxiv.org/abs/0705.0840`.

[11] ———, *A local* Tb *theorem for square functions*, Perspectives in partial differential equations, harmonic analysis and applications (Providence, RI, 2008), Proc. Sympos. Pure Math., vol. 79, Amer. Math. Soc., 2008, pp. 175–185.

[12] S. Hofmann, M. Lacey, and A. McIntosh, *The solution of the Kato problem for divergence form elliptic operators with Gaussian heat kernel bounds*, Ann. of Math. (2) **156** (2002), no. 2, 623–631.

[13] S. Hofmann and A. McIntosh, *The solution of the Kato problem in two dimensions*, Proceedings of the conference on harmonic analysis and PDE (El Escorial, Spain, 2000), Publ. Mat., vol. extra, 2002, pp. 143–160.

[14] T. Hytönen, *The sharp weighted bound for general Calderón-Zygmund operators*, Ann. of Math. (2) **175** (2012), no. 3, 1473–1506.

[15] ———, *The vector-valued non-homogeneous* Tb *theorem*, Int. Math. Res. Notices, posted on 2012, DOI 10.1093/imrn/rns222.

[16] T. Hytönen and H. Martikainen, *On general local* Tb *theorems*, Trans. Amer. Math. Soc. **364** (2012), no. 9, 4819–4846.

[17] T. Hytönen and F. Nazarov, *The local* Tb *theorem with rough test functions* (2012), available at `http://www.arxiv.org/abs/1206.0907`.

[18] M. Lacey and H. Martikainen, *Local* Tb *theorem with* L^2 *testing conditions and general measures: Calderón-Zygmund operators* (2013), available at `http://arxiv.org/abs/1310.8531`.

[19] ———, *Local* Tb *theorem with* L^2 *testing conditions and general measures: Square functions* (2013), available at `http://arxiv.org/abs/1308.4571`.

[20] M. Lacey and A. Vähäkangas, *Non-homogeneous local* $T1$ *theorem: Dual exponents* (2013), available at `http://arxiv.org/abs/1301.5858`.

[21] ———, *On the Local* Tb *Theorem: A Direct Proof under the Duality Assumption*, Proc. Edinb. Math. Soc., to appear (2013), available at `http://arxiv.org/abs/1209.4161`.

[22] H. Martikainen and M. Mourgoglou, *Square functions with general measures*, Proc. Amer. Math. Soc., to appear (2012), available at `http://www.arxiv.org/abs/1212.3684`.

[23] H. Martikainen, M. Mourgoglou, and T. Orponen, *Square functions with general measures II*, Indiana Univ. Math. J., to appear (2013), available at `http://arxiv.org/abs/1305.6865`.

[24] F. Nazarov, S. Treil, and A. Volberg, *Accretive system* Tb*-theorems on nonhomogeneous spaces*, Duke Math. J. **113** (2002), no. 2, 259–312.

[25] E. M. Stein, *Topics in harmonic analysis related to the Littlewood–Paley theory*, Annals of Mathematics Studies, Princeton University Press, Princeton, N.J., 1970.

(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

(M.M.) INSTITUT DES HAUTES ETUDES SCIENTIFIQUES, 91440 BURES-SUR-YVETTE, FRANCE
E-mail address: mourgoglou@ihes.fr