

CARLESON OPERATORS ON DOUBLING METRIC MEASURE SPACES

LARS BECKER, FLORIS VAN DOORN, ASGAR JAMNESHAN,
RAJULA SRIVASTAVA, AND CHRISTOPH THIELE

ABSTRACT. Doubling metric measure spaces provide a natural framework for singular integral operators. In contrast, the study of maximally modulated singular integral operators, the so-called Carleson operators, has largely been limited to Euclidean space with modulation functions such as polynomials defined by algebraic means. We present a general axiomatic approach to modulation functions on doubling metric measure spaces and prove L^p bounds for the corresponding Carleson operators in Theorem 1.1 and Theorem 1.2. This generalizes classical and modern results on Carleson operators. In addition to the proofs presented here, our main results have been computer verified using the language Lean and the library mathlib, as documented in the sibling communication [BdFD⁺25].

CONTENTS

1. Introduction	2
2. Proof of Metric Space Carleson, overview	7
2.1. Choosing parameters and preliminary reductions	8
2.2. The dyadic model operator	9
2.3. Partition of the set of tiles into forests and antichains	11
2.4. The cancellation condition with Hölder regularity	13
2.5. Monotonicity of cube metrics	13
3. Existence of a tile structure	14
4. Organization of the set of tiles	16
4.1. Organization of the tiles	16
4.2. Exceptional set estimates	18
4.3. Auxiliary lemmas	20
4.4. Verification of the forest properties	22
4.5. Verification of the antichain property	24
5. Proof of the Antichain Operator Proposition	26
5.1. The density arguments	26
5.2. The tile correlation lemma	28
5.3. The tile count lemma	29
6. Proof of the Forest Operator Proposition	32
6.1. The pointwise tree estimate	32
6.2. An auxiliary L^2 tree estimate	35
6.3. The quantitative L^2 tree estimate	36
6.4. Almost orthogonality of separated trees	38

6.5. Tiles with large separation	39
6.6. The remaining tiles	43
6.7. Forests	45
7. Proof of the Hölder cancellative condition	47
References	47

1. INTRODUCTION

Doubling metric measure spaces and the more general spaces of homogeneous type of Coifman and Weiss [CW71] are a natural environment to define singular integrals. Thanks to the work of Macias and Segovia [MS79], one can pass back and forth between doubling metric measure spaces and spaces of homogeneous type. We refer to the textbook [Ste93] for an account of these spaces.

A singular integral operator S acts on suitable functions f on the doubling metric measure space X and satisfies

$$\int Sf(x)g(x) d\mu(x) := \int \int K(x, y)f(y)g(x) d\mu(y)d\mu(x)$$

for functions f and g with disjoint supports, where the kernel K satisfies Calderón-Zygmund estimates such as (1.8), (1.9). Estimates (1.8), (1.9), are in terms of the metric distance ρ and the measure distance, that is the measure $\mu(B)$ of the smallest ball B containing two points, and have a natural scaling behaviour.

Much of the theory of singular integrals consists of conditional results in the following sense. Conditioned on boundedness of a singular integral operator in the Hilbert space $L^2(X)$, the theory provides other estimates such as L^p bounds, weighted bounds, and sparse bounds, for the singular integral as well as for related operators such as maximal or nontangential maximal truncations. The initial bound often arises externally from Hilbert space techniques in the setting of the given singular integral. More abstractly, one has $S(1)$ and $S(b)$ theorems that obtain the general $L^2(X)$ bound from testing the $L^2(X)$ estimate on subsets consisting of so-called test functions. In the setting of spaces of homogeneous type, [Chr90] provides a general $S(b)$ theorem.

A variant of singular integral operators beyond the basic variety are the maximally modulated singular integrals T , also called Carleson operators. For disjointly supported f and g , they satisfy

$$\int Tf(x)g(x) d\mu = \int \sup_{\vartheta \in \Theta} \left| \int K(x, y)e^{i\vartheta(y)} f(y)g(x) d\mu(y) \right| d\mu(x)$$

for a set Θ of modulation functions. The classical Carleson operator is defined with the Hilbert kernel $K(x, y) = (x - y)^{-1}$ on the real line and the set Θ of linear functions. Estimates for the classical Carleson operator are provided by the Carleson-Hunt theorem and are closely related to the celebrated theorem of Carleson [Car66] on almost everywhere convergence of Fourier series of functions in L^2 . Besides the translation and dilation symmetry of the Hilbert transform that is typical for singular integral theory,

the classical Carleson operator exhibits an additional modulation symmetry. This modulation symmetry mandates techniques that are often referred to as time-frequency analysis and go beyond the core singular integral techniques. Further examples of Carleson operators in the literature include the Stein-Wainger [SW01] operator with polynomial modulations without linear term, thereby avoiding time-frequency analysis, Lie's quadratic Carleson operator [Lie09] with quadratic polynomials including linear terms, and Lie's general polynomial Carleson operator [Lie20] and its further generalization [Zor21]. Mnatsakanyan [Mna22] has encountered a Carleson operator with non-polynomial albeit still algebraically explicit modulations.

The purpose of the present paper is to generalize all these results by defining classes Θ of modulation functions axiomatically on doubling metric measure spaces and proving bounds for the corresponding Carleson operators. The axioms are in terms of a family of metrics on Θ parameterized by balls in X . The metric related to a ball B controls the relative oscillation of two modulation functions on B as in (1.2), and in many instances it can be chosen equal to the left-hand side of (1.2). The axioms demand a number of doubling properties of the metrics very much in the spirit of doubling metric measure spaces, namely (1.3), (1.5), and (1.6). One additional axiom (1.7) that we call the cancellative axiom is of somewhat more technical nature and replaces techniques of partial integration in the Euclidean setting.

Our main Theorem 1.1 is then of conditional nature roughly in the sense described above. It takes as assumption an $L^2(X)$ estimate on an operator related to the singular integral, the maximally truncated non-tangential operator T_* defined in (1.10). Bounds for this operator follow for standard Calderón-Zygmund kernels that are regular in both variables from bounds in $L^2(X)$ of the singular integral itself. We pose bounds for this stronger operator as assumption, because we only demand regularity of the kernel K in one of two variables. We call such K one-sided kernels. They are natural for Carleson operators in view of the maximal construction in the other variable. The non-tangential construction replaces regularity in the other variable. We also formulate a second theorem, Theorem 1.2, concluding boundedness for the so-called linearized Carleson operator with the supremum over modulations replaced by a choice function Q . Having such a choice function fixed, one can demand a weaker hypothesis on an operator (1.14) adapted to the specific choice function Q . This more technical but stronger version is necessary in some applications such as a possible deduction of the Walsh Carleson theorem [Bil67], where the kernel K is closely linked to the linearizing function Q and one only has the weaker hypothesis.

We proceed to introduce the formal setup for our main theorems. The setup is intentionally worded completely parallel to the sibling communication [BdFD⁺25], see also the end of the introduction for further discussion of the relationship between the sibling communications.

We carry a multi purpose parameter, a natural number

$$a \geq 4$$

in our notation. As a gets larger, both the hypotheses and the conclusions of the main theorems will become weaker.

A doubling metric measure space (X, ρ, μ, a) is a complete and locally compact metric space (X, ρ) equipped with a non-zero locally finite Borel measure μ that satisfies the doubling condition that for all $x \in X$ and all $R > 0$ we have

$$\mu(B(x, 2R)) \leq 2^a \mu(B(x, R)), \quad (1.1)$$

where we have denoted by $B(x, R)$ the open ball of radius R centred at x :

$$B(x, R) := \{y \in X : \rho(x, y) < R\}.$$

A collection Θ of real valued continuous functions on the doubling metric measure space (X, ρ, μ, a) is called compatible, if there is a point $o \in X$ where all the functions are equal to 0 and there exists for each ball $B \subset X$ a metric d_B on Θ such that the following five properties (1.2), (1.3), (1.4), (1.5), and (1.6) are satisfied. For every ball $B \subset X$

$$\sup_{x, y \in B} |\vartheta(x) - \vartheta(y) - \theta(x) + \theta(y)| \leq d_B(\vartheta, \theta). \quad (1.2)$$

For any two balls $B_1 = B(x_1, R)$, $B_2 = B(x_2, 2R)$ in X with $x_1 \in B_2$ and any $\vartheta, \theta \in \Theta$,

$$d_{B_2}(\vartheta, \theta) \leq 2^a d_{B_1}(\vartheta, \theta). \quad (1.3)$$

For any two balls B_1, B_2 in X with $B_1 \subset B_2$ and any $\vartheta, \theta \in \Theta$

$$d_{B_1}(\vartheta, \theta) \leq d_{B_2}(\vartheta, \theta) \quad (1.4)$$

and for any two balls $B_1 = B(x_1, R)$, $B_2 = B(x_2, 2^a R)$ with $B_1 \subset B_2$, and $\vartheta, \theta \in \Theta$,

$$2d_{B_1}(\vartheta, \theta) \leq d_{B_2}(\vartheta, \theta). \quad (1.5)$$

For every ball B in X and every d_B -ball \tilde{B} of radius $2R$ in Θ , there is a collection \mathcal{B} of at most 2^a many d_B -balls of radius R covering \tilde{B} , that is,

$$\tilde{B} \subset \bigcup \mathcal{B}. \quad (1.6)$$

Further, a compatible collection Θ is called cancellative, if for any ball B in X of radius R , any Lipschitz function $\varphi : X \rightarrow \mathbb{C}$ supported on B , and any $\vartheta, \theta \in \Theta$ we have

$$\left| \int_B e(\vartheta(x) - \theta(x)) \varphi(x) d\mu(x) \right| \leq 2^a \mu(B) \|\varphi\|_{\text{Lip}(B)} (1 + d_B(\vartheta, \theta))^{-\frac{1}{a}}, \quad (1.7)$$

where $\|\cdot\|_{\text{Lip}(B)}$ denotes the inhomogeneous Lipschitz norm on B :

$$\|\varphi\|_{\text{Lip}(B)} = \sup_{x \in B} |\varphi(x)| + R \sup_{x, y \in B, x \neq y} \frac{|\varphi(x) - \varphi(y)|}{\rho(x, y)}.$$

A one-sided Calderón–Zygmund kernel K on the doubling metric measure space (X, ρ, μ, a) is a measurable function

$$K : X \times X \rightarrow \mathbb{C}$$

such that for all $x, y', y \in X$ with $x \neq y$, we have

$$|K(x, y)| \leq \frac{2^{a^3}}{V(x, y)} \quad (1.8)$$

and if $2\rho(y, y') \leq \rho(x, y)$, then

$$|K(x, y) - K(x, y')| \leq \left(\frac{\rho(y, y')}{\rho(x, y)} \right)^{\frac{1}{a}} \frac{2^{a^3}}{V(x, y)}, \quad (1.9)$$

where

$$V(x, y) := \mu(B(x, \rho(x, y))).$$

Define the maximally truncated non-tangential singular integral T_* associated with K by

$$T_* f(x) := \sup_{R_1 < R_2} \sup_{\rho(x, x') < R_1} \left| \int_{R_1 < \rho(x', y) < R_2} K(x', y) f(y) d\mu(y) \right|. \quad (1.10)$$

We define the generalized Carleson operator T by

$$Tf(x) := \sup_{\vartheta \in \Theta} \sup_{0 < R_1 < R_2} \left| \int_{R_1 < \rho(x, y) < R_2} K(x, y) f(y) e(\vartheta(y)) d\mu(y) \right|, \quad (1.11)$$

where $e(r) = e^{ir}$.

Our main result is the following restricted weak type estimate for T in the range $1 < q \leq 2$, which by interpolation techniques recovers L^q estimates for the open range $1 < q < 2$.

Theorem 1.1 (metric space Carleson). *For all integers $a \geq 4$ and real numbers $1 < q \leq 2$ the following holds. Let (X, ρ, μ, a) be a doubling metric measure space. Let Θ be a cancellative compatible collection of functions and let K be a one-sided Calderón–Zygmund kernel on (X, ρ, μ, a) . Assume that for every bounded measurable function g on X supported on a set of finite measure we have*

$$\|T_* g\|_2 \leq 2^{a^3} \|g\|_2, \quad (1.12)$$

where T_* is defined in (1.10). Then for all Borel sets F and G in X and all Borel functions $f : X \rightarrow \mathbb{C}$ with $|f| \leq \mathbf{1}_F$, we have, with T defined in (1.11),

$$\left| \int_G Tf d\mu \right| \leq \frac{2^{443a^3}}{(q-1)^6} \mu(G)^{1-\frac{1}{q}} \mu(F)^{\frac{1}{q}}. \quad (1.13)$$

For a Borel function $Q : X \rightarrow \Theta$, and $\vartheta \in \Theta$ and $x \in X$ define

$$R_Q(\vartheta, x) = \sup\{r > 0 : d_{B(x, r)}(\vartheta, Q(x)) < 1\}$$

and define further

$$T_Q^\vartheta f(x) := \sup_{R_1 < R_2} \sup_{\rho(x, x') < R_1} \left| \int_{R_1 < \rho(x', y) < \min\{R_2, R_Q(\vartheta, x')\}} K(x', y) f(y) d\mu(y) \right|. \quad (1.14)$$

Define further the linearized generalized Carleson operator T_Q by

$$T_Q f(x) := \sup_{0 < R_1 < R_2} \left| \int_{R_1 < \rho(x, y) < R_2} K(x, y) f(y) e(Q(x)(y)) d\mu(y) \right|, \quad (1.15)$$

where again $e(r) = e^{ir}$.

Theorem 1.2 (linearised metric Carleson). *For all integers $a \geq 4$ and real numbers $1 < q \leq 2$ the following holds. Let (X, ρ, μ, a) be a doubling metric measure space. Let Θ be a cancellative compatible collection of functions. Let $Q : X \rightarrow \Theta$ be a Borel function with finite range. Let K be a one-sided Calderón–Zygmund kernel on (X, ρ, μ, a) . Assume that for every $\vartheta \in \Theta$ and every bounded measurable function g on X supported on a set of finite measure we have*

$$\|T_Q^\vartheta g\|_2 \leq 2^{a^3} \|g\|_2, \quad (1.16)$$

where T_Q^ϑ is defined in (1.14). Then for all bounded Borel sets F and G in X and all Borel functions $f : X \rightarrow \mathbb{C}$ with $|f| \leq \mathbf{1}_F$, we have, with T_Q defined in (1.15),

$$\left| \int_G T_Q f \, d\mu \right| \leq \frac{2^{443a^3}}{(q-1)^6} \mu(G)^{1-\frac{1}{q}} \mu(F)^{\frac{1}{q}}. \quad (1.17)$$

There is extensive literature on Carleson operators. We mention selected results with particular relevance to the approach in this paper and with focus on the variety of classes of modulation functions. The time frequency analysis required to estimate many Carleson operators goes back to Carleson’s seminal paper [Car66] on convergence and growth of Fourier series. Shortly after, this was extended to the dyadic case with Walsh modulations [Bil67]. The first actual L^p bounds for the classical Carleson operator were observed by Hunt [Hun68]. Somewhat dual approaches to Carleson operators were given in [Fef73], [LT00]. In particular, the approach by Fefferman [Fef73] has found applications in [Lie20] and [Zor21], which were inspirational for the present paper. The possibility of more general classes of modulation functions was elucidated in the work of [Mna22], who proves a Carleson-type theorem for the Malmquist–Takenaka series, which leads to modulation functions related to Blaschke products. A generalization of (1.11) from the previously mentioned Euclidean setting into the anisotropic setting that was suggested in [Zor21] is included in our theory.

Recent interest in polynomial Carleson operators was sparked by extensions of the theory of Stein and Wainger [SW01]. Among these are maximally modulated Radon transforms initiated by the work of Pierce and Yung [PY19], discrete analogs of the results of Stein and Wainger as in [KR22], [KR23], and monomial modulation functions with fractional powers avoiding resonances in [GRY20].

The quest to extend the TT^* methods of Stein and Wainger towards Carleson operators which need time frequency analysis led to various threads. These include estimates involving restricted suprema, [GPR⁺17] and results for a simplified singular kernel motivated by Radon transforms [Ram21]. In [BBL⁺24], [GL24], [HL24], time frequency analysis came in through a hybrid construction between the bilinear Hilbert transform and Carleson operators, albeit with non-resonant frequencies. A basic result with resonant modulations was established in [Bec23].

This paper has a sibling communication containing a computer verification of Theorems 1.1 and 1.2 using the language Lean and the library mathlib. More precisely, the ultimate purpose of the sibling communication

is to provide a computer verification of part of the classical Carleson theorem. It does so as an application of Theorem 1.1, which it therefore also proceeds to verify. The axiomatic setup of our theorems is very suitable for computer verification, hence the route to the classical theorem via the modern generalization is natural.

The sibling communication has seventeen authors, twelve authors in addition to the present ones contributing substantially to the coding in Lean. The coding effort took thirteen months. Over a dozen further experts are acknowledged in the sibling communication for additional contributions. Only recently such speed of formalization has become possible. A mathematical statement coded in Lean is correct with certainty as soon as a proof coded in Lean compiles properly. To verify the mathematical statement written in English, the remaining human task is to verify that the translation of the statement into Lean is correct, which amounts to a tiny fraction of work compared with the classical task of checking correctness of a proof. For easy comparison, the statements of Theorems 1.1 and 1.2 together with their assumptions have been formulated here with much detail and in parallel within the corresponding part of the sibling communication.

The present paper announces the new Theorem 1.1 and Theorem 1.2 and presents a forty page English proof with a level of detail that is common in present research mathematics. This text largely preceeded the formalization and is absent in the sibling communication, but in our opinion is the central piece of communication in the present culture of mathematics. The auxiliary statements along with their headings in this paper match statements in the so-called blueprint, a much longer and more detailed proof text attached to the sibling communication that was guiding Lean experts in the distributed effort to produce the necessary Lean code.

Acknowledgement. L.B., F.v.D., R.S., and C.T. were funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2047/1 – 390685813. L.B., R.S., and C.T. were also supported by SFB 1060. A.J. was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - 547294463.

2. PROOF OF METRIC SPACE CARLESON, OVERVIEW

We first note the quick argument that Theorem 1.2 implies Theorem 1.1. By the monotone convergence theorem, we may restrict the supremum in (1.11) to finitely many values of ϑ . Thus assuming Θ finite, we choose $Q(x)$ to be a maximizer of the supremum in (1.11), so that the left hand side of (1.13) equals the left hand side of (1.17). Observing that assumption (1.16) follows from assumption (1.12) as the operator defined in (1.10) is larger than the operator (1.14), Theorem 1.1 follows from Theorem 1.2.

Our proof of Theorem 1.2 is a refinement of [Zor21], which itself is in the tradition of [Fef73] and [Lie20]. The Carleson operator is broken up into pieces (2.20), parameterized by so-called tiles, which are localized in both metric spaces X and Θ and have nice dyadic properties thanks to a grid structure constructed in Proposition 2.1. The bulk of these pieces is regrouped into collections called trees, which come with a geometric parameter n and two density parameters. In a tree, the modulation parameter ϑ

can be assumed constant so that the hypothesis (1.16) can be applied. It is important to collect almost orthogonal trees together into forests, and the resulting forest bound is formulated in Proposition 2.3.

Outside the bulk, we end up with error terms either collected into antichains or thrown into negligible exceptional sets. Antichains are almost orthogonal collections of individual tiles and estimated in Proposition 2.2 without reference to the assumption (1.16).

The pair (X, Θ) has enough of the geometric properties of the Euclidean phase plane to apply adaptations of the tricks in the Euclidean setting to obtain an efficient decomposition into forests and antichains formulated in Proposition 2.4, so that one obtains the desired bounds when summing the bounds from Propositions 2.3 and 2.2 over the various parameter values.

Each of the above mentioned propositions as well as the short auxiliary Proposition 2.5 is proved in its own section in this paper. These sections can be read independently of each other, as all necessary definitions and statements used across the sections are formulated in the rest of the present section. The strong modularity, at the expense of a more technical overview section, is an accomplishment of the present paper and important for the distributed Lean coding effort.

2.1. Choosing parameters and preliminary reductions. We prove Theorem 1.2. Let a, q be given as in Theorem 1.2. Let a doubling metric measure space (X, ρ, μ, a) , a cancellative compatible collection Θ of functions on X , and a point $o \in X$ with $\vartheta(o) = 0$ for all $\vartheta \in \Theta$ be given. Let further a Borel measurable function $Q : X \rightarrow \Theta$ with finite range, a one-sided Calderón–Zygmund kernel K on X so that for every $\vartheta \in \Theta$ the operator T_Q^ϑ defined in (1.14) satisfies (1.16), Borel sets F, G with finite measure, and a measurable function f with $f \leq \mathbf{1}_F$ be given.

We choose parameters

$$D := 2^{100a^2}, \quad (2.1)$$

$$\kappa := 2^{-10a},$$

and

$$Z := 2^{12a}. \quad (2.2)$$

Let $\psi : \mathbb{R} \rightarrow \mathbb{R}$ be the unique compactly supported, piece-wise linear, continuous function with corners at $\frac{1}{4D}$, $\frac{1}{2D}$, $\frac{1}{4}$, and $\frac{1}{2}$ that satisfies

$$\sum_{s \in \mathbb{Z}} \psi(D^{-s}x) = 1$$

for all $x > 0$. This function vanishes outside $[\frac{1}{4D}, \frac{1}{2}]$, is constant one on $[\frac{1}{2D}, \frac{1}{4}]$, and is Lipschitz with constant $4D$. For $s \in \mathbb{Z}$, we define

$$K_s(x, y) := K(x, y)\psi(D^{-s}\rho(x, y)),$$

so that for each $x, y \in X$ with $x \neq y$ we have

$$K(x, y) = \sum_{s \in \mathbb{Z}} K_s(x, y).$$

As a consequence of (1.8) and (1.9), the functions K_s satisfy

$$|K_s(x, y)| \leq \frac{2^{102a^3}}{\mu(B(x, D^s))} \quad (2.3)$$

and

$$|K_s(x, y) - K_s(x, y')| \leq \frac{2^{127a^3}}{\mu(B(x, D^s))} \left(\frac{\rho(y, y')}{D^s} \right)^{\frac{1}{a}}. \quad (2.4)$$

Furthermore, if $K_s(x, y) \neq 0$, then

$$\frac{1}{4}D^{s-1} \leq \rho(x, y) \leq \frac{1}{2}D^s. \quad (2.5)$$

Up to an error that is controlled by the Hardy-Littlewood maximal function, the left hand side of (1.17) is dominated by

$$\int_G \sup_{\sigma_1 \leq \sigma_2} \left| \sum_{s=\sigma_1}^{\sigma_2} \int K_s(x, y) f(y) e(Q(x)(y)) d\mu(y) \right| d\mu(x).$$

By monotone convergence as $S \rightarrow \infty$, we may and do restrict the supremum to finitely many values $-S \leq \sigma_1, \sigma_2 \leq S$. Define $\sigma_1, \sigma_2 : X \rightarrow [-S, S]$ to be two measurable functions selecting a maximizing pair for the supremum. By Fatou's lemma, we may further assume that F, G are bounded sets. We may and do increase S so that F and G are contained in $B(o, \frac{1}{4}D^S)$. Finally, it is enough to prove that there exists $G' \subset G$ with $2\mu(G') \leq \mu(G)$ such that

$$\begin{aligned} \int_{G \setminus G'} \left| \sum_{s=\sigma_1(x)}^{\sigma_2(x)} \int K_s(x, y) f(y) e(Q(x)(y)) d\mu(y) \right| d\mu(x) \\ \leq \frac{2^{442a^3}}{(q-1)^5} \mu(G)^{1-\frac{1}{q}} \mu(F)^{\frac{1}{q}}. \end{aligned} \quad (2.6)$$

Indeed, applying this iteratively to $G_0 = G$, $G_1 = G'_0$, $G_2 = G'_1$ and so forth and summing the resulting geometric series yields (1.13) and (1.17).

With these reductions done, Theorem 1.2 will follow once we prove (2.6).

2.2. The dyadic model operator. We now define notions of grids on X and on Θ , which we will use to define a dyadic model operator for the left hand side of (2.6).

A grid structure (\mathcal{D}, c, s) on X consists of a finite collection \mathcal{D} of pairs (I, k) of Borel sets in X and integers $k \in [-S, S]$, the projection

$$s : \mathcal{D} \rightarrow [-S, S], \quad (I, k) \mapsto k$$

to the second component which is assumed to be surjective and called scale function, and a function $c : \mathcal{D} \rightarrow X$ called center function such that the five properties (2.7), (2.8), (2.9), (2.10), and (2.11) below hold. We call the elements of \mathcal{D} dyadic cubes. By abuse of notation, we will usually write just I for the cube (I, k) , and we will write $I \subset J$ to mean that for two cubes $(I, k), (J, l) \in \mathcal{D}$ we have $I \subset J$ and $k \leq l$.

For each dyadic cube I and each $-S \leq k < s(I)$ we have

$$I \subset \bigcup_{J \in \mathcal{D}: s(J)=k} J. \quad (2.7)$$

Any two non-disjoint dyadic cubes I, J with $s(I) \leq s(J)$ satisfy

$$I \subset J. \quad (2.8)$$

There exists a $I_0 \in \mathcal{D}$ with $s(I_0) = S$ and $c(I_0) = o$ and for all cubes $J \in \mathcal{D}$, we have

$$J \subset I_0. \quad (2.9)$$

For any dyadic cube I ,

$$c(I) \in B(c(I), \frac{1}{4}D^{s(I)}) \subset I \subset B(c(I), 4D^{s(I)}). \quad (2.10)$$

For any dyadic cube I and any t with $tD^{s(I)} \geq D^{-S}$, recalling κ from (2.1),

$$\mu(\{x \in I : \rho(x, X \setminus I) \leq tD^{s(I)}\}) \leq 2t^\kappa \mu(I). \quad (2.11)$$

A tile structure $(\mathfrak{P}, \mathcal{I}, \Omega, \mathcal{Q}, c, s)$ for a given grid structure (\mathcal{D}, c, s) is a finite set \mathfrak{P} of elements called tiles with five maps

$$\begin{aligned} \mathcal{I}: \mathfrak{P} &\rightarrow \mathcal{D} \\ \Omega: \mathfrak{P} &\rightarrow \mathcal{P}(\Theta) \\ \mathcal{Q}: \mathfrak{P} &\rightarrow \mathcal{Q}(X) \\ c: \mathfrak{P} &\rightarrow X \\ s: \mathfrak{P} &\rightarrow \mathbb{Z} \end{aligned}$$

with \mathcal{I} surjective and $\mathcal{P}(\Theta)$ denoting the power set of Θ such that the six properties (2.12), (2.13), (2.14), (2.15), (2.17), and (2.18) below hold. For each dyadic cube I , the restriction of the map Ω to the set

$$\mathfrak{P}(I) = \{\mathfrak{p} : \mathcal{I}(\mathfrak{p}) = I\} \quad (2.12)$$

is injective and we have the disjoint covering property ($\dot{\cup}$ denotes a disjoint union)

$$\mathcal{Q}(X) \subset \dot{\bigcup}_{\mathfrak{p} \in \mathfrak{P}(I)} \Omega(\mathfrak{p}). \quad (2.13)$$

For any tiles $\mathfrak{p}, \mathfrak{q}$ with $\mathcal{I}(\mathfrak{p}) \subset \mathcal{I}(\mathfrak{q})$ and $\Omega(\mathfrak{p}) \cap \Omega(\mathfrak{q}) \neq \emptyset$ we have

$$\Omega(\mathfrak{q}) \subset \Omega(\mathfrak{p}). \quad (2.14)$$

For each tile \mathfrak{p} ,

$$\mathcal{Q}(\mathfrak{p}) \in B_{\mathfrak{p}}(\mathcal{Q}(\mathfrak{p}), 0.2) \subset \Omega(\mathfrak{p}) \subset B_{\mathfrak{p}}(\mathcal{Q}(\mathfrak{p}), 1), \quad (2.15)$$

where

$$B_{\mathfrak{p}}(\vartheta, R) := \{\theta \in \Theta : d_{\mathfrak{p}}(\vartheta, \theta) < R\},$$

and

$$d_{\mathfrak{p}} := d_{B(c(\mathfrak{p}), \frac{1}{4}D^{s(\mathfrak{p})})}. \quad (2.16)$$

We have for each tile \mathfrak{p}

$$c(\mathfrak{p}) = c(\mathcal{I}(\mathfrak{p})), \quad (2.17)$$

$$s(\mathfrak{p}) = s(\mathcal{I}(\mathfrak{p})). \quad (2.18)$$

Proposition 2.1 (grid existence and tile structure). *There exists a grid structure (\mathcal{D}, c, s) . For a given grid structure (\mathcal{D}, c, s) , there exists a tile structure $(\mathfrak{P}, \mathcal{I}, \Omega, \mathcal{Q}, c, s)$.*

We prove Proposition 2.1 in Section 3.

We fix a grid structure (\mathcal{D}, c, s) and a tile structure $(\mathfrak{P}, \mathcal{I}, \Omega, \mathcal{Q}, c, s)$. We define for $\mathfrak{p} \in \mathfrak{P}$

$$E(\mathfrak{p}) = \{x \in \mathcal{I}(\mathfrak{p}) : Q(x) \in \Omega(\mathfrak{p}), \sigma_1(x) \leq s(\mathfrak{p}) \leq \sigma_2(x)\} \quad (2.19)$$

and

$$T_{\mathfrak{p}}f(x) = \mathbf{1}_{E(\mathfrak{p})}(x) \int K_{s(\mathfrak{p})}(x, y) f(y) e(Q(x)(y) - Q(x)(x)) d\mu(y). \quad (2.20)$$

We also introduce, for any collection $\mathfrak{C} \subset \mathfrak{P}$ of tiles, the notation

$$T_{\mathfrak{C}} = \sum_{\mathfrak{p} \in \mathfrak{C}} T_{\mathfrak{p}}.$$

The covering properties (2.7), (2.9) and (2.13) imply that we can rewrite the left hand side of (2.6) as

$$\int_{G \setminus G'} |T_{\mathfrak{P}}f| d\mu. \quad (2.21)$$

2.3. Partition of the set of tiles into forests and antichains. To estimate (2.21), we decompose the collection of tiles \mathfrak{P} into certain subcollections called antichains and forests, and subsequently apply estimates for the operators associated to the subcollections. We proceed by giving relevant definitions and stating the estimates and the decomposition.

We define the relation

$$\mathfrak{p} \leq \mathfrak{p}' \quad (2.22)$$

on $\mathfrak{P} \times \mathfrak{P}$ meaning $\mathcal{I}(\mathfrak{p}) \subset \mathcal{I}(\mathfrak{p}')$ and $\Omega(\mathfrak{p}') \subset \Omega(\mathfrak{p})$. We further define for $\lambda, \lambda' > 0$ the relation

$$\lambda \mathfrak{p} \lesssim \lambda' \mathfrak{p}' \quad (2.23)$$

on $\mathfrak{P} \times \mathfrak{P}$ meaning $\mathcal{I}(\mathfrak{p}) \subset \mathcal{I}(\mathfrak{p}')$ and

$$B_{\mathfrak{p}'}(\mathcal{Q}(\mathfrak{p}'), \lambda') \subset B_{\mathfrak{p}}(\mathcal{Q}(\mathfrak{p}), \lambda).$$

Occasionally, one of λ or λ' is equal to 1, at which point we omit it in the notation (2.23).

We define for a tile \mathfrak{p} and $\lambda > 0$

$$E_1(\mathfrak{p}) := \{x \in \mathcal{I}(\mathfrak{p}) \cap G : Q(x) \in \Omega(\mathfrak{p})\}, \quad (2.24)$$

$$E_2(\lambda, \mathfrak{p}) := \{x \in \mathcal{I}(\mathfrak{p}) \cap G : Q(x) \in B_{\mathfrak{p}}(\mathcal{Q}(\mathfrak{p}), \lambda)\}. \quad (2.25)$$

Given a subset \mathfrak{P}' of \mathfrak{P} , we define $\mathfrak{P}(\mathfrak{P}')$ to be the set of all $\mathfrak{p} \in \mathfrak{P}$ such that there exist $\mathfrak{p}' \in \mathfrak{P}'$ with $\mathcal{I}(\mathfrak{p}) \subset \mathcal{I}(\mathfrak{p}')$. Define the densities

$$\text{dens}_1(\mathfrak{P}') := \sup_{\mathfrak{p}' \in \mathfrak{P}'} \sup_{\lambda \geq 2} \lambda^{-a} \sup_{\mathfrak{p} \in \mathfrak{P}(\mathfrak{P}'), \lambda \mathfrak{p}' \lesssim \lambda \mathfrak{p}} \frac{\mu(E_2(\lambda, \mathfrak{p}))}{\mu(\mathcal{I}(\mathfrak{p}))}, \quad (2.26)$$

$$\text{dens}_2(\mathfrak{P}') := \sup_{\mathfrak{p}' \in \mathfrak{P}'} \sup_{r \geq 4D^s(\mathfrak{p})} \frac{\mu(F \cap B(c(\mathfrak{p}), r))}{\mu(B(c(\mathfrak{p}), r))}. \quad (2.27)$$

An antichain is a subset \mathfrak{A} of \mathfrak{P} such that for any distinct $\mathfrak{p}, \mathfrak{p}' \in \mathfrak{A}$ we do not have $\mathfrak{p} \leq \mathfrak{p}'$. The following estimate for operators associated to antichains is proved in Section 5.

Proposition 2.2 (antichain operator). *For any antichain \mathfrak{A} and for all $f : X \rightarrow \mathbb{C}$ with $|f| \leq \mathbf{1}_F$ and all $g : X \rightarrow \mathbb{C}$ with $|g| \leq \mathbf{1}_G$*

$$\left| \int \bar{g} T_{\mathfrak{A}} f \, d\mu \right| \leq \frac{2^{117a^3}}{q-1} \text{dens}_1(\mathfrak{A})^{\frac{q-1}{8a^4}} \text{dens}_2(\mathfrak{A})^{\frac{1}{q}-\frac{1}{2}} \|f\|_2 \|g\|_2. \quad (2.28)$$

Let $n \geq 0$. An n -forest is a pair $(\mathfrak{U}, \mathfrak{T})$ where \mathfrak{U} is a subset of \mathfrak{P} and \mathfrak{T} is a map assigning to each $\mathbf{u} \in \mathfrak{U}$ a nonempty set $\mathfrak{T}(\mathbf{u}) \subset \mathfrak{P}$ called tree such that the following properties (2.29), (2.30), (2.31), (2.32), (2.33), and (2.34) hold.

For each $\mathbf{u} \in \mathfrak{U}$ and each $\mathbf{p} \in \mathfrak{T}(\mathbf{u})$ we have $\mathcal{I}(\mathbf{p}) \neq \mathcal{I}(\mathbf{u})$ and

$$4\mathbf{p} \lesssim \mathbf{u}. \quad (2.29)$$

For each $\mathbf{u} \in \mathfrak{U}$ and each $\mathbf{p}, \mathbf{p}'' \in \mathfrak{T}(\mathbf{u})$ and $\mathbf{p}' \in \mathfrak{P}$ we have

$$\mathbf{p}, \mathbf{p}'' \in \mathfrak{T}(\mathbf{u}), \mathbf{p} \leq \mathbf{p}' \leq \mathbf{p}'' \implies \mathbf{p}' \in \mathfrak{T}(\mathbf{u}). \quad (2.30)$$

We have

$$\left\| \sum_{\mathbf{u} \in \mathfrak{U}} \mathbf{1}_{\mathcal{I}(\mathbf{u})} \right\|_{\infty} \leq 2^n. \quad (2.31)$$

We have for every $\mathbf{u} \in \mathfrak{U}$

$$\text{dens}_1(\mathfrak{T}(\mathbf{u})) \leq 2^{4a+1-n}. \quad (2.32)$$

We have for $\mathbf{u}, \mathbf{u}' \in \mathfrak{U}$ with $\mathbf{u} \neq \mathbf{u}'$ and $\mathbf{p} \in \mathfrak{T}(\mathbf{u}')$ with $\mathcal{I}(\mathbf{p}) \subset \mathcal{I}(\mathbf{u})$ that, recalling Z from (2.2),

$$d_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), \mathcal{Q}(\mathbf{u})) > 2^{Z(n+1)}. \quad (2.33)$$

We have for every $\mathbf{u} \in \mathfrak{U}$ and $\mathbf{p} \in \mathfrak{T}(\mathbf{u})$ that

$$B(c(\mathbf{p}), 8D^{s(\mathbf{p})}) \subset \mathcal{I}(\mathbf{u}). \quad (2.34)$$

When discussing the decomposition of the set of tiles, we shall with slight abuse of language identify the forest $(\mathfrak{U}, \mathfrak{T})$ with the set $\bigcup_{\mathbf{u} \in \mathfrak{U}} \mathfrak{T}(\mathbf{u})$ of tiles.

The following estimate for operators associated to forests is proved in Section 6.

Proposition 2.3 (forest operator). *For any $n \geq 0$ and any n -forest $(\mathfrak{U}, \mathfrak{T})$ we have for all $f : X \rightarrow \mathbb{C}$ with $|f| \leq \mathbf{1}_F$ and all $g : X \rightarrow \mathbb{C}$ with $|g| \leq \mathbf{1}_G$*

$$\left| \int \bar{g} \sum_{\mathbf{u} \in \mathfrak{U}} T_{\mathfrak{T}(\mathbf{u})} f \, d\mu \right| \leq 2^{440a^3} 2^{-\frac{q-1}{q}n} \text{dens}_2 \left(\bigcup_{\mathbf{u} \in \mathfrak{U}} \mathfrak{T}(\mathbf{u}) \right)^{\frac{1}{q}-\frac{1}{2}} \|f\|_2 \|g\|_2.$$

The next proposition provides an efficient decomposition into antichains and forests.

Proposition 2.4. *There exists a Borel set G' with $2\mu(G') \leq \mu(G)$ such that the following holds. The set \mathfrak{P}' of all tiles $\mathbf{p} \in \mathfrak{P}$ with $\mathcal{I}(\mathbf{p}) \not\subset G'$ can be decomposed as a disjoint union*

$$\mathfrak{P}' = \bigcup_{n \geq 0} \bigcup_{j=0}^{12(n+2)^2} \bigcup_{\mathbf{u} \in \mathfrak{U}_{n,j}} \mathfrak{T}_{n,j}(\mathbf{u}) \cup \bigcup_{n \geq 0} \bigcup_{j=0}^{Z(n+2)^3} \mathfrak{A}_{n,j}$$

where each $(\mathfrak{U}_{n,j}, \mathfrak{T}_{n,j})$ is an n -forest with

$$\text{dens}_2 \left(\bigcup_{u \in \mathfrak{U}_{n,j}} \mathfrak{T}_{n,j}(u) \right) \leq 2^{2a+5} \frac{\mu(F)}{\mu(G)}$$

and each $\mathfrak{A}_{n,j}$ is an antichain with

$$\text{dens}_1(\mathfrak{A}_{n,j}) \leq 2^{4a+1-n}$$

and

$$\text{dens}_2(\mathfrak{A}_{n,j}) \leq 2^{2a+5} \frac{\mu(F)}{\mu(G)}.$$

The proof of Proposition 2.4 is done in Section 4.

Proposition 2.4 gives a set G' and a decomposition of the set of relevant tiles. We apply the triangle inequality to (2.21) along this decomposition and estimate the pieces by Propositions 2.2 and 2.3. The resulting sum adds to less than the bound (2.6), which completes the proof of Theorem 1.2.

2.4. The cancellation condition with Hölder regularity. To bridge the gap between the Hölder regularity condition (1.9) in Theorem 1.2 and the Lipschitz regularity of the testing functions in the cancellative condition (1.7), we follow [Zor21] to formulate a variant of (1.7) in the following proposition proved in Section 7.

Define

$$\tau := \frac{1}{a}.$$

Define for any open ball B of radius R in X the L^∞ -normalized τ -Hölder norm by

$$\|\varphi\|_{C^\tau(B)} = \sup_{x \in B} |\varphi(x)| + R^\tau \sup_{x, y \in B, x \neq y} \frac{|\varphi(x) - \varphi(y)|}{\rho(x, y)^\tau}.$$

Proposition 2.5 (Holder van der Corput). *Let $z \in X$ and $R > 0$ and set $B = B(z, R)$. Let $\varphi : X \rightarrow \mathbb{C}$ be supported on B and satisfy $\|\varphi\|_{C^\tau(2B)} < \infty$. Let $\vartheta, \theta \in \Theta$. Then*

$$\left| \int e(\vartheta(x) - \theta(x)) \varphi(x) d\mu \right| \leq 2^{7a} \mu(B) \|\varphi\|_{C^\tau(2B)} (1 + d_B(\vartheta, \theta))^{-\frac{1}{2a^2+a^3}}.$$

2.5. Monotonicity of cube metrics. We close this section by recording a technical strengthening of the monotonicity (1.4) of the metrics d_B .

Lemma 2.6 (monotone cube metrics). *Let (\mathcal{D}, c, s) be a grid structure. Denote for cubes $I \in \mathcal{D}$*

$$I^\circ := B(c(I), \frac{1}{4} D^{s(I)}). \quad (2.35)$$

Let $I, J \in \mathcal{D}$ with $I \subset J$. Then for all $\vartheta, \theta \in \Theta$ we have

$$d_{I^\circ}(\vartheta, \theta) \leq d_{J^\circ}(\vartheta, \theta),$$

and if $I \neq J$ then we have

$$d_{I^\circ}(\vartheta, \theta) \leq 2^{-95a} d_{J^\circ}(\vartheta, \theta).$$

Proof. If $s(I) \geq s(J)$ then (2.8) and the assumption $I \subset J$ imply $I = J$.

If $s(J) \geq s(I) + 1$, then from the monotonicity (1.4), (2.1) and (1.5)

$$d_{I^\circ}(\vartheta, \theta) \leq d_{B(c(I), 4D^{s(I)})}(\vartheta, \theta) \leq 2^{-100a} d_{B(c(I), 4D^{s(J)})}(\vartheta, \theta). \quad (2.36)$$

Using (2.10), together with the inclusion $I \subset J$, we obtain

$$B(c(I), 4D^{s(J)}) \subset B(c(J), 8D^{s(J)}).$$

Using this together with the monotonicity property (1.4) and (1.3) in (2.36)

$$d_{I^\circ}(\vartheta, \theta) \leq 2^{-100a} d_{B(c(J), 8D^{s(J)})}(\vartheta, \theta) \leq 2^{-95a} d_{B(c(J), \frac{1}{4}D^{s(J)})}(\vartheta, \theta).$$

This proves the second inequality claimed in the lemma, from which the first follows since $a \geq 4$ and hence $2^{-95a} \leq 1$. \square

3. EXISTENCE OF A TILE STRUCTURE

Here we prove Proposition 2.1. Existence of a grid structure was proved in the generality needed here by Christ [Chr90, §3]. We record this as:

Lemma 3.1 (grid existence). *There exists a grid structure (\mathcal{D}, c, s) .*

The next lemma follows [Zor21, Lemma 2.12] to construct a tile structure.

Lemma 3.2 (tile structure). *For a given grid structure (\mathcal{D}, c, s) , there exists a tile structure $(\mathfrak{P}, \mathcal{I}, \Omega, \mathcal{Q}, c, s)$.*

Proof. Choose a grid structure (\mathcal{D}, c, s) . For each $I \in \mathcal{D}$, fix a set $\mathcal{Z} = \mathcal{Z}(I)$ of maximal cardinality such that

$$\mathcal{Z} \subset Q(X)$$

and such that for any $\vartheta, \theta \in \mathcal{Z}$ with $\vartheta \neq \theta$ we have, recalling definition (2.35),

$$B_{I^\circ}(\vartheta, 0.3) \cap B_{I^\circ}(\theta, 0.3) \cap Q(X) = \emptyset.$$

Note that this is clearly possible, as $Q(X)$ is finite. Maximality implies that for each $I \in \mathcal{D}$, we have

$$Q(X) \subset \bigcup_{z \in \mathcal{Z}(I)} B_{I^\circ}(z, 0.7).$$

We define the set of tiles

$$\mathfrak{P} = \{(I, z) : I \in \mathcal{D}, z \in \mathcal{Z}(I)\},$$

and set

$$\mathcal{I}((I, z)) = I, \quad \mathcal{Q}((I, z)) = z.$$

We further set

$$s(\mathfrak{p}) = s(\mathcal{I}(\mathfrak{p})), \quad c(\mathfrak{p}) = c(\mathcal{I}(\mathfrak{p})).$$

Then (2.17) and (2.18) hold by definition. It remains to construct the map Ω , and verify properties (2.13), (2.14) and (2.15).

We first construct an auxiliary map Ω_1 . For each $I \in \mathcal{D}$, we enumerate

$$\mathcal{Z}(I) = \{z_1, \dots, z_M\}.$$

We define $\Omega_1 : \mathfrak{P} \mapsto \mathcal{P}(\Theta)$ as below. Set

$$\Omega_1((I, z_1)) = B_{I^\circ}(z_1, 0.7) \setminus \bigcup_{z \in \mathcal{Z}(I) \setminus \{z_1\}} B_{I^\circ}(z, 0.3)$$

and then define iteratively

$$\Omega_1((I, z_k)) = B_{I^\circ}(z_k, 0.7) \setminus \bigcup_{z \in \mathcal{Z}(I) \setminus \{z_k\}} B_{I^\circ}(z, 0.3) \setminus \bigcup_{i=1}^{k-1} \Omega_1((I, z_i)).$$

Then the sets $\Omega_1(\mathbf{p})$, $\mathbf{p} \in \mathfrak{P}(I)$ are clearly pairwise disjoint, and an induction argument on k shows that their union still contains $\mathcal{Q}(X)$ and that

$$B_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), 0.3) \subset \Omega_1(\mathbf{p}) \subset B_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), 0.7). \quad (3.1)$$

Now we are ready to define the function Ω . We define for all $\mathbf{p} \in \mathfrak{P}(I_0)$

$$\Omega(\mathbf{p}) = \Omega_1(\mathbf{p}). \quad (3.2)$$

For all other cubes $I \in \mathcal{D}$, there exists a unique parent cube $J \supset I$ with $s(J) = s(I) + 1$, and we may assume that $\Omega(\mathbf{q})$ is already defined for $\mathbf{q} \in \mathfrak{P}(J)$. Then we set for $\mathbf{p} \in \mathfrak{P}(I)$

$$\Omega(\mathbf{p}) = \bigcup_{z \in \mathcal{Z}(J) \cap \Omega_1(\mathbf{p})} \Omega((J, z)) \cup B_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), 0.2). \quad (3.3)$$

We now verify that $(\mathfrak{P}, \mathcal{I}, \Omega, \mathcal{Q}, c, s)$ forms a tile structure.

We first verify (2.15). If $I = I_0$, then (2.15) holds for all $\mathbf{p} \in \mathfrak{P}(I)$ by (3.1) and (3.2). Else, we may assume by induction that (2.15) holds for the parent cube J of I . Suppose that $\vartheta \in \Omega(\mathbf{p})$. By (3.3), $\vartheta \in B_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), 0.2)$ or there exists $z \in \mathcal{Z}(J) \cap \Omega_1(\mathbf{p})$ with $\vartheta \in \Omega(J, z)$, so by (3.1)

$$d_{I^\circ}(\mathcal{Q}(\mathbf{p}), \vartheta) \leq d_{I^\circ}(\mathcal{Q}(\mathbf{p}), z) + d_{I^\circ}(z, \vartheta) \leq 0.7 + d_{I^\circ}(z, \vartheta).$$

By Lemma 2.6 and the induction hypothesis, this is estimated by

$$\leq 0.7 + 2^{-95a} d_{J^\circ}(z, \vartheta) \leq 0.7 + 2^{-95a} \cdot 1 < 1.$$

This shows the second inclusion in (2.15), the first holds by definition.

The disjoint covering property (2.13) holds for $I = I_0$, because it holds for Ω_1 . The covering part of (2.13) then clearly follows by downward induction for all other cubes I from the definition (3.3). For disjointedness, suppose that $\vartheta \in B_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), 0.2)$ and $z \in \mathcal{Z}(J)$ is a point with $\vartheta \in \Omega((J, z))$. Then

$$d_{I^\circ}(\mathcal{Q}(\mathbf{p}), z) \leq d_{I^\circ}(\mathcal{Q}(\mathbf{p}), \vartheta) + d_{I^\circ}(\vartheta, z) \leq 0.2 + 2^{-95a} d_{J^\circ}(\vartheta, z) < 0.3.$$

Thus, by (3.1), $z \in \mathcal{Z}(J) \cap \Omega_1(\mathbf{p})$. By downward induction, it follows that the sets $\Omega(\mathbf{p})$, $\mathbf{p} \in \mathfrak{P}(I)$ are pairwise disjoint.

We turn to the grid condition (2.14). We first prove it when $s(\mathbf{q}) = s(\mathbf{p}) + 1$. Let $I = \mathcal{I}(\mathbf{p})$. Since $I \subset \mathcal{I}(\mathbf{q})$ and $s(\mathbf{q}) = s(\mathbf{p}) + 1$, we have $\mathcal{I}(\mathbf{q}) = J$, where J is the parent cube in the definition (3.3), in particular $\Omega(\mathbf{q}) \subset \Omega(\mathbf{p})$ as needed. If $s(\mathbf{q}) > s(\mathbf{p}) + 1$, the grid condition (2.14) follows by induction on the difference $s(\mathbf{q}) - s(\mathbf{p})$. Indeed, pick \mathbf{p}' with $s(\mathbf{p}') = s(\mathbf{q}) - 1$ and $\mathcal{I}(\mathbf{p}) \subset \mathcal{I}(\mathbf{p}')$ and $\Omega(\mathbf{q}) \subset \Omega(\mathbf{p}')$. This exists by the covering property (2.13) and the part of (2.14) that we already proved. Then $\Omega(\mathbf{p}') \cap \Omega(\mathbf{p}) \neq \emptyset$, so by induction $\Omega(\mathbf{p}') \subset \Omega(\mathbf{p})$. \square

4. ORGANIZATION OF THE SET OF TILES

In this section, we prove Proposition 2.4. The overall proof happens in Section 4.1, where we define the exceptional set G' , the size of which is estimated in Section 4.2, and decompose the set \mathfrak{P} of tiles outside the exceptional set into forests and antichains. Section 4.3 contains auxiliary lemmas needed for the verification of the forest- and antichain properties, which happens in Section 4.4 and Section 4.5.

4.1. Organization of the tiles. In the following definitions, k, n , and j will be nonnegative integers.

We start by sorting tiles \mathbf{p} based on the size of the intersection of the spatial cube $\mathcal{I}(\mathbf{p})$ and its parents with G . Define $\mathcal{C}(G, k)$ to be the set of $I \in \mathcal{D}$ such that there exists a $J \in \mathcal{D}$ with $I \subset J$ and

$$\mu(G \cap J) > 2^{-k-1} \mu(J), \quad (4.1)$$

but there does not exist a $J \in \mathcal{D}$ with $I \subset J$ and

$$\mu(G \cap J) > 2^{-k} \mu(J). \quad (4.2)$$

Let

$$\mathfrak{P}(k) = \{\mathbf{p} \in \mathfrak{P} : \mathcal{I}(\mathbf{p}) \in \mathcal{C}(G, k)\}.$$

We need a notion of density ‘restricted to $\mathfrak{P}(k)$ ’. Define for $\mathfrak{P}' \subset \mathfrak{P}(k)$

$$\text{dens}'_k(\mathfrak{P}') := \sup_{\mathbf{p}' \in \mathfrak{P}'} \sup_{\lambda \geq 2} \lambda^{-a} \sup_{\mathbf{p} \in \mathfrak{P}(k) : \lambda \mathbf{p}' \lesssim \lambda \mathbf{p}} \frac{\mu(E_2(\lambda, \mathbf{p}))}{\mu(\mathcal{I}(\mathbf{p}))}. \quad (4.3)$$

Sorting tiles further by density we define

$$\mathfrak{C}(k, n) := \{\mathbf{p} \in \mathfrak{P}(k) : 2^{4a} 2^{-n} < \text{dens}'_k(\{\mathbf{p}\}) \leq 2^{4a} 2^{-n+1}\}. \quad (4.4)$$

Our goal is to split $\mathfrak{C}(k, n)$ into a controlled number of n -forests. The main difficulty is to achieve separation of the involved trees. This is done following a trick of Fefferman [Fef73]. Define $\mathfrak{M}(k, n)$ to be the set of maximal with respect to \leq tiles $\mathbf{p} \in \mathfrak{P}(k)$ such that

$$\mu(E_1(\mathbf{p})) > 2^{-n} \mu(\mathcal{I}(\mathbf{p})). \quad (4.5)$$

We define for $\mathbf{p} \in \mathfrak{C}(k, n)$

$$\mathfrak{B}(\mathbf{p}) := \{\mathbf{m} \in \mathfrak{M}(k, n) : 100\mathbf{p} \lesssim \mathbf{m}\} \quad (4.6)$$

and

$$\mathfrak{C}_1(k, n, j) := \{\mathbf{p} \in \mathfrak{C}(k, n) : 2^j \leq |\mathfrak{B}(\mathbf{p})| < 2^{j+1}\}$$

and

$$\mathfrak{L}_0(k, n) := \{\mathbf{p} \in \mathfrak{C}(k, n) : |\mathfrak{B}(\mathbf{p})| < 1\}.$$

To increase separation of the trees, we remove the $Z(n+1) + 1$ minimal layers of tiles in each $\mathfrak{C}_1(k, n, j)$. Each minimal layer clearly forms an antichain, so this results in at most $Z(n+1) + 1$ antichains $\mathfrak{L}_1(k, n, j, l)$, where $0 \leq l \leq Z(n+1)$, and a collection of leftover tiles

$$\mathfrak{C}_2(k, n, j) := \mathfrak{C}_1(k, n, j) \setminus \bigcup_{0 \leq l \leq Z(n+1)} \mathfrak{L}_1(k, n, j, l). \quad (4.7)$$

The remaining tile organization will be relative to prospective tree tops, which we define now. Define

$$\mathfrak{U}_1(k, n, j) \quad (4.8)$$

to be the set of all $\mathbf{u} \in \mathfrak{C}_1(k, n, j)$ such that for all $\mathbf{p} \in \mathfrak{C}_1(k, n, j)$ with $\mathcal{I}(\mathbf{u})$ strictly contained in $\mathcal{I}(\mathbf{p})$ we have $B_{\mathbf{u}}(\mathcal{Q}(\mathbf{u}), 100) \cap B_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), 100) = \emptyset$.

We first remove the pairs that are outside the immediate reach of any of the prospective tree tops. Define

$$\mathfrak{L}_2(k, n, j)$$

to be the set of all $\mathbf{p} \in \mathfrak{C}_2(k, n, j)$ such that there does not exist $\mathbf{u} \in \mathfrak{U}_1(k, n, j)$ with $\mathcal{I}(\mathbf{p}) \neq \mathcal{I}(\mathbf{u})$ and $2\mathbf{p} \lesssim \mathbf{u}$. Define

$$\mathfrak{C}_3(k, n, j) := \mathfrak{C}_2(k, n, j) \setminus \mathfrak{L}_2(k, n, j). \quad (4.9)$$

We next remove the $Z(n+1) + 1$ maximal layers in $\mathfrak{C}_3(k, n, j)$, resulting in antichains $\mathfrak{L}_3(k, n, j, l)$, where $0 \leq l \leq Z(n+1)$, and the remainder

$$\mathfrak{C}_4(k, n, j) := \mathfrak{C}_3(k, n, j) \setminus \bigcup_{0 \leq l \leq Z(n+1)} \mathfrak{L}_3(k, n, j, l). \quad (4.10)$$

Finally, we remove tiles close to the boundary of the prospective tree tops, this is to ensure that (2.34) holds. Define

$$\mathcal{L}(\mathbf{u}) \quad (4.11)$$

to be the set of all $I \in \mathcal{D}$ with $I \subset \mathcal{I}(\mathbf{u})$ and $s(I) = s(\mathbf{u}) - Z(n+1) - 1$ and

$$B(c(I), 8D^{s(I)}) \not\subset \mathcal{I}(\mathbf{u}).$$

Define

$$\mathfrak{L}_4(k, n, j)$$

to be the set of all $\mathbf{p} \in \mathfrak{C}_4(k, n, j)$ such that there exists $\mathbf{u} \in \mathfrak{U}_1(k, n, j)$ with $\mathcal{I}(\mathbf{p}) \subset \bigcup \mathcal{L}(\mathbf{u})$, and define

$$\mathfrak{C}_5(k, n, j) := \mathfrak{C}_4(k, n, j) \setminus \mathfrak{L}_4(k, n, j).$$

We define three exceptional sets. The first exceptional set G_1 takes into account the ratio of the measures of F and G . Define $\mathfrak{P}_{F,G}$ to be the set of all $\mathbf{p} \in \mathfrak{P}$ with

$$\text{dens}_2(\{\mathbf{p}\}) > 2^{2a+5} \frac{\mu(F)}{\mu(G)}.$$

Define

$$G_1 := \bigcup_{\mathbf{p} \in \mathfrak{P}_{F,G}} \mathcal{I}(\mathbf{p}).$$

For an integer $\lambda \geq 0$, define $A(\lambda, k, n)$ to be the set of all $x \in X$ such that

$$\sum_{\mathbf{p} \in \mathfrak{M}(k, n)} \mathbf{1}_{\mathcal{I}(\mathbf{p})}(x) > \lambda 2^{n+1} \quad (4.12)$$

and define

$$G_2 := \bigcup_{k \geq 0} \bigcup_{k \leq n} A(2n+6, k, n).$$

Define

$$G_3 := \bigcup_{k \geq 0} \bigcup_{n \geq k} \bigcup_{0 \leq j \leq 2n+3} \bigcup_{\mathbf{p} \in \mathfrak{L}_4(k, n, j)} \mathcal{I}(\mathbf{p}). \quad (4.13)$$

Define $G' = G_1 \cup G_2 \cup G_3$. The following bound of the measure of G' will be proved in Section 4.2.

Lemma 4.1 (exceptional set). *We have*

$$\mu(G') \leq 2^{-1} \mu(G).$$

In Section 4.4, we prove the following lemma. Recall that we defined

$$\mathfrak{P}' = \{\mathfrak{p} \in \mathfrak{P} : \mathcal{I}(\mathfrak{p}) \not\subset G'\}.$$

Lemma 4.2 (forest union). *Let $n \geq k$ and $0 \leq j \leq 2n + 3$. Then the set $\mathfrak{C}_5(k, n, j) \cap \mathfrak{P}'$ is the union of at most $4n + 12$ many n -forests.*

In Section 4.5, we prove the following lemma.

Lemma 4.3 (forest complement). *There exists a decomposition of the set of tiles not contained in any forest into a disjoint union of antichains*

$$\mathfrak{P}' \setminus \bigcup_{k \geq 0} \bigcup_{n \geq k} \bigcup_{0 \leq j \leq 2n+3} \mathfrak{C}_5(k, n, j) = \bigcup_{n \geq 0} \bigcup_{j=0}^{Z(n+2)^3} \mathfrak{A}_{n,j}, \quad (4.14)$$

where each $\mathfrak{A}_{n,j}$ is an antichain with

$$\text{dens}_1(\mathfrak{A}_{n,j}) \leq 2^{4a+1-n}. \quad (4.15)$$

Note that by the definition of G_1 , for all $\mathfrak{C} \subset \mathfrak{P}'$

$$\text{dens}_2(\mathfrak{C}) \leq 2^{2a+5} \frac{\mu(F)}{\mu(G)}.$$

Lemma 4.1, Lemma 4.2 and Lemma 4.3 then prove Proposition 2.4.

4.2. Exceptional set estimates. We prove bounds for G_1 , G_2 and G_3 in (4.16), (4.17) and (4.20) below. Summing the bounds proves Lemma 4.1.

The set G_1 is contained in the set $\{M\mathbf{1}_F > 2^{2a+5} \mu(F)/\mu(G)\}$, so the weak L^1 bound for the Hardy-Littlewood maximal function implies

$$\mu(G_1) \leq 2^{-5} \mu(G). \quad (4.16)$$

We turn to the set G_2 .

Lemma 4.4 (John Nirenberg). *For all integers $k, n, \lambda \geq 0$, we have*

$$\mu(A(\lambda, k, n)) \leq 2^{k+1-\lambda} \mu(G).$$

Proof. By pairwise disjointness of the sets $E_1(\mathfrak{p})$, $\mathfrak{p} \in \mathfrak{M}(n, k)$ and (4.5), we have for all grid cubes J the Carleson packing condition

$$\sum_{\mathfrak{p} \in \mathfrak{M}(n, k): \mathcal{I}(\mathfrak{p}) \subset J} \mu(\mathcal{I}(\mathfrak{p})) \leq 2^n \sum_{\mathfrak{p} \in \mathfrak{M}(n, k): \mathcal{I}(\mathfrak{p}) \subset J} \mu(E_1(\mathfrak{p})) \leq 2^n \mu(J).$$

By the John-Nirenberg inequality, it follows that for all grid cubes J

$$\mu\left(\left\{x \in J : \sum_{\mathfrak{p} \in \mathfrak{M}(n, k): \mathcal{I}(\mathfrak{p}) \subset J} \mathbf{1}_{\mathcal{I}(\mathfrak{p})}(x) > \lambda 2^{n+1}\right\}\right) \leq 2^{-\lambda} \mu(J).$$

Denote the set of maximal cubes J with $\mu(G \cap J) > 2^{-k-1}\mu(J)$ by $\mathcal{M}^*(k)$. Each cube $\mathcal{I}(\mathbf{p})$ with $\mathbf{p} \in \mathfrak{M}(n, k)$ is contained in a cube $J \in \mathcal{M}^*(k)$. Hence, summing the last display over all cubes $J \in \mathcal{M}^*(k)$ gives

$$\mu(A(\lambda, k, n)) = \sum_{J \in \mathcal{M}^*(k)} \mu(A(\lambda, k, n) \cap J) \leq 2^{-\lambda} \sum_{J \in \mathcal{M}^*(k)} \mu(J).$$

Using pairwise disjointness of the cubes in $\mathcal{M}^*(k)$ we conclude

$$\leq 2^{k+1-\lambda} \sum_{J \in \mathcal{M}^*(k)} \mu(J \cap G) \leq 2^{k+1-\lambda} \mu(G). \quad \square$$

Using Lemma 4.4 and summing twice a geometric series yields

$$\mu(G_2) \leq \sum_{0 \leq k} \sum_{k \leq n} \mu(A(2n+6, k, n)) \leq 2^{-2} \mu(G). \quad (4.17)$$

We turn to the set G_3 .

Lemma 4.5 (top tiles). *We have*

$$\sum_{\mathbf{m} \in \mathfrak{M}(k, n)} \mu(\mathcal{I}(\mathbf{m})) \leq 2^{n+k+3} \mu(G). \quad (4.18)$$

Proof. We have for the left-hand side of (4.18)

$$\int \sum_{\mathbf{m} \in \mathfrak{M}(k, n)} \mathbf{1}_{\mathcal{I}(\mathbf{m})}(x) d\mu(x) \leq 2^{n+1} \sum_{\lambda=0}^{|\mathfrak{M}|} \mu(A(\lambda, k, n)).$$

Now the claimed estimate follows from Lemma 4.4. \square

Lemma 4.6 (tree count). *Let $k, n, j \geq 0$. We have for every $x \in X$*

$$\sum_{\mathbf{u} \in \mathfrak{U}_1(k, n, j)} \mathbf{1}_{\mathcal{I}(\mathbf{u})}(x) \leq 2^{-j} 2^{9a} \sum_{\mathbf{m} \in \mathfrak{M}(k, n)} \mathbf{1}_{\mathcal{I}(\mathbf{m})}(x).$$

Proof. Let $x \in X$. For each $\mathbf{u} \in \mathfrak{U}_1(k, n, j)$ with $x \in \mathcal{I}(\mathbf{u})$, as $\mathbf{u} \in \mathfrak{C}_1(k, n, j)$, there are at least 2^j elements $\mathbf{m} \in \mathfrak{M}(k, n)$ with $100\mathbf{u} \lesssim \mathbf{m}$. Hence

$$\mathbf{1}_{\mathcal{I}(\mathbf{u})}(x) \leq 2^{-j} \sum_{\mathbf{m} \in \mathfrak{M}(k, n): 100\mathbf{u} \lesssim \mathbf{m}} \mathbf{1}_{\mathcal{I}(\mathbf{m})}(x). \quad (4.19)$$

For each $\mathbf{m} \in \mathfrak{M}(k, n)$ with $x \in \mathcal{I}(\mathbf{m})$, let $\mathfrak{U}(\mathbf{m})$ be the set of $\mathbf{u} \in \mathfrak{U}_1(k, n, j)$ with $x \in \mathcal{I}(\mathbf{u})$ and $100\mathbf{u} \lesssim \mathbf{m}$. Summing (4.19) over \mathbf{u} gives

$$\sum_{\mathbf{u} \in \mathfrak{U}_1(k, n, j)} \mathbf{1}_{\mathcal{I}(\mathbf{u})}(x) \leq 2^{-j} \sum_{\mathbf{m} \in \mathfrak{M}(k, n)} \sum_{\mathbf{u} \in \mathfrak{U}(\mathbf{m})} \mathbf{1}_{\mathcal{I}(\mathbf{m})}(x).$$

It remains to show that $|\mathfrak{U}(\mathbf{m})| \leq 2^{9a}$. Let $\mathbf{u} \in \mathfrak{U}(\mathbf{m})$, then by definition

$$d_{\mathbf{u}}(\mathcal{Q}(\mathbf{u}), \mathcal{Q}(\mathbf{m})) \leq 100.$$

If \mathbf{u}' is a further element in $\mathfrak{U}(\mathbf{m})$ with $\mathbf{u} \neq \mathbf{u}'$, then

$$\mathcal{Q}(\mathbf{m}) \in B_{\mathbf{u}}(\mathcal{Q}(\mathbf{u}), 100) \cap B_{\mathbf{u}'}(\mathcal{Q}(\mathbf{u}'), 100).$$

By the definition of $\mathfrak{U}_1(k, n, j)$, none of $\mathcal{I}(\mathbf{u}), \mathcal{I}(\mathbf{u}')$ is strictly contained in the other. As both contain x , we have $\mathcal{I}(\mathbf{u}) = \mathcal{I}(\mathbf{u}')$ and in particular $d_{\mathbf{u}} = d_{\mathbf{u}'}$.

The geometric doubling condition (1.6) for $d_{\mathbf{u}}$ implies that every collection of disjoint balls $B_{\mathbf{u}}(\vartheta, 0.2)$ with $\vartheta \in B_{\mathbf{u}}(\mathcal{Q}(\mathbf{m}), 100)$ has size at most 2^{9a} .

Applying this to the balls with centers $\mathcal{Q}(\mathbf{u}'), \mathbf{u}' \in \mathfrak{U}(\mathfrak{m})$, which are disjoint by (2.13), completes the proof. \square

Lemma 4.7 (boundary exception). *Let $\mathcal{L}(\mathbf{u})$ be as defined in (4.11). For each $\mathbf{u} \in \mathfrak{U}_1(k, n, l)$*

$$\mu\left(\bigcup_{I \in \mathcal{L}(\mathbf{u})} I\right) \leq 2 \cdot D^{-\kappa Z(n+1)} \mu(\mathcal{I}(\mathbf{u})).$$

Proof. Let $\mathbf{u} \in \mathfrak{U}_1(k, n, l)$ and $I \in \mathcal{L}(\mathbf{u})$. By definition of $\mathcal{L}(\mathbf{u})$

$$I \subset X(\mathbf{u}) := \{x \in \mathcal{I}(\mathbf{u}) : \rho(x, X \setminus \mathcal{I}(\mathbf{u})) \leq 12D^{s(\mathbf{u})-Z(n+1)-1}\}.$$

By the small boundary property (2.11) and $D \geq 12$

$$\mu(X(\mathbf{u})) \leq 2 \cdot (12D^{-Z(n+1)-1})^\kappa \mu(\mathcal{I}(\mathbf{u})) \leq 2 \cdot D^{-\kappa Z(n+1)} \mu(\mathcal{I}(\mathbf{u})).$$

\square

Lemma 4.8 (third exception). *We have*

$$\mu(G_3) \leq 2^{-4} \mu(G). \quad (4.20)$$

Proof. As each $\mathbf{p} \in \mathfrak{L}_4(k, n, j)$ is contained in $\bigcup \mathcal{L}(\mathbf{u})$ for some $\mathbf{u} \in \mathfrak{U}_1(k, n, l)$, we have

$$\mu\left(\bigcup_{\mathbf{p} \in \mathfrak{L}_4(k, n, j)} \mathcal{I}(\mathbf{p})\right) \leq \sum_{\mathbf{u} \in \mathfrak{U}_1(k, n, j)} \mu\left(\bigcup_{I \in \mathcal{L}(\mathbf{u})} I\right).$$

Using Lemma 4.7, Lemma 4.6 and finally Lemma 4.5, we estimate this further by

$$\begin{aligned} &\leq 2 \sum_{\mathbf{u} \in \mathfrak{U}_1(k, n, j)} D^{-\kappa Z(n+1)} \mu(\mathcal{I}(\mathbf{u})) \leq 2^{9a+1-j} \sum_{\mathbf{m} \in \mathfrak{M}(k, n)} D^{-\kappa Z(n+1)} \mu(\mathcal{I}(\mathbf{m})) \\ &\leq 2^{9a+1-j} D^{-\kappa Z(n+1)} 2^{n+k+3} \mu(G). \end{aligned}$$

Now we estimate G_3 defined in (4.13) by

$$\begin{aligned} \mu(G_3) &\leq \sum_{k \geq 0} \sum_{n \geq k} \sum_{0 \leq j \leq 2n+3} \mu\left(\bigcup_{\mathbf{p} \in \mathfrak{L}_4(k, n, j)} \mathcal{I}(\mathbf{p})\right) \\ &\leq \sum_{k \geq 0} \sum_{n \geq k} \sum_{0 \leq j \leq 2n+3} 2^{9a+4+n+k-j} D^{-\kappa Z(n+1)} \mu(G) \end{aligned}$$

Summing the series, using the definitions of D, Z and κ , proves the lemma. \square

4.3. Auxiliary lemmas. Before proving Lemma 4.2 and Lemma 4.3, we collect some useful properties of \lesssim .

Lemma 4.9 (wiggle order 2). *Let $n, m \geq 1$ and $k > 0$. If $\mathbf{p}, \mathbf{p}' \in \mathfrak{P}$ with $\mathcal{I}(\mathbf{p}) \neq \mathcal{I}(\mathbf{p}')$ and $n\mathbf{p} \lesssim k\mathbf{p}'$ then*

$$(n + 2^{-95a}m)\mathbf{p} \lesssim m\mathbf{p}'. \quad (4.21)$$

Proof. The assumption implies that $\mathcal{I}(\mathbf{p}) \subsetneq \mathcal{I}(\mathbf{p}')$. Let $\vartheta \in B_{\mathbf{p}'}(\mathcal{Q}(\mathbf{p}'), m)$. Then we have by the triangle inequality

$$d_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), \vartheta) \leq d_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), \mathcal{Q}(\mathbf{p}')) + d_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}'), \vartheta)$$

The first summand is bounded by n since

$$\mathcal{Q}(\mathbf{p}') \in B_{\mathbf{p}'}(\mathcal{Q}(\mathbf{p}'), k) \subset B_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), n).$$

Using Lemma 2.6 for the second summand shows

$$d_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), \vartheta) \leq n + 2^{-95a} d_{\mathbf{p}'}(\mathcal{Q}(\mathbf{p}'), \vartheta) < n + 2^{-95a} m.$$

Combined with $\mathcal{I}(\mathbf{p}) \subset \mathcal{I}(\mathbf{p}')$, this yields (4.21). \square

Lemma 4.10 (wiggle order 3). *The following implications hold for all $\mathbf{q}, \mathbf{q}' \in \mathfrak{P}$:*

$$\mathbf{q} \leq \mathbf{q}' \text{ and } \lambda \geq 1.1 \implies \lambda \mathbf{q} \lesssim \lambda \mathbf{q}', \quad (4.22)$$

$$10\mathbf{q} \lesssim \mathbf{q}' \text{ and } \mathcal{I}(\mathbf{q}) \neq \mathcal{I}(\mathbf{q}') \implies 100\mathbf{q} \lesssim 100\mathbf{q}', \quad (4.23)$$

$$2\mathbf{q} \lesssim \mathbf{q}' \text{ and } \mathcal{I}(\mathbf{q}) \neq \mathcal{I}(\mathbf{q}') \implies 4\mathbf{q} \lesssim 500\mathbf{q}'. \quad (4.24)$$

Proof. Claims (4.23) and (4.24) are consequences of Lemma 4.9 and $a \geq 4$. For (4.22), if $\mathcal{I}(\mathbf{q}) = \mathcal{I}(\mathbf{q}')$ then $\mathbf{q} = \mathbf{q}'$ by (2.13) and the order definition (2.22). If $\mathcal{I}(\mathbf{q}) \neq \mathcal{I}(\mathbf{q}')$, then from (2.15) and the order definitions (2.22) and (2.23) it follows that $\mathbf{q} \lesssim 0.2\mathbf{q}'$, and (4.22) follows from Lemma 4.9. \square

We call a collection \mathfrak{A} of tiles convex if

$$\mathbf{p} \leq \mathbf{p}' \leq \mathbf{p}'' \text{ and } \mathbf{p}, \mathbf{p}'' \in \mathfrak{A} \implies \mathbf{p}' \in \mathfrak{A}.$$

With the help of Lemma 4.9 and Lemma 4.10, it is easy to verify that each of the collections $\mathfrak{P}(k)$, $\mathfrak{C}(n, k)$ and $\mathfrak{C}_s(n, k, j)$ for $s = 1, 2, 3, 4, 5$ are convex.

We close this subsection with an estimate for densities.

Lemma 4.11 (dens compare). *For every $k, n \geq 0$ and every $\mathfrak{A} \subset \mathfrak{C}(k, n)$:*

$$\text{dens}_1(\mathfrak{A}) \leq 2^{4a} 2^{-n+1}.$$

Proof. We first show that for all sets $\mathfrak{A} \subset \mathfrak{P}(k)$ it holds

$$\text{dens}_1(\mathfrak{A}) \leq \text{dens}'_k(\mathfrak{A}). \quad (4.25)$$

It suffices to show that for all $\mathbf{p}' \in \mathfrak{A}$ and $\lambda \geq 2$ and $\mathbf{p} \in \mathfrak{P}(\mathfrak{A})$ with $\lambda \mathbf{p}' \lesssim \lambda \mathbf{p}$ we have

$$\frac{\mu(E_2(\lambda, \mathbf{p}))}{\mu(\mathcal{I}(\mathbf{p}))} \leq \sup_{\mathbf{p}'' \in \mathfrak{P}(k): \lambda \mathbf{p}' \lesssim \lambda \mathbf{p}''} \frac{\mu(E_2(\lambda, \mathbf{p}''))}{\mu(\mathcal{I}(\mathbf{p}''))}.$$

But if $\mathbf{p} \in \mathfrak{P}(\mathfrak{A})$ then it satisfies (4.1) and (4.2), so $\mathbf{p} \in \mathfrak{P}(k)$. Thus we can simply take $\mathbf{p}'' = \mathbf{p}$ and the inequality along with (4.25) follows.

Combining (4.25) with definitions (4.3) and (4.4) we have that

$$\text{dens}_1(\mathfrak{C}(n, k)) \leq \text{dens}'_k(\mathfrak{C}(k, n)) = \sup_{\mathbf{p} \in \mathfrak{C}(k, n)} \text{dens}'_k(\{\mathbf{p}\}) \leq 2^{4a} 2^{-n+1}.$$

The lemma follows since dens_1 is increasing with respect to set inclusion. \square

4.4. Verification of the forest properties. We prove Lemma 4.2. Fix $k, n, j \geq 0$. Define

$$\mathfrak{C}_6(k, n, j)$$

to be the set of all tiles $\mathfrak{p} \in \mathfrak{C}_5(k, n, j)$ such that $\mathcal{I}(\mathfrak{p}) \not\subset G'$. Since $\mathfrak{C}_5(k, n, j)$ is convex, so is $\mathfrak{C}_6(k, n, j)$. The following chain of lemmas establishes that the set $\mathfrak{C}_6(k, n, j)$ can be written as a union of a small number of n -forests.

For $\mathfrak{u} \in \mathfrak{U}_1(k, n, j)$, define

$$\mathfrak{T}_1(\mathfrak{u}) := \{\mathfrak{p} \in \mathfrak{C}_1(k, n, j) : \mathcal{I}(\mathfrak{p}) \neq \mathcal{I}(\mathfrak{u}), 2\mathfrak{p} \lesssim \mathfrak{u}\}.$$

Define

$$\mathfrak{U}_2(k, n, j) := \{\mathfrak{u} \in \mathfrak{U}_1(k, n, j) : \mathfrak{T}_1(\mathfrak{u}) \cap \mathfrak{C}_6(k, n, j) \neq \emptyset\}.$$

Define a relation \sim on $\mathfrak{U}_2(k, n, j)$ by setting $\mathfrak{u} \sim \mathfrak{u}'$ for $\mathfrak{u}, \mathfrak{u}' \in \mathfrak{U}_2(k, n, j)$ if $\mathfrak{u} = \mathfrak{u}'$ or there exists \mathfrak{p} in $\mathfrak{T}_1(\mathfrak{u})$ with $10\mathfrak{p} \lesssim \mathfrak{u}'$.

Lemma 4.12 (relation geometry). *If $\mathfrak{u} \sim \mathfrak{u}'$, then $\mathcal{I}(\mathfrak{u}) = \mathcal{I}(\mathfrak{u}')$ and*

$$B_{\mathfrak{u}}(\mathcal{Q}(\mathfrak{u}), 100) \cap B_{\mathfrak{u}'}(\mathcal{Q}(\mathfrak{u}'), 100) \neq \emptyset.$$

Proof. Let $\mathfrak{u}, \mathfrak{u}' \in \mathfrak{U}_2(k, n, j)$ with $\mathfrak{u} \sim \mathfrak{u}'$ and $\mathfrak{u} \neq \mathfrak{u}'$. Then there exists $\mathfrak{p} \in \mathfrak{C}_1(k, n, j)$ such that $\mathcal{I}(\mathfrak{p}) \neq \mathcal{I}(\mathfrak{u})$ and $2\mathfrak{p} \lesssim \mathfrak{u}$ and $10\mathfrak{p} \lesssim \mathfrak{u}'$. Using Lemma 4.10, we deduce that

$$100\mathfrak{p} \lesssim 100\mathfrak{u}, \quad 100\mathfrak{p} \lesssim 100\mathfrak{u}'. \quad (4.26)$$

Now suppose that $B_{\mathfrak{u}}(\mathcal{Q}(\mathfrak{u}), 100)$ and $B_{\mathfrak{u}'}(\mathcal{Q}(\mathfrak{u}'), 100)$ are disjoint. Then $\mathfrak{B}(\mathfrak{u})$ and $\mathfrak{B}(\mathfrak{u}')$ are disjoint, but also $\mathfrak{B}(\mathfrak{u}) \subset \mathfrak{B}(\mathfrak{p})$ and $\mathfrak{B}(\mathfrak{u}') \subset \mathfrak{B}(\mathfrak{p})$, by (4.6), (2.23) and (4.26). Hence,

$$|\mathfrak{B}(\mathfrak{p})| \geq |\mathfrak{B}(\mathfrak{u})| + |\mathfrak{B}(\mathfrak{u}')| \geq 2^j + 2^j = 2^{j+1},$$

which contradicts $\mathfrak{p} \in \mathfrak{C}_1(k, n, j)$. Therefore we must have

$$B_{\mathfrak{u}}(\mathcal{Q}(\mathfrak{u}), 100) \cap B_{\mathfrak{u}'}(\mathcal{Q}(\mathfrak{u}'), 100) \neq \emptyset. \quad (4.27)$$

Since $\mathcal{I}(\mathfrak{p}) \subset \mathcal{I}(\mathfrak{u})$ and $\mathcal{I}(\mathfrak{p}) \subset \mathcal{I}(\mathfrak{u}')$, the cubes $\mathcal{I}(\mathfrak{u})$ and $\mathcal{I}(\mathfrak{u}')$ are nested. Combining this with (4.27) and definition (4.8) of $\mathfrak{U}_1(k, n, j)$, we conclude that $\mathcal{I}(\mathfrak{u}) = \mathcal{I}(\mathfrak{u}')$. \square

Lemma 4.13 (equivalence relation). *For each k, n, j , the relation \sim on $\mathfrak{U}_2(k, n, j)$ is an equivalence relation.*

Proof. Reflexivity holds by definition. For transitivity, pick pairwise distinct $\mathfrak{u}, \mathfrak{u}', \mathfrak{u}'' \in \mathfrak{U}_1(k, n, j)$ with $\mathfrak{u} \sim \mathfrak{u}'$, $\mathfrak{u}' \sim \mathfrak{u}''$. By Lemma 4.12, it follows that $\mathcal{I}(\mathfrak{u}) = \mathcal{I}(\mathfrak{u}') = \mathcal{I}(\mathfrak{u}'')$, that there exists

$$\vartheta \in B_{\mathfrak{u}}(\mathcal{Q}(\mathfrak{u}), 100) \cap B_{\mathfrak{u}'}(\mathcal{Q}(\mathfrak{u}'), 100)$$

and that there exists

$$\theta \in B_{\mathfrak{u}'}(\mathcal{Q}(\mathfrak{u}'), 100) \cap B_{\mathfrak{u}''}(\mathcal{Q}(\mathfrak{u}''), 100).$$

We now estimate for $q \in B_{\mathfrak{u}''}(\mathcal{Q}(\mathfrak{u}''), 1)$

$$\begin{aligned} d_{\mathfrak{u}}(\mathcal{Q}(\mathfrak{u}), q) &\leq d_{\mathfrak{u}}(\mathcal{Q}(\mathfrak{u}), \vartheta) + d_{\mathfrak{u}}(\vartheta, \mathcal{Q}(\mathfrak{u}')) \\ &\quad + d_{\mathfrak{u}}(\mathcal{Q}(\mathfrak{u}'), \theta) + d_{\mathfrak{u}}(\theta, \mathcal{Q}(\mathfrak{u}'')) + d_{\mathfrak{u}}(\mathcal{Q}(\mathfrak{u}''), q). \end{aligned}$$

Using (2.16) and the fact that $\mathcal{I}(\mathbf{u}) = \mathcal{I}(\mathbf{u}') = \mathcal{I}(\mathbf{u}'')$ this equals

$$\begin{aligned} d_{\mathbf{u}}(\mathcal{Q}(\mathbf{u}), \vartheta) + d_{\mathbf{u}'}(\vartheta, \mathcal{Q}(\mathbf{u}')) + d_{\mathbf{u}'}(\mathcal{Q}(\mathbf{u}'), \theta) + d_{\mathbf{u}''}(\theta, \mathcal{Q}(\mathbf{u}'')) + d_{\mathbf{u}''}(\mathcal{Q}(\mathbf{u}''), q) \\ < 100 + 100 + 100 + 100 + 1 < 500. \end{aligned}$$

Since $\mathbf{u} \sim \mathbf{u}'$, there exists some $\mathbf{p} \in \mathfrak{T}_1(\mathbf{u})$ with $10\mathbf{p} \lesssim \mathbf{u}'$. By (4.24), $\mathbf{p} \in \mathfrak{T}_1(\mathbf{u})$ implies $4\mathbf{p} \lesssim 500\mathbf{u}$, from which it follows that $d_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), q) < 4 < 10$. We have shown that $B_{\mathbf{u}''}(\mathcal{Q}(\mathbf{u}''), 1) \subset B_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), 10)$, combining this with $\mathcal{I}(\mathbf{u}'') = \mathcal{I}(\mathbf{u})$ gives $\mathbf{u} \sim \mathbf{u}''$.

For symmetry suppose that $\mathbf{u} \sim \mathbf{u}'$, so $\mathcal{I}(\mathbf{u}) = \mathcal{I}(\mathbf{u}')$ and there exists $\vartheta \in B_{\mathbf{u}}(\mathcal{Q}(\mathbf{u}), 100) \cap B_{\mathbf{u}'}(\mathcal{Q}(\mathbf{u}'), 100)$. We may assume that $\mathbf{u} \neq \mathbf{u}'$. There exists $\mathbf{p} \in \mathfrak{T}_1(\mathbf{u}')$, which then satisfies $2\mathbf{p} \lesssim \mathbf{u}'$ and $\mathcal{I}(\mathbf{p}) \neq \mathcal{I}(\mathbf{u}')$. By (4.24)

$$4\mathbf{p} \lesssim 500\mathbf{u}'. \quad (4.28)$$

If $q \in B_{\mathbf{u}}(\mathcal{Q}(\mathbf{u}), 1)$ then we have, using that $\mathcal{I}(\mathbf{u}) = \mathcal{I}(\mathbf{u}')$ and hence $d_{\mathbf{u}} = d_{\mathbf{u}'}$:

$$d_{\mathbf{u}'}(\mathcal{Q}(\mathbf{u}'), q) \leq d_{\mathbf{u}'}(\mathcal{Q}(\mathbf{u}'), \vartheta) + d_{\mathbf{u}}(\vartheta, \mathcal{Q}(\mathbf{u})) + d_{\mathbf{u}}(\mathcal{Q}(\mathbf{u}), q) < 500.$$

Combined with (4.28) we obtain $B_{\mathbf{u}}(\mathcal{Q}(\mathbf{u}), 1) \subset B_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), 4)$, so

$$10\mathbf{p} \lesssim 4\mathbf{p} \lesssim \mathbf{u} \quad (4.29)$$

and consequently $\mathbf{u}' \sim \mathbf{u}$. \square

Choose a set $\mathfrak{U}_3(k, n, j)$ of representatives for the equivalence classes of \sim in $\mathfrak{U}_2(k, n, j)$. Define for each $\mathbf{u} \in \mathfrak{U}_3(k, n, j)$

$$\mathfrak{T}_2(\mathbf{u}) := \bigcup_{\mathbf{u} \sim \mathbf{u}'} \mathfrak{T}_1(\mathbf{u}') \cap \mathfrak{C}_6(k, n, j). \quad (4.30)$$

It is straightforward to check that each $\mathfrak{T}_2(\mathbf{u})$ is convex, meaning (2.30) holds. By construction

$$\mathfrak{C}_6(k, n, j) = \bigcup_{\mathbf{u} \in \mathfrak{U}_3(k, n, j)} \mathfrak{T}_2(\mathbf{u}).$$

We now check that $(\mathfrak{C}_6(k, n, j), \mathfrak{T}_2)$ satisfies also the forest properties (2.29), (2.32), (2.33) and (2.34).

Lemma 4.14 (forest geometry). *For each $\mathbf{u} \in \mathfrak{U}_3(k, n, j)$, the set $\mathfrak{T}_2(\mathbf{u})$ satisfies (2.29).*

Proof. Let $\mathbf{p} \in \mathfrak{T}_2(\mathbf{u})$. By (4.30), there exists $\mathbf{u}' \sim \mathbf{u}$ with $\mathbf{p} \in \mathfrak{T}_1(\mathbf{u}')$. The proof of (4.29) shows that $4\mathbf{p} \lesssim \mathbf{u}$. Also $\mathcal{I}(\mathbf{p}) \neq \mathcal{I}(\mathbf{u}')$ and by Lemma 4.12 $\mathcal{I}(\mathbf{u}) = \mathcal{I}(\mathbf{u}')$, so $\mathcal{I}(\mathbf{p}) \neq \mathcal{I}(\mathbf{u})$. \square

Lemma 4.15 (forest separation). *For each $\mathbf{u}, \mathbf{u}' \in \mathfrak{U}_3(k, n, j)$ with $\mathbf{u} \neq \mathbf{u}'$ and each $\mathbf{p} \in \mathfrak{T}_2(\mathbf{u})$ with $\mathcal{I}(\mathbf{p}) \subset \mathcal{I}(\mathbf{u}')$ we have*

$$d_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), \mathcal{Q}(\mathbf{u}')) > 2^{Z(n+1)}.$$

Proof. By the definition (4.7) of $\mathfrak{C}_2(k, n, j)$, there exists a tile $\mathbf{p}' \in \mathfrak{C}_1(k, n, j)$ with $\mathbf{p}' \leq \mathbf{p}$ and $s(\mathbf{p}') \leq s(\mathbf{p}) - Z(n+1)$. By Lemma 2.6, we have

$$d_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), \mathcal{Q}(\mathbf{u}')) \geq 2^{95aZ(n+1)} d_{\mathbf{p}'}(\mathcal{Q}(\mathbf{p}'), \mathcal{Q}(\mathbf{u}')). \quad (4.31)$$

Since $\mathbf{p} \in \mathfrak{T}_2(\mathbf{u})$, there exists $\mathbf{v} \sim \mathbf{u}$ with $2\mathbf{p}' \lesssim 2\mathbf{p} \lesssim \mathbf{v}$ and $\mathcal{I}(\mathbf{p}') \neq \mathcal{I}(\mathbf{v})$. Since \mathbf{u}, \mathbf{u}' are not equivalent under \sim , neither are \mathbf{v} and \mathbf{u}' , thus $10\mathbf{p}' \not\lesssim \mathbf{u}'$. This implies that there exists $q \in B_{\mathbf{u}'}(\mathcal{Q}(\mathbf{u}'), 1) \setminus B_{\mathbf{p}'}(\mathcal{Q}(\mathbf{p}'), 10)$.

From $\mathbf{p}' \leq \mathbf{p}$ and $\mathcal{I}(\mathbf{p}') \subset \mathcal{I}(\mathbf{p}) \subset \mathcal{I}(\mathbf{u}')$ and Lemma 2.6 it then follows that

$$\begin{aligned} d_{\mathbf{p}'}(\mathcal{Q}(\mathbf{p}), \mathcal{Q}(\mathbf{u}')) &\geq -d_{\mathbf{p}'}(\mathcal{Q}(\mathbf{p}), \mathcal{Q}(\mathbf{p}')) + d_{\mathbf{p}'}(\mathcal{Q}(\mathbf{p}'), q) - d_{\mathbf{u}'}(q, \mathcal{Q}(\mathbf{u}')) \\ &> -1 + 10 - 1 = 8. \end{aligned}$$

Combining this with (4.31) completes the proof. \square

Lemma 4.16 (forest inner). *For each $\mathbf{u} \in \mathfrak{U}_3(k, n, j)$ and each $\mathbf{p} \in \mathfrak{T}_2(\mathbf{u})$ we have*

$$B(c(\mathbf{p}), 8D^{s(\mathbf{p})}) \subset \mathcal{I}(\mathbf{u}).$$

Proof. Let $\mathbf{p} \in \mathfrak{T}_2(\mathbf{u})$, so in particular $\mathbf{p} \in \mathfrak{C}_4(k, n, j)$. By the definition (4.10) of $\mathfrak{C}_4(k, n, j)$, there exists a tile $\mathbf{q} \in \mathfrak{C}_3(n, k, j)$ with $\mathbf{p} \leq \mathbf{q}$ and $s(\mathbf{p}) \leq s(\mathbf{q}) - Z(n+1)$. By the definition (4.9) of $\mathfrak{C}_3(n, k, j)$, there exists $\mathbf{u}'' \in \mathfrak{U}_1(k, n, j)$ with $2\mathbf{q} \lesssim \mathbf{u}''$ and $s(\mathbf{q}) < s(\mathbf{u}'')$. Then we have in particular that $10\mathbf{p} \lesssim \mathbf{u}''$, so, using transitivity of \sim , we have $\mathbf{u} \sim \mathbf{u}''$. Lemma 4.12 shows that $\mathcal{I}(\mathbf{u}'') = \mathcal{I}(\mathbf{u})$, hence $s(\mathbf{q}) < s(\mathbf{u})$ and $s(\mathbf{p}) \leq s(\mathbf{q}) - Z(n+1) \leq s(\mathbf{u}) - Z(n+1) - 1$.

Let $I \in \mathcal{D}$ be the cube with $s(I) = s(\mathbf{u}) - Z(n+1) - 1$ and $I \subset \mathcal{I}(\mathbf{u})$ and $\mathcal{I}(\mathbf{p}) \subset I$. Since $\mathbf{p} \in \mathfrak{C}_5(k, n, j)$, we have that $I \notin \mathcal{L}(\mathbf{u})$, so $B(c(I), 8D^{s(I)}) \subset \mathcal{I}(\mathbf{u})$. The same then holds for the subcube $\mathcal{I}(\mathbf{p}) \subset I$. \square

It remains to prove the final forest property (2.31). This is accomplished by the following lemma, which implies that $\mathfrak{C}_6(k, n, j)$ is the union of at most $4n + 12$ sub-collections satisfying (2.31), which are then n -forests.

Lemma 4.17 (forest stacking). *It holds for $k \leq n$ that*

$$\sum_{\mathbf{u} \in \mathfrak{U}_3(k, n, j)} \mathbf{1}_{\mathcal{I}(\mathbf{u})} \leq (4n + 12)2^n.$$

Proof. Suppose to the contrary that a point x is contained in more than $(4n + 12)2^n$ cubes $\mathcal{I}(\mathbf{u})$ with $\mathbf{u} \in \mathfrak{U}_3(k, n, j)$. Since $\mathfrak{U}_3(k, n, j) \subset \mathfrak{C}_1(k, n, j)$, for each such \mathbf{u} , there exists $\mathbf{m} \in \mathfrak{M}(k, n)$ such that $100\mathbf{u} \lesssim \mathbf{m}$. We fix such an $\mathbf{m}(\mathbf{u}) := \mathbf{m}$ for each \mathbf{u} , and claim that the map $\mathbf{u} \mapsto \mathbf{m}(\mathbf{u})$ is injective. Indeed, assume for $\mathbf{u} \neq \mathbf{u}'$ there is $\mathbf{m} \in \mathfrak{M}(k, n)$ such that $100\mathbf{u} \lesssim \mathbf{m}$ and $100\mathbf{u}' \lesssim \mathbf{m}$. By (2.8), either $\mathcal{I}(\mathbf{u}) \subset \mathcal{I}(\mathbf{u}')$ or $\mathcal{I}(\mathbf{u}') \subset \mathcal{I}(\mathbf{u})$. By (4.8), the balls $B_{\mathbf{u}}(\mathcal{Q}(\mathbf{u}), 100)$ and $B_{\mathbf{u}'}(\mathcal{Q}(\mathbf{u}'), 100)$ are disjoint. This contradicts $\Omega(\mathbf{m})$ being contained in both sets by (2.15). Thus x is contained in more than $(4n + 12)2^n$ cubes $\mathcal{I}(\mathbf{m})$, $\mathbf{m} \in \mathfrak{M}(k, n)$. Consequently, we have by (4.12) that $x \in A(2n + 6, k, n) \subset G_2$. Let $\mathcal{I}(\mathbf{u})$ be an inclusion minimal cube among the $\mathcal{I}(\mathbf{u}'), \mathbf{u}' \in \mathfrak{U}_3(k, n, j)$ with $x \in \mathcal{I}(\mathbf{u})$. It satisfies

$$\mathcal{I}(\mathbf{u}) \subset \{y : \sum_{\mathbf{u} \in \mathfrak{U}_3(k, n, j)} \mathbf{1}_{\mathcal{I}(\mathbf{u})}(y) > 1 + (4n + 12)2^n\} \subset G_2,$$

thus $\mathfrak{T}_1(\mathbf{u}) \cap \mathfrak{C}_6(k, n, j) = \emptyset$. This contradicts $\mathbf{u} \in \mathfrak{U}_2(k, n, j)$. \square

4.5. Verification of the antichain property. We prove Lemma 4.3. We first claim that the set on the left side of (4.14) equals

$$\begin{aligned} &\mathfrak{P}' \cap \left[\bigcup_{k \geq 0} \bigcup_{n \geq k} [\mathfrak{L}_0(k, n) \cup \bigcup_{0 \leq j \leq 2n+3} \mathfrak{L}_2(k, n, j)] \right. \\ &\quad \left. \cup \bigcup_{k \geq 0} \bigcup_{n \geq k} \bigcup_{0 \leq j \leq 2n+3} \bigcup_{0 \leq l \leq Z(n+1)} [\mathfrak{L}_1(k, n, j, l) \cup \mathfrak{L}_3(k, n, j, l)] \right]. \end{aligned} \quad (4.32)$$

To see the claim, let $\mathbf{p} \in \mathfrak{P}(k) \cap \mathfrak{P}'$. By (4.2),

$$\mu(E_2(\lambda, \mathbf{p}')) \leq 2^{-k} \mu(\mathcal{I}(\mathbf{p}'))$$

whenever $\mathbf{p}' \in \mathfrak{P}(k)$ with $\mathcal{I}(\mathbf{p}) \subset \mathcal{I}(\mathbf{p}')$, so $\text{dens}'_k(\{\mathbf{p}\}) \leq 2^{-k}$. Hence there exists $n \geq k$ with $\mathbf{p} \in \mathfrak{C}(k, n)$. Also, since $\mathcal{I}(\mathbf{p}) \not\subset G'$, there exist at most $1 + (4n + 12)2^n < 2^{2n+4}$ tiles $\mathbf{m} \in \mathfrak{M}(k, n)$ with $\mathbf{p} \leq \mathbf{m}$. It follows that $\mathbf{p} \in \mathfrak{L}_0(k, n)$ or $\mathbf{p} \in \mathfrak{C}_1(k, n, j)$ for some $1 \leq j \leq 2n + 3$. The claim now follows as the construction of the collections $\mathfrak{C}_5(k, n, j)$ consists of removing from $\mathfrak{C}_1(k, n, j)$ the collections $\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3$ and \mathcal{L}_4 with parameters as claimed in (4.32), where we have omitted \mathcal{L}_4 because all tiles $\mathbf{p} \in \mathfrak{L}_4(k, n, j)$ satisfy $\mathcal{I}(\mathbf{p}) \subset G'$ and are therefore not contained in \mathfrak{P}' .

Lemma 4.11 implies the density estimate (4.15) for all terms in the decomposition. Moreover, \mathfrak{L}_1 and \mathfrak{L}_3 were constructed as minimal or maximal sets of tiles, thus they are all antichains. It remains to verify that \mathfrak{L}_0 and \mathfrak{L}_2 can also be decomposed into a small number of antichains.

Lemma 4.18 (L0 antichain). *For all $k \geq 0$ and $n \geq k$, the set $\mathfrak{L}_0(k, n)$ is the union of at most n antichains.*

Proof. It suffices to show that $\mathfrak{L}_0(k, n)$ contains no chain of length $n + 1$. Suppose that we had such a chain $\mathbf{p}_0 \leq \mathbf{p}_1 \leq \dots \leq \mathbf{p}_n$ of pairwise distinct \mathbf{p}_i . Since $\mathbf{p}_n \in \mathfrak{C}(k, n)$, we have that $\text{dens}'_k(\{\mathbf{p}_n\}) > 2^{-n}$, which by definition means that there exists $\mathbf{p}' \in \mathfrak{P}(k)$ and $\lambda \geq 2$ with $\lambda \mathbf{p}_n \leq \lambda \mathbf{p}'$ and

$$\frac{\mu(E_2(\lambda, \mathbf{p}'))}{\mu(\mathcal{I}(\mathbf{p}'))} > \lambda^a 2^{4a} 2^{-n}. \quad (4.33)$$

Let \mathfrak{D} be the set of all $\mathbf{p}'' \in \mathfrak{P}(k)$ such that we have $\mathcal{I}(\mathbf{p}'') = \mathcal{I}(\mathbf{p}')$ and $B_{\mathbf{p}'}(\mathcal{Q}(\mathbf{p}'), \lambda) \cap \Omega(\mathbf{p}'') \neq \emptyset$. From the geometric doubling property (1.6) and the fact (2.15) that the balls $B_{\mathbf{p}'}(\mathcal{Q}(\mathbf{p}''), 0.2)$, $\mathbf{p}'' \in \mathfrak{D}$ are disjoint, it follows that

$$|\mathfrak{D}| \leq 2^{4a} \lambda^a.$$

By the definitions (2.24) and (2.25) we have $E_2(\lambda, \mathbf{p}') \subset \bigcup_{\mathbf{p}'' \in \mathfrak{D}} E_1(\mathbf{p}'')$, thus

$$\sum_{\mathbf{p}'' \in \mathfrak{D}} \frac{\mu(E_1(\mathbf{p}''))}{\mu(\mathcal{I}(\mathbf{p}''))} > 2^{4a} \lambda^a 2^{-n}.$$

Hence there exists a tile $\mathbf{p}'' \in \mathfrak{D}$ with

$$\mu(E_1(\mathbf{p}'')) \geq 2^{-n} \mu(\mathcal{I}(\mathbf{p}'')). \quad (4.34)$$

From (4.33), the inclusion $E_2(\lambda, \mathbf{p}') \subset \mathcal{I}(\mathbf{p}')$ and $a \geq 1$ we obtain

$$2^n \geq 2^{4a} \lambda^a \geq \lambda.$$

Let $\mathbf{m} \in \mathfrak{M}(k, n)$ with $\mathbf{p}'' \leq \mathbf{m}$, it exists by (4.34). From Lemma 2.6 and $a \geq 1$, it holds for all $\vartheta \in B_{\mathbf{m}}(\mathcal{Q}(\mathbf{m}), 1)$ that

$$\begin{aligned} d_{\mathbf{p}_0}(\mathcal{Q}(\mathbf{p}_0), \vartheta) &\leq d_{\mathbf{p}_0}(\mathcal{Q}(\mathbf{p}_0), \mathcal{Q}(\mathbf{p}_n)) + d_{\mathbf{p}_0}(\mathcal{Q}(\mathbf{p}_n), \mathcal{Q}(\mathbf{p}')) + d_{\mathbf{p}_0}(\mathcal{Q}(\mathbf{p}'), \mathcal{Q}(\mathbf{p}'')) \\ &\quad + d_{\mathbf{p}_0}(\mathcal{Q}(\mathbf{p}''), \mathcal{Q}(\mathbf{m})) + d_{\mathbf{p}_0}(\mathcal{Q}(\mathbf{m}), \vartheta) \\ &\leq 1 + 2^{-95an} (d_{\mathbf{p}_n}(\mathcal{Q}(\mathbf{p}_n), \mathcal{Q}(\mathbf{p}')) + d_{\mathbf{p}'}(\mathcal{Q}(\mathbf{p}'), \mathcal{Q}(\mathbf{p}'')) \\ &\quad + d_{\mathbf{p}''}(\mathcal{Q}(\mathbf{p}''), \mathcal{Q}(\mathbf{m})) + d_{\mathbf{m}}(\mathcal{Q}(\mathbf{m}), \vartheta)) \\ &\leq 1 + 2^{-95an} (\lambda + (\lambda + 1) + 1 + 1) \leq 100. \end{aligned}$$

This implies that $100\mathbf{p}_0 \lesssim \mathbf{m}$, a contradiction to $\mathbf{p}_0 \in \mathfrak{L}_0(k, n)$. \square

Lemma 4.19 (L2 antichain). *Each of the sets $\mathfrak{L}_2(k, n, j)$ is an antichain.*

Proof. Suppose that there are $\mathbf{p}_0, \mathbf{p}_1 \in \mathfrak{L}_2(k, n, j)$ with $\mathbf{p}_0 \neq \mathbf{p}_1$ and $\mathbf{p}_0 \leq \mathbf{p}_1$. By Lemma 4.9, it follows that $2\mathbf{p}_0 \lesssim 200\mathbf{p}_1$. Since $\mathfrak{C}_1(k, n, j)$ is finite, there exists a chain $2\mathbf{p}_0 \lesssim 200\mathbf{p}_1 \lesssim \dots \lesssim 200\mathbf{p}_l$ of distinct $\mathbf{p}_i \in \mathfrak{C}_1(k, n, j)$ of maximal length l . Since $2\mathbf{p}_0 \lesssim 200\mathbf{p}_l \lesssim \mathbf{p}_l$ and $\mathbf{p}_0 \in \mathfrak{L}_2(k, n, j)$, we have $\mathbf{p}_l \notin \mathfrak{U}_1(k, n, j)$. So, by definition of $\mathfrak{U}_1(k, n, j)$, there exists $\mathbf{p}_{l+1} \in \mathfrak{C}_1(k, n, j)$ with $\mathcal{I}(\mathbf{p}_l) \subsetneq \mathcal{I}(\mathbf{p}_{l+1})$ and $B_{\mathbf{p}_l}(\mathcal{Q}(\mathbf{p}_l), 100) \cap B_{\mathbf{p}_{l+1}}(\mathcal{Q}(\mathbf{p}_{l+1}), 100) \neq \emptyset$. Using Lemma 2.6, it is straightforward to deduce $200\mathbf{p}_l \lesssim 200\mathbf{p}_{l+1}$. This contradicts maximality of l . \square

5. PROOF OF THE ANTICHAIN OPERATOR PROPOSITION

Let an antichain \mathfrak{A} and functions f, g as in Proposition 2.2 be given. We prove in Section 5.1 two inequalities (5.1) and (5.3), each involving one of the two densities. The claimed estimate (2.28) follows as the product of the $(2 - q)$ -th power of (5.1) and the $(q - 1)$ -st power of (5.3).

The proof of (5.1) will need a careful estimate formulated in Lemma 5.3 of the TT^* correlation between two tile operators. Lemma 5.3 will be proven in Subsection 5.2.

The summation of the contributions of these individual correlations will require a geometric Lemma 5.4 counting the relevant tile pairs. Lemma 5.4 will be proven in Subsection 5.3.

5.1. The density arguments. By the definition of $E(\mathbf{p})$ and the ordering \leq on tiles, the sets $E(\mathbf{p})$, $\mathbf{p} \in \mathfrak{A}$ are pairwise disjoint. This will be used repeatedly below. Set

$$\tilde{q} = \frac{2q}{1+q}.$$

Since $1 < q \leq 2$, we have $1 < \tilde{q} < q \leq 2$.

Lemma 5.1 (dens2 antichain). *We have that*

$$\left| \int \bar{g} T_{\mathfrak{A}} f \, d\mu \right| \leq 2^{111a^3} (q-1)^{-1} \text{dens}_2(\mathfrak{A})^{\frac{1}{q}-\frac{1}{2}} \|f\|_2 \|g\|_2. \quad (5.1)$$

Proof. Let \mathcal{B} be the collection of balls

$$\{B(c(\mathbf{p}), 8D^{s(\mathbf{p})}) : \mathbf{p} \in \mathfrak{A}\}.$$

From disjointness of the $E(\mathbf{p})$, $\mathbf{p} \in \mathfrak{A}$, the triangle inequality and the kernel support (2.5) and upper bound (2.3) it follows that for every x

$$|T_{\mathfrak{A}} f(x)| \leq 2^{107a^3} \sup_{x \in B \in \mathcal{B}} \frac{1}{\mu(B)} \int_B |f| \, d\mu. \quad (5.2)$$

We have $f = \mathbf{1}_F f$. Using Hölder's inequality, that $1 < \tilde{q} \leq 2$ and the definition of dens_2 , we obtain for each $x \in B'$ and each $B' \in \mathcal{B}$

$$\frac{1}{\mu(B')} \int_{B'} |f(y)| \, d\mu(y) \leq \left(M(|f|^{\frac{2\tilde{q}}{3\tilde{q}-2}})(x) \right)^{\frac{3}{2}-\frac{1}{\tilde{q}}} \text{dens}_2(\mathfrak{A})^{\frac{1}{q}-\frac{1}{2}}.$$

Combining this with (5.2) and boundedness of the Hardy-Littlewood maximal function completes the proof. \square

Lemma 5.2 (dens1 antichain). *Set $p := 4a^4$. Then*

$$\left| \int \bar{g} T_{\mathfrak{A}} f \, d\mu \right| \leq 2^{117a^3} \text{dens}_1(\mathfrak{A})^{\frac{1}{2p}} \|f\|_2 \|g\|_2. \quad (5.3)$$

Proof. We have by expanding the square

$$\int \left| \sum_{\mathfrak{p} \in \mathfrak{A}} T_{\mathfrak{p}}^* g(y) \right|^2 d\mu(y) \leq 2 \sum_{\mathfrak{p} \in \mathfrak{A}} \sum_{\mathfrak{p}' \in \mathfrak{A}: s(\mathfrak{p}') \leq s(\mathfrak{p})} \left| \int T_{\mathfrak{p}}^* g(y) \overline{T_{\mathfrak{p}'}^* g(y)} \, d\mu(y) \right|. \quad (5.4)$$

Define for $\mathfrak{p} \in \mathfrak{P}$ the ball

$$B(\mathfrak{p}) := B(c(\mathfrak{p}), 14D^{s(\mathfrak{p})})$$

and the collection of tiles interacting with \mathfrak{p}

$$\mathfrak{A}(\mathfrak{p}) := \{\mathfrak{p}' \in \mathfrak{A} : s(\mathfrak{p}') \leq s(\mathfrak{p}) \wedge \mathcal{I}(\mathfrak{p}') \subset B(\mathfrak{p})\}.$$

Using Lemma 5.3 and the doubling property (1.1), we estimate (5.4) by

$$\leq 2^{232a^3+6a+1} \sum_{\mathfrak{p} \in \mathfrak{A}} \int_{E(\mathfrak{p})} |g|(y) h(\mathfrak{p}) \, d\mu(y) \quad (5.5)$$

with $h(\mathfrak{p})$ defined as

$$\frac{1}{\mu(B(\mathfrak{p}))} \int \sum_{\mathfrak{p}' \in \mathfrak{A}(\mathfrak{p})} (1 + d_{\mathfrak{p}'}(\mathcal{Q}(\mathfrak{p}'), \mathcal{Q}(\mathfrak{p})))^{-1/(2a^2+a^3)} (\mathbf{1}_{E(\mathfrak{p}')} |g|)(y') \, d\mu(y').$$

Note that $p \geq 4$ since $a > 4$. By Hölder, using $|g| \leq \mathbf{1}_G$ and $E(\mathfrak{p}') \subset B(\mathfrak{p})$,

$$h(\mathfrak{p}) \leq \frac{\|g \mathbf{1}_{B(\mathfrak{p})}\|_{p'}}{\mu(B(\mathfrak{p}))} \left\| \sum_{\mathfrak{p}' \in \mathfrak{A}(\mathfrak{p})} (1 + d_{\mathfrak{p}'}(\mathcal{Q}(\mathfrak{p}), \mathcal{Q}(\mathfrak{p}')))^{-1/(2a^2+a^3)} \mathbf{1}_{E(\mathfrak{p}')} \mathbf{1}_G \right\|_p.$$

We apply Lemma 5.4 to estimate this by

$$2^{5a} \text{dens}_1(\mathfrak{A})^{\frac{1}{p}} \frac{\|g \mathbf{1}_{B(\mathfrak{p})}\|_{p'}}{\mu(B(\mathfrak{p}))^{\frac{1}{p'}}} \leq 2^{5a} \text{dens}_1(\mathfrak{A})^{\frac{1}{p}} (M|g|^{p'})^{\frac{1}{p'}},$$

where we used that the $\mathcal{I}(\mathfrak{p}')$ with $\mathfrak{p}' \in \mathfrak{A}(\mathfrak{p})$ are contained in $B(\mathfrak{p})$. Hence we may estimate (5.5) by

$$\leq 2^{232a^3+11a+1} \text{dens}_1(\mathfrak{A})^{\frac{1}{p}} \sum_{\mathfrak{p} \in \mathfrak{A}} \int_{E(\mathfrak{p})} |g|(y) (M|g|^{p'})^{\frac{1}{p'}} \, d\mu(y).$$

Since the sets $E(\mathfrak{p})$ are pairwise disjoint, the lemma now follows from boundedness of the Hardy-Littlewood maximal function. \square

The following TT^* estimate will be proved in Section 5.2.

Lemma 5.3 (tile correlation). *Let $\mathfrak{p}, \mathfrak{p}' \in \mathfrak{P}$ with $s(\mathfrak{p}') \leq s(\mathfrak{p})$. Then*

$$\begin{aligned} & \left| \int T_{\mathfrak{p}'}^* g \overline{T_{\mathfrak{p}}^* g} \right| \\ & \leq 2^{232a^3} \frac{(1 + d_{\mathfrak{p}'}(\mathcal{Q}(\mathfrak{p}'), \mathcal{Q}(\mathfrak{p})))^{-1/(2a^2+a^3)}}{\mu(\mathcal{I}(\mathfrak{p}))} \int_{E(\mathfrak{p}')} |g| \int_{E(\mathfrak{p})} |g|. \end{aligned} \quad (5.6)$$

Moreover, the term (5.6) vanishes unless

$$\mathcal{I}(\mathfrak{p}') \subset B(c(\mathfrak{p}), 14D^{s(\mathfrak{p})}). \quad (5.7)$$

The following lemma will be proved in Section 5.3.

Lemma 5.4 (antichain tile count). *Set $p := 4a^4$ and let p' be the dual exponent of p , that is $1/p + 1/p' = 1$. For every $\vartheta \in \Theta$ and every subset \mathfrak{A}' of \mathfrak{A} we have*

$$\left\| \sum_{\mathfrak{p} \in \mathfrak{A}'} (1 + d_{\mathfrak{p}}(\mathcal{Q}(\mathfrak{p}), \vartheta))^{-1/(2a^2+a^3)} \mathbf{1}_{E(\mathfrak{p})} \mathbf{1}_G \right\|_p \leq 2^{5a} \text{dens}_1(\mathfrak{A})^{\frac{1}{p}} \mu \left(\bigcup_{\mathfrak{p} \in \mathfrak{A}'} I(\mathfrak{p}) \right)^{\frac{1}{p}}. \quad (5.8)$$

5.2. The tile correlation lemma. We start with a geometric estimate for two tiles.

Lemma 5.5. *Let $\mathfrak{p}_1, \mathfrak{p}_2 \in \mathfrak{P}$ with $B(c(\mathfrak{p}_1), 5D^{s(\mathfrak{p}_1)}) \cap B(c(\mathfrak{p}_2), 5D^{s(\mathfrak{p}_2)}) \neq \emptyset$ and $s(\mathfrak{p}_1) \leq s(\mathfrak{p}_2)$. For each $x_1 \in E(\mathfrak{p}_1)$ and $x_2 \in E(\mathfrak{p}_2)$ we have*

$$1 + d_{\mathfrak{p}_1}(\mathcal{Q}(\mathfrak{p}_1), \mathcal{Q}(\mathfrak{p}_2)) \leq 2^{8a} (1 + d_{B(x_1, D^{s(\mathfrak{p}_1)})}(Q(x_1), Q(x_2))).$$

Proof. Let $i \in \{1, 2\}$. By definition (2.19) of E , we have

$$d_{\mathfrak{p}_i}(Q(x_i), \mathcal{Q}(\mathfrak{p}_i)) < 1. \quad (5.9)$$

By the triangle inequality and (2.10) we have $\mathcal{I}(\mathfrak{p}_1) \subset B(c(\mathfrak{p}_2), 14D^{s(\mathfrak{p}_2)})$. Thus, using again (2.10) and the doubling property (1.3),

$$d_{\mathfrak{p}_1}(Q(x_2), \mathcal{Q}(\mathfrak{p}_2)) \leq 2^{6a} d_{\mathfrak{p}_2}(Q(x_2), \mathcal{Q}(\mathfrak{p}_2)) \leq 2^{6a}. \quad (5.10)$$

By the triangle inequality, we obtain from (5.9) and (5.10)

$$1 + d_{\mathfrak{p}_1}(\mathcal{Q}(\mathfrak{p}_1), \mathcal{Q}(\mathfrak{p}_2)) \leq 2 + 2^{6a} + d_{\mathfrak{p}_1}(Q(x_1), Q(x_2)). \quad (5.11)$$

As $x_1 \in \mathcal{I}(\mathfrak{p}_1)$ we have by (2.10) and the triangle inequality

$$\mathcal{I}(\mathfrak{p}_1) \subset B(x_1, 8D^{s(\mathfrak{p}_1)}).$$

Applying monotonicity (1.4) of the metrics d_B and the doubling property (1.3) in (5.11) completes the proof. \square

Now we prove Lemma 5.3.

Proof of Lemma 5.3. The support of K_s , see (2.5), and the triangle inequality imply that the left-hand side of (5.6) vanishes unless $B(c(\mathfrak{p}_1), 5D^{s(\mathfrak{p}_1)}) \cap B(c(\mathfrak{p}_2), 5D^{s(\mathfrak{p}_2)}) \neq \emptyset$. We assume this for the remainder of the proof. Then (5.7) follows from the triangle inequality and the squeezing property (2.10).

To prove (5.6) we expand the left-hand side and apply Fubini and the triangle inequality to bound it from above by

$$\int_{E(\mathfrak{p}_1)} \int_{E(\mathfrak{p}_2)} \mathbf{I}(x_1, x_2) |g(x_1)| |g(x_2)| d\mu(x_1) d\mu(x_2)$$

with

$$\mathbf{I}(x_1, x_2) := \left| \int e(-Q(x_1)(y) + Q(x_2)(y)) \varphi_{x_1, x_2}(y) d\mu(y) \right|$$

and

$$\varphi_{x_1, x_2}(y) := \overline{K_{s_1}(x_1, y)} K_{s_2}(x_2, y).$$

Note that by (2.5) the function φ is supported in $B(x_1, D^{s_1})$ and by (2.3) and (2.4) we have with $\tau = 1/a$

$$\|\varphi_{x_1, x_2}\|_{C^\tau(B(x_1, D^{s_1}))} \leq \frac{2^{231a^3}}{\mu(B(x_1, D^{s_1}))\mu(B(x_2, D^{s_2}))}.$$

We can therefore estimate $\mathbf{I}(x_1, x_2)$ for fixed $x_1 \in E(\mathbf{p}_1)$ and $x_2 \in E(\mathbf{p}_2)$ with the van-der-Corput type estimate from Proposition 2.5 on the ball $B' := B(x_1, D^{s(\mathbf{p}_1)})$. We obtain

$$\mathbf{I}(x_1, x_2) \leq \frac{2^{231a^3+8a}}{\mu(B(x_2, D^{s(\mathbf{p}_2)}))} (1 + d_{B'}(Q(x_1), Q(x_2)))^{-1/(2a^2+a^3)}.$$

Using Lemma 5.5 and $a \geq 1$, we estimate this by

$$\leq \frac{2^{231a^3+8a+1}}{\mu(B(x_2, D^{s(\mathbf{p}_2)}))} (1 + d_{\mathbf{p}_1}(\mathcal{Q}(\mathbf{p}_1), \mathcal{Q}(\mathbf{p}_2)))^{-1/(2a^2+a^3)}. \quad (5.12)$$

As $x_2 \in E(\mathbf{p}_2)$ we have $\rho(x_2, c(\mathbf{p}_2)) \leq 4D^{s(\mathbf{p}_2)}$, thus by the squeezing property (2.10) and the triangle inequality $\mathcal{I}(\mathbf{p}_2) \subset B(x_2, 8D^{s(\mathbf{p}_2)})$. Applying the doubling property (1.1) in the first factor of (5.12), we obtain (5.6). \square

5.3. The tile count lemma. We start with some auxiliary lemmas.

Lemma 5.6 (tile reach). *Let $\vartheta \in \Theta$ and $N \geq 0$ be an integer. Let $\mathbf{p}, \mathbf{p}' \in \mathfrak{P}$ with*

$$d_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), \vartheta) \leq 2^N \quad \text{and} \quad d_{\mathbf{p}'}(\mathcal{Q}(\mathbf{p}'), \vartheta) \leq 2^N. \quad (5.13)$$

Assume $\mathcal{I}(\mathbf{p}) \subset \mathcal{I}(\mathbf{p}')$ and $s(\mathbf{p}) < s(\mathbf{p}')$. Then

$$2^{N+2}\mathbf{p} \lesssim 2^{N+2}\mathbf{p}'. \quad (5.14)$$

Proof. By the monotonicity of Lemma 2.6, we have

$$d_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}'), \vartheta) \leq d_{\mathbf{p}'}(\mathcal{Q}(\mathbf{p}'), \vartheta) \leq 2^N.$$

Together with (5.13) and the triangle inequality, we obtain

$$d_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}'), \mathcal{Q}(\mathbf{p})) \leq 2^{N+1}. \quad (5.15)$$

To show (5.14), pick some

$$\vartheta' \in B_{\mathbf{p}'}(\mathcal{Q}(\mathbf{p}'), 2^{N+2}).$$

By the doubling property (1.3), applied five times, we have

$$d_{B(c(\mathbf{p}'), 8D^{s(\mathbf{p}')})}(\mathcal{Q}(\mathbf{p}'), \vartheta') < 2^{5a+N+2}. \quad (5.16)$$

With the assumption, $\mathcal{I}(\mathbf{p}) \subset \mathcal{I}(\mathbf{p}')$, the squeezing property (2.10), and the triangle inequality, we have

$$B(c(\mathbf{p}), 4D^{s(\mathbf{p}')}) \subseteq B(c(\mathbf{p}'), 8D^{s(\mathbf{p}')}).$$

Together with (5.16) and monotonicity (1.4) of d this implies

$$d_{B(c(\mathbf{p}), 4D^{s(\mathbf{p}')})}(\mathcal{Q}(\mathbf{p}'), \vartheta') < 2^{5a+N+2}.$$

Using $s(\mathbf{p}) < s(\mathbf{p}')$, $D = 2^{100a^2}$, $a \geq 4$, and the doubling property (1.5) gives

$$d_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}'), \vartheta') < d_{B(c(\mathbf{p}), 2^{2-5a^2-2a}D^{s(\mathbf{p}')})}(\mathcal{Q}(\mathbf{p}'), \vartheta') < 2^N.$$

Finally, we obtain with the triangle inequality and (5.15),

$$d_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), \vartheta') < 2^{N+2}.$$

This implies (5.14) and completes the proof of the lemma. \square

For $\vartheta \in \Theta$ and $N \geq 0$ define

$$\mathfrak{A}_{\vartheta, N} := \{\mathbf{p} \in \mathfrak{A} : 2^N \leq 1 + d_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), \vartheta) < 2^{N+1}\}.$$

Lemma 5.7 (stack density). *Let $\vartheta \in \Theta$, $N \geq 0$ and $L \in \mathcal{D}$. Then*

$$\sum_{\mathbf{p} \in \mathfrak{A}_{\vartheta, N} : \mathcal{I}(\mathbf{p}) = L} \mu(E(\mathbf{p}) \cap G) \leq 2^{a(N+5)} \text{dens}_1(\mathfrak{A}) \mu(L). \quad (5.17)$$

Proof. Let ϑ, N, L be given and set

$$\mathfrak{A}' := \{\mathbf{p} \in \mathfrak{A}_{\vartheta, N} : \mathcal{I}(\mathbf{p}) = L\}.$$

Let $\mathbf{p} \in \mathfrak{A}'$. By definition (2.26) of dens_1 with $\lambda = 2$ and the squeezing property (2.15),

$$\mu(E(\mathbf{p}) \cap G) \leq \mu(E_2(2, \mathbf{p})) \leq 2^a \text{dens}_1(\mathfrak{A}') \mu(L). \quad (5.18)$$

By the covering property (1.6), applied $N+4$ times, there is a collection Θ' of at most $2^{a(N+4)}$ elements such that

$$B_{\mathbf{p}}(\vartheta, 2^{N+1}) \subset \bigcup_{\vartheta' \in \Theta'} B_{\mathbf{p}}(\vartheta', 0.2). \quad (5.19)$$

Note that the metrics $d_{\mathbf{p}'}$, $\mathbf{p}' \in \mathfrak{A}'$ are all equal, since they only depend on $\mathcal{I}(\mathbf{p}') = L$. Hence, each $\mathcal{Q}(\mathbf{p}')$ with $\mathbf{p}' \in \mathfrak{A}'$ is contained in the left-hand-side of (5.19). Moreover, the balls $B_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}'), 0.2) \subset \Omega(\mathbf{p}')$ with $\mathbf{p}' \in \mathfrak{A}'$ are pairwise disjoint by (2.13). Hence each $B_{\mathbf{p}}(\vartheta', 0.2)$ contains at most one of the $\mathcal{Q}(\mathbf{p}')$ with $\mathbf{p}' \in \mathfrak{A}'$. It follows that there are at most $2^{a(N+4)}$ elements in \mathfrak{A}' . Adding (5.18) over \mathfrak{A}' proves (5.17). \square

Lemma 5.8 (local antichain density). *Let $\vartheta \in \Theta$ and N be an integer. Let \mathbf{p}_{ϑ} be a tile with $\vartheta \in B_{\mathbf{p}_{\vartheta}}(\mathcal{Q}(\mathbf{p}_{\vartheta}), 2^{N+1})$. Then we have*

$$\sum_{\mathbf{p} \in \mathfrak{A}_{\vartheta, N} : \mathcal{S}(\mathbf{p}_{\vartheta}) < \mathcal{S}(\mathbf{p})} \mu(E(\mathbf{p}) \cap G \cap \mathcal{I}(\mathbf{p}_{\vartheta})) \leq \mu(E_2(2^{N+3}, \mathbf{p}_{\vartheta})). \quad (5.20)$$

Proof. On the left hand side only tiles $\mathbf{p} \in \mathfrak{A}_{\vartheta, N}$ with $\mathcal{I}(\mathbf{p}_{\vartheta}) \subset \mathcal{I}(\mathbf{p})$ contribute. Combining the assumption on \mathbf{p}_{ϑ} with $\mathbf{p} \in \mathfrak{A}_{\vartheta, N}$ and Lemma 5.6, we conclude $2^{N+3}\mathbf{p}_{\vartheta} \lesssim 2^{N+3}\mathbf{p}$ and hence

$$E(\mathbf{p}) \cap G \cap \mathcal{I}(\mathbf{p}_{\vartheta}) \subset E_2(2^{N+3}, \mathbf{p}_{\vartheta}).$$

Using disjointness of the various $E(\mathbf{p})$ with $\mathbf{p} \in \mathfrak{A}$, we obtain (5.20). \square

Lemma 5.9 (global antichain density). *Let $\vartheta \in Q(X)$ and let $N \geq 0$ be an integer. Then we have*

$$\sum_{\mathbf{p} \in \mathfrak{A}_{\vartheta, N}} \mu(E(\mathbf{p}) \cap G) \leq 2^{101a^3 + Na} \text{dens}_1(\mathfrak{A}) \mu(\bigcup_{\mathbf{p} \in \mathfrak{A}} \mathcal{I}(\mathbf{p})).$$

Proof. Fix ϑ and N . The contribution of tiles \mathbf{p} with $s(\mathbf{p}) = -S$ is taken care of by Lemma 5.7. Let \mathfrak{A}' be the set of remaining tiles $\mathbf{p} \in \mathfrak{A}_{\vartheta,N}$ such that $\mathcal{I}(\mathbf{p}) \cap G \neq \emptyset$ and $s(\mathbf{p}) > -S$.

Let \mathcal{L} be the collection of dyadic cubes $I \in \mathcal{D}$ such that $I \subset \mathcal{I}(\mathbf{p})$ for some $\mathbf{p} \in \mathfrak{A}'$ and $\mathcal{I}(\mathbf{p}) \not\subset I$ for all $\mathbf{p} \in \mathfrak{A}'$. Let \mathcal{L}^* be the set of maximal elements in \mathcal{L} with respect to set inclusion. By the grid properties, the elements in \mathcal{L}^* are pairwise disjoint and we have

$$\bigcup \mathcal{L}^* = \bigcup_{\mathbf{p} \in \mathfrak{A}'} \mathcal{I}(\mathbf{p}). \quad (5.21)$$

Using the partition (5.21), it suffices to show that for each $L \in \mathcal{L}^*$

$$\sum_{\mathbf{p} \in \mathfrak{A}'} \mu(E(\mathbf{p}) \cap G \cap L) \leq 2^{101a^3+aN} \text{dens}_1(\mathfrak{A}) \mu(L). \quad (5.22)$$

Fix $L \in \mathcal{L}^*$. By definition of L , there exists an element $\mathbf{p}' \in \mathfrak{A}'$ of minimal scale such that $L \subset \mathcal{I}(\mathbf{p}')$. Let $L' \in \mathcal{D}$ be the unique cube with $s(L') = s(L) + 1$ and $L \subset L'$.

We split the left-hand side of (5.22) as

$$\sum_{\mathbf{p} \in \mathfrak{A}': \mathcal{I}(\mathbf{p}) = L'} \mu(E(\mathbf{p}) \cap G \cap L) + \sum_{\mathbf{p} \in \mathfrak{A}': \mathcal{I}(\mathbf{p}) \neq L'} \mu(E(\mathbf{p}) \cap G \cap L), \quad (5.23)$$

The first term satisfies the required bound by Lemma 5.7 and the doubling property (1.1).

For the second term, note that by the maximality of L in \mathcal{L} , there exists $\mathbf{p}'' \in \mathfrak{A}'$ with $\mathcal{I}(\mathbf{p}'') \subset L'$. If $\mathcal{I}(\mathbf{p}'') = L'$, then we set $\mathbf{p}_\vartheta = \mathbf{p}''$. Otherwise, we use that by the covering property (2.13), there exists a unique \mathbf{p}_ϑ with $\mathcal{I}(\mathbf{p}_\vartheta) = L'$ such that $\vartheta \in \Omega(\mathbf{p}_\vartheta)$, and we take this as the definition of \mathbf{p}_ϑ . Using that $\mathbf{p}'' \in \mathfrak{A}_{\vartheta,N}$ and Lemma 5.6, we conclude in both cases that

$$2^{N+3}\mathbf{p}'' \lesssim 2^{N+3}\mathbf{p}_\vartheta.$$

As $\mathbf{p}'' \in \mathfrak{A}'$, we have by the definition (2.26) of dens_1 that

$$\mu(E_2(2^{N+3}, \mathbf{p}_\vartheta)) \leq 2^{Na+3a} \text{dens}_1(\mathfrak{A}) \mu(L'). \quad (5.24)$$

Now let \mathbf{p} be any tile in the second sum in (5.23). It follows by the dyadic property (2.8) and the definition of L that $L \subset \mathcal{I}(\mathbf{p})$ and $L \neq \mathcal{I}(\mathbf{p})$ and in fact $L' \subset \mathcal{I}(\mathbf{p})$ and $L' \neq \mathcal{I}(\mathbf{p})$, so we conclude $s(L') < s(\mathbf{p})$. By Lemma 5.8, we thus estimate the second term in (5.23) by

$$\sum_{\mathbf{p} \in \mathfrak{A}': \mathcal{I}(\mathbf{p}) \neq L'} \mu(E(\mathbf{p}) \cap G \cap L') \leq \mu(E_2(2^{N+3}, \mathbf{p}_\vartheta)).$$

With (5.24) and the doubling property (1.1), this proves (5.22). \square

We turn to the proof of Lemma 5.4.

Proof of Lemma 5.4. Using that \mathfrak{A} is the disjoint union of the $\mathfrak{A}_{\vartheta,N}$ with $N \geq 0$ and that the sets $E(\mathbf{p})$ with $\mathbf{p} \in \mathfrak{A}$ are pairwise disjoint we estimate the p -th power of (5.8) by

$$\sum_{N \geq 0} 2^{-pN/(2a^2+a^3)} \sum_{\mathbf{p} \in \mathfrak{A}_{\vartheta,N}} \mu(E(\mathbf{p}) \cap G).$$

Using Lemma 5.9, we estimate the last display by

$$\leq \sum_{N \geq 0} 2^{-pN/(2a^2+a^3)+101a^3+Na} \text{dens}_1(\mathfrak{A})\mu(\cup_{\mathfrak{p} \in \mathfrak{A}} \mathcal{I}(\mathfrak{p})). \quad (5.25)$$

Recalling $p = 4a^4$ and using $a \geq 4$, we conclude

$$pN/(2a^2 + a^3) \geq 4a^4N/(3a^3) \geq Na + N.$$

Hence we have for (5.25) the upper bound

$$\leq 2^{101a^3} \sum_{N \geq 0} 2^{-N} \text{dens}_1(\mathfrak{A})\mu(\cup_{\mathfrak{p} \in \mathfrak{A}} \mathcal{I}(\mathfrak{p})).$$

Summing over $N \geq 0$ and taking the p -th root proves the lemma. \square

6. PROOF OF THE FOREST OPERATOR PROPOSITION

After proving a series of auxiliary lemmas, we assemble the proof of Proposition 2.3 in Subsection 6.7. Fix a forest $(\mathfrak{U}, \mathfrak{T})$.

6.1. The pointwise tree estimate. The main result of this subsection is the pointwise estimate for operators associated to sets $\mathfrak{T}(\mathfrak{u})$ stated in Lemma 6.1. For $\mathfrak{u} \in \mathfrak{U}$ and $x \in X$, we define

$$\sigma(\mathfrak{u}, x) := \{s(\mathfrak{p}) : \mathfrak{p} \in \mathfrak{T}(\mathfrak{u}), x \in E(\mathfrak{p})\},$$

$$\bar{\sigma}(\mathfrak{u}, x) := \max \sigma(\mathfrak{T}(\mathfrak{u}), x) \quad \text{and} \quad \underline{\sigma}(\mathfrak{u}, x) := \min \sigma(\mathfrak{T}(\mathfrak{u}), x).$$

By the convexity property (2.30), we have for each $\mathfrak{u} \in \mathfrak{U}$

$$\sigma(\mathfrak{u}, x) = \mathbb{Z} \cap [\underline{\sigma}(\mathfrak{u}, x), \bar{\sigma}(\mathfrak{u}, x)]. \quad (6.1)$$

For a nonempty collection of tiles $\mathfrak{S} \subset \mathfrak{P}$, we define $\mathcal{J}_0(\mathfrak{S})$ to be the collection of all dyadic cubes $J \in \mathcal{D}$ such that $s(J) = -S$ or

$$\mathcal{I}(\mathfrak{p}) \not\subset B(c(J), 100D^{s(J)+1})$$

for all $\mathfrak{p} \in \mathfrak{S}$. We further define $\mathcal{L}_0(\mathfrak{S})$ to be the collection of dyadic cubes $L \in \mathcal{D}$ such that $s(L) = -S$, or there exists $\mathfrak{p} \in \mathfrak{S}$ with $L \subset \mathcal{I}(\mathfrak{p})$ and there exists no $\mathfrak{p} \in \mathfrak{S}$ with $\mathcal{I}(\mathfrak{p}) \subset L$. Let

$$\mathcal{J}(\mathfrak{S}), \mathcal{L}(\mathfrak{S}) \quad (6.2)$$

be the collection of inclusion maximal cubes in $\mathcal{J}_0(\mathfrak{S})$ and $\mathcal{L}_0(\mathfrak{S})$, respectively. Both collections partition the union of all grid cubes:

$$\bigcup_{I \in \mathcal{D}} I = \bigcup_{J \in \mathcal{J}(\mathfrak{S})} J = \bigcup_{L \in \mathcal{L}(\mathfrak{S})} L.$$

For a finite set of pairwise disjoint cubes \mathcal{C} , define the projection operator

$$P_{\mathcal{C}}f(x) := \sum_{J \in \mathcal{C}} \mathbf{1}_J(x) \frac{1}{\mu(J)} \int_J f(y) d\mu(y).$$

We denote by $I_s(x)$ the unique grid cube of scale s containing x . Define for $\vartheta \in \Theta$ the nontangential maximal operator

$$T_{\mathcal{N}}^{\vartheta}f(x) := \sup_{-S \leq s_1} \sup_{x' \in I_{s_1}(x)} \sup_{\substack{s_1 \leq s_2 \leq S \\ D^{s_2-1} \leq R_Q(\vartheta, x')}} \left| \sum_{s=s_1}^{s_2} \int K_s(x', y) f(y) d\mu(y) \right|.$$

Up to boundary terms that are controlled by the Hardy-Littlewood maximal function, this operator is controlled by the maximal operator T_Q^θ defined in (1.14). This implies the estimate

$$\|T_{\mathcal{N}}^\theta f\|_2 \leq 2^{102a^3} \|f\|_2. \quad (6.3)$$

We define also for each $\mathbf{u} \in \mathfrak{U}$ the auxiliary operator

$$S_{1,\mathbf{u}}f(x) := \sum_{I \in \mathcal{D}} \mathbf{1}_I(x) \sum_{\substack{J \in \mathcal{J}(\mathfrak{T}(\mathbf{u})) \\ J \subset B(c(I), 16D^{s(I)}) \\ s(J) \leq s(I)}} \frac{D^{(s(J)-s(I))/a}}{\mu(B(c(I), 16D^{s(I)}))} \int_J |f(y)| d\mu(y).$$

Lemma 6.1 (pointwise tree estimate). *Let $\mathbf{u} \in \mathfrak{U}$ and $L \in \mathcal{L}(\mathfrak{T}(\mathbf{u}))$. Let $x, x' \in L$. Then for all bounded functions f with bounded support*

$$\begin{aligned} & |T_{\mathfrak{T}(\mathbf{u})}[e(-\mathcal{Q}(\mathbf{u}))f](x)| \\ & \leq 2^{129a^3} (M + S_{1,\mathbf{u}}P_{\mathcal{J}(\mathfrak{T}(\mathbf{u}))}|f|(x') + |T_{\mathcal{N}}^{\mathcal{Q}(\mathbf{u})}P_{\mathcal{J}(\mathfrak{T}(\mathbf{u}))}f(x')|. \end{aligned} \quad (6.4)$$

Proof. The left hand side of (6.4) equals

$$\left| \sum_{s \in \sigma(\mathbf{u}, x)} \int e(\mathbf{u}, x, y) K_s(x, y) f(y) d\mu(y) \right|.$$

with

$$e(\mathbf{u}, x, y) := e(-\mathcal{Q}(\mathbf{u})(y) + \mathcal{Q}(x)(y) + \mathcal{Q}(\mathbf{u})(x) - \mathcal{Q}(x)(x)).$$

Using the triangle inequality, we bound this by the sum of three terms:

$$\leq \left| \sum_{s \in \sigma(\mathbf{u}, x)} \int (e(\mathbf{u}, x, y) - 1) K_s(x, y) f(y) d\mu(y) \right| \quad (6.5)$$

$$+ \left| \sum_{s \in \sigma(\mathbf{u}, x)} \int K_s(x, y) P_{\mathcal{J}(\mathfrak{T}(\mathbf{u}))} f(y) d\mu(y) \right| \quad (6.6)$$

$$+ \left| \sum_{s \in \sigma(\mathbf{u}, x)} \int K_s(x, y) (f(y) - P_{\mathcal{J}(\mathfrak{T}(\mathbf{u}))} f(y)) d\mu(y) \right|. \quad (6.7)$$

Unpacking of the definitions shows that (6.6) is bounded by

$$T_{\mathcal{N}}^{\mathcal{Q}(\mathbf{u})} P_{\mathcal{J}(\mathfrak{T}(\mathbf{u}))} f(x').$$

The proof is completed using the bounds for the other two terms proven in Lemma 6.2 and Lemma 6.3. \square

Lemma 6.2 (first tree pointwise). *For all $\mathbf{u} \in \mathfrak{U}$, all $L \in \mathcal{L}(\mathfrak{T}(\mathbf{u}))$, all $x, x' \in L$ and all bounded f with bounded support, we have*

$$(6.5) \leq 10 \cdot 2^{104a^3} M P_{\mathcal{J}(\mathfrak{T}(\mathbf{u}))} |f|(x').$$

Proof. Let $\mathbf{p} \in \mathfrak{T}(\mathbf{u})$ with $s = s(\mathbf{p}) \in \sigma(\mathbf{u}, x)$. If $x, y \in X$ with $K_s(x, y) \neq 0$, then by the support assumption (2.5) we have $\rho(x, y) \leq 1/2D^s$. Hence, by the oscillation control (1.2),

$$|e(\mathbf{u}, x, y) - 1| \leq d_{B(x, 1/2D^s)}(\mathcal{Q}(\mathbf{u}), \mathcal{Q}(x)).$$

Let \mathbf{p}' be a tile with $s(\mathbf{p}') = \bar{\sigma}(\mathbf{u}, x)$ and $x \in E(\mathbf{p}')$. Using the doubling property (1.3) repeatedly, we bound the previous display by

$$d_{B(x, 4D^s)}(\mathcal{Q}(\mathbf{u}), Q(x)) \leq 2^{4a} d_{\mathbf{p}}(\mathcal{Q}(\mathbf{u}), Q(x)) \leq 2^{4a} 2^{s-\bar{\sigma}(\mathbf{u}, x)} d_{\mathbf{p}'}(\mathcal{Q}(\mathbf{u}), Q(x)).$$

Since $\mathcal{Q}(\mathbf{u}) \in B_{\mathbf{p}'}(\mathcal{Q}(\mathbf{p}'), 4)$ by the tree property (2.29) and $Q(x) \in \Omega(\mathbf{p}') \subset B_{\mathbf{p}'}(\mathcal{Q}(\mathbf{p}'), 1)$ by the squeezing property (2.15), the last display is

$$\leq 5 \cdot 2^{4a} 2^{s-\bar{\sigma}(\mathbf{u}, x)}.$$

Using the pointwise kernel bound (2.3), it follows that

$$\begin{aligned} (6.5) &\leq 5 \cdot 2^{103a^3} \sum_{s \in \sigma(x)} 2^{s-\bar{\sigma}(\mathbf{u}, x)} \frac{1}{\mu(B(x, D^s))} \int_{B(x, 0.5D^s)} |f(y)| d\mu(y). \\ &\leq 5 \cdot 2^{103a^3} \sum_{s \in \sigma(x)} 2^{s-\bar{\sigma}(\mathbf{u}, x)} \frac{1}{\mu(B(x, D^s))} \sum_{\substack{J \in \mathcal{J}(\mathfrak{T}(\mathbf{u})) \\ J \cap B(x, 0.5D^s) \neq \emptyset}} \int_J |f(y)| d\mu(y). \end{aligned}$$

This expression does not change if we replace $|f|$ by $P_{\mathcal{J}(\mathfrak{T}(\mathbf{u}))}|f|$. Further, if $J \in \mathcal{J}(\mathfrak{T}(\mathbf{u}))$ with $B(x, 0.5D^s) \cap J \neq \emptyset$ then by the triangle inequality and the definition of \mathcal{J} we obtain $J \subset B(c(\mathbf{p}_s), 16D^s)$. Hence the last display is

$$\leq 5 \cdot 2^{103a^3} \sum_{s \in \sigma(x)} 2^{s-\bar{\sigma}(\mathbf{u}, x)} \frac{1}{\mu(B(x, D^s))} \int_{B(c(\mathbf{p}_s), 16D^s)} P_{\mathcal{J}(\mathfrak{T}(\mathbf{u}))}|f(y)| d\mu(y).$$

Combined with the doubling property (1.1), this completes the estimate for the term (6.5) and thus the lemma. \square

Lemma 6.3 (third tree pointwise). *For all $\mathbf{u} \in \mathfrak{U}$, all $L \in \mathcal{L}(\mathfrak{T}(\mathbf{u}))$, all $x, x' \in L$ and all bounded f with bounded support, we have*

$$(6.7) \leq 2^{128a^3} S_{1, \mathbf{u}} P_{\mathcal{J}(\mathfrak{T}(\mathbf{u}))}|f|(x').$$

Proof. We have for $J \in \mathcal{J}(\mathfrak{T}(\mathbf{u}))$:

$$\begin{aligned} &\int_J K_s(x, y)(1 - P_{\mathcal{J}(\mathfrak{T}(\mathbf{u}))})f(y) d\mu(y) \\ &= \int_J \frac{1}{\mu(J)} \int_J K_s(x, y) - K_s(x, z) d\mu(z) f(y) d\mu(y). \end{aligned}$$

By the kernel regularity (2.4) and the squeezing property (2.10), we have for $y, z \in J$

$$|K_s(x, y) - K_s(x, z)| \leq \frac{2^{127a^3}}{\mu(B(x, D^s))} \left(\frac{8D^{s(J)}}{D^s} \right)^{1/a}.$$

Suppose that $s \in \sigma(\mathbf{u}, x)$. As in the proof of Lemma 6.2, if $K_s(x, y) \neq 0$ for some $y \in J \in \mathcal{J}(\mathfrak{T}(\mathbf{u}))$ then $J \subset B(x, 16D^s)$ and $s(J) \leq s(\mathbf{p})$. Thus, we can estimate (6.7) by

$$2^{127a^3+3/a} \sum_{\mathbf{p} \in \mathfrak{T}} \frac{\mathbf{1}_{E(\mathbf{p})}(x)}{\mu(B(x, D^{s(\mathbf{p})}))} \sum_{\substack{J \in \mathcal{J}(\mathfrak{T}(\mathbf{u})) \\ J \subset B(x, 16D^{s(\mathbf{p})}) \\ s(J) \leq s(\mathbf{p})}} D^{(s(J)-s(\mathbf{p}))/a} \int_J |f|.$$

By (2.13) and definition (2.19), the sets $E(\mathbf{p})$ for tiles \mathbf{p} with $\mathcal{I}(\mathbf{p}) = I$ are pairwise disjoint. It follows from the definition of $\mathcal{L}(\mathfrak{T}(\mathbf{u}))$ that $x \in \mathcal{I}(\mathbf{p})$ if

and only if $x' \in \mathcal{I}(\mathfrak{p})$, thus we can estimate the sum over such $\mathbf{1}_{E(\mathfrak{p})}(x)$ by $\mathbf{1}_I(x')$. Using also the doubling property (1.1), we estimate the last display by

$$\begin{aligned} &\leq 2^{128a^3} \sum_{I \in \mathcal{D}} \frac{\mathbf{1}_I(x')}{\mu(B(c(I), 16D^{s(I)}))} \sum_{\substack{J \in \mathcal{J}(\mathfrak{T}(\mathfrak{u})) \\ J \subset B(x, 16D^{s(I)}) \\ s(J) \leq s(I)}} D^{(s(J)-s(I))/a} \int_J |f| \\ &= 2^{128a^3} S_{1,\mathfrak{u}} P_{\mathcal{J}(\mathfrak{T}(\mathfrak{u}))} |f|(x'). \end{aligned}$$

This completes the proof of the lemma. \square

6.2. An auxiliary L^2 tree estimate. The main result of this subsection is the following estimate on L^2 for operators associated to trees.

Lemma 6.4 (tree projection estimate). *Let $\mathfrak{u} \in \mathfrak{U}$. Then we have for all f, g bounded with bounded support*

$$\left| \int_X \bar{g} T_{\mathfrak{T}(\mathfrak{u})} f \, d\mu \right| \leq 2^{130a^3} \|P_{\mathcal{J}(\mathfrak{T}(\mathfrak{u}))} |f|\|_2 \|P_{\mathcal{L}(\mathfrak{T}(\mathfrak{u}))} |g|\|_2.$$

Proof. Let $L \in \mathcal{L}(\mathfrak{T}(\mathfrak{u}))$. Let $b(x')$ denote the right-hand side of (6.4) in Lemma 6.1. Applying this lemma to $e(\mathcal{Q}(\mathfrak{u}))f$, we obtain for all $y, x' \in L$

$$|T_{\mathfrak{T}(\mathfrak{u})} f(y)| \leq b(x').$$

Hence, choosing a fixed x' in each L ,

$$\left| \int \bar{g}(y) T_{\mathfrak{T}(\mathfrak{u})} f(y) \, d\mu(y) \right| \leq \int_X [P_{\mathcal{L}(\mathfrak{T}(\mathfrak{u}))} |g|(y)] b(y) \, d\mu(y).$$

With Cauchy-Schwarz, this is bounded by $\|P_{\mathcal{L}(\mathfrak{T}(\mathfrak{u}))} |g|\|_2 \|b\|_2$. The bounds for b following from Lemma 6.5 below and (6.3) then complete the proof. \square

Denote $B(I) := B(c(I), 16D^{s(I)})$.

Lemma 6.5 (boundary operator bound). *For all $\mathfrak{u} \in \mathfrak{U}$ and all bounded functions f with bounded support*

$$\|S_{1,\mathfrak{u}} f\|_2 \leq 2^{12a} \|f\|_2.$$

Proof. Let g be a function with $\|g\|_2 = 1$. Then

$$\begin{aligned} &\left| \int \bar{g}(y) S_{1,\mathfrak{u}} f(y) \, d\mu(y) \right| \\ &\leq \sum_{I \in \mathcal{D}} \frac{1}{\mu(B(I))} \int_{B(I)} |g(y)| \, d\mu(y) \times \sum_{\substack{J \in \mathcal{J}(\mathfrak{T}(\mathfrak{u})) \\ J \subset B(I) \\ s(J) \leq s(I)}} D^{(s(J)-s(I))/a} \int_J |f(y)| \, d\mu(y). \end{aligned}$$

Changing the order of summation and using $J \subset B(I)$ to bound the first average integral by $M|g|(y)$ for any $y \in J$, we obtain

$$\leq \sum_{J \in \mathcal{J}(\mathfrak{T}(\mathfrak{u}))} \int_J |f(y)| M |g|(y) \, d\mu(y) \sum_{\substack{I \in \mathcal{D}: J \subset B(I) \\ s(I) \geq s(J)}} D^{(s(J)-s(I))/a}.$$

Using Lemma 6.6 below and summing a geometric series, the last display is bounded by

$$2^{9a+1} \int_X |f(y)| |M|g|(y) \, d\mu(y).$$

This completes the proof using boundedness of M and duality. \square

We used the following simple consequence of the doubling property (1.1), which we do not explicitly prove.

Lemma 6.6 (boundary overlap). *For every cube $I \in \mathcal{D}$, there exist at most 2^{9a} cubes $J \in \mathcal{D}$ with $s(J) = s(I)$ and $B(I) \cap B(J) \neq \emptyset$.*

6.3. The quantitative L^2 tree estimate. This section proves the following bound for tree operators with control by the densities.

Lemma 6.7 (densities tree bound). *Let $\mathbf{u} \in \mathfrak{U}$. Then for all bounded f with bounded support and g with $|g| \leq \mathbf{1}_G$ we have*

$$\left| \int_X \bar{g} T_{\mathfrak{T}(\mathbf{u})} f \, d\mu \right| \leq 2^{181a^3} \text{dens}_1(\mathfrak{T}(\mathbf{u}))^{1/2} \|f\|_2 \|g\|_2.$$

If additionally $|f| \leq \mathbf{1}_F$, then we have

$$\left| \int_X \bar{g} T_{\mathfrak{T}(\mathbf{u})} f \, d\mu \right| \leq 2^{282a^3} \text{dens}_1(\mathfrak{T}(\mathbf{u}))^{1/2} \text{dens}_2(\mathfrak{T}(\mathbf{u}))^{1/2} \|f\|_2 \|g\|_2.$$

Recall that $T_{\mathbf{p}}f$ is supported in $E(\mathbf{p})$. Lemma 6.7 follows immediately from the estimate of Lemma 6.4, Cauchy-Schwarz and Lemmas 6.8 and 6.8 below, controlling the size of the support of the tree operator and its adjoint.

Lemma 6.8 (local dens1 tree bound). *Let $\mathbf{u} \in \mathfrak{U}$ and $L \in \mathcal{L}(\mathfrak{T}(\mathbf{u}))$. Then*

$$\mu(L \cap G \cap \bigcup_{\mathbf{p} \in \mathfrak{T}(\mathbf{u})} E(\mathbf{p})) \leq 2^{101a^3} \text{dens}_1(\mathfrak{T}(\mathbf{u})) \mu(L). \quad (6.8)$$

Proof. We assume there exists $\mathbf{p} \in \mathfrak{T}(\mathbf{u})$ with $L \cap \mathcal{I}(\mathbf{p}) \neq \emptyset$, for otherwise (6.8) is void. Suppose first that there exists such \mathbf{p} with $s(\mathbf{p}) \leq s(L)$. Then $s(\mathbf{p}) = -S$ and $L = \mathcal{I}(\mathbf{p})$ by the definition of \mathcal{L} . Let $\mathbf{q} \in \mathfrak{T}(\mathbf{u})$ be another tile with $E(\mathbf{q}) \cap L \neq \emptyset$. By the grid property we must have $\mathcal{I}(\mathbf{p}) \subset \mathcal{I}(\mathbf{q})$. Using monotonicity of d and the tree property (2.29), we conclude

$$\begin{aligned} d_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), \mathcal{Q}(\mathbf{q})) &\leq d_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), \mathcal{Q}(\mathbf{u})) + d_{\mathbf{p}}(\mathcal{Q}(\mathbf{q}), \mathcal{Q}(\mathbf{u})) \\ &\leq d_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), \mathcal{Q}(\mathbf{u})) + d_{\mathbf{q}}(\mathcal{Q}(\mathbf{q}), \mathcal{Q}(\mathbf{u})) \leq 8. \end{aligned}$$

By definition of $E(\mathbf{q})$ and the triangle inequality, $L \cap G \cap E(\mathbf{q}) \subset E_2(9, \mathbf{p})$. We obtain

$$\mu(L \cap G \cap \bigcup_{\mathbf{q} \in \mathfrak{T}(\mathbf{u})} E(\mathbf{q})) \leq \mu(E_2(9, \mathbf{p})).$$

By the definition of dens_1 , this is bounded by

$$9^a \text{dens}_1(\mathfrak{T}(\mathbf{u})) \mu(\mathcal{I}(\mathbf{p})) = 9^a \text{dens}_1(\mathfrak{T}(\mathbf{u})) \mu(L).$$

Since $a \geq 4$, (6.8) follows in the given case.

Now assume the opposite case that for each $\mathbf{p} \in \mathfrak{T}(\mathbf{u})$ with $L \cap E(\mathbf{p}) \neq \emptyset$, we have $s(\mathbf{p}) > s(L)$. Let L' be the parent cube of L and let $\mathbf{p}'' \in \mathfrak{T}(\mathbf{u})$ with $\mathcal{I}(\mathbf{p}'') \subset L'$. It suffices to show that there exists a tile $\mathbf{p}' \in \mathfrak{P}(\mathfrak{T}(\mathbf{u}))$ with $\mathcal{I}(\mathbf{p}') = L'$, $d_{\mathbf{p}'}(\mathcal{Q}(\mathbf{p}'), \mathcal{Q}(\mathbf{u})) < 4$ and $9\mathbf{p}'' \lesssim 9\mathbf{p}'$. For then, let $\mathbf{q} \in \mathfrak{T}(\mathbf{u})$ with

$L \cap E(\mathbf{q}) \neq \emptyset$. Then $s(\mathbf{q}) \geq s(L')$ and $L' \subset \mathcal{I}(\mathbf{q})$. By the same computation as in the first case we deduce $L \cap G \cap E(\mathbf{q}) \subset E_2(9, \mathbf{p}')$ and

$$\mu(L \cap G \cap \bigcup_{\mathbf{q} \in \mathfrak{T}(\mathbf{u})} E(\mathbf{q})) \leq \mu(E_2(9, \mathbf{p}')) \leq 9^a \text{dens}_1(\mathfrak{T}(\mathbf{u})) \mu(L').$$

This proves (6.8) using the doubling property (1.1).

It remains to show existence of \mathbf{p}' with the required properties. If $\mathcal{I}(\mathbf{p}'') = L'$ we can take $\mathbf{p}' = \mathbf{p}''$. Otherwise, let \mathbf{p}' be the unique tile such that $\mathcal{I}(\mathbf{p}') = L'$ and such that $\Omega(\mathbf{u}) \cap \Omega(\mathbf{p}') \neq \emptyset$. Since $\mathcal{I}(\mathbf{p}') \subset \mathcal{I}(\mathbf{p})$ and $\mathbf{p} \in \mathfrak{T}(\mathbf{u})$, we have $\mathbf{p}' \in \mathfrak{P}(\mathfrak{T}(\mathbf{u}))$. By the tree property (2.29), we have $s(\mathbf{p}') = s(L') \leq s(\mathbf{p}) < s(\mathbf{u})$. By (2.8) and (2.14), we conclude $\Omega(\mathbf{u}) \subset \Omega(\mathbf{p}')$, and hence the distance property required of \mathbf{p}' . The property $9\mathbf{p}'' \lesssim 9\mathbf{p}'$ follows by the triangle inequality, (2.29), Lemma 2.6 and (2.15). This completes the proof. \square

Lemma 6.9 (local dens2 tree bound). *Let $\mathbf{u} \in \mathfrak{U}$ and $J \in \mathcal{J}(\mathfrak{T}(\mathbf{u}))$. Then*

$$\mu(F \cap J) \leq 2^{201a^3} \text{dens}_2(\mathfrak{T}(\mathbf{u})) \mu(J). \quad (6.9)$$

Proof. Suppose first that $s(J) = S \geq 1$. Then J is the maximal cube I_0 as in (2.9) and the definition of $\mathcal{J}(\mathfrak{T}(\mathbf{u}))$ quickly shows $s(J) = -S$, a contradiction. It remains to consider the case $s(J) < S$.

We show the existence of a tile $\mathbf{p} \in \mathfrak{T}(\mathbf{u})$ and an $r \geq 4D^{s(\mathbf{p})}$ such that

$$J \subset B(c(\mathbf{p}), r), \quad \mu(B(c(\mathbf{p}), r)) \leq 2^{200a^3+14a} \mu(J). \quad (6.10)$$

This will imply with the definition (2.27) of dens_2 the desired (6.9) as follows

$$\begin{aligned} \mu(F \cap J) &\leq \mu(F \cap B(c(\mathbf{p}), r)) \\ &\leq \text{dens}_2(\mathfrak{T}(\mathbf{u})) \mu(B(c(\mathbf{p}), r)) \leq 2^{200a^3+14a} \text{dens}_2(\mathfrak{T}(\mathbf{u})) \mu(J). \end{aligned}$$

The grid properties give a $J' \in \mathcal{D}$ with $s(J') = s(J) + 1$ and $J \subset J'$ and $B(c(J'), 204D^{s(J')+1}) \subset B(c(J), 204D^{s(J')+1} + 4D^{s(J')}) \subset B(c(J), 2^8 D^{s(J)+2})$.

The doubling property (1.1), squeezing property (2.10), and $D = 2^{100a^2}$ give

$$\mu(B(c(J'), 204D^{s(J')+1})) \leq 2^{200a^3+10a} \mu(J). \quad (6.11)$$

By definition of $\mathcal{J}(\mathfrak{T}(\mathbf{u}))$, there exists $\mathbf{p} \in \mathfrak{T}(\mathbf{u})$ with

$$\mathcal{I}(\mathbf{p}) \subset B(c(J'), 100D^{s(J')+1}).$$

If $J \subset B(c(\mathbf{p}), 4D^{s(\mathbf{p})})$, then (6.11) gives (6.10) with $r = 4D^{s(\mathbf{p})}$. So assume $J \not\subset B(c(\mathbf{p}), 4D^{s(\mathbf{p})})$. By the triangle inequality,

$$J \subset J' \subset B(c(J'), 4D^{s(J')}) \subset B(c(\mathbf{p}), 104D^{s(J')+1}),$$

so we must have $104D^{s(J')+1} > 4D^{s(\mathbf{p})}$. By the triangle inequality again,

$$B(c(\mathbf{p}), 104D^{s(J')+1}) \subset B(c(J), 204D^{s(J')+1}),$$

so (6.11) proves that \mathbf{p} satisfies (6.10) with $r = 104D^{s(J')+1}$. \square

6.4. Almost orthogonality of separated trees. The main result of this subsection is the almost orthogonality estimate for operators associated to distinct trees in a forest in Lemma 6.10 below. We will deduce it from Lemmas 6.11 and 6.12, which are proven in Subsections 6.5 and 6.6, respectively.

The adjoint of the operator $T_{\mathfrak{p}}$ defined in (2.20) is given by

$$T_{\mathfrak{p}}^* g(x) = \int_{E(\mathfrak{p})} \overline{K_{s(\mathfrak{p})}(y, x)} e(-Q(y)(x) + Q(y)(y)) g(y) d\mu(y). \quad (6.12)$$

For each $\mathfrak{p} \in \mathfrak{P}$, we have

$$T_{\mathfrak{p}}^* g = \mathbf{1}_{B(c(\mathfrak{p}), 5D^s(\mathfrak{p}))} T_{\mathfrak{p}}^* \mathbf{1}_{\mathcal{I}(\mathfrak{p})} g. \quad (6.13)$$

With the tree localization (2.34), we conclude for $\mathfrak{u} \in \mathfrak{U}$ and $\mathfrak{p} \in \mathfrak{T}(\mathfrak{u})$

$$T_{\mathfrak{p}}^* g = \mathbf{1}_{\mathcal{I}(\mathfrak{u})} T_{\mathfrak{p}}^* \mathbf{1}_{\mathcal{I}(\mathfrak{u})} g. \quad (6.14)$$

Lemma 6.7 implies that for all $\mathfrak{u} \in \mathfrak{U}$ and g with $|g| \leq \mathbf{1}_G$ we have

$$\|S_{2,\mathfrak{u}} g\|_2 \leq 2^{182a^3} \|g\|_2, \quad (6.15)$$

$$S_{2,\mathfrak{u}} g := \left| T_{\mathfrak{T}(\mathfrak{u})}^* g \right| + M|g| + |g|.$$

Lemma 6.10 (correlation separated trees). *For any $\mathfrak{u}_1 \neq \mathfrak{u}_2 \in \mathfrak{U}$ and all bounded g_1, g_2 with bounded support, we have*

$$\left| \int_X T_{\mathfrak{T}(\mathfrak{u}_1)}^* g_1 \overline{T_{\mathfrak{T}(\mathfrak{u}_2)}^* g_2} d\mu \right| \leq 2^{512a^3+1-4n} \prod_{j=1}^2 \|S_{2,\mathfrak{u}_j} g_j\|_{L^2(\mathcal{I}(\mathfrak{u}_1) \cap \mathcal{I}(\mathfrak{u}_2))}. \quad (6.16)$$

By (6.14) and the dyadic property (2.8), the left hand side of (6.16) vanishes unless $\mathcal{I}(\mathfrak{u}_1) \subset \mathcal{I}(\mathfrak{u}_2)$ or $\mathcal{I}(\mathfrak{u}_2) \subset \mathcal{I}(\mathfrak{u}_1)$. Without loss of generality we assume $\mathcal{I}(\mathfrak{u}_1) \subset \mathcal{I}(\mathfrak{u}_2)$. Defining

$$\mathfrak{S} := \{\mathfrak{p} \in \mathfrak{T}(\mathfrak{u}_1) \cup \mathfrak{T}(\mathfrak{u}_2) : d_{\mathfrak{p}}(\mathcal{Q}(\mathfrak{u}_1), \mathcal{Q}(\mathfrak{u}_2)) \geq 2^{Zn/2}\},$$

Lemma 6.10 then follows by combining the definition (2.2) of Z with the following two lemmas.

Lemma 6.11 (correlation distant tree parts). *We have for all $\mathfrak{u}_1 \neq \mathfrak{u}_2 \in \mathfrak{U}$ with $\mathcal{I}(\mathfrak{u}_1) \subset \mathcal{I}(\mathfrak{u}_2)$ and all bounded g_1, g_2 with bounded support*

$$\left| \int_X T_{\mathfrak{T}(\mathfrak{u}_1)}^* g_1 \overline{T_{\mathfrak{T}(\mathfrak{u}_2) \cap \mathfrak{S}}^* g_2} d\mu \right| \leq 2^{511a^3} 2^{-Zn/(4a^2+2a^3)} \prod_{j=1}^2 \|S_{2,\mathfrak{u}_j} g_j\|_{L^2(\mathcal{I}(\mathfrak{u}_1))}. \quad (6.17)$$

Lemma 6.12 (correlation near tree parts). *We have for all $\mathfrak{u}_1 \neq \mathfrak{u}_2 \in \mathfrak{U}$ with $\mathcal{I}(\mathfrak{u}_1) \subset \mathcal{I}(\mathfrak{u}_2)$ and all bounded g_1, g_2 with bounded support*

$$\left| \int_X T_{\mathfrak{T}(\mathfrak{u}_1)}^* g_1 \overline{T_{\mathfrak{T}(\mathfrak{u}_2) \setminus \mathfrak{S}}^* g_2} d\mu \right| \leq 2^{232a^3+21a+5} 2^{-\frac{25}{101a} Zn\kappa} \prod_{j=1}^2 \|S_{2,\mathfrak{u}_j} g_j\|_{L^2(\mathcal{I}(\mathfrak{u}_1))}. \quad (6.18)$$

In the proofs of both lemmas, we will need the following observation.

Lemma 6.13 (overlap implies distance). *Let $\mathfrak{u}_1 \neq \mathfrak{u}_2 \in \mathfrak{U}$ with $\mathcal{I}(\mathfrak{u}_1) \subset \mathcal{I}(\mathfrak{u}_2)$. If $\mathfrak{p} \in \mathfrak{T}(\mathfrak{u}_1) \cup \mathfrak{T}(\mathfrak{u}_2)$ with $\mathcal{I}(\mathfrak{p}) \cap \mathcal{I}(\mathfrak{u}_1) \neq \emptyset$, then $\mathfrak{p} \in \mathfrak{S}$. In particular, we have $\mathfrak{T}(\mathfrak{u}_1) \subset \mathfrak{S}$.*

Proof. Suppose first that $\mathbf{p} \in \mathfrak{T}(\mathbf{u}_1)$. Then $\mathcal{I}(\mathbf{p}) \subset \mathcal{I}(\mathbf{u}_1) \subset \mathcal{I}(\mathbf{u}_2)$, by (2.29). We conclude $\mathbf{p} \in \mathfrak{S}$ as follows, where we use the separation condition (2.33), the squeezing property (2.15), (2.29), and $Z = 2^{12a} \geq 4$:

$$\begin{aligned} d_{\mathbf{p}}(\mathcal{Q}(\mathbf{u}_1), \mathcal{Q}(\mathbf{u}_2)) &\geq d_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), \mathcal{Q}(\mathbf{u}_2)) - d_{\mathbf{p}}(\mathcal{Q}(\mathbf{p}), \mathcal{Q}(\mathbf{u}_1)) \\ &\geq 2^{Z(n+1)} - 4 \geq 2^{Zn/2}. \end{aligned}$$

Suppose now that $\mathbf{p} \in \mathfrak{T}(\mathbf{u}_2)$. If $\mathcal{I}(\mathbf{p}) \subset \mathcal{I}(\mathbf{u}_1)$, then the same argument as above with \mathbf{u}_1 and \mathbf{u}_2 swapped shows $\mathbf{p} \in \mathfrak{S}$. If $\mathcal{I}(\mathbf{p}) \not\subset \mathcal{I}(\mathbf{u}_1)$ then, by the dyadic property (2.8), $\mathcal{I}(\mathbf{u}_1) \subset \mathcal{I}(\mathbf{p})$. Pick $\mathbf{p}' \in \mathfrak{T}(\mathbf{u}_1)$, then we have $\mathcal{I}(\mathbf{p}') \subset \mathcal{I}(\mathbf{u}_1) \subset \mathcal{I}(\mathbf{p})$. We conclude $\mathbf{p} \in \mathfrak{S}$ as follows, using the monotonicity Lemma 2.6 and a similar computation as the previous display,

$$d_{\mathbf{p}}(\mathcal{Q}(\mathbf{u}_1), \mathcal{Q}(\mathbf{u}_2)) \geq d_{\mathbf{p}'}(\mathcal{Q}(\mathbf{u}_1), \mathcal{Q}(\mathbf{u}_2)) \geq 2^{Zn/2}. \quad \square$$

6.5. Tiles with large separation. Lemma 6.11 follows from an application of the van der Corput Proposition 2.5 that we will elaborate in Section 6.5.3. To prepare this application, we construct in Section 6.5.1 a suitable partition of unity, and show in Section 6.5.2 the Hölder estimates needed to apply Proposition 2.5.

6.5.1. *A partition of unity.* Define

$$\mathcal{J}' = \{J \in \mathcal{J}(\mathfrak{S}) : J \subset \mathcal{I}(\mathbf{u}_1)\}.$$

This is a partition of $\mathcal{I}(\mathbf{u}_1)$. The definition of \mathcal{J} implies that if two balls

$$B(J) := B(c(J), 8D^{s(J)})$$

and $B(J')$ intersect, then $|s(J) - s(J')| \leq 1$. Therefore, by standard arguments, we obtain the partition of unity of the following lemma.

Lemma 6.14 (Lipschitz partition unity). *There exists a family of functions χ_J , $J \in \mathcal{J}'$ such that*

$$\mathbf{1}_{\mathcal{I}(\mathbf{u}_1)} = \sum_{J \in \mathcal{J}'} \chi_J,$$

and for all $J \in \mathcal{J}'$ and all $y, y' \in \mathcal{I}(\mathbf{u}_1)$

$$0 \leq \chi_J(y) \leq \mathbf{1}_{B(J)}(y),$$

$$|\chi_J(y) - \chi_J(y')| \leq 2^{227a^3} \frac{\rho(y, y')}{D^{s(J)}}.$$

6.5.2. *Hölder estimates for adjoint tree operators.* Let $g_1, g_2 : X \rightarrow \mathbb{C}$ be bounded with bounded support. Define for $J \in \mathcal{J}'$

$$h_J(y) := \chi_J(y) \cdot (e(\mathcal{Q}(\mathbf{u}_1)(y))T_{\mathfrak{T}(\mathbf{u}_1)}^* g_1(y)) \cdot \overline{(e(\mathcal{Q}(\mathbf{u}_2)(y))T_{\mathfrak{T}(\mathbf{u}_2) \cap \mathfrak{S}}^* g_2(y))}.$$

The following main τ -Hölder estimate for h_J holds with $\tau = 1/a$. We use the notation

$$B(J) := B(c(J), 8D^{s(J)}) \quad \text{and} \quad B^\circ(J) := B(c(J), \frac{1}{8}D^{s(J)}).$$

Lemma 6.15 (Holder correlation tree). *We have for all $J \in \mathcal{J}'$ that*

$$\|h_J\|_{C^\tau(2B(J))} \leq 2^{485a^3} \prod_{j=1,2} \left(\inf_{B^\circ(J)} |T_{\mathfrak{T}(\mathbf{u}_j)}^* g_j| + \inf_J M|g_j| \right).$$

Proof. This lemma follows by combining the upper bounds and Hölder bounds from Lemma 6.14 for the first factor of h_J and from Lemma 6.19 and Lemma 6.20 below for the last two factors. \square

Lemma 6.16 (Holder correlation tile). *Let $\mathbf{u} \in \mathfrak{U}$ and $\mathbf{p} \in \mathfrak{T}(\mathbf{u})$. Then for all $y, y' \in X$ and all bounded g with bounded support, we have*

$$\begin{aligned} & |e(\mathcal{Q}(\mathbf{u})(y))T_{\mathbf{p}}^*g(y) - e(\mathcal{Q}(\mathbf{u})(y'))T_{\mathbf{p}}^*g(y')| \\ & \leq \frac{2^{128a^3}}{\mu(B(c(\mathbf{p}), 4D^{s(\mathbf{p})}))} \left(\frac{\rho(y, y')}{D^{s(\mathbf{p})}} \right)^{1/a} \int_{E(\mathbf{p})} |g(x)| d\mu(x). \end{aligned} \quad (6.19)$$

Proof. We will assume $y, y' \in B(c(\mathbf{p}), 5D^{s(\mathbf{p})})$. Else at least one summand on the left of (6.19) vanishes, and the proof is easy. Then we have $\rho(y, y') \leq 10D^{s(\mathbf{p})}$. We estimate the left hand side of (6.19) by

$$\begin{aligned} & \int_{E(\mathbf{p})} |g(x) \overline{K_{s(\mathbf{p})}(x, y)}| |e(\mathcal{Q}(\mathbf{u})(y) - \mathcal{Q}(\mathbf{u})(y') - Q(x)(y) + Q(x)(y')) - 1| d\mu(x) \\ & + \int_{E(\mathbf{p})} |g(x)| |\overline{K_{s(\mathbf{p})}(x, y)} - \overline{K_{s(\mathbf{p})}(x, y')}| d\mu(x). \end{aligned} \quad (6.20)$$

Let $k \in \mathbb{Z}$ be such that $2^{ak}\rho(y, y') \leq 10D^{s(\mathbf{p})}$ but $2^{a(k+1)}\rho(y, y') > 10D^{s(\mathbf{p})}$. In particular, $k \geq 0$. By the oscillation estimate (1.2), followed by the doubling properties (1.5) and (1.3), we have

$$\begin{aligned} & |\mathcal{Q}(\mathbf{u})(y) - \mathcal{Q}(\mathbf{u})(y') - Q(x)(y) + Q(x)(y')| \leq d_{B(y, 1.6\rho(y, y'))}(Q(x), \mathcal{Q}(\mathbf{u})) \\ & \leq 2^{6a-k}d_{\mathbf{p}}(Q(x), \mathcal{Q}(\mathbf{u})) \leq 5 \cdot 2^{6a-k} \leq 10 \cdot 2^{6a} \left(\frac{\rho(y, y')}{10D^{s(\mathbf{p})}} \right)^{1/a}. \end{aligned}$$

Together with the kernel bound (2.3) and the doubling property (1.1) this gives the needed estimate for the first summand in (6.20). The second summand is estimated similarly using the kernel regularity (2.4) and (1.1). \square

Lemma 6.17 (limited scale impact). *Let $\mathbf{p} \in \mathfrak{T}(\mathbf{u}_2) \setminus \mathfrak{S}$, $J \in \mathcal{J}'$ and suppose that*

$$B(\mathcal{I}(\mathbf{p})) \cap B^\circ(J) \neq \emptyset. \quad (6.21)$$

Then

$$s(J) \leq s(\mathbf{p}) \leq s(J) + 3. \quad (6.22)$$

Proof. For the first inequality in (6.22), assume to get a contradiction that $s(\mathbf{p}) < s(J)$. Since $\mathbf{p} \notin \mathfrak{S}$, we have by Lemma 6.13 that $\mathcal{I}(\mathbf{p}) \cap \mathcal{I}(\mathbf{u}_1) = \emptyset$. Since $B(c(J), \frac{1}{4}D^{s(J)}) \subset \mathcal{I}(J) \subset \mathcal{I}(\mathbf{u}_1)$, this implies

$$\rho(c(J), c(\mathbf{p})) \geq \frac{1}{4}D^{s(J)}.$$

On the other hand by our assumption

$$\rho(c(J), c(\mathbf{p})) \leq \frac{1}{8}D^{s(J)} + 8D^{s(\mathbf{p})}.$$

Thus $D^{s(\mathbf{p})} \geq 64^{-1}D^{s(J)}$, contradicting the definition (2.1) of D and $a \geq 4$.

For the second inequality in (6.22), assume to get a contradiction that $s(\mathbf{p}) > s(J) + 3$. Let $J' \in \mathcal{D}$ with $J \subset J'$ and $s(J') = s(J) + 1$, and

$\mathfrak{p}' \in \mathfrak{S}$ such that $\mathcal{I}(\mathfrak{p}') \subset B(c(J'), 100D^{s(J)+2})$. By (6.21) and the triangle inequality,

$$B(c(J'), 100D^{s(J)+3}) \subset B(c(\mathfrak{p}), 10D^{s(\mathfrak{p})}). \quad (6.23)$$

Using the definition of \mathfrak{S} , we have

$$2^{Zn/2} \leq d_{\mathfrak{p}'}(\mathcal{Q}(\mathfrak{u}_1), \mathcal{Q}(\mathfrak{u}_2)) \leq d_{B(c(J'), 100D^{s(J)+2})}(\mathcal{Q}(\mathfrak{u}_1), \mathcal{Q}(\mathfrak{u}_2)).$$

By the doubling property (1.5) and (6.23) and the definition of \mathfrak{S} , this is

$$\leq 2^{-100a} d_{B(c(\mathfrak{p}), 10D^{s(\mathfrak{p})})}(\mathcal{Q}(\mathfrak{u}_1), \mathcal{Q}(\mathfrak{u}_2)) \leq 2^{-94a} d_{\mathfrak{p}}(\mathcal{Q}(\mathfrak{u}_1), \mathcal{Q}(\mathfrak{u}_2)) \leq 2^{Zn/2-94a}.$$

This is a contradiction, hence the second inequality in (6.22) follows. \square

Lemma 6.18 (local tree control). *For all $J \in \mathcal{J}'$ and all bounded g with bounded support,*

$$\sup_{B^\circ(J)} |T_{\mathfrak{T}(\mathfrak{u}_2) \setminus \mathfrak{S}}^* g| \leq 2^{104a^3} \inf_J M |g|. \quad (6.24)$$

Proof. Since $T_{\mathfrak{p}}^*$ is supported on $B(c(\mathfrak{p}), 5D^{s(\mathfrak{p})})$, the triangle inequality and Lemma 6.17 bound the left hand side of (6.24) by

$$\sup_{B^\circ(J)} \sum_{\substack{\mathfrak{p} \in \mathfrak{T}(\mathfrak{u}_2) \setminus \mathfrak{S} \\ B(\mathcal{I}(\mathfrak{p})) \cap B^\circ(J) \neq \emptyset}} |T_{\mathfrak{p}}^* g| \leq \sum_{s=s(J)}^{s(J)+3} \sum_{\substack{\mathfrak{p} \in \mathfrak{P}, s(\mathfrak{p})=s \\ B(\mathcal{I}(\mathfrak{p})) \cap B^\circ(J) \neq \emptyset}} \sup_{B^\circ(J)} |T_{\mathfrak{p}}^* g|. \quad (6.25)$$

If $x \in E(\mathfrak{p})$ and $B(\mathcal{I}(\mathfrak{p})) \cap B^\circ(J) \neq \emptyset$, then

$$B(c(J), 16D^{s(\mathfrak{p})}) \subset B(x, 32D^{s(\mathfrak{p})}).$$

Together with the doubling property (1.1) and the kernel bounds (2.3) we bound (6.25) by

$$2^{103a^3} \sum_{s=s(J)}^{s(J)+3} \sum_{\substack{\mathfrak{p} \in \mathfrak{P}, s(\mathfrak{p})=s \\ B(\mathcal{I}(\mathfrak{p})) \cap B^\circ(J) \neq \emptyset}} \frac{1}{\mu(B(c(J), 16D^s))} \int_{E(\mathfrak{p})} |g| d\mu.$$

Since the $E(\mathfrak{p})$ in the inner sum are pairwise disjoint and contained in $B(c(J), 16D^{s(\mathfrak{p})})$, the last display is bounded by

$$2^{103a^3} \sum_{s=s(J)}^{s(J)+3} \frac{1}{\mu(B(c(J), 16D^s))} \int_{B(c(J), 16D^s)} |g| d\mu \leq \inf_{x' \in J} 2^{103a^3+2} M |g|.$$

\square

Lemma 6.19 (global tree control 1). *Let $\mathfrak{C}_1 = \mathfrak{T}(\mathfrak{u}_1)$ and $\mathfrak{C}_2 = \mathfrak{T}(\mathfrak{u}_2) \cap \mathfrak{S}$. Then for $i = 1, 2$ and $J \in \mathcal{J}'$ and bounded g with bounded support, we have*

$$\sup_{2B(J)} |T_{\mathfrak{C}_i}^* g| \leq \inf_{B^\circ(J)} |T_{\mathfrak{C}_i}^* g| + 2^{128a^3+4a+3} \inf_J M |g| \quad (6.26)$$

and for all $y, y' \in 2B(J)$

$$\begin{aligned} & |e(\mathcal{Q}(\mathfrak{u}_i)(y)) T_{\mathfrak{C}_i}^* g(y) - e(\mathcal{Q}(\mathfrak{u}_i)(y')) T_{\mathfrak{C}_i}^* g(y')| \\ & \leq 2^{128a^3+4a+1} \left(\frac{\rho(y, y')}{D^{s(J)}} \right)^{\frac{1}{a}} \inf_J M |g|. \end{aligned} \quad (6.27)$$

Proof. Note that (6.26) follows from (6.27). By the triangle inequality, Equation (6.13) and Lemma 6.16, we have for all $y, y' \in 2B(J)$

$$\begin{aligned} & |e(\mathcal{Q}(\mathbf{u}_i)(y))T_{\mathfrak{C}_i}^*g(y) - e(\mathcal{Q}(\mathbf{u}_i)(y'))T_{\mathfrak{C}_i}^*g(y')| \\ & \leq \sum_{\substack{\mathbf{p} \in \mathfrak{C}_i \\ B(\mathcal{I}(\mathbf{p})) \cap 2B(J) \neq \emptyset}} |e(\mathcal{Q}(\mathbf{u}_i)(y))T_{\mathbf{p}}^*g(y) - e(\mathcal{Q}(\mathbf{u}_i)(y'))T_{\mathbf{p}}^*g(y')| \\ & \leq 2^{128a^3} \rho(y, y')^{1/a} \sum_{\substack{\mathbf{p} \in \mathfrak{C}_i \\ B(\mathcal{I}(\mathbf{p})) \cap 2B(J) \neq \emptyset}} \frac{D^{-s(\mathbf{p})/a}}{\mu(B(c(\mathbf{p}), 4D^{s(\mathbf{p})}))} \int_{E(\mathbf{p})} |g| \, d\mu. \end{aligned}$$

For tiles $\mathbf{p} \in \mathfrak{C}_i$ with $B(\mathcal{I}(\mathbf{p})) \cap 2B(J) \neq \emptyset$ and $s(\mathbf{p}) < s(J)$, we have $\mathcal{I}(\mathbf{p}) \subset B(c(J), 100D^{s(J)+1})$. Since $\mathbf{p} \in \mathfrak{C}_i \subset \mathfrak{S}$, it follows from the definition of \mathcal{J}' that $s(J) = -S$, which contradicts $s(\mathbf{p}) < s(J)$. Further, for each $s \geq s(J)$, the sets $E(\mathbf{p})$ for $\mathbf{p} \in \mathfrak{P}$ with $s(\mathbf{p}) = s$ are pairwise disjoint and contained in $B(c(J), 32D^s)$. With (1.1), we then estimate the previous display by

$$\begin{aligned} & \leq 2^{128a^3} \rho(y, y')^{1/a} \sum_{S \geq s \geq s(J)} D^{-s/a} \frac{2^{4a}}{\mu(B(c(J), 32D^s))} \int_{B(c(J), 32D^s)} |g| \, d\mu \\ & \leq 2^{128a^3+4a+1} \left(\frac{\rho(y, y')}{D^{s(J)}} \right)^{1/a} \inf_J M |g|. \quad \square \end{aligned}$$

Combining Lemma 6.19 and Lemma 6.18 also proves the following lemma.

Lemma 6.20 (global tree control 2). *We have for all $J \in \mathcal{J}'$ and all bounded g with bounded support*

$$\sup_{2B(J)} |T_{\mathfrak{S}(u_2) \cap \mathfrak{S}}^*g| \leq \inf_{B^\circ(J)} |T_{\mathfrak{S}(u_2)}^*g| + 2^{129a^3+4a+4} \inf_J M |g|.$$

6.5.3. The van der Corput estimate.

Lemma 6.21 (lower oscillation bound). *For all $J \in \mathcal{J}'$, we have that*

$$d_{B(J)}(\mathcal{Q}(\mathbf{u}_1), \mathcal{Q}(\mathbf{u}_2)) \geq 2^{-201a^3} 2^{Zn/2}.$$

Proof. Let J' be the parent cube of J . By definition of \mathcal{J}' and the triangle inequality, there exists $\mathbf{p} \in \mathfrak{S}$ such that

$$\mathcal{I}(\mathbf{p}) \subset B(c(J'), 100D^{s(J')+1}) \subset B(c(J), 128D^{s(J)+2}).$$

Thus, by definition of \mathfrak{S} :

$$2^{Zn/2} \leq d_{\mathbf{p}}(\mathcal{Q}(\mathbf{u}_1), \mathcal{Q}(\mathbf{u}_2)) \leq d_{B(c(J), 128D^{s(J)+2})}(\mathcal{Q}(\mathbf{u}_1), \mathcal{Q}(\mathbf{u}_2)).$$

The lemma follows using the doubling property (1.3) and $a \geq 4$. \square

Proof of Lemma 6.11. By the triangle inequality, the left hand side of (6.17) is at most

$$\leq \sum_{J \in \mathcal{J}'} \left| \int_{B(J)} e(\mathcal{Q}(\mathbf{u}_2)(y) - \mathcal{Q}(\mathbf{u}_1)(y)) h_J(y) \, d\mu(y) \right|.$$

The van der Corput Proposition 2.5 estimates this by

$$\leq 2^{7a} \sum_{J \in \mathcal{J}'} \mu(2B(J)) \|h_J\|_{C^\tau(B(J))} (1 + d_{B(J)}(\mathcal{Q}(\mathbf{u}_1), \mathcal{Q}(\mathbf{u}_2)))^{-1/(2a^2+a^3)}.$$

With Lemma 6.15, Lemma 6.21 and $a \geq 4$, the last display is estimated by

$$\leq 2^{485a^3+201} 2^{-Zn/(4a^2+2a^3)} \sum_{J \in \mathcal{J}'} \mu(B(J)) \prod_{j=1}^2 \left(\inf_{B^c(J)} |T_{\mathfrak{T}(u_j)}^* g_j| + \inf_J M|g_j| \right).$$

Using the doubling property (1.1), a summand with fixed J is controlled by

$$2^{6a} \int_J \prod_{j=1}^2 (|T_{\mathfrak{T}(u_j)}^* g_j|(x) + M|g_j|(x)) d\mu(x).$$

The lemma follows by summing over $J \in \mathcal{J}'$ and using Cauchy-Schwarz. \square

6.6. The remaining tiles. Differently from the previous subsection, define

$$\mathcal{J}' := \{J \in \mathcal{J}(\mathfrak{T}(u_1)) : J \subset \mathcal{I}(u_1)\}.$$

The collection \mathcal{J}' is a partition of $\mathcal{I}(u_1)$.

In this section, we prove Lemma 6.12. By Lemma 6.4 and the support property (6.13), we estimate the left side of (6.18) by

$$\leq 2^{130a^3} \|g_1 \mathbf{1}_{\mathcal{I}(u_1)}\|_2 \|P_{\mathcal{J}'} |T_{\mathfrak{T}(u_2) \setminus \mathfrak{S}}^* g_2|\|_2.$$

This reduces Lemma 6.12 to the following lemma.

Lemma 6.22 (bound for tree projection). *We have*

$$\|P_{\mathcal{J}'} |T_{\mathfrak{T}(u_2) \setminus \mathfrak{S}}^* g_2|\|_2 \leq 2^{102a^3+21a+5} 2^{-\frac{25}{101a} Zn\kappa} \|\mathbf{1}_{\mathcal{I}(u_1)} M|g_2|\|_2. \quad (6.28)$$

Proof. Expanding the definition of $P_{\mathcal{J}'}$, we have for the left side of (6.28)

$$\begin{aligned} & \left(\sum_{J \in \mathcal{J}'} \frac{1}{\mu(J)} \left| \int_J \sum_{\mathfrak{p} \in \mathfrak{T}(u_2) \setminus \mathfrak{S}} T_{\mathfrak{p}}^* g_2 d\mu(y) \right|^2 \right)^{1/2} \\ & \leq \sum_{s \geq s_1} \left(\sum_{J \in \mathcal{J}'} \frac{1}{\mu(J)} \left| \int_J \sum_{\substack{\mathfrak{p} \in \mathfrak{T}(u_2) \setminus \mathfrak{S} \\ s(\mathfrak{p})=s(J)-s \\ J \cap B(\mathcal{I}(\mathfrak{p})) \neq \emptyset}} T_{\mathfrak{p}}^* g_2 d\mu(y) \right|^2 \right)^{1/2}. \end{aligned} \quad (6.29)$$

Here we have restricted the summation set using Lemma 6.23 below with $s_1 := \frac{Zn}{202a^3}$ and used that by the support property (6.13), the integral over J vanishes if $J \cap B(\mathcal{I}(\mathfrak{p})) = \emptyset$. We also used Minkowski's inequality.

Since for each $I \in \mathcal{D}$ the sets $E(\mathfrak{p})$ with $\mathfrak{p} \in \mathfrak{P}(I)$ are disjoint, it follows from the doubling property (1.1) and the kernel upper bound (2.3) that

$$\left| \int_J \sum_{\substack{\mathfrak{p} \in \mathfrak{T}(u_2) \setminus \mathfrak{S} \\ \mathcal{I}(\mathfrak{p})=I \\ J \cap B(\mathcal{I}(\mathfrak{p})) \neq \emptyset}} T_{\mathfrak{p}}^* g_2 d\mu \right| \leq 2^{103a^3} \int_J \mathbf{1}_{B(I)} M|g_2| d\mu.$$

By Lemma 6.13, we have $\mathcal{I}(\mathfrak{p}) \cap \mathcal{I}(u_1) = \emptyset$ for all $\mathfrak{p} \in \mathfrak{T}(u_2) \setminus \mathfrak{S}$. Thus we can estimate (6.29) by

$$2^{103a^3} \sum_{s \geq s_1} \left(\sum_{J \in \mathcal{J}'} \frac{1}{\mu(J)} \left| \int_J \sum_{\substack{I \in \mathcal{D}, s(I)=s(J)-s \\ I \cap \mathcal{I}(u_1)=\emptyset \\ J \cap B(I) \neq \emptyset}} M|g_2| \mathbf{1}_{B(I)} d\mu \right|^2 \right)^{\frac{1}{2}},$$

which by Cauchy-Schwarz and Lemma 6.24 below is bounded by

$$\leq 2^{103a^3} \sum_{s \geq s_1} \left(\sum_{J \in \mathcal{J}'} \int_J (M|g_2|)^2 2^{14a+1} (8D^{-s})^\kappa \right)^{\frac{1}{2}}.$$

Summing a geometric series using $s_1 = \frac{Zn}{202a^3}$ and using that \mathcal{J}' is a partition of $\mathcal{I}(\mathbf{u}_1)$ completes the proof. \square

Lemma 6.23 (thin scale impact). *If $\mathbf{p} \in \mathfrak{T}(\mathbf{u}_2) \setminus \mathfrak{S}$ and $J \in \mathcal{J}'$ with $B(\mathcal{I}(\mathbf{p})) \cap B(J) \neq \emptyset$, then*

$$s(\mathbf{p}) \leq s(J) - \frac{Zn}{202a^3}.$$

Proof. Assume to the contrary that $s(\mathbf{p}) > s(J) - s_1$ with $s_1 := \frac{Zn}{202a^3}$. Then

$$\rho(c(\mathbf{p}), c(J)) \leq 8D^{s(J)} + 8D^{s(\mathbf{p})} \leq 16D^{s(\mathbf{p})+s_1}.$$

Let J' be the parent of J . By definition (6.2) of \mathcal{J} , there is $\mathbf{p}' \in \mathfrak{T}(\mathbf{u}_1)$ with

$$\mathcal{I}(\mathbf{p}') \subset B(c(J'), 100D^{s(J')+1}) \subset B(c(\mathbf{p}), 128D^{s(\mathbf{p})+s_1+1}). \quad (6.30)$$

Since $\mathcal{I}(\mathbf{u}_1) \subset \mathcal{I}(\mathbf{u}_2)$, we have by the forest properties (2.33) and (2.29)

$$d_{\mathbf{p}'}(\mathcal{Q}(\mathbf{u}_1), \mathcal{Q}(\mathbf{u}_2)) > 2^{Z(n+1)} - 4 \geq 2^{Z(n+1)-1}.$$

It follows by (6.30) and the monotonicity property (1.4) that

$$2^{Z(n+1)-1} \leq d_{B(c(\mathbf{p}), 128D^{s(\mathbf{p})+s_1+1})}(\mathcal{Q}(\mathbf{u}_1), \mathcal{Q}(\mathbf{u}_2)).$$

Using the doubling property (1.3) and $\mathbf{p} \notin \mathfrak{S}$, we estimate this further by

$$\leq 2^{9a+100a^3(s_1+2)} d_{\mathbf{p}}(\mathcal{Q}(\mathbf{u}_1), \mathcal{Q}(\mathbf{u}_2)) \leq 2^{9a+100a^3(s_1+2)} 2^{Zn/2}.$$

The last two displays give the following contradiction to the definition of s_1 :

$$Zn/2 + Z - 1 \leq 9a + 100a^3(s_1 + 2). \quad \square$$

Lemma 6.24 (square function count). *For $J \in \mathcal{J}'$ and $s \geq 0$, we have*

$$\frac{1}{\mu(J)} \int_J \left(\sum_{\substack{I \in \mathcal{D}, s(I)=s(J)-s \\ I \cap \mathcal{I}(\mathbf{u}_1)=\emptyset \\ J \cap B(I) \neq \emptyset}} \mathbf{1}_{B(I)} \right)^2 d\mu \leq 2^{14a+1} (8D^{-s})^\kappa.$$

Proof. Since $J \in \mathcal{J}'$ we have $J \subset \mathcal{I}(\mathbf{u}_1)$. Thus, if $B(I) \cap J \neq \emptyset$, then

$$B(I) \cap J \subset \{x \in J : \rho(x, X \setminus J) \leq 8D^{-s} D^{s(J)}\}.$$

Furthermore, for each s , the balls $B(I)$ with $s(I) = s$ have overlap at most 2^{7a} by the doubling property (1.1). Combining this with the small boundary property (2.11), the lemma follows. \square

6.7. Forests. In this subsection, we complete the proof of Proposition 2.3 from the results of the previous subsections.

Define an n -row to be an n -forest $(\mathfrak{U}, \mathfrak{T})$, such that in addition the sets $\mathcal{I}(\mathbf{u}), \mathbf{u} \in \mathfrak{U}$ are pairwise disjoint. By iteratively selecting the trees with inclusion maximal top tiles \mathbf{u} , we can decompose the forest $(\mathfrak{U}, \mathfrak{T})$ into a disjoint union of at most 2^n many n -rows

$$(\mathfrak{U}_j, \mathfrak{T}_j) := (\mathfrak{U}_j, \mathfrak{T}|_{\mathfrak{U}_j}).$$

We set $\mathfrak{R}_j = \cup_{\mathbf{u} \in \mathfrak{U}_j} \mathfrak{T}(\mathbf{u})$. The support property (6.13) implies that $T_{\mathfrak{T}(\mathbf{u})}$ with $\mathbf{u} \in \mathfrak{U}_j$ maps the corresponding summand of the orthogonal direct sum

$$\bigoplus_{\mathbf{u} \in \mathfrak{U}_j} L^2(\mathcal{I}(\mathbf{u})) \subset L^2(X),$$

into itself and annihilates all other summands. Hence, the operator norm of the operator associated to a row is the maximum of the norms of the corresponding tree operators. With Lemma 6.7 and the density assumption (2.32) this gives the following row estimate.

Lemma 6.25 (row bound). *For each $1 \leq j \leq 2^n$ and each bounded g with bounded support with $|g| \leq \mathbf{1}_G$, we have*

$$\|T_{\mathfrak{R}_j}^* g\|_2 \leq 2^{182a^3} 2^{-n/2} \|g\|_2$$

and

$$\|\mathbf{1}_F T_{\mathfrak{R}_j}^* g\|_2 \leq 2^{283a^3} 2^{-n/2} \text{dens}_2(\bigcup_{\mathbf{u} \in \mathfrak{U}} \mathfrak{T}(\mathbf{u}))^{1/2} \|g\|_2.$$

We further have the following TT^* estimate for distinct rows.

Lemma 6.26 (row correlation). *For all $1 \leq j < j' \leq 2^n$ and for all g_1, g_2 with $|g_i| \leq \mathbf{1}_G$, it holds that*

$$\left| \int T_{\mathfrak{R}_j}^* g_1 \overline{T_{\mathfrak{R}_{j'}}^* g_2} d\mu \right| \leq 2^{876a^3+1-4n} \|g_1\|_2 \|g_2\|_2. \quad (6.31)$$

Proof. Using the support property (6.13) and Lemma 6.10 first and then Cauchy Schwarz, we bound the left hand side of (6.31) by

$$\begin{aligned} & \sum_{\mathbf{u} \in \mathfrak{U}_j} \sum_{\mathbf{u}' \in \mathfrak{U}_{j'}} \left| \int T_{\mathfrak{T}_j(\mathbf{u})}^* (\mathbf{1}_{\mathcal{I}(\mathbf{u})} g_1) \overline{T_{\mathfrak{T}_{j'}(\mathbf{u}')}^* (\mathbf{1}_{\mathcal{I}(\mathbf{u}')} g_2)} d\mu \right| \\ & \leq 2^{512a^3+1-4n} \sum_{\mathbf{u} \in \mathfrak{U}_j} \sum_{\mathbf{u}' \in \mathfrak{U}_{j'}} \|S_{2,\mathbf{u}}(\mathbf{1}_{\mathcal{I}(\mathbf{u})} g_1)\|_{L^2(\mathcal{I}(\mathbf{u}'))} \|S_{2,\mathbf{u}'}(\mathbf{1}_{\mathcal{I}(\mathbf{u}')} g_2)\|_{L^2(\mathcal{I}(\mathbf{u}))} \\ & \leq 2^{512a^3+1-4n} \left(\sum_{\substack{\mathbf{u} \in \mathfrak{U}_j \\ \mathbf{u}' \in \mathfrak{U}_{j'}}} \|S_{2,\mathbf{u}}(\mathbf{1}_{\mathcal{I}(\mathbf{u})} g_1)\|_{L^2(\mathcal{I}(\mathbf{u}'))}^2 \sum_{\substack{\mathbf{u} \in \mathfrak{U}_j \\ \mathbf{u}' \in \mathfrak{U}_{j'}}} \|S_{2,\mathbf{u}'}(\mathbf{1}_{\mathcal{I}(\mathbf{u}')} g_2)\|_{L^2(\mathcal{I}(\mathbf{u}))}^2 \right)^{\frac{1}{2}}. \end{aligned}$$

By the estimate (6.15) on $S_{2,\mathbf{u}}$ and pairwise disjointness of the sets $\mathcal{I}(\mathbf{u})$ for $\mathbf{u} \in \mathfrak{U}_j$ and of the sets $\mathcal{I}(\mathbf{u}')$ for $\mathbf{u}' \in \mathfrak{U}_{j'}$, the last display is controlled by the right hand side of (6.31), completing the proof of the lemma. \square

Define for $1 \leq j \leq 2^n$

$$E_j := \bigcup_{u \in \mathfrak{U}_j} \bigcup_{p \in \mathfrak{T}(u)} E_1(p).$$

The separation condition (2.33) for trees in a forest implies that the sets E_j are pairwise disjoint.

Proof of Proposition 2.3. Recalling the expression (6.12) for T_p^* , we have for each j

$$T_{\mathfrak{R}_j}^* g = \sum_{u \in \mathfrak{U}_j} \sum_{p \in \mathfrak{T}(u)} T_p^* g = \sum_{u \in \mathfrak{U}_j} \sum_{p \in \mathfrak{T}(u)} T_p^* \mathbf{1}_{E_j} g = T_{\mathfrak{R}_j}^* \mathbf{1}_{E_j} g.$$

Using this we can write

$$\left\| \sum_{j=1}^{2^n} T_{\mathfrak{R}_j}^* g \right\|_2^2 = \sum_{j=1}^{2^n} \int_X |T_{\mathfrak{R}_j}^* \mathbf{1}_{E_j} g|^2 + \sum_{j=1}^{2^n} \sum_{j'=1, j' \neq j}^{2^n} \int_X \overline{T_{\mathfrak{R}_j}^* \mathbf{1}_{E_j} g} T_{\mathfrak{R}_{j'}}^* \mathbf{1}_{E_{j'}} g \, d\mu.$$

We use Lemma 6.25 to estimate each term in the first sum, and Lemma 6.26 to bound each term in the second sum.

$$\leq 2^{566a^3-n} \sum_{j=1}^{2^n} \|\mathbf{1}_{E_j} g\|_2^2 + 2^{876a^3+1-4n} \sum_{j=1}^{2^n} \sum_{j'=1}^{2^n} \|\mathbf{1}_{E_j} g\|_2 \|\mathbf{1}_{E_{j'}} g\|_2.$$

By Cauchy-Schwarz in the second two sums and disjointness of the sets E_j , this is at most

$$2^{876a^3+1} (2^{-n} + 2^n 2^{-4n}) \sum_{j=1}^n \|\mathbf{1}_{E_j} g\|_2^2 \leq 2^{876a^3+2-n} \|g\|_2^2.$$

Taking adjoints and square roots, it follows that for all f

$$\left\| \sum_{u \in \mathfrak{U}} \sum_{p \in \mathfrak{T}(u)} T_p f \right\|_2 \leq 2^{439a^3 - \frac{n}{2}} \|f\|_2. \quad (6.32)$$

On the other hand, we have by disjointness of the sets E_j

$$\left\| \sum_{u \in \mathfrak{U}} \sum_{p \in \mathfrak{T}(u)} T_p f \right\|_2 = \left\| \sum_{j=1}^{2^n} \mathbf{1}_{E_j} T_{\mathfrak{R}_j} f \right\|_2 = \left(\sum_{j=1}^{2^n} \|\mathbf{1}_{E_j} T_{\mathfrak{R}_j} f\|_2^2 \right)^{1/2}.$$

If $|f| \leq \mathbf{1}_F$ then we obtain from Lemma 6.25

$$\begin{aligned} &\leq 2^{283a^3} \text{dens}_2 \left(\bigcup_{u \in \mathfrak{U}} \mathfrak{T}(u) \right)^{\frac{1}{2}} 2^{-\frac{n}{2}} \left(\sum_{j=1}^{2^n} \|f\|_2^2 \right)^{\frac{1}{2}} = \\ &= 2^{283a^3} \text{dens}_2 \left(\bigcup_{u \in \mathfrak{U}} \mathfrak{T}(u) \right)^{\frac{1}{2}} \|f\|_2. \end{aligned} \quad (6.33)$$

Proposition 2.3 follows by taking the product of the $(2 - \frac{2}{q})$ -th power of (6.32) and the $(\frac{2}{q} - 1)$ -st power of (6.33). \square

7. PROOF OF THE HÖLDER CANCELLATIVE CONDITION

We use the following standard approximation lemma, which we will not prove here. Recall that $\tau = 1/a$.

Lemma 7.1 (Lipschitz Holder approximation). *Let $z \in X$ and $R > 0$. Let $\varphi : X \rightarrow \mathbb{C}$ be a function supported in the ball $B := B(z, R)$ with finite norm $\|\varphi\|_{C^\tau(2B)}$. Let $0 < t \leq 1$. There exists a function $\tilde{\varphi} : X \rightarrow \mathbb{C}$, supported in $B(z, 2R)$, such that for every $x \in X$*

$$|\varphi(x) - \tilde{\varphi}(x)| \leq t^\tau \|\varphi\|_{C^\tau(2B)} \quad (7.1)$$

and

$$\|\tilde{\varphi}\|_{\text{Lip}(B(z, 2R))} \leq 2^{4a} t^{-1-a} \|\varphi\|_{C^\tau(2B)}. \quad (7.2)$$

We turn to the proof of Proposition 2.5.

Proof of Proposition 2.5. Let $z \in X$ and $R > 0$ and set $B = B(z, R)$. Let φ be given as in Proposition 2.5. Set

$$t := (1 + d_B(\vartheta, \theta))^{-\frac{\tau}{2+a}}$$

and define $\tilde{\varphi}$ as in Lemma 7.1. Let ϑ and θ be in Θ . Then

$$\begin{aligned} & \left| \int e(\vartheta(x) - \theta(x)) \varphi(x) \, d\mu(x) \right| \\ & \leq \left| \int e(\vartheta(x) - \theta(x)) \tilde{\varphi}(x) \, d\mu(x) \right| + \left| \int e(\vartheta(x) - \theta(x)) (\varphi(x) - \tilde{\varphi}(x)) \, d\mu(x) \right| \end{aligned} \quad (7.3)$$

Using the cancellative condition (1.7) of Θ on the ball $B(z, 2R)$, the first term in (7.3) is bounded above by

$$2^a \mu(B(z, 2R)) \|\tilde{\varphi}\|_{\text{Lip}(B(z, 2R))} (1 + d_{B(z, 2R)}(\vartheta, \theta))^{-\tau}.$$

With the doubling condition (1.1), the inequality (7.2), and the monotonicity $d_B \leq d_{B(z, 2R)}$, the last display is bounded above by

$$2^{6a} t^{-1-a} \mu(B) \|\varphi\|_{C^\tau(B)} (1 + d_B(\vartheta, \theta))^{-\tau}.$$

The second term in (7.3) is by (7.1) bounded by

$$\mu(B(z, 2R)) t^\tau \|\varphi\|_{C^\tau(B)} \leq 2^a \mu(B) t^\tau \|\varphi\|_{C^\tau(B)}.$$

The proposition now follows by adding these two estimates. \square

REFERENCES

- [BBL⁺24] C. Benea, F. Bernicot, V. Lie, and M. Vitturi. The non-resonant bilinear Hilbert-Carleson operator. *Adv. Math.*, 458:Paper No. 109939, 136, 2024 (cited on page 6).
- [BdFD⁺25] L. Becker et al. A blueprint for the formalization of Carleson's theorem on convergence of fourier series, 2025. arXiv: 2405.06423 [math.CA] (cited on pages 1, 3).
- [Bec23] L. Becker. A degree one Carleson operator along the paraboloid, 2023. arXiv: 2312.01134 [math.CA] (cited on page 6).
- [Bil67] P. Billard. Sur la convergence presque partout des séries de Fourier-Walsh des fonctions de l'espace $L^2(0, 1)$. *Studia Math.*, 28:363–388, 1966/67 (cited on pages 3, 6).

- [Car66] L. Carleson. On convergence and growth of partial sums of Fourier series. *Acta Math.*, 116:135–157, 1966 (cited on pages 2, 6).
- [Chr90] M. Christ. A T(b) theorem with remarks on analytic capacity and the Cauchy integral. *Colloquium Mathematicum*, 2(60-61):601–628, 1990 (cited on pages 2, 14).
- [CW71] R. R. Coifman and G. Weiss. *Analyse harmonique non-commutative sur certains espaces homogènes*. Lecture Notes in Mathematics, Vol. 242. Springer-Verlag, Berlin-New York, 1971. Étude de certaines intégrales singulières (cited on page 2).
- [Fef73] C. Fefferman. Pointwise convergence of Fourier series. *Ann. of Math. (2)*, 98:551–571, 1973 (cited on pages 6, 7, 16).
- [GL24] A. Gaitan and V. Lie. Non-zero to zero curvature transition: operators along hybrid curves with no quadratic (quasi-)resonances, 2024. arXiv: 2402.03451 [math.CA] (cited on page 6).
- [GPR⁺17] S. Guo, L. B. Pierce, J. Roos, and P.-L. Yung. Polynomial Carleson operators along monomial curves in the plane. *J. Geom. Anal.*, 27(4):2977–3012, 2017 (cited on page 6).
- [GRY20] S. Guo, J. Roos, and P.-L. Yung. Sharp variation-norm estimates for oscillatory integrals related to Carleson’s theorem. *Anal. PDE*, 13(5):1457–1500, 2020 (cited on page 6).
- [HL24] M. Hsu and V. Lie. On a Carleson-Radon transform (the non-resonant setting), 2024. arXiv: 2411.01660 [math.CA] (cited on page 6).
- [Hun68] R. A. Hunt. On the convergence of Fourier series. In *Orthogonal Expansions and their Continuous Analogues (Proc. Conf., Edwardsville, Ill., 1967)*, pages 235–255. Southern Illinois Univ. Press, Carbondale, IL, 1968 (cited on page 6).
- [KR22] B. Krause and J. Roos. Discrete analogues of maximally modulated singular integrals of Stein-Wainger type. *J. Eur. Math. Soc. (JEMS)*, 24(9):3183–3213, 2022 (cited on page 6).
- [KR23] B. Krause and J. Roos. Discrete analogues of maximally modulated singular integrals of Stein-Wainger type: ℓ^p bounds for $p > 1$. *J. Funct. Anal.*, 285(10):Paper No. 110123, 19, 2023 (cited on page 6).
- [Lie09] V. Lie. The (weak- L^2) boundedness of the quadratic Carleson operator. *Geom. Funct. Anal.*, 19(2):457–497, 2009 (cited on page 3).
- [Lie20] V. Lie. The polynomial Carleson operator. *Ann. of Math. (2)*, 192(1):47–163, 2020 (cited on pages 3, 6, 7).
- [LT00] M. Lacey and C. Thiele. A proof of boundedness of the Carleson operator. *Math. Res. Lett.*, 7(4):361–370, 2000 (cited on page 6).
- [Mna22] G. Mnatsakanyan. On almost-everywhere convergence of Malmquist Takenaka series. *J. Funct. Anal.*, 282(12):Paper No. 109461, 33, 2022 (cited on pages 3, 6).

- [MS79] R. A. Macías and C. Segovia. Lipschitz functions on spaces of homogeneous type. *Adv. in Math.*, 33(3):257–270, 1979 (cited on page 2).
- [PY19] L. B. Pierce and P.-L. Yung. A polynomial Carleson operator along the paraboloid. *Rev. Mat. Iberoam.*, 35(2):339–422, 2019 (cited on page 6).
- [Ram21] J. P. G. Ramos. The Hilbert transform along the parabola, the polynomial Carleson theorem and oscillatory singular integrals. *Math. Ann.*, 379(1-2):159–185, 2021 (cited on page 6).
- [Ste93] E. M. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, volume 43 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1993, pages xiv+695. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III (cited on page 2).
- [SW01] E. M. Stein and S. Wainger. Oscillatory integrals related to Carleson’s theorem. *Math. Res. Lett.*, 8(5-6):789–800, 2001 (cited on pages 3, 6).
- [Zor21] P. Zorin-Kranich. Maximal polynomial modulations of singular integrals. *Adv. Math.*, 386:Paper No. 107832, 40, 2021 (cited on pages 3, 6, 7, 13, 14).

MATHEMATICAL INSTITUTE, UNIVERSITY OF BONN, ENDENICHER ALLEE 60, 53115, BONN, GERMANY

Email address: `becker@math.uni-bonn.de`

MATHEMATICAL INSTITUTE, UNIVERSITY OF BONN, ENDENICHER ALLEE 60, 53115, BONN, GERMANY

Email address: `vdoorn@math.uni-bonn.de`

MATHEMATICAL INSTITUTE, UNIVERSITY OF BONN, ENDENICHER ALLEE 60, 53115, BONN, GERMANY

Email address: `ajamnesh@math.uni-bonn.de`

MATHEMATICAL INSTITUTE, UNIVERSITY OF BONN, ENDENICHER ALLEE 60, 53115, BONN, GERMANY, AND

MAX PLANCK INSTITUTE FOR MATHEMATICS, VIVATSGASSE 7, 53111, BONN, GERMANY.

Email address: `rajulas@math.uni-bonn.de`

MATHEMATICAL INSTITUTE, UNIVERSITY OF BONN, ENDENICHER ALLEE 60, 53115, BONN, GERMANY

Email address: `thiele@math.uni-bonn.de`