Publ. Mat. **68** (2024), 445–506
DOI: 10.5565/PUBLMAT6822406

NEW LOCAL $T1$ THEOREMS ON NON-HOMOGENEOUS SPACES

Paco Villarroya

Abstract: We develop new local $T1$ theorems to characterize Calderón–Zygmund operators that extend boundedly or compactly on $L^p(\mathbb{R}^n,\mu)$, with μ a measure of power growth.

The results, whose proofs do not require random grids, have weaker hypotheses than previously known local $T1$ theorems since they only require a countable collection of testing functions. Moreover, a further extension of this work allows the use of testing functions supported on cubes of different dimensions.

As a corollary, we describe the measures μ of the complex plane for which the Cauchy integral defines a compact operator on $L^p(\mathbb{C},\mu)$.

2020 Mathematics Subject Classification: 42B20, 42C40, 47B07, 47G10, 28C05.

Key words: Calderón–Zygmund operator, compact operator, non-doubling Radon measures, Cauchy integral.

1. Introduction

The $T1$ theorem characterizes the boundedness of Calderón–Zygmund operators T in terms of the functions $T1$ and T^*1. On the other hand, the local $T1$ theorem attains a similar characterization using the action of T and T^* over a system of indicator functions $(\chi_Q)_{Q\in\mathcal{Q}}$ of all cubes with edges parallel to the coordinate axes.

The idea of a local $T1$ theorem was first introduced in 1990 by M. Christ [**3**] in connection with the geometric description of removable compact sets for bounded analytic functions (known as Pailenvé's problem). His motivation was that, in principle, finding a system of local testing functions should be easier than identifying a single function over which the operator behaves well globally. This approach was shown to be right at the turn of the century when F. Nazarov, S. Treil, and A. Volberg proved the first local $T1$ and Tb theorems for non-doubling measures in [**18**] and [**21**] (see also [**20**] and [**22**]). Since then, research work on this subject has been continuously growing with special focus on more general criteria of boundedness ([**1**], [**13**], [**12**]), variants that apply to new settings ([**11**], [**10**], [**14**]) and applications to PDEs ([**9**]).The articles [**7**], [**8**], [**2**], [**16**] and the books [**4**], [**5**], [**15**], [**24**] provide detailed accounts of the evolution of this theory.

A few years ago, papers [**25**], [**26**] presented global $T1$ and Tb theorems characterizing the compactness of Calderón–Zygmund operators. These results can be used to prove the compactness of many double layer potential operators [**26**]. In turn, this allows us via Fredholm theory ([**6**]) to deduce invertibility of the Laplacian on a large class of domains. Following this line of research, the current paper introduces a local $T1$ theorem for non-doubling measures, that is, a criterion of boundedness and compactness that relies on the action of the operator over a family of indicator functions of dyadic cubes (Theorems 4.1 and 4.2).

Most proofs of $T1$ theorems on non-doubling spaces employ randomization methods ([**18**], [**20**], [**21**], [**22**], [**12**], [**13**], [**14**]). The reason for this is the fact that, when using the kernel decay, estimates of the dual pair $\langle T\chi_I, \chi_J\rangle$ grow logarithmically with both the distance between the cubes I, J and the ratio between their side lengths. To overcome this issue, the probabilistic approach considers grids of general cubes rather than only the grid of dyadic cubes. In the space of all these grids, cubes with close boundaries and very different sizes are rare, and thus they can be assigned a small probability. Then, by averaging among all grids, the contribution of such cubes can be made arbitrarily small. The costs of this method include a delicate technique of decomposition called surgery, and the requirement by hypothesis of a non-countable family of testing functions (one for each cube in $\mathbb{R}^n$). The recent work [**17**] develops a local Tb theorem for non-doubling measures without the use of random grids. But the result still requires the use of a non-numerable family of testing measures.

We introduce a new proof approach that does not use random grids and instead explicitly addresses the contribution of close cubes with very different sizes. To work with such cubes, the method proceeds, broadly speaking, as follows: in the dual pair $\langle Tf, g\rangle$, we decompose each argument function as the sum of two functions with supports in the interior of dyadic cubes and in their boundaries respectively. Then we estimate the dual pair with functions supported inside nearby cubes by summing telescoping sums and using that the measure of the part of each open cube close to the boundary can be made arbitrarily small. Finally, for the dual pair over functions supported on the boundary we take translated cubes so that the functions are supported on the interior of these translated cubes and we apply a recursion process. This work is mostly carried out at the end of the proof of Theorem 4.2 in Subsections 10.6 to 10.8.

The new theorem allows the use of a countable family of testing functions and opens up the possibility of extending the results to new settings such as manifolds and fractal sets.

Regarding compactness, we provide an application to the Cauchy integral operator with a non-doubling Radon measure μ, which is defined by

$$C_\mu f(z) = \int_{\mathbb{C}} \frac{f(w)}{w-z}\, d\mu(w).$$

It is known that if the measure is defined by the indicator function of a unit line segment S, that is, $d\mu_S = \chi_S\, d\mathcal{H}^1$, with $\mathcal{H}^1$ the one-dimensional Hausdorff measure in $\mathbb{C}$, then C_{μ_S} is bounded but not compact on $L^2(\mu_S)$. On the other hand, if the measure is defined by the indicator function of a unit square Q, that is, $d\mu_Q = \chi_Q\, dm$, with m the Lebesgue measure in $\mathbb{C}$, then C_{μ_Q} is compact on $L^2(\mu_Q)$. Theorem 4.4 describes for which measures μ of the complex plane the Cauchy integral operator C_μ can be compactly extended on $L^2(\mu)$.

The outline of the paper is as follows. In Sections 2, 3, and 4 we introduce some notation, define the class of operators under study, and state the main results respectively. In Section 5, we study a smooth truncation of the kernel, while in Section 6 we describe the Haar wavelet system. Section 7 contains two technical results on some auxiliary functions. Section 8 focuses on obtaining estimates for the action of the operator over Haar wavelets. In Section 9 we deal with the paraproducts, and Section 10 is devoted to the proof of the main result, Theorem 4.2.

Acknowledgements. I want to express my appreciation to 김자영, and my gratitude to Professors Oscar Blasco, Anthony Carbery, Neil Lyall, and Gerardo Mendoza for their unreserved support. I also thank the reviewers for their valuable work, which helped to improve the manuscript.

2. Notation

2.1. Cubes and dyadic cubes. Let $\mathcal{C}$ be the family of cubes in $\mathbb{R}^n$ defined by tensor products of intervals of the same length, namely, $I = \prod_{i=1}^n [r_i, r_i + l)$, with $r_i, l \in \mathbb{R}$.

For each cube $I \in \mathcal{C}$, we denote its center by $c(I)$, its side length by $\ell(I)$, and its boundary in the Euclidean topology of $\mathbb{R}^n$ by ∂I.

Let $\mathcal{D}_1$ be the family of dyadic cubes $I = \prod_{i=1}^n 2^{-k}[j_i, j_i + 1)$, with $j_i, k \in \mathbb{Z}$. Let $\tilde{\mathcal{D}}_1$ be the family of open dyadic cubes $I = \prod_{i=1}^n 2^{-k}(j_i, j_i + 1)$, with $j_i, k \in \mathbb{Z}$.

Now, let $\lambda_1 = 0$ and $\lambda_2, \dots, \lambda_n \in \mathbb{R}^+$ such that $\lambda_i \in \mathbb{R}^+ \setminus \big(\bigcup_{j=1}^{i-1}(\lambda_j + \mathbb{Q})\big)$. Without loss of generality, we can assume $\lambda_1 < \lambda_2 < \cdots < \lambda_n < 1$. Let $a_i = \lambda_i(1, \dots, 1) \in \mathbb{R}^n$. For $i \in \{1, \dots, n\}$, we define the families of cubes

$$\mathcal{T}_i\mathcal{D} = a_i + \mathcal{D}_1 = \{a_i + I : I \in \mathcal{D}_1\}, \tag{1}$$

with $a_i + I \in \mathcal{C}$ such that $c(a_i + I) = a_i + c(I)$ and $\ell(a_i + I) = \ell(I)$.

For each grid $\mathcal{T}_i\mathcal{D}$, we define its first quadrant as $\mathbb{R}^{n,+}_i = a_i + \mathbb{R}^{n,+}$, where $\mathbb{R}^{n,+} = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n : x_i \geq 0\}$.

We write any particular instance of the families of cubes $\mathcal{T}_i\mathcal{D}$ (and $\mathcal{D}^r_\partial$, defined later in this section) simply as $\mathcal{D}$. We often call these cubes dyadic, using the term loosely. We also denote any particular instance of the families of cubes $\mathcal{T}_i\tilde{\mathcal{D}}$, which is defined in a similar way as $\mathcal{T}_i\mathcal{D}$, simply by $\tilde{\mathcal{D}}$.

Given a measurable set $\Omega \subset \mathbb{R}^n$, let $\mathcal{D}(\Omega)$ be the family of dyadic cubes $I \in \mathcal{D}$ such that $I \subsetneq \Omega$.

For $\lambda > 0$, we write λI for the cube such that $c(\lambda I) = c(I)$ and $\ell(\lambda I) = \lambda \ell(I)$. We write $\mathbb{B} = [-1/2, 1/2)^n$ and $\mathbb{B}_\lambda = \lambda\mathbb{B}$. We also denote by $\lambda\mathcal{D}$ the family of cubes λI with $I \in \mathcal{D}$.

For $I \in \mathcal{D}$, we define the children of I, denoted by $\mathrm{ch}(I)$, as the collection of dyadic cubes $I' \subset I$ such that $\ell(I') = \ell(I)/2$, and the parent of I as $I_p \in \mathcal{D}$, that is, the only cube in $\mathcal{D}$ such that $I \in \mathrm{ch}(I_p)$. If $\Omega \in \mathcal{D}$ and $I \in \mathcal{D}(\Omega)$, then $I_p \subseteq \Omega$.

We define the friends of I, denoted by I^{fr}, as the collection of cubes $J \in \mathcal{D}$ such that $\ell(J) = \ell(I)$ and $\mathrm{dist}(I, J) = 0$, where $\mathrm{dist}(I, J)$ denotes the Euclidean set distance between I and J.

2.2. Pairs of cubes: eccentricity and relative distances. Given $I, J \in \mathcal{C}$, if $\ell(J) \leq \ell(I)$, we write $I \wedge J = J$, $I \vee J = I$, while if $\ell(I) < \ell(J)$, we write $I \wedge J = I$, $I \vee J = J$.

We define $\langle I, J\rangle$ as the only cube containing $I \cup J$ with the smallest possible side length and such that $\sum_{i=1}^n c(\langle I, J\rangle)_i$ is minimum. We note that $\ell(\langle I, J\rangle) \approx \mathrm{dist}(I, J) + \ell(I \vee J)$.

Let $[I, J]$ be the unique cube satisfying $\ell([I, J]) = \mathrm{dist}(I, J)$, $\lambda[I, J] \cap I \neq \emptyset$, and $\lambda[I, J] \cap J \neq \emptyset$ for any $\lambda > 1$, and such that $\sum_i c([I, J])_i$ is minimum.

We define the eccentricity and the relative distance of I and J as

$$\mathrm{ec}(I, J) = \frac{\ell(I \wedge J)}{\ell(I \vee J)}, \quad \mathrm{rdist}(I, J) = 1 + \frac{\mathrm{dist}(I, J)}{\ell(I \vee J)}.$$

We define the inner boundary of I as $\mathfrak{D}_I = \bigcup_{I' \in \mathrm{ch}(I)} \partial I'$, and the inner relative distance of J and I by

$$\mathrm{inrdist}(I, J) = 1 + \frac{\mathrm{dist}(I \wedge J, \mathfrak{D}_{I \vee J})}{\ell(I \wedge J)}.$$

2.3. Lagom cubes. For $M \in \mathbb{N}$, we define $\mathcal{C}_M$ as the family of cubes in $\mathcal{C}$ such that $2^{-M} \leq \ell(I) \leq 2^M$ and $\mathrm{rdist}(I, \mathbb{B}_{2^M}) \leq M$. We call the cubes in $\mathcal{C}_M$ lagom cubes.

We write $\mathcal{D}_M = \mathcal{C}_M \cap \mathcal{D}$, $\mathcal{D}_M^c = \mathcal{D} \setminus \mathcal{D}_M$, $\mathcal{D}_M(\Omega) = \mathcal{D}_M \cap \mathcal{D}(\Omega)$, and $\mathcal{D}_M^c(\Omega) = \mathcal{D}_M^c \cap \mathcal{D}(\Omega)$.

2.4. Grids of dyadic cubes of lower dimensions. We write $\mathcal{D}_\partial^n = \mathcal{D}_1$. Let $\partial\mathcal{D}^n$ be the set defined by the union of ∂I for all $I \in \mathcal{D}_1$. This set is the union of countably many affine Euclidean spaces of dimension $n-1$. Then let $\mathcal{D}^{n-1}$ be the family of dyadic cubes in $\partial\mathcal{D}^n$, namely, the cubes of dimension $n-1$ of the form

$$I = \prod_{i=1}^{l-1} 2^{-k}[j_i, j_i+1) \times \alpha_l \times \prod_{i=l+1}^{n} 2^{-k}[j_i, j_i+1),$$

where $l \in \{1, \dots, n\}$, $j_i, k \in \mathbb{Z}$, and $\alpha_l \in \{2^{-k}j_l, 2^{-k}(j_l+1)\}$. We use the convention that $\prod_{i=1}^{-1} 2^{-k}[j_i, j_i+1) = \alpha_1$ and $\prod_{i=n+1}^{n} 2^{-k}[j_i, j_i+1) = \alpha_n$.

We continue recursively. For $0 < r < n$, we define $\partial\mathcal{D}^{r+1}$ as the union of ∂I for all $I \in \mathcal{D}^{r+1}$, where ∂I denotes the border of I in the Euclidean topology of $\mathbb{R}^{r+1}$. In this way $\partial\mathcal{D}^{r+1}$ is the union of countably many affine Euclidean spaces of dimension r. Finally, we define $\mathcal{D}^r$ as the family of r-dimensional dyadic cubes in $\partial\mathcal{D}^{r+1}$.

3. Measure, kernel, and operator

3.1. Non-homogeneous measure. We describe the class of measures for which the theory applies.

Definition 3.1. Let μ be a Radon measure on $\mathbb{R}^n$, which without loss of generality we assume to be positive.

We say that μ has power growth if there is $0 < \alpha \leq n$ such that

$$\mu(I) \lesssim \ell(I)^\alpha \tag{2}$$

for all $I \in \mathcal{C}$.

We now define three densities of the measure: for $I \in \mathcal{C}$, let

$$\rho(I) = \frac{\mu(I)}{\ell(I)^\alpha},$$

$$\rho_{\mathrm{in}}(I) = \sup_{\substack{t \in I \\ 0 < \lambda < \ell(I)}} \frac{\mu(I \cap B(t,\lambda))}{\lambda^\alpha} = \sup_{\substack{t \in I \\ \lambda > 0}} \frac{\mu(I \cap B(t,\lambda))}{\lambda^\alpha},$$

where $B(t,\lambda) = \{x \in \mathbb{R}^n : |t-x| < \lambda\}$, and given $0 < \delta \leq 1$,

$$\rho_{\mathrm{out}}(I) = \sum_{m \geq 1} \frac{\mu(mI)}{\ell(mI)^\alpha} \frac{1}{m^{\frac{\delta}{2}+1}}.$$

We denote

$$\rho_\mu(I) = \rho_{\mathrm{in}}(I) + \rho_{\mathrm{out}}(I). \tag{3}$$

Remark 3.2. In the definition of ρ_{in} one can substitute the balls $B(t,\lambda)$ by dyadic cubes just by taking the smallest $Q \in \mathcal{D}(I)$ with $B(t,\lambda) \subset Q$.

The sum in the definition of ρ_{out} is comparable to

$$\int_1^\infty \frac{\mu(tI)}{\ell(tI)^\alpha} \frac{dt}{t^{\frac{\delta}{2}+1}} \approx \sum_{k \geq 0} 2^{-k\frac{\delta}{2}} \frac{\mu(2^k I)}{\ell(2^k I)^\alpha}.$$

If μ satisfies the power growth for n-dimensional cubes of $\mathcal{D}$, then it satisfies the same power growth for r-dimensional cubes of $\mathcal{D}_\partial^r$. To show this, we note that each r-dimensional dyadic cube $I \in \mathcal{D}_\partial^r$ is on the border of an n-dimensional dyadic cube $Q \in \mathcal{D}$ with the same side length and then $\mu(I) \leq \mu(Q) \lesssim \ell(Q)^\alpha = \ell(I)^\alpha$.

3.2. Compact Calderón–Zygmund kernel and its associated operator. We now describe the class of kernels and operators for which the theory applies.

Throughout the paper, $|\cdot|$ denotes the norm on $l^p(\mathbb{R}^n)$ for any $1 \leq p \leq \infty$. However, to simplify notation, in some inequalities we reason as if we were using $l^\infty(\mathbb{R}^n)$. For other norms, the results hold just by replacing $\leq$ with $\lesssim$.

Definition 3.3. Let μ be a positive Radon measure on $\mathbb{R}^n$ with power growth $0 < \alpha \leq n$.

A function $K\colon (\mathbb{R}^n \times \mathbb{R}^n) \setminus \{(t,x) \in \mathbb{R}^n \times \mathbb{R}^n : t = x\} \to \mathbb{C}$ is a Calderón–Zygmund kernel if it is bounded on compact sets of its domain and there exist $0 < \delta \leq 1$ and bounded functions $L, S, D\colon [0,\infty) \to [0,\infty)$ satisfying

$$|K(t,x) - K(t',x')| \lesssim \left(\frac{(|t-t'| + |x-x'|)}{|t-x|}\right)^\delta \frac{F(t,x)}{|t-x|^\alpha}, \tag{4}$$

with $F(t,x) = L(|t-x|)S(|t-x|)D(|t+x|)$, whenever $2(|t-t'| + |x-x'|) < |t-x|$.

We say that $K\, d\mu \times d\mu$ is a compact Calderón–Zygmund kernel if (4) holds and

$$\lim_{\ell(I)\to\infty} L(\ell(I))\rho_\mu(I) = \lim_{\ell(I)\to 0} S(\ell(I))\rho_\mu(I) = \lim_{\mathrm{rdist}(I,\mathbb{B})\to\infty} D(\mathrm{rdist}(I,\mathbb{B}))\rho_\mu(I) = 0. \tag{5}$$

The purpose of the function F in (4) is to estimate the tails of sums indexed by dyadic cubes, especially in the theorems on compactness. When summing over an infinite collection of cubes, the tails are defined by those cubes that are very large, or very small, or those with intermediate size but at a large distance from the origin. The functions L, S, D are used to control the contributions of large, small, and distant cubes respectively.

Remark 3.4. Since a dilation of a function satisfying any of the limits in (5) satisfies the same limit, that is, $\mathcal{D}_\lambda(L\rho_\mu)(a) = L(\lambda^{-1}a)\rho_\mu(\lambda^{-1}a)$ also satisfies the first limit, we omit universal constants in the arguments of the functions.

We note that, without loss of generality, L and D can be assumed to be non-increasing, while S can be assumed to be non-decreasing. Otherwise, we define

$$\begin{aligned}
\mathcal{L}(r) &= \sup\{L(s) : s \geq r\},\\
\mathcal{D}(r) &= \sup\{D(s) : s \geq r\},\\
\mathcal{S}(r) &= \sup\{S(s) : 0 \leq s \leq r\}.
\end{aligned}$$

The functions $\mathcal{L}$, $\mathcal{D}$ are non-increasing, $\mathcal{S}$ is non-decreasing, and, since they are greater than the original functions L, D, and S respectively, they satisfy (4). Then one can use the newly defined functions instead.

Notation 3.5. *Given three cubes $I_1, I_2, I_3 \in \mathcal{C}$, we denote*

$$F(I_1, I_2, I_3) = L(\ell(I_1))S(\ell(I_2))D(\mathrm{rdist}(I_3, \mathbb{B}))$$

and $F(I) = F(I,I,I)$. Then the limits in (5) *can be compactly written as*

$$\lim_{M\to\infty} \sup_{I \in \mathcal{D}_M^c} F(I)\rho_\mu(I) = 0. \tag{6}$$

The two expressions are equivalent because $\mathcal{D}_M^c$ contains the cubes that are large, small, or distant. Therefore, when $\ell(I)$ tends to infinity, since the functions S and D are bounded, we have

$$\begin{aligned}\lim_{\ell(I)\to\infty} F(I)\rho_\mu(I) &= \lim_{\ell(I)\to\infty} L(\ell(I))S(\ell(I))D(\mathrm{rdist}(I,\mathbb{B}))\rho_\mu(I)\\ &\leq \|S\|_\infty \|D\|_\infty \lim_{\ell(I)\to\infty} L(\ell(I))\rho_\mu(I) = 0.\end{aligned}$$

Similar reasoning applies when either $\ell(I)$ tends to zero or $\mathrm{rdist}(I,\mathbb{B})$ *tends to infinity.*

Given $0 < \delta \leq 1$, we define

$$\tilde{D}(I) = \sum_{k\geq 0} 2^{-k\frac{\delta}{2}} D(\mathrm{rdist}(2^k I, \mathbb{B})). \tag{7}$$

If D satisfies (5)*, then by Lebesgue's dominated convergence theorem so does $\tilde{D}$.*

In [**25**] it was proved, when $\alpha = n = 1$, that the smoothness condition (4) and the mild assumption $\lim_{|t-x|\to\infty} K(t,x) = 0$ imply the following pointwise decay condition:

$$|K(t,x)| \lesssim \frac{F(t,x)}{|t-x|^\alpha}, \tag{8}$$

with $F(t,x) = L(|t-x|)S(|t-x|)D(|t+x|)$. This is also the case when $F \equiv 1$.

Definition 3.6. A linear operator T is associated with a Calderón–Zygmund kernel K if the representation

$$Tf(x) = \int_{\mathbb{R}^n} f(t)K(t,x)\,d\mu(t) \tag{9}$$

holds for all functions f bounded and compactly supported, and $x \notin \operatorname{supp} f$.

By (8) and the properties of f and x, the previous integral is absolutely convergent with

$$\int_{\mathbb{R}^n} |f(t)K(t,x)|\,d\mu(t) \lesssim \|f\|_{L^\infty(\mu)} \frac{\mu(\operatorname{supp} f)}{\mathrm{dist}(x, \operatorname{supp} f)^\alpha}.$$

4. Statements of main results

We denote Kronecker's delta by δ: $\delta(x) = 1$ if $x = 0$ and $\delta(x) = 0$ otherwise. We denote by $\lfloor x \rfloor$ the floor function, that is, the greatest integer less than or equal to x.

Theorem 4.1. *Let μ be a positive Radon measure on $\mathbb{R}^n$ with power growth $0 < \alpha \leq n$. Let T be a linear operator with a Calderón–Zygmund kernel and measure μ as in* (9)*. Let $1 < p < \infty$, and $k = n - \lfloor\alpha\rfloor + \delta(\alpha - \lfloor\alpha\rfloor)$.*

Then the following statements are equivalent:

(i) *T extends to a bounded operator on $L^p(\mu)$.*

(ii) *There exist k grids of n-dimensional cubes, $\mathcal{T}_i\mathcal{D}$ as in* (1) *with $i \in \{1, 2, \ldots, k\}$, such that the testing condition*

$$\|\chi_I T\chi_I\|_{L^2(\mu)} + \|\chi_I T^*\chi_I\|_{L^2(\mu)} \lesssim \mu(I)^{\frac{1}{2}} \tag{10}$$

holds for all $I \in \mathcal{T}_i\mathcal{D} \cup 4\mathcal{T}_i\mathcal{D}$.

(iii) *For the k grids $\mathcal{D}_\partial^r$ of r-dimensional cubes as defined in Subsection 2.4 with $r \in \{n, n-1, \ldots, n-k+1\}$, T satisfies* (10) *for all $I \in \mathcal{D}_\partial^r \cup 4\mathcal{D}_\partial^r$.*

Theorem 4.1 follows from the proof of the result below.

Theorem 4.2. *Let μ be a positive Radon measure on $\mathbb{R}^n$ with power growth $0 < \alpha \leq n$. Let T be a linear operator with a Calderón–Zygmund kernel and measure μ as in* (9). *Let $1 < p < \infty$, and $k = n - \lfloor \alpha \rfloor + \delta(\alpha - \lfloor \alpha \rfloor)$.*

Then the following statements are equivalent:

(i) *T extends to a compact operator on $L^p(\mu)$.*

(ii) *$K\, d\mu \times d\mu$ is a compact Calderón–Zygmund kernel and there exist k grids of n-dimensional cubes, $\mathcal{T}_i\mathcal{D}$ as in* (1) *with $i \in \{1, 2, \dots, k\}$, such that*

$$\|\chi_I T \chi_I\|_{L^2(\mu)} + \|\chi_I T^* \chi_I\|_{L^2(\mu)} \lesssim \mu(I)^{\frac{1}{2}} F_T(I) \tag{11}$$

for all $I \in \mathcal{T}_i\mathcal{D} \cup 4\mathcal{T}_i\mathcal{D}$, with F_T bounded and satisfying

$$\lim_{M \to \infty} \sup_{I \in (\mathcal{T}_i\mathcal{D})_M^c \cup 4(\mathcal{T}_i\mathcal{D})_M^c} F_T(I) = 0. \tag{12}$$

(iii) *$K\, d\mu \times d\mu$ is a compact Calderón–Zygmund kernel and for the k grids of r-dimensional cubes $\mathcal{D}_\partial^r$ as in Subsection 2.4 with $r \in \{n, n-1, \dots, n-k+1\}$ we have that T satisfies* (11) *for all $I \in \mathcal{D}_\partial^r \cup 4\mathcal{D}_\partial^r$, with F_T bounded and satisfying*

$$\lim_{M \to \infty} \sup_{I \in (\mathcal{D}_\partial^r)_M^c \cup 4(\mathcal{D}_\partial^r)_M^c} F_T(I) = 0.$$

To be used in forthcoming results and the proof of Theorem 4.2, we provide the following notation:

Notation 4.3. *Let L, S, $\tilde{D}$ be the functions of Notation 3.5. We denote*

$$F_K(I, J) = L(\ell(I \wedge J)S(\ell(I \wedge J))(D(\mathrm{rdist}(\langle I, J\rangle, \mathbb{B})) + \tilde{D}(\mathrm{inrdist}(I, J)I \wedge J)) \tag{13}$$

and $F_K(I) = F_K(I, I)$.

Let ρ_μ, F_K, F_T be as defined in (3), (13), *and* (11) *respectively. We write Kronecker's delta by δ. We define*

$$F_\mu(I, J) = \sup_{\substack{R \subset I \\ S \subset J}} F_K(R, S)\rho_\mu(R \vee S) + F_T(I)\delta(I, J) \tag{14}$$

and $F_\mu(I) = F_\mu(I, I)$.

We apply the previous result to the Cauchy integral operator

$$C_\mu(f)(z) = \int \frac{f(w)}{w - z}\, d\mu(w).$$

The cases when $0 < \alpha \leq 2$ with $\alpha \neq 1$ are already known. For $\alpha = 1$, we obtain the following result.

Theorem 4.4. *Let μ be a positive Radon measure on the complex plane $\mathbb{C}$ such that $\mu(I) \lesssim \ell(I)$ for each $I \in \mathcal{D}$. Let $1 < p < \infty$. Then the following statements are equivalent:*

(i) *C_μ is bounded on $L^p(\mu)$.*

(ii) *There exist two grids of two-dimensional cubes, $\mathcal{T}_i\mathcal{D}$ with $i \in \{1, 2\}$ as defined in* (1), *such that the testing condition*

$$\|\chi_I C_\mu \chi_I\|_{L^2(\mu)} \lesssim \mu(I)^{\frac{1}{2}} \tag{15}$$

holds for all $I \in \mathcal{T}_i\mathcal{D} \cup 4\mathcal{T}_i\mathcal{D}$.

(iii) *For the grids of dyadic squares $\mathcal{D}_\partial^2$ and dyadic line segments $\mathcal{D}_\partial^1$ as defined in Subsection 2.4, we have that* (15) *holds for all $I \in \mathcal{D}_\partial^r \cup 4\mathcal{D}_\partial^r$ with $r \in \{2, 1\}$.*

Furthermore, the following statements are also equivalent:

(i') *C_μ is compact on $L^p(\mu)$.*

(ii') *There exist two grids of two-dimensional cubes, $\mathcal{T}_i\mathcal{D}$ with $i \in \{1,2\}$ as defined in* (1)*, such that* (15) *holds and*

$$\lim_{M\to\infty} \sup_{I\in(\mathcal{T}_i\mathcal{D})^c_M} \rho_\mu(I) = \lim_{M\to\infty} \sup_{I\in(\mathcal{T}_i\mathcal{D})^c_M\cup 4(\mathcal{T}_i\mathcal{D})^c_M} \frac{\|\chi_I C_\mu \chi_I\|_{L^2(\mu)}}{\mu(I)^{\frac{1}{2}}} = 0.$$

(iii') *For the grids of dyadic squares $\mathcal{D}^2_\partial$ and dyadic line segments $\mathcal{D}^1_\partial$ as defined in Subsection 2.4, we have that* (15) *holds and*

$$\lim_{M\to\infty} \sup_{I\in(\mathcal{D}^r_\partial)^c_M} \rho_\mu(I) = \lim_{M\to\infty} \sup_{I\in(\mathcal{D}^r_\partial)^c_M\cup 4(\mathcal{D}^r_\partial)^c_M} \frac{\|\chi_I C_\mu \chi_I\|_{L^2(\mu)}}{\mu(I)^{\frac{1}{2}}} = 0.$$

We end this section with a technical remark. By duality and density, to prove the boundedness of T it is enough to show

$$|\langle Tf, g\rangle_\mu| \lesssim \|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)} \tag{16}$$

for all functions f, g that are bounded and compactly supported on a cube Q, with implicit bound in (16) independent of f, g, and Q. The next lemma proves that we can make one more assumption on f and g.

Lemma 4.5. *To prove the boundedness of T it is enough to show $|\langle Tf, g\rangle_\mu| \lesssim \|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)}$ for all functions f, g that, in addition to the previous properties, are supported on the interior of the first quadrant of each grid $\mathcal{T}_i\mathcal{D}$.*

Proof: Let $f_a(x) = f(x-a)$, $K_a(t,x) = K(t-a, x-a)$, $\mu_a(A) = \mu(-a+A)$, T_a be the operator with associated kernel K_a and associated measure μ_a as in (9), $F_{a,\mu}(I) = F_\mu(-a+I)$, and $F_{a,T}(I) = F_T(-a+I)$.

We note three facts that can be directly proved from the given definitions:

- K_a is a Calderón–Zygmund kernel satisfying the smoothness condition (4) with the same constant as K;
- μ_a is a Radon measure of power growth satisfying (2) with the same exponent α and constant as μ;
- $F_{a,\mu}$, $F_{a,T}$ are bounded with the same bounds as F_μ and F_T respectively, and satisfy the limits stated in (6) and (12) respectively.

Furthermore, we have $\|f_a\|^2_{L^2(\mu_a)} = \int_{\mathbb{R}^n} |f(t-a)|^2\, d\mu(t-a) = \|f\|^2_{L^2(\mu)}$, $|\langle f, g\rangle_\mu| = |\langle f_a, g_a\rangle_{\mu_a}|$, and for $x \notin a + \sup f$,

$$\begin{aligned}(Tf)_a(x) = Tf(x-a) &= \int_{\mathbb{R}^n} f(t)K(t, x-a)\, d\mu(t)\\ &= \int_{\mathbb{R}^n} f(t-a)K(t-a, x-a)\, d\mu(t-a) = T_a(f_a)(x).\end{aligned}$$

Then

$$\begin{aligned}\|\chi_I T_a \chi_I\|_{L^2(\mu_a)} &= \|(\chi_I)_{-a}(T_a\chi_I)_{-a}\|_{L^2(\mu)} = \|\chi_{-a+I} T \chi_{-a+I}\|_{L^2(\mu)}\\ &\lesssim F_T(-a+I)\mu(-a+I)^{\frac{1}{2}} = F_{a,T}(I)\mu_a(I)^{\frac{1}{2}}.\end{aligned}$$

Similarly for $(T^*)_a$. This implies that T_a satisfies the testing conditions (11) with the same constant as T.

Therefore, if we prove that

$$|\langle\tilde T(f_a), g_a\rangle_{\tilde\mu}| \lesssim \|f_a\|_{L^2(\tilde\mu)}\|g_a\|_{L^2(\tilde\mu)} \tag{17}$$

for all operators $\tilde T$ with the same constants as T and all measures $\tilde\mu$ with the same exponent and constant as μ, we will have

$$\begin{aligned}|\langle Tf, g\rangle_\mu| &= |\langle (Tf)_a, g_a\rangle_{\mu_a}| = |\langle T_a(f_a), g_a\rangle_{\mu_a}|\\ &\lesssim \|f_a\|_{L^2(\mu_a)}\|g_a\|_{L^2(\mu_a)} = \|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)}.\end{aligned}$$

We now note that if $\operatorname{supp} f \subset Q$, then $\operatorname{supp} f_a \subset a+Q$. We then define $a = \lambda(1,\dots,1)$ with λ large enough such that $\lambda > 10\max|\lambda_i|$ for all λ_i as in (1). In this way $a+Q$, and thus the supports of f_a and g_a which are the argument functions in (17), are all contained in the interior of the first quadrant of each grid $\mathcal{T}_i\mathcal{D}$. □

5. Truncated operators

In this section we define and study the properties of a particular type of smooth truncations of Calderón–Zygmund operators. We start with a technical result.

Lemma 5.1. *Let $I\in\mathcal{C}$, and $x\in I$. Then*

$$\int_I \frac{1}{|t-x|^{\alpha-1}}\,d\mu(t) \lesssim \ell(I)\rho_{\rm in}(I).$$

Proof: For $k\ge 0$, let $S_k = \{t\in I/|t-x|\le 2^{-k}\ell(I)\}$ and $C_k = S_k\setminus S_{k+1} = \{t\in I : 2^{-(k+1)}\ell(I) < |t-x| \le 2^{-k}\ell(I)\}$. Then

$$\begin{aligned}\int_I \frac{1}{|t-x|^{\alpha-1}}\,d\mu(t) &\le \sum_{k\ge 0}\frac{2^{(\alpha-1)(k+1)}}{\ell(I)^{\alpha-1}}\mu(S_k\setminus S_{k+1})\\ &\lesssim \ell(I)^{1-\alpha}\sum_{k\ge0}2^{(\alpha-1)k}(\mu(S_k)-\mu(S_{k+1})).\end{aligned}$$

To prove the result, we write $I_R = \ell(I)^{1-\alpha}\sum_{k=0}^R 2^{(\alpha-1)k}(a_k - a_{k+1})$, where $a_k = \mu(S_k)$. Then we bound I_R uniformly on R. By Abel's formula, we have

$$I_R = \ell(I)^{1-\alpha}\Big(a_0 - a_{R+1}2^{(\alpha-1)R} + \sum_{k=1}^R a_k(2^{(\alpha-1)k} - 2^{(\alpha-1)(k-1)})\Big).$$

Since $a_0\le\mu(I) = \rho(I)\ell(I)^\alpha \le \rho_{\rm in}(I)\ell(I)^\alpha$, for the first term we have $\ell(I)^{1-\alpha}a_0 \le \ell(I)\rho_{\rm in}(I)$.

Similarly, since

$$a_{R+1} = \mu(S_{R+1}) \le \mu(I\cap B(x, 2^{-(R+1)}\ell(I))) \le \rho_{\rm in}(I)2^{-(R+1)\alpha}\ell(I)^\alpha,$$

the absolute value of the second term can be bounded by

$$\ell(I)^{1-\alpha}a_{R+1}2^{(\alpha-1)R} \lesssim \ell(I)\rho_{\rm in}(I)2^{-R} \le \ell(I)\rho_{\rm in}(I).$$

Meanwhile, since $a_k = \mu(S_k) \le \rho_{\rm in}(I)2^{-k\alpha}\ell(I)^\alpha$, the absolute value of the last term is bounded by

$$\ell(I)^{1-\alpha}\sum_{k=1}^R a_k 2^{(\alpha-1)k}(1-2^{-(\alpha-1)}) \lesssim \ell(I)\rho_{\rm in}(I)\sum_{k=1}^R 2^{-k} \lesssim \ell(I)\rho_{\rm in}(I). \qquad □$$

We now define the following smooth truncation of an operator associated with a Calderón–Zygmund kernel.

Definition 5.2. Let ϕ be a smooth function such that $0 \leq \phi(x) \leq 1$, $\sup \phi \subset [-2,2]$, $\phi(x) = 1$ for all $|x| < 1$ and $0 \leq |\phi'(x)| \leq 2$.

Let $Q = [-2^r, 2^r]^n$ such that $\ell(Q) > 4$ and $\gamma = \sqrt{n}2^{-s}$ such that $0 < \gamma \leq 1$. We define the kernel

$$K_{\gamma,Q}(t,x) = K(t,x)\left(1 - \phi\left(\frac{|t-x|}{\gamma}\right)\right)\phi\left(\frac{4|t|}{\ell(Q)}\right)\phi\left(\frac{4|x|}{\ell(Q)}\right).$$

Let $T_{\gamma,Q}$ be the operator with kernel $K_{\gamma,Q}$.

In the next result we prove that $T_{\gamma,Q}$ is bounded with a bound that depends on γ and Q, while it has a Calderón–Zygmund kernel whose estimates are uniform on γ and Q. Later we show that $T_{\gamma,Q}$ satisfies a variation of the testing condition, which is close to being uniform on γ and Q. That is enough for our purposes in the proof of the main result.

Lemma 5.3. *The operator $T_{\gamma,Q}$ is bounded with bounds depending on γ and Q. Moreover, $K_{\gamma,Q}$ is a Calderón–Zygmund kernel with parameter $0 < \delta \leq 1$ and constant independent of γ and Q.*

Proof: We first show that $K_{\gamma,Q}$ is a bounded function: by (8),

$$|K_{\gamma,Q}(t,x)| \lesssim \frac{1}{|t-x|^\alpha}\left(1 - \phi\left(\frac{|t-x|}{\gamma}\right)\right) \leq \frac{1}{\gamma^\alpha}.$$

The last inequality holds because when $|t-x| \leq \gamma$ we have $\phi\big(\frac{|t-x|}{\gamma}\big) = 1$ and so, the second factor is zero. Then, since $K_{\gamma,Q}$ is bounded and supported on $Q \times Q$, for $f, g \in L^2(\mu)$ we have by the Cauchy–Schwarz inequality

$$\begin{aligned} |\langle T_{\gamma,Q} f, g\rangle| &= \left|\iint K_{\gamma,Q}(t,x) f(t) g(x)\, d\mu(t)\, d\mu(x)\right| \\ &\leq \left(\int_Q \int_Q |K_{\gamma,Q}(t,x)|^2\, d\mu(t)\, d\mu(x)\right)^{\frac{1}{2}} \|f\|_{L^2(\mu)} \|g\|_{L^2(\mu)} \\ &\lesssim \frac{\mu(Q)}{\gamma^\alpha} \|f\|_{L^2(\mu)} \|g\|_{L^2(\mu)}. \end{aligned}$$

This proves that the operator $T_{\gamma,Q}$ is bounded.

We now show that $K_{\gamma,Q}$ is a Calderón–Zygmund kernel. We prove the appropriate estimate for $A = |K_{\gamma,Q}(t,x) - K_{\gamma,Q}(t',x)|$, the work for $|K_{\gamma,Q}(t,x) - K_{\gamma,Q}(t,x')|$ being similar. Let t, t', x be such that $2|t-t'| < |t-x|$. We note that this inequality implies $|t-x| \leq |t'-x| + |t-t'| \leq |t'-x| + |t-x|/2$ and so, $|t-x| \leq 2|t'-x|$. Then

$$\begin{aligned} A \leq{}& |K(t,x) - K(t',x)|\phi\left(\frac{4|t|}{\ell(Q)}\right)\phi\left(\frac{4|x|}{\ell(Q)}\right)\left(1 - \phi\left(\frac{|t-x|}{\gamma}\right)\right) \\ &+ |K(t',x)|\left|\phi\left(\frac{4|t|}{\ell(Q)}\right) - \phi\left(\frac{4|t'|}{\ell(Q)}\right)\right|\phi\left(\frac{4|x|}{\ell(Q)}\right)\left(1 - \phi\left(\frac{|t-x|}{\gamma}\right)\right) \\ &+ |K(t',x)|\phi\left(\frac{4|t'|}{\ell(Q)}\right)\phi\left(\frac{4|x|}{\ell(Q)}\right)\left|\phi\left(\frac{|t-x|}{\gamma}\right) - \phi\left(\frac{|t'-x|}{\gamma}\right)\right|. \end{aligned}$$

Since $2|t-t'| < |t-x|$ we can use the kernel smoothness condition (4) and the fact that ϕ is bounded, to estimate the first term by a constant times

$$\left(\frac{|t-t'|}{|t-x|}\right)^\delta \frac{F(t,x)}{|t-x|^\alpha} = \frac{|t-t'|^\delta}{|t-x|^{\alpha+\delta}} F(t,x).$$

If the second term is non-zero, then $x \in Q$, and either $t \in Q$ or $t' \in Q$. If $t \in Q$, we have $|t-x| < |t| + |x| \leq \ell(Q)$. Meanwhile, if $t' \in Q$, we get $|t-x| \leq 2|t'-x| \leq 2(|t'|+|x|) \leq 2\ell(Q)$. Then, by the kernel decay condition (8), the fact that ϕ is bounded, and the mean value theorem on ϕ with bounded derivative, the second term can be estimated by a constant times

$$\frac{F(t,x)}{|t-x|^\alpha}\frac{||t|-|t'||}{\ell(Q)} \lesssim \frac{F(t,x)}{|t-x|^\alpha}\frac{|t-t'|}{|t-x|} \leq \frac{|t-t'|^\delta}{|t-x|^{\alpha+\delta}}F(t,x).$$

If the third term is non-zero, then $t', x \in Q$, and either $|t-x| < 2\gamma$ or $|t'-x| < 2\gamma$. In the latter case we have $|t-x| \leq 2|t'-x| < 4\gamma$. Then, by using again the kernel decay, that ϕ is bounded and the mean value theorem on ϕ, we can estimate this third term by a constant times

$$\frac{F(t,x)}{|t-x|^\alpha}\frac{||t-x|-|t'-x||}{\gamma} \lesssim \frac{F(t,x)}{|t-x|^\alpha}\frac{|t-t'|}{|t-x|} \leq \frac{|t-t'|^\delta}{|t-x|^{\alpha+\delta}}F(t,x). \qquad \square$$

For the next result, we denote

$$F_{\gamma,Q}(I) = (S(\gamma) + L(\ell(Q)^{\frac{1}{2}}) + D(\ell(Q)^{\frac{1}{2}}))\chi_{[\gamma,10\ell(Q)]}(\ell(I))\chi_{[0,10\ell(Q)]}(|c(I)|).$$

Lemma 5.4. *The operator $T_{\gamma,Q}$ satisfies the following testing condition: for $I \in \mathcal{D}(10Q)$,*

$$\|\chi_I T_{\gamma,Q}\chi_I\|_{L^2(\mu)} \lesssim G(I)\mu(I)^{\frac{1}{2}},$$

where $G(I) = F_\mu(I) + F_{\gamma,Q}(I)\rho_{\text{in}}(I)$ and the implicit bound is independent of γ and Q.

Remark 5.5. When $F \equiv 1$, since $F_{\gamma,Q}(I) \lesssim 1$, the result implies that $T_{\gamma,Q}$ satisfies the testing condition uniformly on γ and Q.

When $\lim_{M\to\infty}\sup_{I\in\mathcal{D}_M^c\cup 4\mathcal{D}_M^c} F_\mu(I) = 0$, due to the factors $\chi_{[\gamma,10\ell(Q)]}(\ell(I))$ and $\chi_{[0,10\ell(Q)]}(|c(I)|)$, we also have $\lim_{M\to\infty}\sup_{I\in\mathcal{D}_M^c\cup 4\mathcal{D}_M^c} G(I) = 0$.

Moreover, since $\lim_{\gamma\to 0}\lim_{\ell(Q)\to\infty} F_{\gamma,Q}(I)\rho_{\text{in}}(I) = 0$, one can make the function $G(I)$ arbitrarily close to $F_\mu(I)$.

Proof: By symmetry of the kernel $K_{\gamma,Q}$ with respect to the variables t, x, it is clear that the computations to prove the testing condition on $T_{\gamma,Q}$ also work for $T^*_{\gamma,Q}$. Therefore, we write the calculations only for $T_{\gamma,Q}$. We are going to show that for all $I \in \mathcal{D}\cup 4\mathcal{D}$ we have $\|\chi_I T_{\gamma,Q}\chi_I\|^2_{L^2(\mu)} \lesssim G(I)^2\mu(I)$, with G bounded and such that $\lim_{M\to\infty}\sup_{I\in\mathcal{D}_M^c\cup 4\mathcal{D}_M^c} G(I) = 0$.

We also note that if $\ell(I) < \gamma$, then $\chi_I T_{\gamma,Q}\chi_I \equiv 0$. Therefore, we assume $\gamma \leq \ell(I)$. With this and the fact that $I \subset 10Q$, we have $\chi_{[\gamma,10\ell(Q)]}(\ell(I)) = \chi_{[0,10\ell(Q)]}(|c(I)|) = 1$. Then we do not need to work with such factors.

Since $K_{\gamma,Q}$ is bounded, we can write $\|\chi_I T_{\gamma,Q}\chi_I\|^2_{L^2(\mu)}$ as

$$\int_I \bigg|\int_I K(t,x)\Big(1-\phi\Big(\frac{|t-x|}{\gamma}\Big)\Big)\phi\Big(\frac{4|t|}{\ell(Q)}\Big)\,d\mu(t)\bigg|^2 \phi\Big(\frac{4|x|}{\ell(Q)}\Big)^2 d\mu(x) = \text{Int}.$$

By the mean value theorem, there exists $\xi \in \big(\frac{4|t|}{\ell(Q)}, \frac{4|x|}{\ell(Q)}\big)$ such that

$$\phi\Big(\frac{4|t|}{\ell(Q)}\Big) = \phi\Big(\frac{4|x|}{\ell(Q)}\Big) + \phi'(\xi)\frac{4(|t|-|x|)}{\ell(Q)}. \tag{18}$$

Accordingly, we bound Int by the sum of two terms:

$$\mathrm{Int}_1 = \int_I \Big| \int_I K(t,x)\Big(1-\phi\Big(\frac{|t-x|}{\gamma}\Big)\Big)\, d\mu(t)\Big|^2 \phi\Big(\frac{4|x|}{\ell(Q)}\Big)^4 d\mu(x)$$

and

$$\mathrm{Int}_2 = \int_I \Big| \int_I K(t,x)\Big(1-\phi\Big(\frac{|t-x|}{\gamma}\Big)\Big)\phi'(\xi)\frac{4(|t|-|x|)}{\ell(Q)}\, d\mu(t)\Big|^2 \phi\Big(\frac{4|x|}{\ell(Q)}\Big)^2 d\mu(x), \tag{19}$$

which we bound in different ways.

Since $0 \le \phi(x) \le 1$, $\sup \phi \subset [-2,2]$, and $\phi(x) = 1$ for all $|x| < 1$, we estimate Int_1 by a constant times

$$\int_I \Big(\int_{\substack{t\in I\\ \gamma \le |t-x| \le 2\gamma}} |K(t,x)|\, d\mu(t)\Big)^2 d\mu(x) + \int_I \Big| \int_{\substack{t\in I\\ 2\gamma \le |t-x|}} K(t,x)\, d\mu(t)\Big|^2 d\mu(x).$$

By the kernel decay and the fact that $|t-x| \le 2\gamma$ implies $F(t,x) \lesssim S(|t-x|) \lesssim S(\gamma)$, the first term can be bounded by a constant times

$$\begin{aligned} \int_I \Big(\int_{\substack{t\in I\\ \gamma < |t-x| \le 2\gamma}} \frac{F(t,x)}{|t-x|^\alpha}\, d\mu(t)\Big)^2 d\mu(x) &\le S(\gamma)^2 \int_I \frac{\mu(I\cap B(x,2\gamma))^2}{\gamma^{2\alpha}}\, d\mu(x) \\ &\le S(\gamma)^2 \rho_{\mathrm{in}}(I)^2 \mu(I) \lesssim G(I)^2\mu(I). \end{aligned} \tag{20}$$

To deal with the second term, we denote $D_x = \{t \in I : |t-x| \le 2\gamma\}$ and $D_x^c = I\backslash D_x$. Since $\chi_{D_x^c} = \chi_I - \chi_{D_x}$, the second term can be written as

$$\int_I |T(\chi_{D_x^c})(x)|^2\, d\mu(x) \lesssim \int_I |T(\chi_I)(x)|^2\, d\mu(x) + \int_I |T(\chi_{D_x})(x)|^2\, d\mu(x).$$

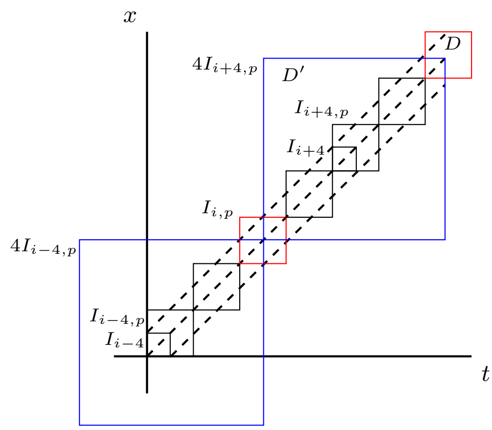
The new first term equals

$$\|\chi_I T \chi_I\|_{L^2(\mu)}^2 \lesssim F_T(I)^2\mu(I) \le G(I)^2\mu(I),$$

where we have used the testing condition (11) and $F_T \le F_\mu \le G$.

For the second term, we denote $D = \{(t,x) \in I\times I : |t-x| \le 2\gamma\}$, and $g(x) = T(\chi_{D_x})(x)$ for $x \in I$. Then we are going to prove that $\|\chi_I\, g\|_{L^2(\mu)} \lesssim G(I)\mu(I)^{\frac12}$ or, equivalently, that for all $\Phi_I \in L^2(\mu)$ with support on I and $\|\Phi_I\|_{L^2(\mu)} \le \mu(I)^{\frac12}$ we have $|\langle \Phi_I, g\rangle| \lesssim G(I)\mu(I)$.

Let $I_i \in \mathcal{D}_1(I)$ such that $I_i \times I_i \subset D$ is maximal inside D with respect to the inclusion. Therefore, $\ell(I_i)$ is the same for all cubes and it is comparable to γ. In fact, $\ell(I_i) = 2\gamma/\sqrt{n}$. Let $I_{i,p}$ be the parent of I_i and let $4I_{i,p} \in \mathcal{C}$ such that $c(4I_{i,p}) = c(I_{i,p})$, and $\ell(4I_{i,p}) = 4\ell(I_{i,p})$. Then we can choose two subcollections of bi-cubes $I_{i,p}\times I_{i,p}$ and $4I_{i+4,p}\times 4I_{i+4,p}$, satisfying the following: they completely cover the union of D and a set $D' \subset \{(t,x)\in I\times I) : \gamma < |t-x| \le 8\gamma\}$, the cubes $4I_{i,p}$ are pairwise disjoint, and the intersection $I_{i,p}\cap 4I_{j,p}$ is either empty or it is a cube that belongs to $\mathrm{ch}(I_{i,p})$. In some cases that intersection is exactly I_i.



We denote the indexes in $\mathbb{Z}^n$ corresponding to the bi-cubes in each collection as $\mathcal{O}$ and $\mathcal{E}$ respectively. Therefore, we can write

$$\chi_D(t,x) = \sum_{i\in\mathcal{O}} \chi_{I_{i,p}}(t)\chi_{I_{i,p}}(x) + \sum_{i\in\mathcal{E}} \chi_{4I_{i,p}}(t)\chi_{4I_{i,p}}(x) - \sum_{\substack{i/I_i\subset I_{j,p}\\ j\in\mathcal{O}}} \chi_{I_i}(t)\chi_{I_i}(x) - \chi_{D'}(t,x).$$

With this, we have

$$\begin{aligned}|\langle \Phi_I, g\rangle| \le \sum_{i\in\mathcal{O}} |\langle \Phi_I\chi_{I_{i,p}}, T\chi_{I_{i,p}}\rangle| + \sum_{i\in\mathcal{E}} |\langle \Phi_I\chi_{4I_{i,p}}, T\chi_{4I_{i,p}}\rangle| \\ + \sum_{\substack{i/Q_i\subset Q_{j,p}\\ j\in\mathcal{O}}} |\langle \Phi_I\chi_{I_i}, T\chi_{I_i}\rangle| + \|\Phi_I\|_{L^2(\mu)}\|T\chi_{D'}\|_{L^2(\mu)}.\end{aligned}$$

Now, as we did in (20), we can estimate the last term by a constant times

$$\mu(I)^{\frac12}\bigg(\int_I\bigg(\int_{\substack{t\in I\\ \gamma<|t-x|\le 8\gamma}} |K(t,x)|\,d\mu(t)\bigg)^2 d\mu(x)\bigg)^{\frac12} \lesssim G(I)\mu(I).$$

On the other hand, by the testing condition for T on the cubes $I_{i,p}$ and Cauchy's inequality, the first term is bounded by

$$\begin{aligned}\sum_{i\in\mathcal{O}} \|\Phi_I\chi_{I_{i,p}}\|_{L^2(\mu)}\|\chi_{I_{i,p}}T\chi_{I_{i,p}}\|_{L^2(\mu)} &\lesssim \sum_{i\in\mathcal{O}} \|\Phi_I\chi_{I_{i,p}}\|_{L^2(\mu)} F_T(I_{i,p})\mu(I_{i,p})^{\frac12} \\ &\le \bigg(\sum_{i\in\mathcal{O}} \|\Phi_I\chi_{I_{i,p}}\|^2_{L^2(\mu)}\bigg)^{\frac12}\bigg(\sum_{i\in\mathcal{O}}\mu(I_{i,p})\bigg)^{\frac12} S(\gamma).\end{aligned}$$

We have used that in $I_{i,p}$ we have $|t-x|\le\gamma$, and so we get $F_T(I_{i,p})\lesssim S_T(\gamma)\le G(I)$, where from (12), S_T satisfies $\lim_{\gamma\to 0} S_T(\gamma)=0$.

Now, since the cubes $I_{i,p}\subset I$ have the same side length, they are pairwise disjoint. Then $\sum_{i\in\mathcal{O}}\mu(I_{i,p})\le\mu(I)$ and

$$\sum_{i\in\mathcal{O}} \|\Phi_I\chi_{I_{i,p}}\|^2_{L^2(\mu)} = \sum_{i\in\mathcal{O}} \int_{I_{i,p}} |\Phi_I(x)|^2\,d\mu(x) \le \|\Phi_I\|^2_{L^2(\mu)} \le \mu(I).$$

With this, we get

$$\sum_{i\in\mathcal{O}} \|\Phi_I\chi_{I_{i,p}}\|_{L^2(\mu)}\|\chi_{I_{i,p}}T\chi_{I_{i,p}}\|_{L^2(\mu)} \le G(I)\mu(I).$$

Similar computations using the testing condition for T on the cubes $4I_{i,p}$ (which are also pairwise disjoint) and I_i respectively prove the corresponding inequalities for the second and third terms. This finishes the work to estimate Int_1.

To deal with the second term Int_2, we note the following fact. If $t, x \in 2^{-1}Q$, then $\phi\big(\frac{4|t|}{\ell(Q)}\big) = \phi\big(\frac{4|x|}{\ell(Q)}\big) = 1$ and so for $|t| \neq |x|$ we have, from equality (18), that the corresponding ξ satisfies $\phi'(\xi) = 0$. This implies that the integrand in (19) is zero. Moreover, when $|t| = |x|$ we have $\phi'(\xi) = 0$ by definition, and so the integrand is again zero. In other words, if the integrand in (19) is non-zero, then $(t,x) \in (Q\times Q)\setminus(2^{-1}Q\times 2^{-1}Q)$ and $|t| \neq |x|$. In that region we have $|x-t| + |x+t| > \max(|x-t|, |x+t|) > c\,\ell(Q)$, with $c = \sqrt{n}/4$. With this,

- If $|x-t| > c\,\ell(Q)^{\frac{1}{2}}$, then $F(t,x) \lesssim L(|t-x|) \lesssim L(\ell(Q)^{\frac{1}{2}}) \leq F_{\gamma,Q}(I)$.
- If $|x-t| \leq c\,\ell(Q)^{\frac{1}{2}}$, then

$$|x+t| > c\,\ell(Q) - |t-x| > c\,\ell(Q)^{\frac{1}{2}}(\ell(Q)^{\frac{1}{2}} - 1) \gtrsim c\,\ell(Q)/2,$$

since $\ell(Q) > 4$ implies $\ell(Q)^{\frac{1}{2}} - 1 > \ell(Q)^{\frac{1}{2}}/2$. With this and (8),

$$F(t,x) \lesssim D\bigg(1 + \frac{|t+x|}{1+|t-x|}\bigg) \lesssim D(\ell(Q)^{\frac{1}{2}}) \leq F_{\gamma,Q}(I).$$

Now we reason as follows. Since $0 \leq \phi(x) \leq 1$, $0 \leq |\phi'(x)| \leq 2$, and $\phi(x) = 1$ for $|x| \leq 1$, the second term is bounded by

$$\begin{aligned}
\mathrm{Int}_2 &\lesssim \int_I \bigg(\int_I |K(t,x)|\bigg|1 - \phi\bigg(\frac{|t-x|}{\gamma}\bigg)\bigg|\frac{||t|-|x||}{\ell(Q)}\chi_Q(t)\,d\mu(t)\bigg)^2 \phi\bigg(\frac{4|x|}{\ell(Q)}\bigg)^2 d\mu(x)\\
&\lesssim \int_I \bigg(\frac{1}{\ell(Q)}\int_{\substack{t\in I\\ |t-x|>\gamma}} |K(t,x)||t-x|\chi_Q(t)\,d\mu(t)\bigg)^2 d\mu(x)\\
&\lesssim \int_I \bigg(\frac{1}{\ell(Q)}\int_{\substack{t\in I\\ |t-x|>\gamma}} \frac{F(t,x)}{|t-x|^{\alpha}}|t-x|\chi_Q(t)\,d\mu(t)\bigg)^2 d\mu(x)\\
&\lesssim F_{\gamma,Q}(I)^2 \int_I \bigg(\frac{1}{\ell(Q)}\int_{\substack{t\in I\\ |t-x|>\gamma}} \frac{1}{|t-x|^{\alpha-1}}\,d\mu(t)\bigg)^2 d\mu(x).
\end{aligned}$$

By Lemma 5.1, the last expression is bounded by a constant times

$$F_{\gamma,Q}(I)^2 \int_I \bigg(\frac{1}{\ell(Q)}\rho_{\mathrm{in}}(I)\ell(I)\bigg)^2 d\mu(x) \lesssim F_{\gamma,Q}(I)^2\rho_{\mathrm{in}}(I)^2\mu(I) \lesssim G(I)^2\mu(I),$$

where we have used that since $I \subset 10Q$, we have $\ell(I) \lesssim \ell(Q)$. □

6. Haar wavelet systems and the characterization of compactness

6.1. The Haar wavelet system.

Definition 6.1. Let μ be a measure on $\mathbb{R}^n$. For $R \in \mathcal{D}\cup\tilde{\mathcal{D}}$ with $\mu(R) \neq 0$ we denote the average $\langle f\rangle_R = \mu(R)^{-1}\int_Q f(x)\,d\mu(x)$. For $R \in \mathcal{D}$ with $\mu(R) = 0$, we set $\langle f\rangle_R = 0$.

We define the averaging operator by $E_R f = \langle f\rangle_R \chi_R$ and the difference operator by

$$\Delta_R f = \bigg(\sum_{I\in \mathrm{ch}(R)} E_I f\bigg) - E_R f = \sum_{I\in\mathrm{ch}(R)} (\langle f\rangle_I - \langle f\rangle_R)\chi_I. \tag{21}$$

For $k \in \mathbb{Z}$, we define

$$E_k f = \sum_{\substack{R \in \mathcal{D} \\ \ell(R) = 2^{-k}}} E_R f \quad \textit{and} \quad \Delta_k f = E_k f - E_{k-1} f = \sum_{\substack{R \in \mathcal{D} \\ \ell(R) = 2^{-k}}} \Delta_R f.$$

Definition 6.2 (Haar wavelets)**.** Let $I \in \mathcal{D} \cup \tilde{\mathcal{D}}$. For $\mu(I) \neq 0$ we define the Haar wavelet function associated with I by

$$\psi_I = \mu(I)^{\frac{1}{2}} \left(\frac{1}{\mu(I)} \chi_I - \frac{1}{\mu(I_p)} \chi_{I_p} \right),$$

where $I_p \in \mathcal{D}$ is such that $I \in \operatorname{ch}(I_p)$. For $\mu(I) = 0$ we set $\psi_I = 0$.

Lemma 6.3. *For $R \in \mathcal{D} \cup \tilde{\mathcal{D}}$ and f locally integrable we have*

$$\Delta_R f = \sum_{I \in \operatorname{ch}(R)} \langle f, \psi_I \rangle \psi_I$$

almost everywhere with respect to μ.

Proof: If $\mu(R) = 0$, then $\mu(I) = 0$ for every $I \in \operatorname{ch}(R)$. With this, both $\Delta_R = 0$ and $\psi_I = 0$, and so the equality is trivial.

For $\mu(R) \neq 0$, from (21) and $\mu(R)\langle f \rangle_R = \sum_{I \in \operatorname{ch}(R)} \mu(I) \langle f \rangle_I$ we have

$$\Delta_R f = \sum_{I \in \operatorname{ch}(R)} \langle f \rangle_I \chi_I - \langle f \rangle_R \chi_R = \sum_{I \in \operatorname{ch}(R)} \langle f \rangle_I \left(\chi_I - \frac{\mu(I)}{\mu(R)} \chi_R \right) = \sum_{I \in \operatorname{ch}(R)} \mu(I)^{\frac{1}{2}} \langle f \rangle_I \psi_I,$$

where the last equality holds even for those terms for which $\mu(I) = 0$ since in that case $\langle f \rangle_I = 0$.

Also from (21) we have for each $I \in \operatorname{ch}(R)$

$$\langle \Delta_R f \rangle_I = \langle f \rangle_I - \langle f \rangle_R \tag{22}$$

and so

$$\Delta_R f = \sum_{I \in \operatorname{ch}(R)} \mu(I)^{\frac{1}{2}} \langle \Delta_R f \rangle_I \psi_I + \langle f \rangle_R \sum_{I \in \operatorname{ch}(R)} \mu(I)^{\frac{1}{2}} \psi_I.$$

For the first term, we compute the coefficients: for $\mu(I) = 0$, we have $\psi_I = 0$ and so $\mu(I)^{\frac{1}{2}} \langle \Delta_R f \rangle_I \psi_I = 0 = \langle f, \psi_I \rangle \psi_I$. Meanwhile, for $\mu(I) \neq 0$, we can use (22) to write

$$\mu(I)^{\frac{1}{2}} \langle \Delta_R f \rangle_I = \mu(I)^{\frac{1}{2}} \int f(x) \left(\frac{\chi_I(x)}{\mu(I)} - \frac{\chi_R(x)}{\mu(R)} \right) d\mu(x) = \langle f, \psi_I \rangle.$$

We now denote by R' the union of cubes $I \in \operatorname{ch}(R)$ such that $\mu(I) \neq 0$. Then for the second term we have

$$\begin{aligned} \sum_{I \in \operatorname{ch}(R)} \mu(I)^{\frac{1}{2}} \psi_I &= \sum_{\substack{I \in \operatorname{ch}(R) \\ \mu(I) \neq 0}} \mu(I)^{\frac{1}{2}} \psi_I = \sum_{\substack{I \in \operatorname{ch}(R) \\ \mu(I) \neq 0}} \left(\chi_I - \frac{\mu(I)}{\mu(R)} \chi_R \right) \\ &= \chi_{R'} - \frac{\mu(R')}{\mu(R)} \chi_R = -\chi_{R \setminus R'} = 0 \end{aligned} \tag{23}$$

almost everywhere since $\mu(R \setminus R') = 0$. □

Lemma 6.4. *Let f be bounded and compactly supported on $Q \in \mathcal{D}$. Then*

$$\int f(x) g(x) \, d\mu(x) = \lim_{M \to \infty} \int \Bigg(\sum_{\substack{I \in \mathcal{D}(Q) \\ 2^{-M} \leq \ell(I)}} \langle f, \psi_I \rangle \psi_I(x) + E_Q f(x) \Bigg) g(x) \, d\mu(x) \tag{24}$$

for g bounded and compactly supported on Q.

Proof: We first note that on the right-hand side of (24) we can write

$$\sum_{\substack{I\in\mathcal{D}(Q)\\ 2^{-M}\leq\ell(I)<\log\ell(Q)}}\langle f,\psi_I\rangle\psi_I=\sum_{\substack{R\in\mathcal{D}\\ 2^{-(M-1)}\leq\ell(R)\leq\log\ell(Q)}}\sum_{I\in\mathrm{ch}(R)}\langle f,\psi_I\rangle\psi_I.$$

Since f is bounded and compactly supported, we have by Lemma 6.3

$$\sum_{I\in\mathrm{ch}(R)}\langle f,\psi_I\rangle\psi_I=\Delta_R f$$

μ-almost everywhere. Then we can write the right-hand side of (24) as

$$\lim_{M\to\infty}\int\sum_{\substack{R\in\mathcal{D}(Q)\\ 2^{-(M-1)}\leq\ell(R)\leq\log\ell(Q)}}\sum_{I\in\mathrm{ch}(R)}\langle f,\psi_I\rangle\psi_I(x)g(x)\,d\mu(x)$$

$$=\lim_{M\to\infty}\int_Q\sum_{-M\leq k\leq\log\ell(Q)}\Delta_k f(x)g(x)\,d\mu(x).$$

Now we choose $M\in\mathbb{N}$ such that $2^{-M}\leq\ell(Q)$. For $x\in Q$, we select $J\in\mathcal{D}$ such that $x\in J\subset Q$, and $\ell(J)=2^{-M}$. Then, by summing a telescopic series, we get

$$\chi_Q(x)\sum_{-M\leq k\leq\log\ell(Q)}\Delta_k f(x)=\langle f\rangle_J\chi_J(x)-\langle f\rangle_Q\chi_Q(x)=\chi_Q(x)(E_Mf(x)-E_Qf(x)).$$

That is,

$$\chi_Q(x)\bigg(\sum_{-M\leq k\leq\log\ell(Q)}\Delta_k f(x)+E_Qf(x)\bigg)=\chi_Q(x)E_Mf(x).$$

With this,

$$\lim_{M\to\infty}\int\bigg(\sum_{-M\leq k\leq\log\ell(Q)}\Delta_k f(x)+E_Qf(x)\bigg)g(x)\,d\mu=\lim_{M\to\infty}\int_Q E_Mf(x)g(x)\,d\mu(x).$$

Since f is locally integrable, by Lebesgue's differentiation theorem we have that E_Mf converges to f pointwise almost everywhere with respect to μ when M tends to infinity. Moreover, since

$$|E_Mf(x)g(x)|\lesssim\|f\|_{L^\infty(\mu)}\|g\|_{L^\infty(\mu)}\chi_Q(x)$$

we can use Lebesgue's dominated convergence theorem to conclude the result. □

Similar work shows the validity of the following result:

Lemma 6.5. *Recall that $\tilde{\mathcal{D}}$ denotes the family of open dyadic cubes and $\partial\mathcal{D}$ denotes the union of the borders of all dyadic cubes. For $I\in\mathcal{D}$, we denote $\tilde{I}=I\setminus\partial I\in\tilde{\mathcal{D}}$.*

Let f be integrable and compactly supported on $Q\in\mathcal{D}$. We denote $f_1=f-f\chi_{\partial\mathcal{D}}$ and let $\tilde{E}_Qf$ be the average operator as in Definition 6.1 but with cubes in $\tilde{\mathcal{D}}$. Then the equality

$$\int g(x)f_1(x)\,d\mu(x)=\int g(x)\bigg(\sum_{I\in\tilde{\mathcal{D}}(Q)}\langle f,\psi_I\rangle\psi_{\tilde{I}}(x)+\tilde{E}_Qf(x)\bigg)\,d\mu(x)$$

holds for g bounded and compactly supported on Q.

6.2. A variation of the Haar wavelet system. We now define a new Haar wavelet system and show that the analog of Lemma 6.3 holds. These wavelets will be used when dealing with the paraproducts.

Definition 6.6. Let $Q \in \mathcal{C}$, $J_p \in \mathcal{D}$. We denote $c_{J_p} = c(J_p)$. For $I \in \mathcal{D}$ with $\mu(I) \neq 0$, we define

$$\psi_{I,J_p}^{\mathrm{full}}(t) = \mu(I)^{\frac{1}{2}} \left(\frac{\chi_I(c_{J_p})}{\mu(I)} - \frac{\chi_{I_p}(c_{J_p})}{\mu(I_p)} \right) \chi_Q(t).$$

If $\mu(I) = 0$, we define $\psi_{I,J_p}^{\mathrm{full}} \equiv 0$.

We omit the dependence of $\psi_{I,J_p}^{\mathrm{full}}$ on the cube Q. We note that $\psi_{I,J_p}^{\mathrm{full}} = 0$ if $J_p \cap I_p = \emptyset$, and that $\psi_{I,J_p}^{\mathrm{full}} \chi_I = \psi_I \chi_I$ when $J_p \subseteq I$.

We define the localized averaging operators by $\hat{E}_R(f) = \langle f \rangle_R \chi_R(c_{J_p}) \chi_Q$ and the corresponding localized differences

$$\hat{\Delta}_R f = \Bigg(\sum_{I \in \mathrm{ch}(R)} \hat{E}_I f \Bigg) - \hat{E}_R f = \Bigg(\sum_{I \in \mathrm{ch}(R)} \langle f \rangle_I \chi_I(c_{J_p}) \Bigg) \chi_Q - \langle f \rangle_R \chi_R(c_{J_p}) \chi_Q.$$

The following result is the analog of Lemma 6.3 for the localized difference operator.

Lemma 6.7. *Let $R, J_p \in \mathcal{D}$, with $\ell(J_p) < \ell(R)$ and $\mu(J_p) \neq 0$. Then*

$$\hat{\Delta}_R(f) = \sum_{I \in \mathrm{ch}(R)} \langle f, \psi_I \rangle \psi_{I,J_p}^{\mathrm{full}}$$

for f bounded and compactly supported.

Proof: If $\mu(R) = 0$, then both sides of the equality are zero. If $\mu(R) \neq 0$, we reason as follows. Since

$$\mu(R) \langle f \rangle_R = \int_R f \, d\mu = \sum_{\substack{I \in \mathrm{ch}(R) \\ \mu(I) \neq 0}} \int_I f \, d\mu = \sum_{\substack{I \in \mathrm{ch}(R) \\ \mu(I) \neq 0}} \mu(I) \langle f \rangle_I,$$

we have

$$\hat{\Delta}_R(f) = \sum_{\substack{I \in \mathrm{ch}(R) \\ \mu(I) \neq 0}} \langle f \rangle_I \left(\chi_I(c_{J_p}) - \frac{\mu(I)}{\mu(R)} \chi_R(c_{J_p}) \right) \chi_Q = \sum_{\substack{I \in \mathrm{ch}(R) \\ \mu(I) \neq 0}} \mu(I)^{\frac{1}{2}} \langle f \rangle_I \psi_{I,J_p}^{\mathrm{full}}.$$

Now, by (22), we have $\langle f \rangle_I = \langle \Delta_R f \rangle_I + \langle f \rangle_R$ and so

$$\hat{\Delta}_R(f) = \sum_{I \in \mathrm{ch}(R)} \mu(I)^{\frac{1}{2}} \langle \Delta_R f \rangle_I \psi_{I,J_p}^{\mathrm{full}} + \langle f \rangle_R \sum_{\substack{I \in \mathrm{ch}(R) \\ \mu(I) \neq 0}} \mu(I)^{\frac{1}{2}} \psi_{I,J_p}^{\mathrm{full}}.$$

We have as before that $\mu(I)^{\frac{1}{2}} \langle \Delta_R f \rangle_I = \langle f, \psi_I \rangle$.

On the other hand, now let R' be the union of cubes $I \in \mathrm{ch}(R)$ such that $\mu(I) \neq 0$. Since $\mu(R) = \sum_{\substack{I \in \mathrm{ch}(R) \\ \mu(I) \neq 0}} \mu(I)$, we have

$$\begin{aligned} \sum_{\substack{I \in \mathrm{ch}(R) \\ \mu(I) \neq 0}} \mu(I)^{\frac{1}{2}} \psi_{I,J_p}^{\mathrm{full}} &= \sum_{\substack{I \in \mathrm{ch}(R) \\ \mu(I) \neq 0}} \left(\chi_I(c_{J_p}) - \frac{\mu(I)}{\mu(R)} \chi_R(c_{J_p}) \right) \chi_Q \\ &= (\chi_{R'}(c_{J_p}) - \chi_R(c_{J_p})) \chi_Q = -\chi_{R \setminus R'}(c_{J_p}) \chi_Q. \end{aligned}$$

With this,

$$\hat{\Delta}_R(f) = \sum_{I \in \mathrm{ch}(I_p)} \langle f, \psi_I \rangle \psi_{I,J_p}^{\mathrm{full}} - \langle f \rangle_R \chi_{R \setminus R'}(c_{J_p}) \chi_Q.$$

If $\chi_{R\backslash R'}(c_{J_p}) \neq 0$, then $c(J_p) \in R\backslash R'$, which implies $R\cap J_p \neq \emptyset$. Since $\ell(J_p) < \ell(R)$, we have $J_p \subsetneq R$ and so $J_p \subset I$ for some $I \in \mathrm{ch}\,(R)$. Since $\mu(J_p) \neq 0$, we deduce $\mu(I) \neq 0$, which implies $J_p \subset I \subset R'$. But this is contradictory with $c(J_p) \in R \setminus R'$ and so $\chi_{R\backslash R'}(c_{J_p}) = 0$. □

6.3. Orthogonality and Bessel inequality of the Haar wavelet systems. The following lemma summarizes the orthogonality properties of the Haar wavelets.

Lemma 6.8. *Let $I, J \in \mathcal{D}$ or $I, J \in \tilde{\mathcal{D}}$. Then $\int \psi_I(x)\, d\mu(x) = 0$. If $\mu(I) = 0$, then $\langle \psi_I, \psi_J \rangle = 0$, while if $\mu(I) \neq 0$, then*

$$\langle \psi_I, \psi_J \rangle = \delta(I_p, J_p)\mu(I)^{\frac{1}{2}}\mu(J)^{\frac{1}{2}}\left(\frac{\delta(I,J)}{\mu(I)} - \frac{1}{\mu(I_p)}\right), \tag{25}$$

where we denote $\delta(I,J) = 1$ if $I = J$ and zero otherwise. In addition, if $\mu(I) \neq 0$, we have $\|\psi_I\|_{L^q(\mu)} \lesssim \mu(I)^{-\frac{1}{2}+\frac{1}{q}}$.

Proof: The first equality is trivial. Equality (25) is also trivial when $I_p \cap J_p = \emptyset$. When $I_p \subsetneq J_p$, ψ_J is constant on the support of ψ_I and so the dual pair is zero due to the mean zero of ψ_I. Symmetrically, we have the same result when $J_p \subsetneq I_p$.

On the other hand, for $J_p = I_p$, we have

$$\begin{aligned}\langle \psi_I, \psi_J \rangle &= \mu(I)^{\frac{1}{2}}\mu(J)^{\frac{1}{2}} \int \left(\frac{\chi_I(x)}{\mu(I)} - \frac{\chi_{I_p}(x)}{\mu(I_p)}\right)\left(\frac{\chi_J(x)}{\mu(J)} - \frac{\chi_{J_p}(x)}{\mu(J_p)}\right) d\mu(x)\\ &= \mu(I)^{\frac{1}{2}}\mu(J)^{\frac{1}{2}} \frac{1}{\mu(I)}\left(\frac{\mu(I\cap J)}{\mu(J)} - \frac{\mu(I)}{\mu(I_p)}\right).\end{aligned}$$

For $I \neq J$, since $I \cap J = \emptyset$, we have

$$\langle \psi_I, \psi_J \rangle = -\frac{\mu(I)^{\frac{1}{2}}\mu(J)^{\frac{1}{2}}}{\mu(I_p)},$$

while for $I = J$, we get

$$\langle \psi_I, \psi_J \rangle \mu(I)\left(\frac{1}{\mu(I)} - \frac{1}{\mu(I_p)}\right).$$

On the other hand, for $\mu(I) \neq 0$,

$$\begin{aligned}\|\psi_I\|_{L^q(\mu)} &\leq \mu(I)^{\frac{1}{2}}\left(\frac{1}{\mu(I)}\|\chi_I\|_{L^q(\mu)} + \frac{1}{\mu(I_p)}\|\chi_{I_p}\|_{L^q(\mu)}\right)\\ &= \mu(I)^{\frac{1}{2}}\left(\frac{1}{\mu(I)^{\frac{1}{q'}}} + \frac{1}{\mu(I_p)^{\frac{1}{q'}}}\right) \leq 2\mu(I)^{-\frac{1}{2}+\frac{1}{q}}\end{aligned}$$

since $\mu(I) \leq \mu(I_p)$. □

Although these Haar wavelets do not constitute an orthogonal system of functions, they still satisfy Parseval's identity, as we see in the next lemma.

Lemma 6.9. *For $f \in L^2(\mu)$ supported on $Q \in \mathcal{D}$, we have*

$$\|f\|_{L^2(\mu)}^2 = \sum_{I \in \mathcal{D}(Q)} |\langle f, \psi_I \rangle|^2 + \|E_Q f\|_{L^2(\mu)}^2.$$

Proof: We have from Lemma 6.4

$$\begin{aligned}\|f\|^2_{L^2(\mu)} &= \int_{\mathbb{R}^n} f(x)f(x)\,d\mu(x)\\ &= \lim_{M\to\infty}\int_Q f(x)\Bigg(\sum_{\substack{I\in\mathcal{D}(Q)\\ 2^{-M}\le\ell(I)}}\langle f,\psi_I\rangle\psi_I(x)+E_Qf(x)\Bigg)\,d\mu(x)\\ &= \lim_{M\to\infty}\sum_{\substack{I\in\mathcal{D}(Q)\\ 2^{-M}\le\ell(I)}}\langle f,\psi_I\rangle\langle f,\psi_I\rangle+\langle f,\langle f\rangle_Q\chi_Q\rangle\\ &= \sum_{I\in\mathcal{D}(Q)}|\langle f,\psi_I\rangle|^2+\langle f\rangle_Q^2\mu(Q). \qquad\square\end{aligned}$$

Corollary 6.10. *For $f\in L^2(\mu)$ supported on $Q\in\mathcal{D}$, we have*

$$\|f-E_Qf\|^2_{L^2(\mu)}=\sum_{I\in\mathcal{D}(Q)}|\langle f,\psi_I\rangle|^2.$$

With this, $\|f\|^2_{L^2(\mu)}=\|E_Qf\|^2_{L^2(\mu)}+\|f-E_Qf\|^2_{L^2(\mu)}$.

Proof: By Lemma 6.9,

$$\|f-E_Qf\|^2_{L^2(\mu)}=\sum_{I\in\mathcal{D}(Q)}|\langle f-E_Qf,\psi_I\rangle|^2+\langle f-E_Qf\rangle_Q^2\mu(Q)=\sum_{I\in\mathcal{D}(Q)}|\langle f,\psi_I\rangle|^2,$$

since $I\subset Q$ implies $\langle E_Qf,\psi_I\rangle=0$, and $f-E_Qf$ has zero mean. $\square$

6.4. Characterization of compactness. In this subsection, we explain how to use the Haar wavelets to characterize compactness on $L^2(\mu)$ of Calderón–Zygmund operators.

Definition 6.11. Let $(\psi_I)_{I\in\mathcal{D}}$ be a Haar wavelet system of $L^2(\mu)$. For every $M\in\mathbb{N}$ and $Q\in\mathcal{C}$ we define the lagom projection operator by

$$P_Mf=\sum_{I\in\mathcal{D}_M(Q)}\langle f,\psi_I\rangle\psi_I,$$

where $\langle f,\psi_I\rangle=\int_{\mathbb{R}^n}f(x)\psi_I(x)\,d\mu(x)$. We omit from the notation the dependence of the operator with respect to Q. We also define $P_M^\perp f=f-P_Mf$. We note that $P_M^*f=P_Mf$.

Remark 6.12. When we deal with boundedness, we can consider $M=0$ and so $P_Mf=0$ and $P_M^\perp f=f$.

Lemma 6.13. *For $f\in L^2(\mu)$ supported on $Q\in\mathcal{C}$,*

$$\|P_M^\perp f-E_Qf\|^2_{L^2(\mu)}=\sum_{I\in\mathcal{D}_M^c(Q)}|\langle f,\psi_I\rangle|^2.$$

Proof: By Parseval's identity as in Corollary 6.10 we have

$$\|P_M^\perp f-E_Qf\|^2_{L^2(\mu)}=\|f-P_Mf-E_Qf\|^2_{L^2(\mu)}=\sum_{\substack{I\in\mathcal{D}(Q)\\ \mu(I)\neq 0}}|\langle f-P_Mf,\psi_I\rangle|^2. \tag{26}$$

Now, for $I \in \mathcal{D}$ with $\mu(I) \neq 0$,

$$\langle P_M f, \psi_I \rangle = \sum_{J \in \mathcal{D}_M(Q)} \langle f, \psi_J \rangle \langle \psi_J, \psi_I \rangle = \sum_{J \in \mathrm{ch}(I)} \langle f, \psi_J \rangle \langle \psi_J, \psi_I \rangle, \tag{27}$$

since from (25), $\langle \psi_J, \psi_I \rangle = 0$ if $I_p \neq J_p$.

With this, we reason by considering two cases.

(a) If $I \in \mathcal{D}^c_M(Q)$, we have $I_p \neq J_p$ for all $J \in \mathcal{D}_M(Q)$. Then, from (27), $\langle P_M f, \psi_I \rangle = 0$ and thus $\langle f - P_M f, \psi_I \rangle = \langle f, \psi_I \rangle$.

(b) If $I \in \mathcal{D}_M(Q)$, by using again (25), we have

$$\begin{aligned}\langle P_M f, \psi_I \rangle &= \sum_{J \in \mathrm{ch}(I_p)} \langle f, \psi_J \rangle \mu(I)^{\frac{1}{2}} \mu(J)^{\frac{1}{2}} \left(\frac{\delta(I,J)}{\mu(I)} - \frac{1}{\mu(I_p)} \right) \\ &= \langle f, \psi_I \rangle - \frac{\mu(I)^{\frac{1}{2}}}{\mu(I_p)} \sum_{J \in \mathrm{ch}(I_p)} \langle f, \psi_J \rangle \mu(J)^{\frac{1}{2}}.\end{aligned} \tag{28}$$

By (23), we know $\sum_{J \in \mathrm{ch}(I_p)} \mu(J)^{\frac{1}{2}} \psi_J = 0$ a.e. with respect to μ. Then

$$\sum_{J \in \mathrm{ch}(I_p)} \langle f, \psi_J \rangle \mu(J)^{\frac{1}{2}} = \int f(x) \sum_{J \in \mathrm{ch}(I_p)} \mu(J)^{\frac{1}{2}} \psi_J(x)\, d\mu(x) = 0.$$

Then from (28) we get $\langle P_M f, \psi_I \rangle = \langle f, \psi_I \rangle$ and thus $\langle f - P_M f, \psi_I \rangle = 0$.

With both results, we end the equality in (26) as follows:

$$\|P^\perp_M f - E_Q f\|^2_{L^2(\mu)} = \sum_{I \in \mathcal{D}^c_M(Q)} |\langle f, \psi_I \rangle|^2. \qquad \square$$

Remark 6.14. Previous work also shows that

$$\|P_M f\|^2_{L^2(\mu)} = \sum_{I \in \mathcal{D}_M(Q)} |\langle f, \psi_I \rangle|^2 \leq \|f\|_{L^2(\mu)}$$

and so $\|P_M\|_{L^2(\mu) \to L^2(\mu)} \leq 1$ and $\|P^\perp_M\|_{L^2(\mu) \to L^2(\mu)} \leq 1$.

Corollary 6.15. *Let $f \in L^2(\mu)$ supported on $Q \in \mathcal{C}$. Then*

$$\lim_{M \to \infty} \|P^\perp_M f - E_Q f\|_{L^2(\mu)} = 0. \tag{29}$$

Proof: By Lemma 6.9 we have $\sum_{I \in \mathcal{D}(Q)} |\langle f, \psi_I \rangle|^2 \leq \|f\|^2_{L^2(\mu)} < \infty$. Then, by Lemma 6.13,

$$\lim_{M \to \infty} \|P^\perp_M f - E_Q f\|^2_{L^2(\mu)} = \lim_{M \to \infty} \sum_{I \in \mathcal{D}^c_M(Q)} |\langle f, \psi_I \rangle|^2 = 0. \qquad \square$$

The following result, which is implicitly obtained in Lemma 6.13, will be used later during the work with paraproducts.

Lemma 6.16. *Let $f \in L^2(\mu)$ be supported on $Q \in \mathcal{C}$. Then for $I \in \mathcal{D}^c_M(Q)$*

$$\langle P^\perp_M f, \psi_I \rangle = \langle f, \psi_I \rangle,$$

while for $I \in \mathcal{D}_M(Q)$

$$\langle P^\perp_M f, \psi_I \rangle = 0.$$

Proof: By definition of $P_M^\perp$

$$\langle P_M^\perp f, \psi_I\rangle = \langle f, \psi_I\rangle - \langle P_M f, \psi_I\rangle,$$

while by definition of P_M and the orthogonality Lemma 6.8,

$$\langle P_M f, \psi_I\rangle = \sum_{J\in\mathcal{D}_M(Q)} \langle f, \psi_J\rangle\langle \psi_J, \psi_I\rangle = \sum_{\substack{J\in\mathcal{D}_M(Q)\\ J_p=I_p}} \langle f, \psi_J\rangle\langle \psi_J, \psi_I\rangle.$$

If $I \in \mathcal{D}_M(Q)^c$, there is no $J \in \mathcal{D}_M(Q)$ such that $J_p = I_p$ and so the last expression is zero. Then $\langle P_M^\perp f, \psi_I\rangle = \langle f, \psi_I\rangle$.

On the other hand, if $I \in \mathcal{D}_M(Q)$, we have by Lemma 6.8

$$\begin{aligned}\langle P_M f, \psi_I\rangle &= \langle f, \psi_I\rangle \mu(I)\left(\frac{1}{\mu(I)} - \frac{1}{\mu(I_p)}\right) - \sum_{\substack{J\in\mathcal{D}_M(Q)\\ J_p=I_p\\ J\neq I}} \langle f, \psi_J\rangle \frac{\mu(I)^{\frac12}\mu(J)^{\frac12}}{\mu(I_p)}\\ &= \langle f, \psi_I\rangle - \sum_{\substack{J\in\mathcal{D}_M(Q)\\ J_p=I_p}} \langle f, \psi_J\rangle \frac{\mu(I)^{\frac12}\mu(J)^{\frac12}}{\mu(I_p)}.\end{aligned}$$

Then, by the definition of $P_M^\perp$ first and ψ_J later, we have

$$\begin{aligned}\langle P_M^\perp f, \psi_I\rangle &= \sum_{\substack{J\in\mathcal{D}_M(Q)\\ J_p=I_p}} \langle f, \psi_J\rangle \frac{\mu(I)^{\frac12}\mu(J)^{\frac12}}{\mu(I_p)}\\ &= \frac{\mu(I)^{\frac12}}{\mu(I_p)} \int f(x) \sum_{\substack{J\in\mathcal{D}_M(Q)\\ J_p=I_p}} \mu(J)\left(\frac{\chi_J(x)}{\mu(J)} - \frac{\chi_{I_p}(x)}{\mu(I_p)}\right) d\mu(x)\\ &= \frac{\mu(I)^{\frac12}}{\mu(I_p)} \left(\sum_{\substack{J\in\mathcal{D}_M(Q)\\ J_p=I_p}} \int_J f(x)\, d\mu(x) - \int_{I_p} f(x)\, d\mu(x) \sum_{\substack{J\in\mathcal{D}_M(Q)\\ J_p=I_p}} \frac{\mu(J)}{\mu(I_p)}\right)\\ &= \frac{\mu(I)^{\frac12}}{\mu(I_p)} \left(\int_{I_p} f(x)\, d\mu(x) - \int_{I_p} f(x)\, d\mu(x)\right) = 0.\end{aligned}$$

□

6.5. Representation of operators. We now show how to decompose bounded operators in terms of wavelet systems.

Lemma 6.17. *Let T be a bounded operator on $L^2(\mu)$. Let $(\psi_I)_{I\in\mathcal{D}}$ be the Haar wavelet system. Then*

$$\langle Tf, g\rangle = \sum_{I,J\in\mathcal{D}(Q)} \langle f, \psi_I\rangle\langle g, \psi_J\rangle\langle T\psi_I, \psi_J\rangle + \langle T(E_Q f), g\rangle + \langle T(f - E_Q f), E_Q g\rangle \tag{30}$$

for all f, g bounded and compactly supported on $Q \in \mathcal{D}$.

Remark 6.18. A symmetric way to write the previous equality is

$$\langle Tf - E_Q f, g - E_Q g\rangle = \sum_{I,J\in\mathcal{D}(Q)} \langle f,\psi_I\rangle\langle g,\psi_J\rangle\langle T\psi_I,\psi_J\rangle.$$

Proof: Let P_M be the lagom projection related to the Haar wavelet frame. Since T is bounded, we have

$$\begin{aligned}
&|\langle Tf,g\rangle - \langle TP_M f, P_M g\rangle - \langle T(E_Q f), g\rangle - \langle Tf, E_Q g\rangle + \langle TE_Q f, E_Q g\rangle|\\
&\quad = |\langle T(f - P_M f - E_Q f), g\rangle + \langle TP_M f, g - P_M g - E_Q g\rangle\\
&\qquad\qquad - \langle T(f - P_M f - E_Q f), E_Q g\rangle|\\
&\quad \le \|T\|\|P_M^\perp f - E_Q f\|_{L^2(\mu)}\|g\|_{L^2(\mu)} + \|T\|\|P_M f\|_{L^2(\mu)}\|P_M^\perp g - E_Q g\|_{L^2(\mu)}\\
&\qquad\qquad + \|T\|\|P_M^\perp f - E_Q f\|_{L^2(\mu)}\|E_Q g\|_{L^2(\mu)}.
\end{aligned}$$

Now, since by (29) we have that $\|P_M^\perp f - E_Q f\|_{L^2(\mu)}$ and $\|P_M^\perp g - E_Q g\|_{L^2(\mu)}$ tend to zero, so does the left-hand side of the previous chain of inequalities. □

As explained in [**25**], to prove the compactness of an operator on $L^2(\mu)$ it suffices to show that $\langle TP_M^\perp f, P_M^\perp g\rangle$ tends to zero when M tends to infinity uniformly for all functions f, g in the unit ball of $L^2(\mu)$. For that, we need a representation of this dual pair.

Corollary 6.19. *With the same hypotheses of Lemma 6.17, let P_M be the lagom projection related to the Haar wavelet frame. Then*

$$\begin{aligned}
\langle P_M^\perp TP_M^\perp f, g\rangle = &\sum_{I,J\in\mathcal{D}_M^c(Q)} \langle f,\psi_I\rangle\langle g,\psi_J\rangle\langle T\psi_I,\psi_J\rangle\\
&+ \langle T(E_Q f), P_M^\perp g\rangle + \langle T(P_M^\perp f - E_Q f), E_Q g\rangle
\end{aligned}$$

for all f, g bounded and compactly supported on $Q\in\mathcal{D}$.

Proof: We can write

$$\langle P_M^\perp TP_M^\perp f, g\rangle = \langle TP_M^\perp f, P_M^\perp g\rangle = \langle Tf,g\rangle - \langle Tf, P_M g\rangle - \langle TP_M f, g\rangle + \langle TP_M f, P_M g\rangle.$$

By (30) and the facts that $E_Q(P_M f) = 0$ and $E_Q(E_Q f) = E_Q f$, we have

$$\begin{aligned}
\langle P_M^\perp TP_M^\perp f, g\rangle = &\Bigg(\sum_{I,J\in\mathcal{D}(Q)} - \sum_{\substack{I\in\mathcal{D}(Q)\\ J\in\mathcal{D}_M(Q)}} - \sum_{\substack{I\in\mathcal{D}_M(Q)\\ J\in\mathcal{D}(Q)}} + \sum_{I,J\in\mathcal{D}_M(Q)}\Bigg)\langle f,\psi_I\rangle\langle g,\psi_J\rangle\langle T\psi_I,\psi_J\rangle\\
&+ \langle T(E_Q f), g\rangle + \langle T(f - E_Q f), E_Q g\rangle - \langle T(E_Q f), P_M g\rangle - \langle TP_M f, E_Q g\rangle\\
= &\Bigg(\sum_{\substack{I\in\mathcal{D}(Q)\\ J\in\mathcal{D}_M^c(Q)}} - \sum_{\substack{I\in\mathcal{D}_M(Q)\\ J\in\mathcal{D}_M^c(Q)}}\Bigg)\langle f,\psi_I\rangle\langle g,\psi_J\rangle\langle T\psi_I,\psi_J\rangle\\
&+ \langle T(E_Q f), g - P_M g\rangle - \langle T(f - E_Q f - P_M f), E_Q g\rangle\\
= &\sum_{\substack{I\in\mathcal{D}_M^c(Q)\\ J\in\mathcal{D}_M^c(Q)}} \langle f,\psi_I\rangle\langle g,\psi_J\rangle\langle T\psi_I,\psi_J\rangle\\
&+ \langle T(E_Q f), P_M^\perp g\rangle + \langle T(P_M^\perp f - E_Q f), E_Q g\rangle.
\end{aligned}$$

□

7. Technical results on the functions F and F_μ

We prove two technical results on the properties of F, defined in (4) of Definition 3.3, and F_μ, defined in (14), when acting over cubes of $\mathcal{D}_M^c$.

7.1. An extra property of F. Recall the auxiliary functions L, S, D, and F provided in Definition 3.3 and Notation 3.5.

In [**25**], it was proved that the smoothness condition (4) implies the modified smoothness condition (31), which we will often use:

$$|K(t,x)-K(t',x')| \lesssim \frac{(|t-t'|+|x-x'|)^\delta}{|t-x|^\delta}\frac{F(t,x,t',x')}{|t-x|^\alpha}, \tag{31}$$

whenever $2(|t-t'|+|x-x'|)<|t-x|<|t'-x'|$, with $0<\delta<1$ and

$$F(t,x,t',x')=L_1(|t-x'|)S_1(|t-t'|+|x-x'|)D_1\left(1+\frac{|t+x'|}{1+|t-x'|}\right),$$

where L_1, S_1, D_1 satisfy the limits in (5).

7.2. A technical lemma on F. Now we state and prove the mentioned technical results.

Lemma 7.1. *Let $I_p, J_p \in \mathcal{D}$ such that $\ell(J_p)\leq \ell(I_p)$ and $\operatorname{dist}(I_p,J_p)\geq \ell(J_p)$. Let $t\in I_p$, $x\in J_p$, $c_{J_p}=c(J_p)$, and*

$$F(t,x)=L(|t-c_{J_p}|)S(|x-c_{J_p}|)D\left(1+\frac{|t+c_{J_p}|}{1+|t-c_{J_p}|}\right).$$

Then

$$F(t,x)\leq L(\ell([I_p,J_p]))S(\ell(J_p))D(\operatorname{rdist}(\langle I_p,J_p\rangle,\mathbb{B})).$$

Proof: Since L is non-increasing, S is non-decreasing, $|t-c_{J_p}|>\operatorname{dist}(I_p,J_p)=\ell([I_p,J_p])$ and $|x-c_{J_p}|\leq \ell(J_p)/2$, we get

$$F(t,x)\leq L(\ell([I_p,J_p]))S(\ell(J_p))D\left(1+\frac{|t+c_{J_p}|}{1+|t-c_{J_p}|}\right).$$

From $t\in I_p$, $c_{J_p}\in J_p$, and $I_p\cap J_p=\emptyset$, we get $|t-c_{J_p}|\leq \operatorname{dist}(I_p,J_p)+\ell(I_p)+\ell(J_p)\leq 2\ell(\langle I_p,J_p\rangle)$. Then, since $|t+c_{J_p}|\geq 2|c_{J_p}|-|t-c_{J_p}|$, we have

$$\begin{aligned}2\left(1+\frac{|t+c_{J_p}|}{1+|t-c_{J_p}|}\right)&\geq 2+\frac{|t+c_{J_p}|}{1+|t-c_{J_p}|}\\&\geq 2+\frac{2|c_{J_p}|}{1+|t-c_{J_p}|}-\frac{|t-c_{J_p}|}{1+|t-c_{J_p}|}\\&\geq 1+\frac{|c_{J_p}|}{1+\ell(\langle I_p,J_p\rangle)}.\end{aligned}$$

Now we bound below the numerator in the last expression as follows: since $|c(I_p)|-|c(J_p)|\leq |c(I_p)-c(J_p)|\leq \ell(\langle I_p,J_p\rangle)$, we have

$$\begin{aligned}1+\ell(\langle I_p,J_p\rangle)+|c_{J_p}|&\geq 1+\frac{\ell(\langle I_p,J_p\rangle)}{2}+\frac{|c(I_p)|-|c(J_p)|}{2}+|c(J_p)|\\&\geq \frac{1}{2}\left(1+\ell(\langle I_p,J_p\rangle)+\frac{1}{2}|c(I_p)+c(J_p)|\right).\end{aligned}$$

Then

$$|c_{J_p}| \geq \frac{1}{4}|c(I_p)+c(J_p)| - \frac{1}{2}(1+\ell(\langle I_p,J_p\rangle))$$

and so,

$$\begin{aligned}
1+\frac{|c_{J_p}|}{1+\ell(\langle I_p,J_p\rangle)} &\geq 1+\frac{1}{2}\frac{|c(I_p)+c(J_p)|/2}{1+\ell(\langle I_p,J_p\rangle)}-\frac{1}{2}\\
&= \frac{1}{2}\left(1+\frac{|c(I_p)+c(J_p)|/2}{1+\ell(\langle I_p,J_p\rangle)}\right)\\
&\geq \frac{1}{3}\left(\frac{3}{2}+\frac{|c(I_p)+c(J_p)|/2}{1+\ell(\langle I_p,J_p\rangle)}\right).
\end{aligned}$$

Now, since $(c(I_p)+c(J_p))/2 \in \langle I_p,J_p\rangle$, we have $|(c(I_p)+c(J_p))/2 - c(\langle I_p,J_p\rangle)| \leq \ell(\langle I_p,J_p\rangle)/2$ and so we can bound below the previous expression by

$$\begin{aligned}
\frac{1}{3}\left(\frac{3}{2}+\frac{|c(\langle I_p,J_p\rangle)|}{1+\ell(\langle I_p,J_p\rangle)}-\frac{1}{2}\right) &\geq \frac{1}{3}\left(1+\frac{|c(\langle I_p,J_p\rangle)|}{2\max(\ell(\langle I_p,J_p\rangle),1)}\right)\\
&= \frac{1}{6}\left(2+\frac{|c(\langle I_p,J_p\rangle)|}{\max(\ell(\langle I_p,J_p\rangle),1)}\right)\\
&\gtrsim 1+\frac{|c(\langle I_p,J_p\rangle)|+\max(\ell(\langle I_p,J_p\rangle),1)}{\max(\ell(\langle I_p,J_p\rangle),1)}\\
&\geq 1+\frac{\operatorname{dist}(\langle I_p,J_p\rangle,\mathbb{B})}{\max(\ell(\langle I_p,J_p\rangle),1)} = \operatorname{rdist}(\langle I_p,J_p\rangle,\mathbb{B}),
\end{aligned}$$

with $\mathbb{B}=[-1/2,1/2]^n$.

Finally, by using that D is non-increasing, we get

$$F(t,x) \leq L(\ell([I_p,J_p]))S(\ell(J_p))D(\operatorname{rdist}(\langle I_p,J_p\rangle,\mathbb{B})). \qquad \square$$

7.3. A technical lemma about F_μ. Recall the definition given in (14):

$$F_\mu(I,J)=\sup_{\substack{R\subset I\\ S\subset J}} F_K(R,S)\rho_\mu(R\vee S)+F_T(I)\delta(I,J),$$

and $F_\mu(I)=F_\mu(I,I)$, where ρ_μ is defined in (3), F_K is defined in (13), F_T is given in (11), and δ is Kronecker's delta.

Lemma 7.2. *By* (5)*, given $\epsilon>0$, we can take $M>0$ so that*

(i) *if $\ell(I)>2^M$, then $L(2^M)\rho_\mu(I)<\epsilon$,*

(ii) *if $\ell(I)<2^{-M}$, then $S(2^{-M})\rho_\mu(I)<\epsilon$,*

(iii) *if $\operatorname{rdist}(I,\mathbb{B}_{2^M})>M^{\frac14}$, then $\tilde{D}(M^{\frac14})\rho_\mu(I)<\epsilon$, and*

(iv) *if $I\in\mathcal{D}_M^c$, then $F_T(I)<\epsilon$.*

Then for all $I\in\mathcal{D}_{2M}^c$ and $J\in\mathcal{D}_M^c$ we have that either $F_\mu(I,J)\lesssim\epsilon$, or $|\log(\operatorname{ec}(I,J))|\gtrsim \log M$, or $\operatorname{rdist}(I,J)\gtrsim M^{\frac18}$.

Proof: We start with $F_T(I)\delta(I,J)$ since the proof is trivial in this case: from $I=J\in\mathcal{D}_{2M}^c\subset\mathcal{D}_M^c$ we have $F_T(I)<\epsilon$ by the choice of M.

We continue with F_K. Since $I\in\mathcal{D}_{2M}^c$, we consider three cases:

(a) When $\ell(I)<2^{-2M}$, we have $\ell(I\wedge J)<2^{-2M}$. Since $J\in\mathcal{D}_M^c$, we distinguish two cases:

(a.1) If $\ell(J) < 2^{-M}$, then we have $\ell(I \vee J) < 2^{-M}$ and so we get $F_\mu(I,J) \lesssim S(\ell(I \wedge J))\rho_\mu(I \vee J) \leq S(2^{-M})\rho_\mu(I \vee J) < \epsilon$.

(a.2) If $\ell(J) \geq 2^{-M}$, then

$$\mathrm{ec}(I,J) = \frac{\ell(I \wedge J)}{\ell(I \vee J)} = \frac{\ell(I)}{\ell(J)} \leq \frac{2^{-2M}}{2^{-M}} = 2^{-M}$$

and thus, $\log \mathrm{ec}(I,J) \leq -M$.

(b) When $\ell(I) > 2^{2M}$, since $J \in \mathcal{D}_M^c$ we distinguish two cases:

(b.1) When $\ell(J) > 2^M$, we get $\ell(I \vee J) \geq \ell(I \wedge J) > 2^M$. Therefore $F_\mu(I,J)\rho_\mu(I \vee J) \lesssim L(\ell(I \wedge J))\rho_\mu(I \vee J) \leq L(2^M)\rho_\mu(I \vee J) < \epsilon$.

(b.2) When $\ell(J) \leq 2^M$, we have that

$$\mathrm{ec}(I,J) = \frac{\ell(I \wedge J)}{\ell(I \vee J)} = \frac{\ell(J)}{\ell(I)} < \frac{2^M}{2^{2M}} = 2^{-M}$$

and thus, $\log \mathrm{ec}(I,J) \leq -M$.

(c) When $2^{-2M} \leq \ell(I) \leq 2^{2M}$ with $\mathrm{rdist}(I, \mathbb{B}_{2^{2M}}) > 2M$, we have $|c(I)| > (2M-1)2^{2M}$. We fix $\alpha = \frac{1}{8}$, $\beta = \gamma = \frac{1}{4}$. We distinguish two more cases:

(c.1) When $\ell(J) > (2M)^\alpha 2^{2M}$, since $\alpha > 0$ we have

$$\mathrm{ec}(I,J) = \frac{\ell(I)}{\ell(J)} < \frac{2^{2M}}{(2M)^\alpha 2^{2M}} \lesssim M^{-\frac{1}{8}},$$

which implies $\log \mathrm{ec}(I,J) \lesssim -\log M$.

(c.2) When $\ell(J) \leq (2M)^\alpha 2^{2M}$, we have $\ell(I \vee J) < (2M)^\alpha 2^{2M}$. Now:

(c.2.1) When $\mathrm{rdist}(\langle I,J\rangle, \mathbb{B}) > (2M)^\beta$, we also have $\mathrm{rdist}(I \vee J, \mathbb{B}) > (2M)^\beta$. Then

$$F_K(I,J)\rho_\mu(I \vee J) \lesssim \tilde{D}(\mathrm{rdist}(\langle I,J\rangle, \mathbb{B}))\rho_\mu(I \vee J) \leq \tilde{D}(M^\beta)\rho_\mu(I \vee J) < \epsilon.$$

(c.2.2) When $\mathrm{rdist}(\langle I,J\rangle, \mathbb{B}) \leq (2M)^\beta$, we get $|c(\langle I,J\rangle)| \leq (2M)^\beta(1 + \ell(\langle I,J\rangle))$. Then, we examine the last two cases:

- When $\ell(\langle I,J\rangle) > (2M)^\gamma 2^{2M}$, we get

$$\mathrm{rdist}(I,J) = \frac{\ell(\langle I,J\rangle)}{\ell(I \vee J)} > \frac{(2M)^\gamma 2^{2M}}{(2M)^\alpha 2^{2M}} \gtrsim M^{\gamma-\alpha} = M^{\frac{1}{8}}.$$

- When $\ell(\langle I,J\rangle) \leq (2M)^\gamma 2^{2M}$, we have instead

$$\begin{aligned}|c(I) - c(J)| &> |c(I)| - |c(\langle I,J\rangle) - c(J)| - |c(\langle I,J\rangle)| \\ &\geq |c(I)| - 2^{-1}\ell(\langle I,J\rangle) - (2M)^\beta(1 + \ell(\langle I,J\rangle)) \\ &\geq (2M-1)2^{2M} - (2M)^\gamma 2^{2M} - (2M)^\beta(1 + (2M)^\gamma 2^{2M}) \\ &\gtrsim (M - M^\gamma - M^{\beta+\gamma})2^{2M} \gtrsim (M - 2M^{\frac{1}{2}})2^{2M} \geq 2^{-1}M2^{2M}\end{aligned}$$

for $M > 2$. Then

$$\mathrm{rdist}(I,J) \geq \frac{|c(I) - c(J)|}{\ell(I \vee J)} \gtrsim \frac{M2^{2M}}{(2M)^\alpha 2^{2M}} \gtrsim M^{1-\alpha} = M^{\frac{7}{8}} > M^{\frac{1}{8}}. \qquad \square$$

Definition 7.3. As shown in the proof, $F_\mu(I,J) \lesssim \epsilon$ holds when either $\ell(I \wedge J) > 2^M$, or $\ell(I \vee J) < 2^{-M}$, or $\mathrm{rdist}(\langle I,J\rangle, \mathbb{B}) > M^{1/8}$. For this reason, we denote by $\mathcal{F}_M$ the family of ordered pairs (I,J) with $I,J \in \mathcal{D}_M^c$ satisfying some of these three inequalities.

8. The operator acting on bump functions

In this section we estimate the dual pair $\langle T\psi_I, \psi_J\rangle$ in terms of the space and frequency location of the argument functions. The computations are carried out in two different propositions.

8.1. The operator acting on bump functions with disjoint supports.

Proposition 8.1. *Let T be a linear operator with compact Calderón–Zygmund kernel K and parameters $0<\delta<1$, $0<\alpha\le n$. Let $\theta\in(0,1)$ and $I,J\in\mathcal{D}$ be such that $\mathrm{dist}(I_p,J_p)>0$ and $\mathrm{ec}(I,J)^\theta(\mathrm{inrdist}(I_p,J_p)-1)>1$. Then*

$$|\langle T\psi_I,\psi_J\rangle| \lesssim \mathrm{inrdist}(I_p,J_p)^{-(\alpha+\delta)}\frac{\mu(I)^{\frac12}\mu(J)^{\frac12}}{\ell(I\wedge J)^\alpha}F_1(I,J),$$

with $F_1(I,J)=L(\ell([I_p,J_p]))S(\ell(I_p\wedge J_p))D(\mathrm{rdist}(\langle I_p,J_p\rangle,\mathbb{B}))$.

Proof: By symmetry we can assume $\ell(J)\le\ell(I)$. Let $e\in\mathbb{N}$ such that $\mathrm{ec}(I,J)^{-1}=\ell(I)/\ell(J)=2^e\ge 1$. Then

$$\frac{\mathrm{dist}(I_p,J_p)}{\ell(J_p)}=\frac{\mathrm{dist}(J_p,\mathfrak{D}_{I_p})}{\ell(J_p)}=\mathrm{inrdist}(I_p,J_p)-1>\mathrm{ec}(I,J)^{-\theta}=2^{e\theta},$$

that is, $\mathrm{dist}(I_p,J_p)>2^{e\theta}\ell(J_p)\ge\ell(J_p)$. We can then use the kernel representation of T and the zero mean of ψ_J to write

$$\langle T\psi_I,\psi_J\rangle=\iint \psi_I(t)\psi_J(x)(K(t,x)-K(t,c_{J_p}))\,d\mu(t)\,d\mu(x),$$

with $c_{J_p}=c(J_p)$. Since $\psi_I=\mu(I)^{\frac12}(\mu(I)^{-1}\chi_I-\mu(I_p)^{-1}\chi_{I_p})$ and similarly for ψ_J, we have

$$\begin{aligned}|\langle T\psi_I,\psi_J\rangle| &\lesssim \mu(I)^{\frac12}\mu(J)^{\frac12}\sum_{R\in\{I,I_p\}}\sum_{S\in\{J,J_p\}}\mu(R)^{-1}\mu(S)^{-1}\\ &\qquad\times\int_S\int_R|K(t,x)-K(t,c_{J_p})|\,d\mu(t)\,d\mu(x).\end{aligned}\tag{32}$$

We fix $R\in\{I,I_p\}$ and $S\in\{J,J_p\}$. In the domain of integration of the double integral we have $t\in R\subset I_p$, $x\in S\subset J_p$, and so

$$2|x-c_{J_p}|\le\ell(J_p)<\mathrm{dist}(I_p,J_p)\le|t-c_{J_p}|.$$

Then, by the smoothness condition of a compact Calderón–Zygmund kernel (31), the double integral in (32) is bounded by

$$\int_S\int_R\frac{|x-c_{J_p}|^\delta}{|t-x|^{\alpha+\delta}}F(t,x)\,d\mu(t)\,d\mu(x),$$

with $F(t,x)=L(|t-c_{J_p}|)S(|x-c_{J_p}|)D\big(1+\frac{|t+c_{J_p}|}{1+|t-c_{J_p}|}\big)$. Now, by Lemma 7.1, the previous expression can be bounded by a constant times

$$\frac{\ell(J_p)^\delta}{\mathrm{dist}(S,R)^{\alpha+\delta}}\mu(R)\mu(S)L(\ell([I_p,J_p]))S(\ell(J))D(\mathrm{rdist}(\langle I_p,J_p\rangle,\mathbb{B})).$$

Since $R\subset I_p$ and $S\subset J_p$, we have $\mathrm{dist}(S,R)\ge\mathrm{dist}(I_p,J_p)$. Furthermore, since $\mathrm{dist}(I_p,J_p)\ge\ell(J_p)$, we have

$$\mathrm{dist}(I_p,J_p)\ge 2^{-1}(\mathrm{dist}(I_p,J_p)+\ell(J_p)).$$

With this and $\ell(J_p)=2\ell(J)$, we can continue the bound in (32) as

$$
\begin{aligned}
|\langle T\psi_I,\psi_J\rangle| &\lesssim \mu(I)^{\frac12}\mu(J)^{\frac12}\sum_{R\in\{I,I_p\}}\sum_{S\in\{J,J_p\}}\frac{\ell(J_p)^\delta}{\operatorname{dist}(I_p,J_p)^{\alpha+\delta}}F_1(I,J)\\
&\lesssim \left(\frac{\ell(J_p)}{\ell(J_p)+\operatorname{dist}(I_p,J_p)}\right)^{\alpha+\delta}\frac{\mu(I)^{\frac12}\mu(J)^{\frac12}}{\ell(J)^\alpha}F_1(I,J). \qquad \Box
\end{aligned}
\tag{33}
$$

Remark 8.2. When $\ell(I_p\vee J_p)\le \operatorname{dist}(I_p,J_p)$ we will use the weaker inequality

$$
|\langle T\psi_I,\psi_J\rangle| \lesssim \operatorname{ec}(I,J)^\delta \operatorname{rdist}(I_p,J_p)^{-(\alpha+\delta)}\frac{\mu(I)^{\frac12}\mu(J)^{\frac12}}{\ell(I\vee J)^\alpha}F_1(I,J),
\tag{34}
$$

which we now justify.

From $\operatorname{dist}(I_p,J_p)\le\operatorname{dist}(I_p,J_p)+\ell(J_p)$, $\ell(J_p)=2\ell(J)$, and (33) we get

$$
|\langle T\psi_I,\psi_J\rangle| \lesssim \mu(I)^{\frac12}\mu(J)^{\frac12}\left(\frac{\ell(J_p)}{\ell(J_p)+\operatorname{dist}(I_p,J_p)}\right)^{\delta}\frac{1}{\operatorname{dist}(I_p,J_p)^\alpha}F_1(I,J).
$$

By assuming $\ell(J)\le\ell(I)$, we have $\ell(I_p)=\ell(I_p\vee J_p)\le\operatorname{dist}(I_p,J_p)$. Then

$$
\operatorname{dist}(I_p,J_p)\ge 2^{-1}(\operatorname{dist}(I_p,J_p)+\ell(I_p)).
$$

With this and $\ell(I_p)=2\ell(I)$, we continue the previous estimate as follows:

$$
\begin{aligned}
|\langle T\psi_I,\psi_J\rangle| &\lesssim \left(\frac{\ell(J_p)}{\ell(J_p)+\operatorname{dist}(I_p,J_p)}\right)^{\delta}\left(\frac{\ell(I_p)}{\ell(I_p)+\operatorname{dist}(I_p,J_p)}\right)^{\alpha}\frac{\mu(I)^{\frac12}\mu(J)^{\frac12}}{\ell(I)^\alpha}F_1(I,J)\\
&= \operatorname{inrdist}(I_p,J_p)^{-\delta}\operatorname{rdist}(I_p,J_p)^{-\alpha}\frac{\mu(I)^{\frac12}\mu(J)^{\frac12}}{\ell(I)^\alpha}F_1(I,J).
\end{aligned}
$$

Finally,

$$
\begin{aligned}
\operatorname{inrdist}(I_p,J_p)^{-\delta} &\lesssim \left(\frac{\ell(J)}{\operatorname{dist}(I_p,J_p)}\right)^{\delta}\lesssim\left(\frac{\ell(J)}{\ell(I)}\right)^{\delta}\left(\frac{\ell(I_p)}{\ell(I_p)+\operatorname{dist}(I_p,J_p)}\right)^{\delta}\\
&= \operatorname{ec}(I,J)^\delta\operatorname{rdist}(I_p,J_p)^{-\delta}.
\end{aligned}
$$

Remark 8.3. We also note that, from $\operatorname{dist}(I_p,J_p)\le\operatorname{dist}(I,J)\le\operatorname{dist}(I_p,J_p)+\ell(I_p)$, we have

$$
\frac13\left(1+\frac{\operatorname{dist}(I,J)}{\ell(I)}\right)\le 1+\frac{\operatorname{dist}(I_p,J_p)}{\ell(I_p)}\le 1+\frac{\operatorname{dist}(I,J)}{\ell(I)},
$$

that is, $\operatorname{rdist}(I_p,J_p)\approx\operatorname{rdist}(I,J)$.

8.2. The operator acting on bump functions with non-disjoint supports. For the next result, recall the following notation introduced in Definition 6.6. For $I_p,J_p\in\mathcal{D}$, $Q\in 3\mathcal{D}$ with $I_p,J_p\subset 3^{-1}Q$, we write

$$
\psi_{I,J}^{\text{full}}(t)=\mu(I)^{\frac12}(\varphi_I(c_{J_p})-\varphi_{I_p}(c_{J_p}))\chi_Q(t),
$$

with $\varphi_I=\frac{1}{\mu(I)}\chi_I$, $c_{J_p}=c(J_p)$.

Proposition 8.4. *Let T be a linear operator with compact Calderón–Zygmund kernel K and parameter $0<\delta<1$. Let $I,J\in\mathcal{D}$ be such that* $\operatorname{dist}(I_p,J_p)=0$ *and* $\operatorname{ec}(I,J)^{\theta}(\operatorname{inrdist}(I_p,J_p)-1)>1$. *Then*

$$|\langle T(\psi_I-\psi_{I,J}^{\rm full}),\psi_J\rangle|\lesssim \operatorname{inrdist}(I_p,J_p)^{-\delta}\sum_{R\in\{I,I_p\}}\Big(\frac{\mu(R\cap J)}{\mu(R)}\Big)^{\frac12}F_{2,\mu}(I,J)$$

$$+\operatorname{inrdist}(I_p,J_p)^{-(\alpha+\delta)}\frac{\mu(I)^{\frac12}\mu(J)^{\frac12}}{\ell(I\wedge J)^{\alpha}}\chi_{I_p\setminus I}(c_{J_p})F_3(I,J),$$

where

$$F_{2,\mu}(I,J)=L(\ell(I\wedge J))S(\ell(I\wedge J))\sum_{k\geq 0}2^{-k\delta}\frac{\mu(2^kK)}{\ell(2^kK)}D(\operatorname{rdist}(2^kK,\mathbb{B}))$$

and

$$F_3(I,J)=L(\ell(I\wedge J))S(\ell(I\wedge J))\sum_{k\geq 0}2^{-k\delta}D(\operatorname{rdist}(2^kK,\mathbb{B})),$$

with $K=\operatorname{inrdist}(I_p,J_p)(I\wedge J)$.

Proof: We assume $\ell(J)\leq\ell(I)$. Let $e\in\mathbb{N}$ such that $2^e=\ell(I)/\ell(J)\geq 1$. Since $\operatorname{ec}(I,J)^{\theta}(\operatorname{inrdist}(I_p,J_p)-1)>1$, we have

$$\frac{\operatorname{dist}(J_p,\mathfrak{D}_{I_p})}{\ell(J_p)}=\operatorname{inrdist}(I_p,J_p)-1>\operatorname{ec}(I,J)^{-\theta}=2^{e\theta}.$$

Then $\operatorname{dist}(J_p,\partial I_p)>2^{e\theta}\ell(J_p)\geq\ell(J_p)$, which together with $\operatorname{dist}(I_p,J_p)=0$ imply $3J_p\subsetneq I_p$ and $\ell(J)\leq\ell(I)/8$. Therefore $3J_p\subseteq I'$ for some $I'\in\operatorname{ch}(I_p)$.

Now we note that

$$\psi_I(t)-\psi_{I,J}^{\rm full}(t)=\mu(I)^{\frac12}[\varphi_I(t)-\varphi_I(c_{J_p})\chi_Q(t)-\varphi_{I_p}(t)+\varphi_{I_p}(c_{J_p})\chi_Q(t)]. \tag{35}$$

Then for $t\in 3J_p\subsetneq I_p$ we have $\varphi_R(t)\chi_{3J_p}(t)=\varphi_R(c_{J_p})\chi_{3J_p}(t)$ for $R\in\{I,I_p\}$ and so $\psi_I(t)-\psi_{I,J}^{\rm full}(t)=0$. With this and (35),

$$\psi_I-\psi_{I,J}^{\rm full}=(\psi_I-\psi_{I,J}^{\rm full})(1-\chi_{3J_p})=\psi_I^{\rm out}.$$

We denote the last expression by $\psi_I^{\rm out}$, which is supported on $(I_p\cup Q)\setminus 3J_p$. Since $\operatorname{dist}((I_p\cup Q)\setminus 3J_p,J_p)\geq\ell(J_p)$, we can apply the reasoning we used in Proposition 8.4 with some variations. We describe the argument again because we aim for slightly different estimates.

Since $J_p\subseteq I'$ for some $I'\in\operatorname{ch}(I_p)$, we have for $t\in I'$ that $\varphi_R(t)=\varphi_R(c_{J_p})$ with $R\in\{I,I_p\}$, and so $\psi_I^{\rm out}(t)\equiv 0$. That is, $\psi_I^{\rm out}(t)\neq 0$ implies $t\in((I_p\cup Q)\setminus I')\cap(3J_p)^c$. Then

$$|t-c(J_p)|\geq\frac{\ell(J_p)}{2}+\operatorname{dist}(I_p\setminus I',J_p)=\frac{\ell(J_p)}{2}+\operatorname{dist}(J_p,\mathfrak{D}_{I_p})\geq\frac12\operatorname{inrdist}(I_p,J_p)\ell(J_p).$$

Now we prove the following inequalities: for $J_p\subset I_p$,

(a) if $J_p\subset I$, then $|\psi_I^{\rm out}|\lesssim\mu(I)^{\frac12}\frac{1}{\mu(I)}\chi_{Q\setminus I}$,

(b) if $J_p\cap I=\emptyset$, then $|\psi_I^{\rm out}|\lesssim\mu(I)^{\frac12}\big(\frac{1}{\mu(I)}\chi_I+\frac{1}{\mu(I_p)}\chi_{Q\setminus I}\big)$.

(a) If $I' = I$, since $J_p \subset I \subset I_p$, we showed $\psi_I^{\rm out}(t) = 0$ for all $t \in I$. Moreover, for $t \in I_p \setminus I$ we have $\varphi_I(t) = 0$ and $\varphi_{I_p}(t) = \varphi_{I_p}(c_{J_p})$. Then from (35),

$$\psi_I^{\rm out}(t) = \mu(I)^{\frac{1}{2}}[-\varphi_I(c_{J_p})\chi_Q(t)] = -\mu(I)^{\frac{1}{2}}\frac{1}{\mu(I)}\chi_{Q\setminus I}(t)$$

since $t \notin I$. Finally, for $t \in Q \setminus I_p$ we have $\varphi_I(t) = \varphi_{I_p}(t) = 0$ and from (35),

$$|\psi_I^{\rm out}(t)| = \mu(I)^{\frac{1}{2}}| - \varphi_I(c_{J_p})\chi_Q(t) + \varphi_{I_p}(c_{J_p})\chi_Q(t)| \leq \mu(I)^{\frac{1}{2}}\frac{2}{\mu(I)}\chi_{Q\setminus I}(t)$$

since $\mu(I) \leq \mu(I_p)$.

(b) On the other hand, if $I' \neq I$, we have $I' \cap I = \emptyset$ and so, since $J_p \subset I' \subset I_p$, for $t \in I$ we get $\varphi_I(c_{J_p}) = 0$ and $\varphi_{I_p}(t) = \varphi_{I_p}(c_{J_p})$. With this,

$$\psi_I^{\rm out}(t) = \mu(I)^{\frac{1}{2}}\varphi_I(t) = \mu(I)^{\frac{1}{2}}\frac{1}{\mu(I)}\chi_I(t).$$

Meanwhile, for $t \in I_p \setminus I$ we have $\varphi_I(t) = \varphi_I(c_{J_p}) = 0$ and $\varphi_{I_p}(t) = \varphi_{I_p}(c_{J_p})$ and so we get from (35)

$$\psi_I^{\rm out}(t) = \mu(I)^{\frac{1}{2}}[-\varphi_{I_p}(t) + \varphi_{I_p}(c_{J_p})\chi_Q(t)] = 0.$$

Finally, for $t \in Q \setminus I_p$ we have $\varphi_I(t) = \varphi_{I_p}(t) = \varphi_I(c_{J_p}) = 0$ and so

$$\psi_I^{\rm out}(t) = \mu(I)^{\frac{1}{2}}\varphi_{I_p}(c_{J_p})\chi_Q(t) \leq \mu(I)^{\frac{1}{2}}\frac{1}{\mu(I_p)}\chi_{Q\setminus I_p}(t).$$

This finishes the proof of the two inequalities. We can write these inequalities in a unified way as follows:

$$\begin{aligned}|\psi_I^{\rm out}(t)| &\lesssim \mu(I)^{\frac{1}{2}}\Big(\frac{1}{\mu(I)}\chi_{Q\setminus I}(t)\chi_I(c_{J_p}) + \frac{1}{\mu(I)}\chi_I(t)\chi_{I_p\setminus I}(c_{J_p})\\ &\qquad\qquad + \frac{1}{\mu(I_p)}\chi_{Q\setminus I}(t)\chi_{I_p\setminus I}(c_{J_p})\Big)\\ &\lesssim \mu(I)^{\frac{1}{2}}\Big(\frac{1}{\mu(I)}\chi_I(c_{J_p}) + \frac{1}{\mu(I_p)}\chi_{I_p}(c_{J_p}) + \frac{1}{\mu(I)}\chi_I(t)\chi_{I_p\setminus I}(c_{J_p})\Big).\end{aligned}\tag{36}$$

Moreover, for $t \in (I_p \cup Q) \setminus 3J_p$ we have $|t - c(J_p)| \geq 3\ell(J_p)/2 > \ell(J_p)$. We then decompose the support of $\psi_I^{\rm out}$ as follows. Let $\Delta_k = \{t \in (I_p \cup Q) \setminus 3J_p : 2^{k-1}\ell(J_p) < |t - c(J_p)| \leq 2^k\ell(J_p)\} \subset (2^{k+1}J_p) \setminus (2^kJ_p)$. Then

$$(I_p \cup Q) \setminus (3J_p) \subset \bigcup_{k=m_0}^{m_1} \Delta_k,$$

with $m_0 = \log \operatorname{inrdist}(I_p, J_p)$ and $m_1 = \log \frac{\ell(I_p)+\ell(Q)}{\ell(J_p)} + 1$. In this way we can write

$$\psi_I^{\rm out} = \sum_{k=m_0}^{m_1} \Phi_k,$$

where $\Phi_k = \psi_I^{\rm out}(\chi_{2^{k+1}J_p} - \chi_{2^kJ_p})$. We note that, since $J_p \subset 3J$, we have $\operatorname{supp}\Phi_k \subseteq \Delta_k \subseteq 2^{k+1}J_p \subset 2^{k+3}J$ and so $\mu(\Delta_k) \leq \mu(2^{k+3}J)$. Moreover, Δ_k is included in the difference of two concentric cubes with side lengths $2^k\ell(J_p)$ and $2^{k+1}\ell(J_p)$. Then, although Δ_k is not a cube, we denote $\ell(\Delta_k) = 2^{k+1}\ell(J_p)$ and $c(\Delta_k) = c(J_p)$.

The plan is now to estimate $|\langle T\Phi_k, \psi_J\rangle|$. Since $\Delta_k \cap J_p = \emptyset$, we use the kernel representation and the zero mean of ψ_J to write

$$|\langle T\Phi_k, \psi_J\rangle| \leq \int_{J_p}\int_{\Delta_k} |\psi_I^{\rm out}(t)||\psi_J(x)||K(t,x) - K(t, c_{J_p})|\, d\mu(t)\, d\mu(x).$$

With (36) and $\psi_J = \mu(J)^{\frac{1}{2}}\big(\frac{1}{\mu(J)}\chi_J - \frac{1}{\mu(J_p)}\chi_{J_p}\big)$, we have

$$\begin{aligned}
|\langle T\Phi_k, \psi_J\rangle| \lesssim \mu(I)^{\frac{1}{2}}\mu(J)^{\frac{1}{2}} &\sum_{R\in\{I,I_p\}}\sum_{S\in\{J,J_p\}} \frac{\chi_R(c_{J_p})}{\mu(R)}\frac{1}{\mu(S)} \\
&\times \int_S\int_{\Delta_k} |K(t,x)-K(t,c_{J_p})|\,d\mu(t)\,d\mu(x) \\
&+ \mu(I)^{\frac{1}{2}}\mu(J)^{\frac{1}{2}}\sum_{S\in\{J,J_p\}} \frac{\chi_{I_p\setminus I}(c_{J_p})}{\mu(I)}\frac{1}{\mu(S)} \\
&\times \int_S\int_{I\cap\Delta_k} |K(t,x)-K(t,c_{J_p})|\,d\mu(t)\,d\mu(x).
\end{aligned} \tag{37}$$

We now estimate the two double integrals on the right-hand side of (37), starting with the first one, which we denote by Int. We fix $R\in\{I,I_p\}$ and $S\in\{J,J_p\}$. For $t\in\Delta_k$ we have $|t-c(J_p)|>2^{k-1}\ell(J_p)\geq \ell(J_p)$. For $x\in S\subset J_p$ we have $|x-c(J_p)|\leq \ell(J_p)/2$. With both things

$$|x-c(J_p)|\leq \ell(J_p)/2\leq 2^{k-1}\ell(J_p)/2<|t-c(J_p)|/2.$$

Then we can use the smoothness property (31), to write

$$\mathrm{Int}\leq \int_S\int_{\Delta_k}\frac{|x-c(J_p)|^\delta}{|t-c(J_p)|^{\alpha+\delta}}F(t,x)\,d\mu(t)\,d\mu(x), \tag{38}$$

with $F(t,x)=L(|t-c(J_p)|)S(|x-c(J_p)|)D\big(1+\frac{|t+c(J_p)|}{1+|t-c(J_p)|}\big)$. Since L is non-increasing, S is non-decreasing, $2^k\ell(J)\geq|t-c(J_p)|>2^{k-1}\ell(J_p)\geq\ell(J_p)=2\ell(J)$ and $|x-c(J_p)|\leq \ell(J_p)/2=\ell(J)$, we have

$$F(t,x)\leq L(\ell(J))S(\ell(J))D\bigg(1+\frac{|t+c(J_p)|}{1+|t-c(J_p)|}\bigg).$$

On the other hand, $|t+c(J_p)|\geq 2|c(J_p)|-|t-c(J_p)|$, which implies

$$2\bigg(1+\frac{|t+c(J_p)|}{1+|t-c(J_p)|}\bigg)\geq 2+\frac{2|c(J_p)|}{1+|t-c(J_p)|}-\frac{|t-c(J_p)|}{1+|t-c(J_p)|}\geq 1+\frac{|c(J_p)|}{1+2^k\ell(J_p)}.$$

Moreover, since $\Delta_k\subset 2^{k+3}J$, $\ell(\Delta_k)=2^{k+2}\ell(J)$, and $c(\Delta_k)=c(J_p)$, we have

$$1+\frac{|c(J_p)|}{1+2^k\ell(J_p)}\gtrsim 1+\frac{|c(\Delta_k)|}{1+\ell(\Delta_k)}\gtrsim \mathrm{rdist}(\Delta_k,\mathbb{B})\geq \mathrm{rdist}(2^{k+3}J,\mathbb{B}),$$

the meaning of $\mathrm{rdist}(\Delta_k,\mathbb{B})$ being clear even though Δ_k is not a cube. Then, since D is non-increasing, we get

$$F(t,x)\leq L(\ell(J))S(\ell(J))D(\mathrm{rdist}(2^kJ,\mathbb{B}))=F(J,J,2^kJ).$$

With this and $\Delta_k\subset 2^{k+3}J$, we continue the bound in (38) as

$$\mathrm{Int}\lesssim\frac{\ell(J)^\delta}{(2^k\ell(J))^{\delta+\alpha}}\mu(2^{k+3}J)\mu(S)F(J,J,2^kJ)\lesssim 2^{-k\delta}\frac{\mu(2^{k+3}J)}{(2^{k+3}\ell(J))^\alpha}\mu(S)F(J,J,2^kJ).$$

For the second double integral on the right-hand side of (37), which we denote by Int', we can apply the same reasoning with the only difference of integrating over $I\cap\Delta_k\subset\Delta_k$ instead of the whole Δ_k. With this we obtain

$$\mathrm{Int}'\lesssim\frac{\ell(J)^\delta}{(2^k\ell(J))^{\delta+\alpha}}\mu(I)\mu(S)F(J,J,2^kJ)\lesssim 2^{-k\delta}\frac{1}{(2^k\ell(J))^\alpha}\mu(I)\mu(S)F(J,J,2^kJ).$$

Then we continue the estimate in (37) as follows: since $\mu(I) \le \mu(R)$,

$$
\begin{aligned}
|\langle T\Phi_k, \psi_J\rangle| &\lesssim 2^{-k\delta} F(J,J,2^kJ)\mu(I)^{\frac12}\mu(J)^{\frac12} \\
&\qquad \times \Bigg(\sum_{\substack{R\in\{I,I_p\}\\ S\in\{J,J_p\}}} \frac{\chi_R(c_{J_p})}{\mu(R)} \frac{\mu(2^{k+3}J)}{(2^{k+3}\ell(J))^\alpha} + \sum_{S\in\{J,J_p\}} \frac{\chi_{I_p\setminus I}(c_{J_p})}{(2^k\ell(J))^\alpha} \Bigg) \\
&\lesssim 2^{-k\delta} F(J,J,2^kJ) \Bigg(\sum_{R\in\{I,I_p\}} \left(\frac{\mu(J)\chi_R(c_{J_p})}{\mu(R)} \right)^{\frac12} \frac{\mu(2^{k+3}J)}{(2^{k+3}\ell(J))^\alpha} \\
&\qquad + \chi_{I_p\setminus I}(c_{J_p}) \frac{\mu(I)^{\frac12}\mu(J)^{\frac12}}{(2^k\ell(J))^\alpha} \Bigg) \\
&\lesssim \sum_{R\in\{I,I_p\}} \left(\frac{\mu(J\cap R)}{\mu(R)} \right)^{\frac12} 2^{-k\delta} F(J,J,2^kJ) \frac{\mu(2^{k+3}J)}{(2^{k+3}\ell(J))^\alpha} \\
&\qquad + \chi_{I_p\setminus I}(c_{J_p}) \frac{\mu(I)^{\frac12}\mu(J)^{\frac12}}{\ell(J)^\alpha} 2^{-k(\alpha+\delta)} F(J,J,2^kJ).
\end{aligned}
$$

The last inequality holds because $\chi_R(c_{J_p}) \neq 0$ if and only if $J_p \cap R \neq \emptyset$. Moreover, since $\ell(J) \le \ell(I) \le \ell(R)$, we have $J \subset J_p \subset R$ and so $J = J \cap R$.

Now, using that $F(J,J,2^kJ) = L(\ell(J))L(\ell(J))D(\mathrm{rdist}(2^kJ,\mathbb{B}))$ and summing in k, we have that $|\langle T\psi_I^{\mathrm{out}}, \psi_J\rangle|$ can be bounded by

$$
\begin{aligned}
&\sum_{R\in\{I,I_p\}} \left(\frac{\mu(J\cap R)}{\mu(R)} \right)^{\frac12} L(\ell(J))S(\ell(J)) \sum_{k\ge m_0} 2^{-k\delta} \frac{\mu(2^{k+3}J)}{(2^{k+3}\ell(J))^\alpha} D(\mathrm{rdist}(2^kJ,\mathbb{B})) \\
&\qquad + \chi_{I_p\setminus I}(c_{J_p}) \frac{\mu(I)^{\frac12}\mu(J)^{\frac12}}{\ell(J)^\alpha} L(\ell(J))S(\ell(J)) \sum_{k\ge m_0} 2^{-k(\alpha+\delta)} D(\mathrm{rdist}(2^kJ,\mathbb{B})).
\end{aligned}
$$

We denote $\lambda = \mathrm{inrdist}(I_p, J_p)$. Since $m_0 = \log \mathrm{inrdist}(I_p, J_p)$, the last factor in each term is respectively bounded by

$$
\begin{aligned}
&2^{-m_0\delta} \sum_{k\ge 0} 2^{-k\delta} \frac{\mu(2^{k+3}2^{m_0}J)}{\ell(2^{k+3}2^{m_0}J)^\alpha} D(\mathrm{rdist}(2^k(2^{m_0}J),\mathbb{B})) \\
&\quad \lesssim \mathrm{inrdist}(I_p,J_p)^{-\delta} \sum_{k\ge 0} 2^{-k\frac{\delta}{2}} D(\mathrm{rdist}(2^k\lambda J,\mathbb{B})) \sum_{k\ge 0} 2^{-k\frac{\delta}{2}} \frac{\mu(2^{k+3}\lambda J)}{(2^{k+3}\ell(\lambda J))^\alpha} \\
&\quad \lesssim \mathrm{inrdist}(I_p,J_p)^{-\delta} \tilde{D}(\mathrm{rdist}(\lambda J,\mathbb{B}))\rho_{\mathrm{out}}(\lambda J),
\end{aligned}
$$

and

$$
\begin{aligned}
&2^{-m_0(\alpha+\delta)} \sum_{k\ge 0} 2^{-k(\alpha+\delta)} D(\mathrm{rdist}(2^k(2^{m_0}J),\mathbb{B})) \\
&\quad \lesssim \mathrm{inrdist}(I_p,J_p)^{-(\alpha+\delta)} \sum_{k\ge 0} 2^{-k(\alpha+\delta)} D(\mathrm{rdist}(2^k\lambda J,\mathbb{B})) \\
&\quad \lesssim \mathrm{inrdist}(I_p,J_p)^{-(\alpha+\delta)} \tilde{D}(\mathrm{rdist}(\lambda J,\mathbb{B})),
\end{aligned}
$$

with $\tilde{D}$ defined in (7). This ends the proof. □

9. Paraproducts

The proof of Theorem 4.2 is divided into two parts: we first deal with the part associated to $\psi_{I,J}^{\text{full}}$, which corresponds to the paraproduct case in the classical proof. Then we use the estimates of the bump lemma for the remaining part.

In this section we cover the paraproduct part, which requires the use of the classical Carleson embedding theorem.

Lemma 9.1 (Carleson embedding theorem)**.** *Let $(a_I)_{I\in\mathcal{D}}$ be a collection of non-negative numbers. Then*

$$\sum_{I\in\mathcal{D}} a_I|\langle f\rangle_I|^2 \lesssim \sup_{I\in\mathcal{D}}\bigg(\frac{1}{\mu(I)}\sum_{J\in\mathcal{D}(I)} a_J\bigg)\|f\|_{L^2(\mu)}^2$$

for all $f\in L^2(\mu)$.

The following proposition deals with the paraproduct part of the operator. The proof provided follows the ideas developed in [**18**].

Proposition 9.2 (Paraproducts)**.** *Let $Q\in\mathcal{D}$ and $\theta\in(0,1)$ be fixed. We define the following bilinear forms: for f, g bounded functions with* $\operatorname{supp} f\cup\operatorname{supp} g\subset Q$*,*

$$\Pi(f,g)=\sum_{I\in\mathcal{D}(Q)}\sum_{\substack{J\in\mathcal{D}(I_p)\\ \operatorname{inrdist}(J_p,I_p)>\lambda_\theta}}\langle f,\psi_I\rangle\langle g,\psi_J\rangle\langle T\psi_{I,J}^{\text{full}},\psi_J\rangle,$$

$$\Pi'(f,g)=\sum_{J\in\mathcal{D}(Q)}\sum_{\substack{I\in\mathcal{D}(J_p)\\ \operatorname{inrdist}(I_p,J_p)>\lambda_\theta}}\langle f,\psi_I\rangle\langle g,\psi_J\rangle\langle T\psi_I,\psi_{J,I}^{\text{full}}\rangle,$$

with $\lambda_\theta=1+\operatorname{ec}(I,J)^{-\theta}$.

Then given $\epsilon>0$ there exist $M_0\in\mathbb{N}$ independent of Q and functions f, g such that for all $M>M_0$

$$|\Pi(P_M^\perp f,P_M^\perp g)|+|\Pi'(P_M^\perp f,P_M^\perp g)|\lesssim\epsilon\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)}.$$

Proof: By symmetry, we only need to work with Π. By Notation 4.3 and the properties (6), (12), given $\epsilon>0$, we can choose $M_0\in\mathbb{N}$ such that for all $M>M_0$ and all $I\in\mathcal{D}_M(Q)^c$ we have

$$F_\mu(I)<\epsilon. \tag{39}$$

By Lemma 6.16 and Fubini's theorem, we can rewrite $\Pi(P_M^\perp f,P_M^\perp g)$ as

$$\begin{aligned}\Pi(P_M^\perp f,P_M^\perp g)&=\sum_{I\in\mathcal{D}(Q)}\sum_{\substack{J\in\mathcal{D}(I_p)\\ \operatorname{inrdist}(J_p,I_p)>\lambda_\theta}}\langle P_M^\perp f,\psi_I\rangle\langle P_M^\perp g,\psi_J\rangle\langle T\psi_{I,J}^{\text{full}},\psi_J\rangle\\ &=\sum_{I\in\mathcal{D}_M^c(Q)}\sum_{\substack{J\in\mathcal{D}_M^c(I_p)\\ \operatorname{inrdist}(J_p,I_p)>\lambda_\theta}}\langle f,\psi_I\rangle\langle g,\psi_J\rangle\langle T\psi_{I,J}^{\text{full}},\psi_J\rangle\\ &=\sum_{J\in\mathcal{D}_M^c(Q)}\langle g,\psi_J\rangle\bigg\langle T\bigg(\sum_{\substack{I\in\mathcal{D}_M^c(Q)\\ J_p\subset I_p,\operatorname{inrdist}(J_p,I_p)>\lambda_\theta}}\langle f,\psi_I\rangle\psi_{I,J}^{\text{full}}\bigg),\psi_J\bigg\rangle.\end{aligned}$$

We can assume that the first sum in the previous expression only contains terms for which $\mu(J_p)\neq 0$ since otherwise $\psi_J\equiv 0$. Moreover, for fixed $J\in\mathcal{D}$, and each cube I satisfying the condition $J_p\subset I_p$, all cubes $I'\in\operatorname{ch}(I_p)$ also satisfy the same condition. In other words, for each cube I in the sum, all its siblings are also in the sum.

Then, since in the second sum we have $\ell(J_p) < \ell(I_p)$, by Lemma 6.7,

$$\sum_{I\in \mathrm{ch}(I_p)} \langle f,\psi_I\rangle \psi_{I,J}^{\mathrm{full}} = \hat{\Delta}_{I_p}(f) = \Bigg(\sum_{I\in \mathrm{ch}(I_p)} \hat{E}_I f\Bigg) - \hat{E}_{I_p} f,$$

where $\hat{E}_I f = \langle f\rangle_I \chi_I(c_{J_p})\chi_Q$.

The inner sum takes place under the condition $\mathrm{inrdist}(J_p, I_p) > \lambda_\theta \geq 2$. For $J_p \in \mathcal{D}_M^c(Q)$, let λ be the smallest integer such that $\mathrm{inrdist}(J_p, I_p) > \lambda$. Also let $I_J \in \mathcal{D}(Q)$ be the smallest cube such that $J_p \subset I_J$ and $\mathrm{inrdist}(J_p, I_J) > \lambda$. If such a cube does not exist, we then define $I_J = \emptyset$.

We now add and subtract the following term:

$$\mathrm{Ad}_1 = \sum_{\substack{I\in \mathcal{D}_M^c(Q)\\ J_p\subset I_p, \lambda(I_p,J_p)\leq \lambda_\theta}} \langle f,\psi_I\rangle \psi_{I,J}^{\mathrm{full}}$$

to obtain

$$\sum_{\substack{I\in \mathcal{D}_M^c(Q)\\ J_p\subset I_p, \lambda(\widehat{I},\widehat{J})>\lambda_\theta}} \langle f,\psi_I\rangle \psi_{I,J}^{\mathrm{full}} = \sum_{\substack{I\in \mathcal{D}_M^c(Q)\\ J_p\subseteq I_p}} \langle f,\psi_I\rangle \psi_{I,J}^{\mathrm{full}} - \mathrm{Ad}_1.$$

The second term $-\mathrm{Ad}_1$, together with a symmetric expression Ad_2 containing cubes such that $I_p \subset J_p$, will be estimated in Subsection 10.6. We now focus on the first term.

Now, for fixed $J \in \mathcal{D}_M^c(Q)$, since all siblings of each cube in the sum are also contained in the sum, we obtain a telescoping sum and so

$$\begin{aligned}
\sum_{\substack{I\in \mathcal{D}_M^c(Q)\\ J_p\subset I_p}} \langle f,\psi_I\rangle \psi_{I,J_p}^{\mathrm{full}} &= \sum_{\substack{I_p\in \mathcal{D}_M^c(Q)\\ J_p\subset I_p}} \sum_{I'\in \mathrm{ch}(I_p)} \langle f,\psi_{I'}\rangle \psi_{I',J_p}^{\mathrm{full}}\\
&= \sum_{\substack{I_p\in \mathcal{D}_M^c(Q)\\ J_p\subset I_p}} \hat{\Delta}_{I_p}(f)\\
&= \sum_{R\in \mathrm{ch}(J_p)} \hat{E}_R f - \hat{E}_Q f\\
&= \sum_{R\in \mathrm{ch}(J_p)\cup\{Q\}} \alpha_R \langle f\rangle_R \chi_Q,
\end{aligned}$$

with $|\alpha_R| = 1$.

We denote $J_0 = Q$. The cardinality of $\mathrm{ch}(J_p)$ is 2^n and so we can enumerate the family in a uniform way in accordance with their position inside J_p: $\mathrm{ch}(J_p) = \{J_j\}_{j=1}^{2^n}$. We then write $\alpha_j = \alpha_{J_j}$. With this,

$$\begin{aligned}
\Pi(P_M^\perp f, P_M^\perp g) &= \sum_{J\in \mathcal{D}_M^c(Q)} \langle g,\psi_J\rangle \Bigg\langle T\Bigg(\sum_{R\in \mathrm{ch}(I_J)} \hat{E}_R f - \hat{E}_Q f\Bigg), \psi_J\Bigg\rangle\\
&= \sum_{j=0}^{2^n} \alpha_j \sum_{J\in \mathcal{D}_M^c(Q)} \langle f\rangle_{J_j} \langle g,\psi_J\rangle \langle T\chi_Q, \psi_J\rangle.
\end{aligned}$$

With this, the boundedness of $\Pi(P_M^\perp f, P_M^\perp g)$ follows once we obtain a uniform estimate for each fixed j:

$$\Pi_j(P_M^\perp f, P_M^\perp g) = \sum_{J\in\mathcal{D}_M^c(Q)} \langle g,\psi_J\rangle\langle f\rangle_{J_j}\langle T\chi_Q,\psi_J\rangle.$$

For $j=0$, we have $I_j=Q$ and so

$$\Pi_0(P_M^\perp f, P_M^\perp g) = \langle f\rangle_Q\langle T\chi_Q, g\rangle,$$

which can be estimated by using the testing condition (11): since f, g are compactly supported on Q,

$$\begin{aligned}|\Pi_0(P_M^\perp f, P_M^\perp g)| &\leq |\langle f\rangle_Q|\|\chi_Q T\chi_Q\|_{L^2(\mu)}\|g\|_{L^2(\mu)}\\ &\lesssim \mu(Q)^{-1}\|f\|_{L^2(\mu)}\mu(Q)^{\frac12}F_T(Q)\mu(Q)^{\frac12}\|g\|_{L^2(\mu)}\\ &\leq F_T(Q)\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)} < \epsilon\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)}.\end{aligned}$$

For any other term, by Cauchy's inequality and Lemma 6.9,

$$\begin{aligned}|\Pi_j(P_M^\perp f, P_M^\perp g)| &\leq \Big(\sum_{J\in\mathcal{D}_M^c(Q)} |\langle f\rangle_{J_j}|^2|\langle T\chi_Q,\psi_J\rangle|^2\Big)^{\frac12}\Big(\sum_{J\in\mathcal{D}_M^c(Q)}|\langle g,\psi_J\rangle|^2\Big)^{\frac12}\\ &\lesssim \Big(\sum_{J\in\mathcal{D}_M^c(Q)} |\langle f\rangle_{J_j}|^2|\langle T\chi_Q,\psi_J\rangle|^2\Big)^{\frac12}\|g\|_{L^2(\mu)}.\end{aligned}$$

Then by Fubini's theorem and Carleson's Lemma 9.1,

$$|\Pi_j(P_M^\perp f, P_M^\perp g)| \lesssim \sup_{R\in\mathcal{D}_M^c(Q)}\Big(\mu(R)^{-1}\sum_{J\in\mathcal{D}_M^c(R)}|\langle T\chi_Q,\psi_J\rangle|^2\Big)^{\frac12}\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)}.$$

We now prove that for all $R\in\mathcal{D}_M^c(Q)$,

$$\sum_{J\in\mathcal{D}_M^c(R)}|\langle T\chi_Q,\psi_J\rangle|^2 \lesssim \epsilon\mu(R).$$

For $R\in\mathcal{D}_M^c(Q)$ with $\ell(R)<2^{-M}$ or $\mathrm{rdist}(R,\mathbb{B}_{2^M})>M$, we construct $\mathcal{W}(R)$ a Whitney decomposition of R defined by the maximal (with respect to the inclusion) cubes $S\in\mathcal{D}(Q)$ such that $5S\subset R$. The cubes S in $\mathcal{W}(R)$ form a partition of R and for any cube $J\in\mathcal{D}_M^c(R)$ there exists $S\in\mathcal{W}(R)$ such that $J_p\subset S$. Then we can write

$$\begin{aligned}\sum_{J\in\mathcal{D}_M^c(R)}|\langle T\chi_Q,\psi_J\rangle|^2 &\leq \sum_{S\in\mathcal{W}(R)}\sum_{J\in\mathcal{D}_M^c(S)}|\langle T\chi_Q,\psi_J\rangle|^2\\ &\lesssim \sum_{S\in\mathcal{W}(R)}\Big(\sum_{J\in\mathcal{D}_M^c(S)}|\langle T\chi_{4S},\psi_J\rangle|^2+\sum_{J\in\mathcal{D}_M^c(S)}|\langle T(\chi_{Q\setminus 4S}),\psi_J\rangle|^2\Big).\end{aligned}\tag{40}$$

We will later deal with each term in different ways.

On the other hand, for $R\in\mathcal{D}_M^c(Q)$ with $\ell(R)>2^M$, we start by defining the same Whitney decomposition $\mathcal{W}(R)$ of R as before. But then we decompose each $S\in\mathcal{W}(R)$ as follows:

$$S = \bigcup_{\substack{\bar S\in\mathcal{D}(S)\\ \ell(\bar S)=2^{-(M+2)}}}\bar S.$$

This ensures that $\bar{S} \in \mathcal{D}_M^c(Q)$. Then, similarly to before, we write

$$\sum_{J \in \mathcal{D}_M^c(R)} |\langle T\chi_Q, \psi_J\rangle|^2 \leq \sum_{S \in \mathcal{W}(R)} \sum_{\substack{\bar{S} \in \mathcal{D}(S) \\ \ell(\bar{S}) = 2^{-(M+2)}}} \sum_{J \in \mathcal{D}_M^c(\bar{S})} |\langle T\chi_Q, \psi_J\rangle|^2$$

$$\lesssim \sum_{S \in \mathcal{W}(R)} \Bigg(\sum_{\substack{\bar{S} \in \mathcal{D}(S) \\ \ell(\bar{S}) = 2^{-(M+2)}}} \sum_{J \in \mathcal{D}_M^c(\bar{S})} |\langle T\chi_{4\bar{S}}, \psi_J\rangle|^2$$

$$+ \sum_{\substack{\bar{S} \in \mathcal{D}(S) \\ \ell(\bar{S}) = 2^{-(M+2)}}} \sum_{J \in \mathcal{D}_M^c(\tilde{S})} |\langle T(\chi_{Q\setminus 4\bar{S}}), \psi_J\rangle|^2 \Bigg).$$

As before, we also deal with each term differently. We will show in detail the case when $\ell(R) < 2^{-M}$ or $\mathrm{rdist}(R, \mathbb{B}_{2^M}) > M$, and only the small differences of the case when $\ell(R) > 2^M$.

In the first case, we start by estimating the inner double sum in the first term of (40) for each $S \in \mathcal{W}(R)$. For this, we use Lemma 6.9 and the testing condition (11):

$$\sum_{J \in \mathcal{D}_M^c(S)} |\langle T\chi_{4S}, \psi_J\rangle|^2 \leq |\langle \chi_{4S} T\chi_{4S}, \psi_J\rangle|^2 \lesssim \|\chi_{4S} T\chi_{4S}\|_{L^2(\mu)}^2 \lesssim \mu(4S) F_\mu(4S)^2.$$

Since $4S \subset R \in \mathcal{D}_M^c(Q)$ with $\ell(R) < 2^{-M}$ or $\mathrm{rdist}(R, \mathbb{B}_{2^M}) > M$, we have $\ell(4S) \leq 2^{-M}$ or $\mathrm{rdist}(4S, \mathbb{B}_{2^M}) > M$. Moreover, since $S \in \mathcal{D}(Q)$, we have that $4S$ is the union of $3^n + 14n$ cubes in $S' \in \mathcal{D}(Q)$ with $\ell(S') \in \{\ell(S), \ell(S)/2\}$. Therefore, by (39), $F_\mu(4S) \lesssim \sup\{F_\mu(K) : K \in \mathcal{D}_M^c\} < \epsilon$. Then, by summing in $S \in \mathcal{W}(R)$, we have

$$\sum_{S \in \mathcal{W}(R)} \sum_{J \in \mathcal{D}_M^c(S)} |\langle T\chi_{4S}, \psi_J\rangle|^2 \lesssim \epsilon^2 \sum_{S \in \mathcal{W}(R)} \mu(4S) \lesssim \epsilon^2 \mu(R).$$

In the last inequality we have used that, since $5S \subset R$ and the cubes S are disjoint by maximality, the cubes $4S$ can only overlap a uniform number of times. We prove this claim.

Since $5S \subset R$ and $6S \not\subset R$, we have $\ell(S)/2 \leq \mathrm{dist}(4S, \partial R) \leq \ell(S)$. Then, for all S such that $x \in 4S$, we have $\mathrm{dist}(x, \partial R) \geq \mathrm{dist}(4S, \partial R) \geq \ell(S)/2$ and $\mathrm{dist}(x, \partial R) \leq 4\ell(S) + \mathrm{dist}(4S, \partial R) \leq 5\ell(S)$.

Now we reason as follows. For each $x \in R$, by disjointness there is at most one cube $S_0 \in W(R)$ such that $x \in S_0 \subset 4S_0$. Then, for any other $S \in W(R)$ such that $x \in 4S$, we have that

- If $\ell(S) \leq \ell(S_0)/32$, from $\mathrm{dist}(x, \partial R) \geq \ell(S_0)/2$ we get $\ell(S) \leq \mathrm{dist}(x, \partial R)/8$. That is, $\mathrm{dist}(x, \partial R) \geq 8\ell(S)$, which implies $x \notin 4S$ since otherwise we proved that $\mathrm{dist}(x, \partial R) \leq 5\ell(S)$.
- If $\ell(S) \geq 32\ell(S_0)$, from $\mathrm{dist}(x, \partial R) \leq 5\ell(S_0)$ we deduce $\ell(S) \geq 32\,\mathrm{dist}(x, \partial R)/5$. That is, $\mathrm{dist}(x, \partial R) \leq \ell(S)/4$, which implies $x \notin 4S$ since otherwise we proved that $\mathrm{dist}(x, \partial R) \geq \ell(S)/2$.

Therefore, there are up to 12 different side lengths of S for which $x \in 4S$. In addition, since the cubes S are disjoint, for each $x \in R$ there are up to 3^n cubes S of a fixed side length such that $x \in 4S$. With this, we get that there are in total up to $12 \cdot 3^n$ different cubes S such that $x \in 4S$. This finishes the work to estimate the first term of (40).

For the second term of (40), we reason as follows. Let $S \in \mathcal{W}(R)$, $J \in \mathcal{D}_M^c(S)$ be fixed. Since $J_p \subset S$ and ψ_J has mean zero, we can write

$$\begin{aligned}|\langle T(\chi_{Q\setminus 4S}), \psi_J\rangle| &= |\langle T(\chi_{Q\setminus 4S}) - T(\chi_{Q\setminus 4S})(c_{J_p}), \psi_J\rangle| \\ &\leq \int_{J_p}\int_{Q\setminus 4S} |K(t,x) - K(t,c_{J_p})||\psi_J(x)|\,d\mu(t)\,d\mu(x).\end{aligned}$$

Since $x \in J_p$, we have $|x - c_{J_p}| \leq \ell(J_p)/2$. And since $t \in Q \setminus 4S$, $x \in J_p \subset S$, we have $|t - x| > \ell(S)$. Then $2|x - c_{J_p}| \leq \ell(J_p) \leq \ell(S) < |t - x|$. Therefore, we can use the smoothness kernel condition (31) to write

$$|\langle T(\chi_{Q\setminus 4S}), \psi_J\rangle| \leq \int_{J_p}\int_{Q\setminus 4S} \frac{|x - c_{J_p}|^\delta}{|t-x|^{\alpha+\delta}} F(t,x)|\psi_J(x)|\,d\mu(t)\,d\mu(x),$$

where $F(t,x) = L(|t-x|)S(|x-c_{J_p}|)D\big(1 + \frac{|t+c_{J_p}|}{1+|t-c_{J_p}|}\big)$. Moreover, $|t - x| \geq |t - c_{J_p}| - |x - c_{J_p}| \geq |t - c_{J_p}| - |t-x|/2$, that is, $|t - x| \geq 2|t - c_{J_p}|/3$. With this, we have

$$\begin{aligned}|\langle T(\chi_{Q\setminus 4S}), \psi_J\rangle| &\lesssim L(\ell(S))S(\ell(S)) \int_{J_p} |\psi_J(x)|\,d\mu(x) \\ &\qquad \times \int_{Q\setminus 4S} \frac{\ell(J)^\delta}{|t - c_{J_p}|^{\alpha+\delta}} D\left(1 + \frac{|t + c_{J_p}|}{1 + |t - c_{J_p}|}\right) d\mu(t) \\ &\lesssim \ell(J)^\delta \mu(J_p)^{\frac{1}{2}} L(\ell(S))S(\ell(S)) \int_{Q\setminus 4S} \frac{D\big(1 + \frac{|t+c_{J_p}|}{1+|t-c_{J_p}|}\big)}{|t - c_{J_p}|^{\alpha+\delta}}\,d\mu(t).\end{aligned}$$

Since $J \subset S$, for $t \in Q \setminus 4S$ we have $\operatorname{dist}(t, J) \geq \ell(S)$. Then we decompose

$$Q \setminus 4S = \bigcup_{i=1}^{\log \frac{\ell(Q)}{\ell(S)}} S_i,$$

where $S_i = \{t \in Q \setminus 4S : 2^{i-1}\ell(S) < |t - c_{J_p}| \leq 2^i \ell(S)\} \subset 2^{i+1}S$. Note that $S_i \subset B(c_i, 2^i\ell(S))$, with $c_i \in S_i$, and $c(S_i) = c_{J_p}$. Moreover, since $|t - c_{J_p}| + |t + c_{J_p}| \geq 2|c_{J_p}|$, we have

$$2\left(1 + \frac{|t + c_{J_p}|}{1 + |t - c_{J_p}|}\right) \geq 1 + \frac{2|c_{J_p}|}{1 + |t - c_{J_p}|} \geq 1 + \frac{|c(2^iS)|}{1 + 2^i\ell(S)}.$$

Then $D\big(1 + \frac{|t+c_{J_p}|}{1+|t-c_{J_p}|}\big) \lesssim D\big(1 + \frac{|c(2^iS)|}{1+2^i\ell(S)}\big) \lesssim D(\operatorname{rdist}(2^iS, \mathbb{B}))$. With this,

$$\begin{aligned}\int_{Q\setminus 4S} \frac{D\big(1 + \frac{|t+c_{J_p}|}{1+|t-c_{J_p}|}\big)}{|t - c_{J_p}|^{\alpha+\delta}}\,d\mu(t) &\lesssim \sum_{i\geq 1} \frac{D(\operatorname{rdist}(2^iS, \mathbb{B}))}{(2^i\ell(S))^{\alpha+\delta}} \mu(S_i) \\ &\lesssim \frac{1}{\ell(S)^\delta} \sum_{i\geq 1} \frac{1}{2^{i\delta}} \frac{\mu(2^{i+1}S)}{(2^{i+1}\ell(S))^\alpha} D(\operatorname{rdist}(2^iS, \mathbb{B})) \\ &\lesssim \frac{1}{\ell(S)^\delta} \tilde{D}(S)\rho_\mu(S).\end{aligned}$$

Then

$$|\langle T(\chi_{Q\setminus 4S}),\psi_J\rangle| \lesssim \left(\frac{\ell(J)}{\ell(S)}\right)^{\delta}\mu(J_p)^{\frac{1}{2}}L(\ell(S))S(\ell(S))\tilde{D}(S)\rho_\mu(S)$$
$$\leq \left(\frac{\ell(J)}{\ell(S)}\right)^{\delta}\mu(J_p)^{\frac{1}{2}}F_\mu(S) \leq \epsilon\left(\frac{\ell(J)}{\ell(S)}\right)^{\delta}\mu(J_p)^{\frac{1}{2}}.$$

The last inequality is due to the fact that, as we saw before, $S\in\mathcal{D}_M^c(Q)$, and then by (39) $F_\mu(S)\lesssim\sup\{F_\mu(K): K\in\mathcal{D}_M^c(Q)\}<\epsilon$. Now, we parametrize the cubes J according to their relative size with respect to S: $\ell(J)=2^{-k}\ell(S)$. To sum in J, we use $J_p\subset S$ together with the fact that the cubes with fixed side length are disjoint. In this way,

$$\sum_{J\in\mathcal{D}_M^c(S)}|\langle T(\chi_{Q\setminus 4S}),\psi_J\rangle|^2 \lesssim \epsilon^2\sum_{J\in\mathcal{D}(S)}\left(\frac{\ell(J)}{\ell(S)}\right)^{2\delta}\mu(J_p)$$
$$\lesssim \epsilon^2\sum_{k\geq 1}2^{-k2\delta}\sum_{\substack{J\in\mathcal{D}(S)\\ \ell(J)=2^{-k}\ell(S)}}\mu(J_p)$$
$$\leq \epsilon^2\sum_{k\geq 1}2^{-k2\delta}\mu(S)\lesssim\mu(S)\epsilon^2.$$

Summing now over the cubes S in $\mathcal{W}(R)$, we finally get

$$\sum_{S\in\mathcal{W}(R)}\sum_{J\in\mathcal{D}_M^c(S)}|\langle T(\chi_{Q\setminus 4S}),\psi_J\rangle|^2\lesssim\epsilon^2\sum_{S\in\mathcal{W}(R)}\mu(S)\lesssim\epsilon^2\mu(R).$$

When $\ell(R)>2^M$, the reasoning is similar with very few modifications. In this case we have for the first term

$$\sum_{\substack{\bar{S}\in\mathcal{D}(S)\\ \ell(\bar{S})=2^{-(M+2)}}}\sum_{J\in\mathcal{D}_M^c(\bar{S})}|\langle T\chi_{4\bar{S}},\psi_J\rangle|^2 \leq \sum_{\substack{\tilde{S}\in\mathcal{D}(S)\\ \ell(\bar{S})=2^{-(M+2)}}}|\langle\chi_{4\bar{S}}T\chi_{4S},\psi_J\rangle|^2$$
$$\lesssim \sum_{\substack{\bar{S}\in\mathcal{D}(S)\\ \ell(\bar{S})=2^{-(M+2)}}}\|\chi_{4\bar{S}}T\chi_{4\bar{S}}\|^2_{L^2(\mu)}$$
$$\lesssim \sum_{S\in\mathcal{W}(R)}\sum_{\substack{\bar{S}\in\mathcal{D}(S)\\ \ell(\bar{S})=2^{-(M+2)}}}\mu(4\bar{S})F_\mu(4\bar{S})^2.$$

Since $\ell(4\bar{S})=2^{-M}$ we have $F_\mu(4\bar{S})<\epsilon$. Then

$$\sum_{S\in\mathcal{W}(R)}\sum_{\substack{\bar{S}\in\mathcal{D}(S)\\ \ell(\bar{S})=2^{-(M+2)}}}\sum_{J\in\mathcal{D}_M^c(\bar{S})}|\langle T\chi_{4\bar{S}},\psi_J\rangle|^2\lesssim\epsilon^2\sum_{S\in\mathcal{W}(R)}\sum_{\substack{\bar{S}\in\mathcal{D}(S)\\ \ell(\bar{S})=2^{-(M+2)}}}\mu(4\bar{S})$$
$$\lesssim\epsilon^2\sum_{S\in\mathcal{W}(R)}\mu(4S)\lesssim\epsilon^2\mu(R),$$

and we continue as before. We estimate the second term in a similar way.

To finish the proof, we still need to prove that $\Pi(P_M^\perp f, P_M^\perp g)$ belongs to the class of operators for which the theory applies. In particular, we must show that the integral representation of Definition 3.6 holds with a kernel satisfying the Definition 3.3 of a compact Calderón–Zygmund kernel. This work is independent of the measure μ and it can be done in exactly the same way as it was performed in [**25**]. □

10. L^p compactness

In this section we develop the proof of the main result, Theorem 4.2. But first we prove a technical lemma showing that the regions sufficiently close to the border of an open dyadic cube have arbitrarily small measure.

Notation 10.1. *For $N \in \mathbb{N}$, we define the following two collections of dyadic cubes: $\mathcal{D}(Q)_{\geq N} = \{I \in \mathcal{D}(Q) : \ell(I) \geq 2^{-N}\ell(Q)\}$ and $\mathcal{D}(Q)_N = \{I \in \mathcal{D}(Q) : \ell(I) = 2^{-N}\ell(Q)\}$.*

Lemma 10.2. *Let μ be a positive Radon measure in $\mathbb{R}^n$ with power growth $0 < \alpha \leq n$. Let $Q \in \mathcal{D}$, $N_0, M \in \mathbb{N}$, and $\theta \in (0,1)$ be fixed.*

Let $I \in \mathcal{D}(Q)_{\geq N_0}$ with $\ell(I) = 2^{-k_I}\ell(Q)$, $0 \leq k_I \leq N_0$. For $k \geq k_I$ let $C_k(I)$ be the union of the interior of all cubes $R \in \mathcal{D}(3I)$ such that $\ell(R) = 2^{-k}\ell(Q) \leq \ell(I)$ and $\mathrm{inrdist}(R, I) < 1 + \mathrm{ec}(I,R)^{-\theta}$. Finally, let $C_k = \bigcup_{I \in \mathcal{D}(Q)_{\geq N_0}} C_k(I)$.

Then for each $\epsilon > 0$ there exist $k_0 \in \mathbb{N}$ such that $\mu(C_k) < \epsilon$ for all $k > k_0$.

Proof: We start by noting that the family of cubes $\mathcal{D}(Q)_{\geq N_0}$ has cardinality less than $2^{(N_0+1)n}$. Let $I \in \mathcal{D}(Q)_{\geq N_0}$ be fixed.

Recall that $\mathfrak{D}_I = \bigcup_{I' \in \mathrm{ch}(I)} \partial I'$. Then, for each cube R in the definition of $C_k(I)$, the condition $\mathrm{inrdist}(R, I) - 1 \leq \mathrm{ec}(I,R)^{-\theta}$ implies

$$\frac{\mathrm{dist}(R, \mathfrak{D}_I)}{\ell(R)} = \mathrm{inrdist}(R, I) - 1 \leq \left(\frac{\ell(R)}{\ell(I)}\right)^{-\theta},$$

that is,

$$\mathrm{dist}(R, \mathfrak{D}_I) \leq \left(\frac{\ell(R)}{\ell(I)}\right)^{1-\theta} \ell(I).$$

Since $\ell(I) = 2^{-k_I}\ell(Q)$ and $\ell(R) = 2^{-k}\ell(Q)$ then

$$\mathrm{dist}(R, \mathfrak{D}_I) \leq 2^{-(k-k_I)(1-\theta)}\ell(I).$$

Now, for $j \geq k$ we define the set

$$D_j(I) = \{x \in 3I : 2^{-(j-k_I+1)(1-\theta)}\ell(I) < \mathrm{dist}(x, \mathfrak{D}_I) \leq 2^{-(j-k_I)(1-\theta)}\ell(I)\}.$$

Then the sets $(D_j(I))_{j \geq k_I}$ are pairwise disjoint. Moreover, for each $k > k_I$, we have $\bigcup_{j \geq k} D_j(I) \subset \overline{C_k(I)} \subset \bigcup_{j \geq k-1} D_j(I)$, where $\overline{C_k(I)}$ is the union of the topological closures of the cubes $R \in C_k(I)$. Then

$$\sum_{j \geq k} \mu(D_j(I)) \leq \mu(\overline{C_k(I)}) \leq \mu(\overline{C_k}) \leq \mu(\overline{Q}) \lesssim \ell(Q)^\alpha < \infty.$$

Therefore, for any $\epsilon > 0$ there exists $k_{0,I} \geq k_I$ dependent on I such that

$$\sum_{j \geq k} \mu(D_j(I)) < 2^{-(N_0+1)n}\epsilon$$

for all $k \geq k_{0,I}$.

Now let $k_0 = \max\{k_{0,I} : I \in \mathcal{D}(Q)_{\geq N_0}\}$. Since $C_{k+1} \subset C_k$ for each $k \in \mathbb{N}$, we have for all $k > k_0$

$$\begin{aligned}
\mu(C_k) \leq \mu(C_{k_0}) &\leq \sum_{I \in \mathcal{D}(Q)_{\geq N_0}} \mu(C_{k_0}(I)) \\
&\leq \sum_{I \in \mathcal{D}(Q)_{\geq N_0}} \mu\Big(\bigcup_{j \geq k_0 - 1 \geq k_{0,I}-1} D_j(I)\Big) \leq \sum_{I \in \mathcal{D}(Q)_{\geq N_0}} \sum_{j \geq k_{0,I}} \mu(D_j(I)) \\
&< \sum_{I \in \mathcal{D}(Q)_{\geq N_0}} 2^{-(N_0+1)n}\epsilon < \epsilon.
\end{aligned}$$

$\square$

Finally, we proceed with the proof of the main result of the paper, Theorem 4.2. Since the proof is long, we divide it into several subsections.

10.1. Proof of Theorem 4.2. Preliminaries. The necessity of the hypotheses can be shown in a similar way as it was done in [**25**]. Then we focus on their sufficiency.

Once boundedness is proved on $L^2(\mu)$, a classical argument that applies to Calderón–Zygmund operators allows us to extend the result to weak estimates from $L^1(\mu)$ to $L^{1,\infty}(\mu)$ (see [**19**], for example). Then by a standard interpolation argument one can prove boundedness on $L^p(\mu)$ for all $1 < p < \infty$. Moreover, as shown in [**25**], we can deduce compactness on $L^p(\mu)$ for all $1 < p < \infty$ by interpolation between compactness on $L^2(\mu)$ and boundedness on $L^p(\mu)$. For all this, we only focus on the case $p = 2$.

Let $Q = [-2^r, 2^r]^n \in \mathcal{C}$ with $c(Q) = 0$, $\ell(Q) > 4$, with side length large enough so that

$$F_\mu(Q) + F_T(Q) + F_{T^*}(Q) < \epsilon.$$

Since $Q \notin \mathcal{D}$, this requires extending the definition of F_μ, F_T, F_{T^*} to Q. This is possible by using, for example, that Q is the union of 2^n cubes in $\mathcal{D}$. Moreover, the smallness described by the previous inequality is deduced by the large side length of Q. Let $\gamma = \sqrt{n}2^{-s}$ such that $0 < \gamma \leq 1 < \ell(Q)$, $N_0 = \log \frac{6\ell(Q)^{\alpha+2}}{\gamma^{\alpha+1}}$, and $T_{\gamma,Q}$ be the truncated operator of Definition 5.2.

We start by considering the dyadic grid $\mathcal{D} = \mathcal{D}^1$ as denoted in Subsection 2.1. Let $(\psi_I)_{I\in\mathcal{D}}$ be the Haar wavelets frame of Definition 6.2 and P_M be the lagom projection operators related to that system as given in Definition 6.11. We also fix the parameter $\theta = \frac{\alpha}{\alpha+\delta/2} \in (0,1)$.

We aim to prove that $T_{\gamma,Q}$ is uniformly compact on $L^2(\mu)$ with bounds independent of $Q \in \mathcal{C}$ and $0 < \gamma < \ell(Q)$. By the comments at the end of Subsection 6.4, we need to show that for any $\epsilon > 0$ there exists $M_0 \in \mathbb{N}$ (independent of Q and γ) such that $\|P_M^\perp T_{\gamma,Q} P_M^\perp\|_2 \lesssim \epsilon$ for all $M > M_0$, with implicit constant independent of Q and γ (it may depend on δ and the constants appearing in the kernel smooth condition and the testing conditions). By duality and the fact that the kernel $K_{\gamma,Q}$ is supported on Q, this is equivalent to showing that

$$|\langle T_{\gamma,Q} P_M^\perp f, P_M^\perp g\rangle| \lesssim \epsilon$$

for all $M > M_0$, all f, g functions in the unit ball of $L^2(\mu)$, bounded, and compactly supported on Q. Moreover, by Lemma 4.5, we can assume that f, g are supported on the interior of the first quadrant of each grid, which we denote by $\mathbb{R}_i^{n,+}$. In that case, there is $Q_i \in \mathcal{T}_i\mathcal{D}$ such that $\operatorname{supp} f \cup \operatorname{supp} g \subset Q_i$ such that $Q \cap \mathbb{R}_i^{n,+} \subset Q_i \subset 10Q$. Therefore, when using the representation result Corollary 6.19 and Parseval's identity of Lemma 6.9, we should write Q_i each time. However, since Q_i and Q play a similar role, we will just write Q each time. This is equivalent to assuming that the original cube is contained in the first quadrant of each grid and $Q \in \mathcal{T}_i\mathcal{D}$ for all grids.

Then let f, g be fixed functions as described and satisfying

$$\|P_M^\perp T_{\gamma,Q} P_{2M}^\perp\|_2 \leq 2|\langle P_M^\perp T_{\gamma,Q} P_{2M}^\perp f, g\rangle|.$$

Let $0<\epsilon<((\|f\|_{L^\infty(\mu)}+\|g\|_{L^\infty(\mu)})\mu(2Q)^{\frac{1}{2}})^{-4}$ be fixed. Let M_0 be such that for all $M>M_0$ we have $M^{-\frac{\delta}{8}}+M^{-\alpha(\frac{\alpha+\delta}{\alpha+\delta/2}-1)}+M^{-\frac{\alpha\delta}{\alpha+\delta/2}}<\epsilon$ and

$$\sup_{\substack{J\in\mathcal{D}_M^c(Q)\\(I,J)\in\mathcal{F}_M}} F_\mu(I,J)+F_\mu(J,I)+F_T(J)+F_{T^*}(J)<\epsilon, \tag{41}$$

where $\mathcal{F}_M$ is given in Definition 7.3.

Then, for fixed $\epsilon>0$ and chosen $M_0\in\mathbb{N}$, we are going to prove that

$$|\langle T_{\gamma,Q}P_{2M}^\perp f,P_M^\perp g\rangle|\lesssim\epsilon^{1/4} \tag{42}$$

for all $M>M_0$, which is also enough for our purposes. To simplify notation, from now on we denote the operator $T_{\gamma,Q}$ simply by T.

10.2. Discretization of the operator. By Corollary 6.19, we have that

$$\begin{aligned}\langle TP_{2M}^\perp f,P_M^\perp g\rangle &= \sum_{I\in\mathcal{D}_{2M}^c(Q)}\sum_{J\in\mathcal{D}_M^c(Q)}\langle f,\psi_I\rangle\langle g,\psi_J\rangle\langle T\psi_I,\psi_J\rangle\\ &\quad+\langle T(E_Qf),P_M^\perp g\rangle+\langle T(P_{2M}^\perp f-E_Qf),E_Qg\rangle.\end{aligned} \tag{43}$$

Similarly to Lemma 10.2, we set up the following notation: for $N\in\mathbb{N}$, let $\mathcal{D}_M^c(Q)_{\geq N}=\{I\in\mathcal{D}_M^c(Q):\ell(I)\geq 2^{-N}\ell(Q)\}$; for $I\in\mathcal{D}(Q)_{\geq N_0}$ and $k>N_0$, $C_k(I)$ denotes the union of the interior of all cubes $R\in\mathcal{D}(Q)$ such that $\ell(R)=2^{-k}\ell(Q)\leq\ell(I)$ and $\operatorname{inrdist}(R,I)-1\leq\operatorname{ec}(I,R)^{-\theta}$; finally, $C_k=\bigcup_{I\in\mathcal{D}(Q)_{\geq N_0}}C_k(I)$.

Now, by Lemma 10.2 and the implicit limit in the equality at (43), we can choose $N_1>N_0+M$ so that for all $N>N_1$,

$$\mu(C_N)^{\frac{1}{2}}\|f\|_{L^\infty(\mu)}\|g\|_{L^\infty(\mu)}2^{N_0(n+3)}<\epsilon, \tag{44}$$

and

$$\begin{aligned}|\langle TP_{2M}^\perp f,P_M^\perp g\rangle| &\leq 2\Bigg|\sum_{I\in\mathcal{D}_{2M}^c(Q)_{\geq N}}\sum_{J\in\mathcal{D}_M^c(Q)_{\geq N}}\langle f,\psi_I\rangle\langle g,\psi_J\rangle\langle T\psi_I,\psi_J\rangle\\ &\qquad+\langle T(E_Qf),P_M^\perp g\rangle+\langle T(P_{2M}^\perp f-E_Qf),E_Qg\rangle\Bigg|\\ &\leq 2|\langle TP_{2M}^\perp P_{M_N}f,P_M^\perp P_{M_N}g\rangle|\\ &\quad+2|\langle T(E_Qf),P_M^\perp g\rangle|+2|\langle T(P_{2M}^\perp f-E_Qf),E_Qg\rangle|,\end{aligned}$$

with $M_N=N-\log\ell(Q)$. We note that $P_M^\perp P_{M_N}$ is not the zero operator since $M_N>N>N_0+M>M$. Again, to simplify notation, we often stop writing the conditions $I\in\mathcal{D}_{2M}^c(Q)_{\geq N}$, $J\in\mathcal{D}_M^c(Q)_{\geq N}$ in the sums. We will recover this notation whenever needed.

We briefly deal with the last two terms in the previous expression. We only work with $\langle T(E_Qf),P_M^\perp g\rangle$ since the work for the last term is similar. By definition, the function $P_M^\perp g$ is compactly supported on Q. Then, by the testing condition (11), Cauchy's inequality, and (41), we have

$$\begin{aligned}|\langle T(E_Qf),P_M^\perp g\rangle| &\leq |\langle f\rangle_Q|\|\chi_QT\chi_Q\|_{L^2(\mu)}\|P_M^\perp g\|_{L^2(\mu)}\\ &\lesssim\mu(Q)^{-1}\|f\|_{L^2(\mu)}\mu(Q)^{\frac{1}{2}}F_T(Q)\mu(Q)^{\frac{1}{2}}\|g\|_{L^2(\mu)}\lesssim\epsilon.\end{aligned}$$

Now we further decompose the term $\langle TP_{2M}^{\perp}P_{M_N}f, P_M^{\perp}P_{M_N}g\rangle$. For this, we fix such $N > N_1$. We denote by $\partial\mathcal{D}(Q)$ the union of ∂I for all $I \in \mathcal{D}(Q)_{\geq N}$. Then we decompose the argument functions as $P_{2M}^{\perp}P_{M_N}f = f_1 + f_{1,\partial}$, $P_M^{\perp}P_{M_N}g = g_1 + g_{1,\partial}$, where $f_{1,\partial} = (P_{2M}^{\perp}P_{M_N}f)\chi_{\partial D(Q)}$ and $g_{1,\partial} = (P_M^{\perp}P_{M_N}g)\chi_{\partial D(Q)}$. With this,

$$\langle TP_{2M}^{\perp}P_{M_N}f, P_M^{\perp}P_{M_N}g\rangle = \langle TP_{2M}^{\perp}P_{M_N}f, g_1\rangle + \langle Tf_1, g_{1,\partial}\rangle + \langle Tf_{1,\partial}, g_{1,\partial}\rangle. \tag{45}$$

We note that

$$g_1 = \sum_{J\in\mathcal{D}_M^c(Q)_{\geq N}} \langle g, \psi_J\rangle\psi_{\tilde{J}},$$

where $\tilde{J}$ is the interior of J and so it is an open cube. Likewise for f_1. Therefore, when we deal with

$$\langle TP_{2M}^{\perp}P_{M_N}f, g_1\rangle = \sum_{I\in\mathcal{D}_{2M}^c(Q)_{\geq N}} \sum_{J\in\mathcal{D}_M^c(Q)_{\geq N}} \langle f, \psi_I\rangle\langle g, \psi_J\rangle\langle T\psi_I, \psi_{\tilde{J}}\rangle,$$

we have that the cubes $\tilde{J} \in \tilde{\mathcal{D}}$ are all open. Similarly, when we later deal with

$$\langle Tf_1, g_{1,\partial}\rangle = \sum_{I\in\mathcal{D}_{2M}^c(Q)_{\geq N}} \sum_{J\in\mathcal{D}_M^c(Q)_{\geq N}} \langle f, \psi_I\rangle\langle g, \psi_J\rangle\langle T\psi_{\tilde{I}}, \psi_J\rangle,$$

we will have that $\tilde{I} \in \tilde{\mathcal{D}}$ are all open cubes.

However, we will start our work without reflecting this distinction in the notation since it is only useful at the end of the argument. That is, although the work to prove (42) starts with the first term $\langle TP_{2M}^{\perp}P_{M_N}f, g_1\rangle$, since the same argument will also work for the second term $\langle Tf_1, g_{1,\partial}\rangle$, we write each term simply as $\langle P_M^{\perp}TP_{2M}^{\perp}f, g\rangle$ and we will make distinctions only at the end of the proof. We hope this license will not cause any confusion.

In view of the rates of decay stated in Propositions 8.1 and 8.4, we parametrize the sums according to eccentricity, relative distance, and inner relative distance of the cubes as follows. For fixed $e \in \mathbb{Z}$, $m \in \mathbb{N}$, and every given dyadic cube J, we define the family

$$J_{e,m} = \{I \in \mathcal{D}_{2M}^c(Q) : \ell(I) = 2^e\ell(J), m \leq \operatorname{rdist}(I_p, J_p) < m+1\}.$$

For $m = 1$ and $1 \leq k \leq 2^{-\min(e,0)-2} - 1$, we also define

$$J_{e,1,k} = J_{e,1} \cap \{I \in \mathcal{D}_{2M}^c(Q) : k \leq \operatorname{inrdist}(I_p, J_p) < k+1\}.$$

The cardinality of $J_{e,m}$ is comparable to $2^{-\min(e,0)n}nm^{n-1}$, while the cardinality of $J_{e,1,k}$ is comparable to $n(2^{-\min(e,0)}-k)^{(n-1)\frac{\min(e,0)}{e}}$. By symmetry, we have $I \in J_{e,m}$ if and only if $J \in I_{-e,m}$ and, similarly, $I \in J_{e,1,k}$ if and only if $J \in I_{-e,1,k}$.

In accordance with the previous parametrization, we divide the double sum in (43) into three parts D_i, N_i, and B_6 (distant or disjoint cubes, nested cubes, and borderline cubes). Then we add and subtract the paraproducts P_i into the second part. Specifically, we write

$$\begin{aligned}
&\sum_{I\in\mathcal{D}^c_{2M}(Q)_{\geq N}}\ \sum_{J\in\mathcal{D}^c_{M}(Q)_{\geq N}}\langle f,\psi_I\rangle\langle g,\psi_J\rangle\langle T\psi_I,\psi_J\rangle\\
&\qquad=\sum_{e\in\mathbb{Z}}\sum_{m\geq 2}\sum_{J}\sum_{I\in J_{e,m}}\langle f,\psi_I\rangle\langle g,\psi_J\rangle\langle T\psi_I,\psi_J\rangle\\
&\qquad\quad+\sum_{e\in\mathbb{Z}}\sum_{k=2^{\theta|e|}+1}^{2^{|e|}}\sum_{J}\sum_{\substack{I\in J_{e,1,k}\\ \mathrm{dist}(I_p,J_p)>0}}\langle f,\psi_I\rangle\langle g,\psi_J\rangle\langle T\psi_I,\psi_J\rangle\\
&\qquad\quad+\sum_{e\geq 0}\sum_{k=2^{\theta|e|}+1}^{2^{|e|-2}}\sum_{I}\sum_{\substack{J_p\subset I_p\\ J\in I_{-e,1,k}}}\langle f,\psi_I\rangle\langle g,\psi_J\rangle\langle T(\psi_I-\psi^{\mathrm{full}}_{I,J_p}),\psi_J\rangle\\
&\qquad\quad+\sum_{e<0}\sum_{k=2^{\theta|e|}+1}^{2^{|e|-2}}\sum_{J}\sum_{\substack{I_p\subset J_p\\ I\in J_{e,1,k}}}\langle f,\psi_I\rangle\langle g,\psi_J\rangle\langle T\psi_I,\psi_J-\psi^{\mathrm{full}}_{J,I_p}\rangle\\
&\qquad\quad+\Pi(P^{\perp}_{2M}f,P^{\perp}_{M}g)+\Pi'(P^{\perp}_{2M}f,P^{\perp}_{M}g)\\
&\qquad\quad+\sum_{e\in\mathbb{Z}}\sum_{k=1}^{2^{\theta|e|}}\Bigg(\sum_{I}\sum_{J\in I_{e,1,k}}+\sum_{J}\sum_{I\in J_{-e,1,k}}\Bigg)\langle f,\psi_I\rangle\langle g,\psi_J\rangle\langle T\psi_I,\psi_J\rangle\\
&\qquad=D_1+D_2+N_2+N_3+P_4+P_5+B_6.
\end{aligned}$$

The terms P_4 and P_5 are the paraproduct bilinear forms, which are bounded by Proposition 9.2. The terms D_1, D_2 correspond to the distant cubes and the barely disjoint cubes respectively, which we estimate by using the inequalities of Remark 8.2 and Proposition 8.1 respectively. The terms N_2, N_3 correspond to the nested cubes, for which we use the estimate of Proposition 8.4. By symmetry we only need to work with N_2. Finally, the term B_6 corresponds to borderline cubes.

10.3. Distant cubes. The term D_1 contains the cubes for which $m\geq 2$ and so, by (34) in Remark 8.2, we have

$$|\langle T\psi_I,\psi_J\rangle|\lesssim\frac{2^{-|e|\delta}}{m^{\alpha+\delta}}\frac{\mu(I)^{\frac12}\mu(J)^{\frac12}}{\ell(I\vee J)^{\alpha}}F_1(I,J),$$

where $F_1(I,J)$ is given in Proposition 8.1. Then

$$|D_1|\lesssim\sum_{\substack{e\in\mathbb{Z}\\ m\geq 2}}\frac{2^{-|e|\delta}}{m^{\alpha+\delta}}\sum_{J}\sum_{I\in J_{e,m}}\frac{\mu(I)^{\frac12}\mu(J)^{\frac12}}{\ell(I\vee J)^{\alpha}}|\langle f,\psi_I\rangle||\langle g,\psi_J\rangle|F_1(I,J).\tag{46}$$

To estimate this last quantity, we divide the study into two cases: $(I,J)\in\mathcal{F}_M$ and $(I,J)\notin\mathcal{F}_M$.

(a) In the first case, to simplify the argument, we assume $\ell(J) \le \ell(I)$, that is, $e \ge 0$. The other case follows by symmetry. Then $I \vee J = I$ and, by Cauchy's inequality, we can bound the terms in (46) corresponding to this case by

$$\begin{aligned}(47)\quad &\sum_{e\ge 0} 2^{-e\delta}\Big(\sum_I \sup_{\substack{J\in\mathcal{D}^c_M(Q)\\(I,J)\in\mathcal{F}_M}} F_1(I,J)|\langle f,\psi_I\rangle|^2 \sum_{m\ge 2}\frac{1}{m^{\alpha+\delta}}\frac{1}{\ell(I)^\alpha}\sum_{J\in I_{-e,m}}\mu(J)\Big)^{\frac12}\\ &\quad\times\Big(\sum_J \sup_{\substack{I\in\mathcal{D}^c_{2M}(Q)\\(I,J)\in\mathcal{F}_M}} F_1(I,J)|\langle g,\psi_J\rangle|^2 \sum_{m\ge 2}\frac{1}{m^{\alpha+\delta}}\frac{1}{2^{e\alpha}\ell(J)^\alpha}\sum_{I\in J_{e,m}}\mu(I)\Big)^{\frac12}.\end{aligned}$$

We note that the cubes $J \in I_{-e,m}$ are pairwise disjoint, and that also the cubes $I \in J_{e,m}$ are pairwise disjoint. Let $J_e \in \mathcal{D}$ such that $J \subset J_e$ and $\ell(J_e) = 2^e\ell(J) = \ell(I)$. Then we have

$$\sum_{J\in I_{-e,m}} \mu(J) \lesssim \mu(mI \setminus (m-1)I),$$

$$\sum_{I\in J_{e,m}} \mu(I) \lesssim \mu(mJ_e \setminus (m-1)J_e).$$

We start with the first factor of (47), whose inner sum can be written as

$$\frac{1}{\ell(I)^\alpha}\sum_{m\ge 2}\frac{1}{m^{\alpha+\delta}}\sum_{J\in I_{-e,m}}\mu(J) \lesssim \lim_{R\to\infty}\frac{1}{\ell(I)^\alpha}\sum_{m=2}^{R}\frac{\mu(mI)-\mu((m-1)I))}{m^{\alpha+\delta}}.$$

Now, we write $a_m = \mu(mI)$ and use Abel's formula to get

$$\begin{aligned}\frac{1}{\ell(I)^\alpha}\sum_{m=2}^{R}\frac{a_m - a_{m-1}}{m^{\alpha+\delta}} &= \frac{a_R}{R^{\alpha+\delta}\ell(I)^\alpha} - \frac{a_1}{2^{\alpha+\delta}\ell(I)^\alpha}\\ &\quad + \frac{1}{\ell(I)^\alpha}\sum_{m=2}^{R-1} a_m\Big(\frac{1}{m^{\alpha+\delta}} - \frac{1}{(m+1)^{\alpha+\delta}}\Big).\end{aligned}$$

For the first term we have

$$\frac{a_R}{\ell(I)^\alpha R^{\alpha+\delta}} = \frac{\mu(RI)}{\ell(RI)^\alpha}\frac{1}{R^\delta} \lesssim \frac{1}{R^\delta} \le \rho_{\rm out}(I)$$

for R sufficiently large, where we recall that $\rho_{\rm out}(I) = \sum_{m\ge 1}\frac{\mu(mI)}{\ell(mI)^\alpha} m^{-(\frac{\delta}{2}+1)}$. The second term is bounded in a similar way:

$$\frac{a_1}{2^{\alpha+\delta}\ell(I)^\alpha} = \frac{1}{2^\delta}\frac{\mu(I)}{\ell(I)^\alpha} = \frac{1}{2^\delta}\rho(I) \le \rho_{\rm out}(I).$$

The last term is bounded by

$$\begin{aligned}\frac{1}{\ell(I)^\alpha}\sum_{m=2}^{R-1} a_m \frac{(m+1)^{\alpha+\delta} - m^{\alpha+\delta}}{(m+1)^{\alpha+\delta}m^{\alpha+\delta}} &\lesssim \sum_{m=2}^{R-1}\frac{\mu(mI)}{m^\alpha\ell(I)^\alpha}\frac{(m+1)^{\alpha+\delta-1}}{(m+1)^{\alpha+\delta}m^\delta}\\ &\lesssim \sum_{m=2}^{R-1}\frac{\mu(mI)}{\ell(mI)^\alpha}\frac{1}{m^{\delta+1}} \le \rho_{\rm out}(I).\end{aligned}$$

We finish the work with the first factor by noting that $F_1(I,J)\rho_{\rm out}(I) \le F_\mu(I,J) < \epsilon$, since $(I,J)\in\mathcal{F}_M$. For the second factor we can use similar calculations to obtain

$$\sum_{m\ge 2}\frac{1}{m^{\alpha+\delta}}\frac{1}{2^{e\alpha}\ell(J)^\alpha}\sum_{I\in J_{e,m}}\mu(I)\lesssim \rho_{\rm out}(J_e).$$

However, now J_e does not belong to $\mathcal{D}_M^c(Q)$ in general and so the only inequality we can use is $F_1(I,J)\rho_{\rm out}(J_e)\lesssim 1$.

With both things and Lemma 6.9, we conclude that the terms in D_1 corresponding to both cases ($e\ge 0$ and $e\le 0$) can be bounded by a constant times

$$\sum_{e\ge 0}2^{-e\delta}\Big(\sum_{I\in\mathcal{D}_{2M}^c(Q)}\sup_{\substack{J\in\mathcal{D}_M^c(Q)\\(I,J)\in\mathcal{F}_M}}F_\mu(I,J)|\langle f,\psi_I\rangle|^2\Big)^{\frac12}\Big(\sum_{J\in\mathcal{D}_M^c(Q)}|\langle g,\psi_J\rangle|^2\Big)^{\frac12}$$

$$\lesssim \epsilon^{\frac12}\sum_{e\ge 0}2^{-e\delta}\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)}\lesssim \epsilon^{\frac12}.$$

(b) We now study the case when $(I,J)\notin\mathcal{F}_M$, that is, when $I\in\mathcal{D}_{2M}^c(Q)$, $J\in\mathcal{D}_M^c(Q)$ are such that $F_\mu(I,J)\ge\epsilon$. By Lemma 7.2, we have that $|\log(\mathrm{ec}(I,J))|\gtrsim \log M$, or $\mathrm{rdist}(I,J)\gtrsim M^{\frac18}$. Then, instead of the smallness of F_μ, in this case we use that the size and location of the cubes I and J are such that either their eccentricity or their relative distance are extreme.

We fix $e_M\in\{0,\log M\}$, $m_M\in\{M^{\frac18},1\}$ such that $e_M=0$ implies $m_M=M^{\frac18}$. Then, by the calculations made in the subcase (a.1) and $F_\mu(I,J)\lesssim 1$, we can bound the relevant part of (46) by a constant times

$$\sum_{|e|\ge e_M}\sum_{m\ge m_M}\frac{2^{-|e|\delta}}{m^{\alpha+\delta}}\sum_J\sum_{I\in J_{e,m}}\frac{\mu(I)^{\frac12}\mu(J)^{\frac12}}{\ell(I\vee J)^\alpha}|\langle f,\psi_I\rangle||\langle g,\psi_J\rangle|F_1(I,J)$$

$$\lesssim \sum_{|e|\ge e_M}2^{-|e|\delta}\Big(\sum_I|\langle f,\psi_I\rangle|^2\sum_{m\ge m_M}\frac{1}{m^{\alpha+\delta}}\frac{\mu(mI\setminus(m-1)I)}{\ell(I)^\alpha}\Big)^{\frac12} \tag{48}$$

$$\times\Big(\sum_J|\langle g,\psi_J\rangle|^2\sum_{m\ge m_M}\frac{1}{m^{\alpha+\delta}}\frac{\mu(mJ_e\setminus(m-1)J_e)}{\ell(J_e)^\alpha}\Big)^{\frac12}.$$

Now, by Abel's inequality as in case (a) and $\rho(I)=\frac{\mu(I)}{\ell(I)^\alpha}\lesssim 1$, we have

$$\sum_{m\ge m_M}\frac{\mu(mI\setminus(m-1)I)}{m^{\alpha+\delta}\ell(I)^\alpha}\lesssim\lim_{R\to\infty}\Big(\frac{1}{R^\delta}+\frac{1}{m_M^\delta}+\sum_{m=m_M+1}^{R}\frac{1}{m^{\delta+1}}\Big)\lesssim m_M^{-\delta},$$

and similarly for the second factor. With this, Lemma 6.9, and the choice of M, expression (48) is bounded by

$$\sum_{|e|\ge e_M}2^{-|e|\delta}m_M^{-\delta}\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)}\lesssim 2^{-e_M\delta}m_M^{-\delta}\lesssim M^{-\frac{\delta}{8}}<\epsilon.$$

10.4. Disjoint cubes. The term D_2 contains cubes for which $m=1$, $k\ge 1+2^{|e|\theta}$, and $I_p\cap J_p=\emptyset$. Then, by Proposition 8.1, we have

$$|\langle T\psi_I,\psi_J\rangle|\lesssim \frac{1}{k^{\alpha+\delta}}\frac{\mu(I)^{\frac12}\mu(J)^{\frac12}}{\ell(I\wedge J)^{\alpha}}F_1(I,J),$$

with $F_1(I,J)$ as before. Therefore,

$$|D_2|\lesssim \sum_{e\in\mathbb{Z}}\sum_{k=2^{|e|\theta}}^{2^{|e|}}\sum_{J}\sum_{I\in J_{e,1,k}}\frac{1}{k^{\alpha+\delta}}\frac{\mu(I)^{\frac12}\mu(J)^{\frac12}}{\ell(I\wedge J)^{\alpha}}|\langle f,\psi_I\rangle||\langle g,\psi_J\rangle|F_1(I,J). \tag{49}$$

Again, to estimate this last quantity, we divide the study into the same two cases as before: $(I,J)\in\mathcal{F}_M$ and $(I,J)\notin\mathcal{F}_M$.

(a) For the first case, we assume again $\ell(J)\le\ell(I)$. In this case, $I\wedge J=J$ and $\ell(I)=2^e\ell(J)$. Moreover, since $e\ge 0$ we have that for each I and each $k\in\{2^{\theta e},\dots,2^e\}$ the cardinality of $I_{-e,1,k}$ is at most $n(2^{e-1}-2^{e\theta})^{n-1}$. On the other hand, for each J there is only a quantity comparable to n cubes I such that $m=1$ and there is only one $k\ge 2^{e\theta}$ such that $J_{e,1,k}$ is not empty. Then we can consider that this parameter k, which we now denote as k_J, is completely determined by J.

With this and Cauchy's inequality, we can bound the terms in (49) corresponding to this case by

$$\begin{aligned}
&\sum_{e\ge 0}\Bigg(\sum_{k=2^{e\theta}}^{2^e}\sum_{I}\sum_{J\in I_{-e,1,k}}F_1(I,J)|\langle f,\psi_I\rangle|^2\frac{1}{k^{\alpha+\delta}}\frac{1}{\ell(J)^{\alpha}}\mu(J)\Bigg)^{\frac12}\\
&\qquad\times\Bigg(\sum_{k=2^{e\theta}}^{2^e}\sum_{I}\sum_{J\in I_{-e,1,k}}F_1(I,J)|\langle g,\psi_J\rangle|^2\frac{1}{k^{\alpha+\delta}}\frac{1}{\ell(J)^{\alpha}}\mu(I)\Bigg)^{\frac12}\\
&\le\sum_{e\ge 0}\Bigg(\sum_{I}\sup_{\substack{J\in\mathcal{D}^c_M(Q)\\(I,J)\in\mathcal{F}_M}}F_1(I,J)|\langle f,\psi_I\rangle|^2\sum_{k=2^{e\theta}}^{2^e}\frac{1}{k^{\alpha+\delta}}\frac{2^{e\alpha}}{\ell(I)^{\alpha}}\sum_{J\in I_{-e,1,k}}\mu(J)\Bigg)^{\frac12}\\
&\qquad\times\Bigg(\sum_{J}\sup_{\substack{I\in\mathcal{D}^c_M(Q)\\(I,J)\in\mathcal{F}_M}}F_1(I,J)|\langle g,\psi_J\rangle|^2\frac{1}{k_J^{\alpha+\delta}}\frac{1}{\ell(J)^{\alpha}}\sum_{I\in J_{e,1,k_J}}\mu(I)\Bigg)^{\frac12}.
\end{aligned}\tag{50}$$

Now, for each cube $I\in\mathcal{D}^c_{2M}(Q)$ and $k\in\{2^{\theta e},\dots,2^e\}$, since for all $J\in I_{-e,1,k}$ we have that $\ell(J)=2^{-e}\ell(I)$ is fixed, we denote by $I_k\in\mathcal{C}$ the cube such that $c(I_k)=c(I)$ and $\ell(I_k)=(1+k2^{-e+1})\ell(I)\le 3\ell(I)$. With this,

$$\bigcup_{J\in I_{-e,1,k}}J\subset\{t\in 3I: k2^{-e}\ell(I)<\operatorname{dist}(t,I)\le(k+1)2^{-e}\ell(I)\}\subset I_k\setminus I_{k+1}\subset 3I.$$

Now, since the cubes $J\in I_{-e,1,k}$ are pairwise disjoint

$$\sum_{J\in I_{-e,1,k}}\mu(J)\lesssim\mu(I_k\setminus I_{k+1}).$$

On the other hand, since the cardinality of $J_{e,1,k}$ is comparable to n and the cubes $I\in J_{e,1,k}$ are disjoint and included in $3J_e$, we have

$$\sum_{I\in J_{e,1,k}}\mu(I)\le\mu(3J_e).$$

Then the expression in (50) is bounded by a constant times

$$\sum_{e\geq 0}\Big(\sum_{I\in\mathcal{D}_{2M}^c(Q)}\sup_{\substack{J\in\mathcal{D}_M^c(Q)\\ \langle I,J\rangle\in\mathcal{I}_M^c(Q)}}F_1(I,J)|\langle f,\psi_I\rangle|^2\sum_{k=2^{\theta e}}^{2^e}\frac{1}{k^{\alpha+\delta}}\frac{2^{e\alpha}}{\ell(I)^\alpha}\mu(I_k\setminus I_{k+1})\Big)^{\frac12}$$

(51)

$$\times\Big(\sum_{J\in\mathcal{D}_M^c(Q)}\sup_{\substack{I\in\mathcal{D}_M^c(Q)\\ \langle I,J\rangle\in\mathcal{I}_M^c}}F_1(I,J)|\langle g,\psi_J\rangle|^2 2^{-e\theta(\alpha+\delta)}\frac{\mu(3J_e)}{2^{-e\alpha}\ell(3J_e)^\alpha}\Big)^{\frac12}.$$

We start by working on the first factor of (51). As before, we write $a_k=\mu(I_k)$ and evaluate the inner sum by using Abel's formula:

$$\sum_{k=2^{\theta e}}^{2^e}\frac{a_k-a_{k+1}}{k^{\alpha+\delta}}=\frac{a_{2^{\theta e}}}{2^{(\alpha+\delta)\theta e}}-\frac{a_{2^e+1}}{2^{(\alpha+\delta)e}}+\sum_{k=2^{\theta e}+1}^{2^e}a_k\Big(\frac{1}{k^{\alpha+\delta}}-\frac{1}{(k-1)^{\alpha+\delta}}\Big).$$

For the first term we have

$$\frac{a_{2^{\theta e}}}{2^{(\alpha+\delta)\theta e}}\leq\frac{\mu(I_{2^{\theta e}})}{2^{(\alpha+\delta)\theta e}}\leq\rho(I_{2^{\theta e}})\ell(I_{2^{\theta e}})^\alpha 2^{-(\alpha+\delta)\theta e}\leq\rho_{\rm in}(3I)\ell(I)^\alpha 2^{-(\alpha+\delta)\theta e}.$$

Similarly, the absolute value of the second term can be bounded by

$$\frac{a_{2^{e-2}}}{2^{(\alpha+\delta)e}}\leq\frac{\mu(I_{2^{e-2}})}{2^{(\alpha+\delta)e}}\leq\rho(3I)\ell(I)^\alpha 2^{-(\alpha+\delta)e}.$$

The absolute value of the last term is bounded by

$$\begin{aligned}\sum_{k=2^{\theta e}+1}^{2^e}a_k\frac{k^{\alpha+\delta}-(k-1)^{\alpha+\delta}}{(k-1)^{\alpha+\delta}k^{\alpha+\delta}}&\lesssim\sum_{k=2^{\theta e}+1}^{2^e}\mu(I_k)\frac{k^{\alpha+\delta-1}}{(k-1)^{\alpha+\delta}k^{\alpha+\delta}}\\&\lesssim\sum_{k=2^{\theta e}+1}^{2^e}\rho(I_k)\ell(I_k)^\alpha\frac{1}{k^{\alpha+\delta+1}}\\&\lesssim\rho_{\rm in}(3I)\ell(I)^\alpha\sum_{k=2^{\theta e}}^{2^e}\frac{1}{k^{\alpha+\delta+1}}\\&\lesssim\rho_{\rm in}(3I)\ell(I)^\alpha 2^{-(\alpha+\delta)\theta e}.\end{aligned}$$

From the three inequalities, $\ell(I)=2^e\ell(J)$, and the fact that the cardinality of $I^{\rm fr}$ is 3^n, we get

$$\frac{2^{e\alpha}}{\ell(I)^\alpha}\sum_{k=2^{\theta e}}^{2^e}\frac{\mu(I_k\setminus I_{k+1})}{k^\delta}\lesssim\rho_{\rm in}(3I)2^{-e(\theta(\alpha+\delta)-\alpha)}.$$

On the other hand, for the second factor in (51), we have

$$2^{-e\theta(\alpha+\delta)}\frac{\mu(3J_e)}{2^{-e\alpha}\ell(3J_e)^\alpha}\leq 2^{-e(\theta(\alpha+\delta)-\alpha)}\rho_{\rm in}(3J_e).$$

With this, the inequalities $F_1(I,J)\rho_{\rm in}(3I)\lesssim F_\mu(I,J)\lesssim \epsilon$, $F_1(I,J)\rho_{\rm in}(3J_e)\lesssim F_\mu(I,J)\lesssim \epsilon$, and Lemma 6.9, the terms in N_2 corresponding to this case can be bounded by a constant times

$$\sum_{e\geq 0}2^{-e(\theta(\alpha+\delta)-\alpha)}\Big(\sum_{I\in\mathcal{D}_{2M}^c(Q)}\sup_{\substack{J\in\mathcal{D}_M^c(Q)\\ \langle I,J\rangle\in\mathcal{I}_M^c(Q)}}F_\mu(I,J)|\langle f,\psi_I\rangle|^2\Big)^{\frac12}$$

$$\times\Big(\sum_{J\in\mathcal{D}_M^c(Q)}\sup_{\substack{J\in\mathcal{D}_M^c(Q)\\ \langle I,J\rangle\in\mathcal{I}_M^c(Q)}}F_\mu(I,J)|\langle g,\psi_J\rangle|^2\Big)^{\frac12}$$

$$\lesssim \epsilon\sum_{e\geq 0}2^{-e(\theta(\alpha+\delta)-\alpha)}\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)}\lesssim\epsilon,$$

by the choice of $0<\frac{\alpha}{\alpha+\delta}<\theta=\frac{\alpha}{\alpha+\frac{\delta}{2}}<1$.

(b) We now study the case when $(I,J)\notin\mathcal{F}_M$ and so, as before, instead of the smallness of F_μ, we use that either the eccentricity or the relative distance between I and J is extreme.

As in case (b) of the previous subsection, we fix $e_M\in\{0,\log M\}$, $m_M\in\{M^{\frac18},1\}$ such that $e_M=0$ implies $m_M=M^{\frac18}$. But, since $m=1$, we have that $m_M\leq m=1<M^{\frac18}$, which implies $m_M=1$ and so $e_M=\log M$.

Then, by Lemma 6.9, the calculations made in the previous case (a) and $F_\mu(I,J)\lesssim 1$, we can bound the relevant part of (46) by a constant times

$$\sum_{|e|\geq e_M}\sum_{k=2^{|e|\theta}}^{2^{|e|}}\sum_J\sum_{I\in J_{e,1,k}}\frac{1}{k^{\alpha+\delta}}\frac{\mu(I)^{\frac12}\mu(J)^{\frac12}}{\ell(I\wedge J)^\alpha}|\langle f,\psi_I\rangle||\langle g,\psi_J\rangle|F_1(I,J)$$

$$\lesssim\sum_{|e|\geq e_M}\Big(\sum_{I\in\mathcal{D}_{2M}^c(Q)}\sup_{\substack{J\in\mathcal{D}_M^c(Q)\\ \langle I,J\rangle\in\mathcal{I}_M^c(Q)}}F_1(I,J)|\langle f,\psi_I\rangle|^2\sum_{k=2^{\theta e}}^{2^e}\frac{1}{k^{\alpha+\delta}}\frac{2^{e\alpha}}{\ell(I)^\alpha}\mu(I_k\setminus I_{k+1})\Big)^{\frac12}$$

$$\times\Big(\sum_{J\in\mathcal{D}_M^c(Q)}\sup_{\substack{I\in\mathcal{D}_M^c(Q)\\ \langle I,J\rangle\in\mathcal{I}_M^c(Q)}}F_1(I,J)|\langle g,\psi_J\rangle|^2 2^{-e\theta(\alpha+\delta)}\frac{\mu(3J_e)}{2^{-e\alpha}\ell(3J_e)^\alpha}\Big)^{\frac12}$$

$$\lesssim\sum_{|e|\geq e_M}2^{-|e|(\theta(\alpha+\delta)-\alpha)}\Big(\sum_{I\in\mathcal{D}_{2M}^c(Q)}\sup_{\substack{J\in\mathcal{D}_M^c(Q)\\ \langle I,J\rangle\in\mathcal{I}_M^c(Q)}}F_\mu(I,J)|\langle f,\psi_I\rangle|^2\Big)^{\frac12}$$

$$\times\Big(\sum_{J\in\mathcal{D}_M^c(Q)}\sup_{\substack{I\in\mathcal{D}_M^c(Q)\\ \langle I,J\rangle\in\mathcal{I}_M^c(Q)}}F_\mu(I,J)|\langle g,\psi_J\rangle|^2\Big)^{\frac12}$$

$$\lesssim\sum_{|e|\geq\log M}2^{-|e|\alpha(\frac{\alpha+\delta}{\alpha+\delta/2}-1)}\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)}\lesssim M^{-\alpha(\frac{\alpha+\delta}{\alpha+\delta/2}-1)}<\epsilon,$$

by the choices of θ and M.

10.5. Nested cubes. The term N_2 contains the cubes for which we have $2^{\theta|e|} \leq k \leq 2^{|e|-2}$ with $\theta = \frac{\alpha}{\alpha+\delta/2}$ and $I_p \cap J_p \neq \emptyset$. By Proposition 8.4, when $m=1$ and $k \geq 2^{\theta|e|}$, we have

$$|\langle T(\psi_I - \psi_{I,J}^{\text{full}}), \psi_J\rangle| \lesssim k^{-\delta} \sum_{R\in\{I,I_p\}} \left(\frac{\mu(R\cap J)}{\mu(R)}\right)^{\frac{1}{2}} F_{2,\mu}(I,J) + k^{-(\alpha+\delta)} \frac{\mu(I)^{\frac{1}{2}}\mu(J)^{\frac{1}{2}}}{\ell(I\wedge J)^{\alpha}} F_3(I,J),$$

with $F_{2,\mu}$ and F_3 given in Proposition 8.4. The second term can be bounded using the same approach as we used in the previous subsection when we worked with the term D_2 since the only difference between the estimates obtained in these two cases is the last factor, which is given by F_3 instead of F_1. Then we focus on the first term.

In N_2 we have $e \geq 0$, which implies $\ell(J) \leq \ell(I)$. Moreover, $F_{2,\mu} \leq F_\mu$ and so the terms corresponding to this case can be bounded by a constant times

$$\sum_{e\geq 0} \sum_{k=2^{\theta e}}^{2^{e-2}} \sum_{\substack{J \\ J_p\subset I_p}} \sum_{I\in J_{e,1,k}} |\langle f,\psi_I\rangle||\langle g,\psi_J\rangle| \sum_{R\in\{I,I_p\}} \left(\frac{\mu(R\cap J)}{\mu(R)}\right)^{\frac{1}{2}} k^{-\delta} F_\mu(I,J).$$

As before, we distinguish two cases: $(I,J)\in\mathcal{F}_M$ and $(I,J)\notin\mathcal{F}_M$.

(a) In the first case, we have that $F_\mu(I,J) < \epsilon$. Moreover, since $e\geq 0$ the cardinality of $J_{e,1,k}$ is comparable to n and there is only one $k \geq 2^{e\theta}$ such that $J_{e,1,k}$ is not empty, that is, k_J is completely determined by J. Then, by Cauchy's inequality again, we can bound the terms of N_2 corresponding to this case by a constant times

$$\epsilon \sum_{e\geq 0} \left(\sum_{I\in\mathcal{D}_{2M}^c(Q)} |\langle f,\psi_I\rangle|^2 \sum_{k=2^{\theta e}}^{2^{e-2}} k^{-\delta} \sum_{R\in\{I,I_p\}} \sum_{J\in I_{-e,1,k}} \frac{\mu(R\cap J)}{\mu(R)} \right)^{\frac{1}{2}} \times \left(\sum_{J\in\mathcal{D}_M^c(Q)} |\langle g,\psi_J\rangle|^2 k_J^{-\delta}\right)^{\frac{1}{2}}. \tag{52}$$

For fixed $I\in\mathcal{D}_{2M}^c(Q)$, $k\in\{2^{\theta e},\dots,2^{e-2}\}$, and all $J\in I_{-e,1,k}$, since $\ell(J) = 2^{-e}\ell(I)$ is fixed, we define $I_k \in \mathcal{C}$ to be the cube such that $c(I_k) = c(I_p)$ and $\ell(I_k) = (1 - k2^{-e})\ell(I_p) \leq \ell(I_p)$. With this, we have

$$\bigcup_{J\in I_{-e,1,k}} J \subset \{t\in I_p : k2^{-e}\ell(I) < \operatorname{dist}(t,\mathfrak{D}_{I_p}) \leq (k+1)2^{-e}\ell(I)\} \subset I_k\setminus I_{k+1} \subset I_p.$$

Moreover, since the cubes $J\in I_{-e,1,k}$ are pairwise disjoint, we have for $R\in\{I,I_p\}$

$$\sum_{J\in I_{-e,1,k}} \mu(R\cap J) \lesssim \mu(R\cap(I_k\setminus I_{k+1})).$$

Then the expression (52) can be bounded by a constant times

$$\epsilon \sum_{e\geq 0} \left(\sum_{I\in\mathcal{D}_{2M}^c(Q)} |\langle f,\psi_I\rangle|^2 \sum_{R\in\{I,I_p\}} \sum_{k=2^{\theta e}}^{2^{e-2}} k^{-\delta} \frac{\mu(R\cap I_k) - \mu(R\cap I_{k+1})}{\mu(R)} \right)^{\frac{1}{2}} \times \left(\sum_{J\in\mathcal{D}_M^c(Q)} |\langle g,\psi_J\rangle|^2 k_J^{-\delta}\right)^{\frac{1}{2}}.$$

As before, we write $a_k = \mu(R \cap I_k)$ and evaluate the inner sum of the first factor by using Abel's formula:

$$\frac{1}{\mu(R)} \sum_{k=2^{\theta e}}^{2^{e-2}} \frac{a_k - a_{k+1}}{k^\delta} = \frac{a_{2^{\theta e}}}{2^{\theta\delta e}\mu(R)} - \frac{a_{2^{e-2}+1}}{2^{(e-2)\delta}\mu(R)} + \frac{1}{\mu(R)} \sum_{k=2^{\theta e}+1}^{2^{e-2}} a_k \left(\frac{1}{k^\delta} - \frac{1}{(k-1)^\delta}\right).$$

For the first term we have

$$\frac{a_{2^{\theta e}}}{2^{\theta\delta e}\mu(R)} \leq \frac{\mu(R \cap I_{2^{\theta e}})}{2^{\delta\theta e}\mu(R)} \lesssim 2^{-\delta\theta e}.$$

Similarly, the absolute value of the second term can be bounded by

$$\frac{a_{2^{e-2}+1}}{2^{(e-2)\delta}\mu(R)} \lesssim \frac{\mu(R \cap I_{2^{e-2}+1})}{2^{\delta e}\mu(R)} \leq 2^{-\delta e}.$$

The absolute value of the last term is bounded by

$$\begin{aligned} \frac{1}{\mu(R)} \sum_{k=2^{\theta e}+1}^{2^{e-2}} a_k \frac{k^\delta - (k-1)^\delta}{(k-1)^\delta k^\delta} &\lesssim \sum_{k=2^{\theta e}+1}^{2^{e-2}} \frac{\mu(R \cap I_k)}{\mu(R)} \frac{k^{\delta-1}}{(k-1)^\delta k^\delta} \\ &\lesssim \sum_{k \geq 2^{\theta e}+1} \frac{1}{(k-1)^{\delta+1}} \lesssim 2^{-\delta\theta e}. \end{aligned}$$

For the second factor, we just use that $k_J \geq 2^{\theta e}$. With this, the fact that the cardinality of $\mathrm{ch}(I_p)$ is 2^n, and Lemma 6.9, we bound the terms in N_2 corresponding to this case by

$$\begin{aligned} \epsilon \sum_{e \geq 0} \Big(2^{-\theta\delta e} \sum_{I \in \mathcal{D}_{2M}^c(Q)} |\langle f, \psi_I\rangle|^2\Big)^{\frac{1}{2}} \Big(2^{-\theta\delta e} \sum_{J \in \mathcal{D}_M^c(Q)} |\langle g, \psi_J\rangle|^2\Big)^{\frac{1}{2}} \\ \lesssim \epsilon \sum_{e \geq 0} 2^{-\theta\delta e} \|f\|_{L^2(\mu)} \|g\|_{L^2(\mu)} \lesssim \epsilon, \end{aligned}$$

since $0 < \theta$.

(b) When $(I, J) \notin \mathcal{F}_M$, as in case (b) of the previous subsection, we fix $e_M = \log M$. Then, by the calculations made in the case (a) and $F_\mu(I, J) \lesssim 1$, we bound the relevant part of (46) by a constant times

$$\begin{aligned} &\sum_{|e| \geq e_M} \sum_{k=2^{\theta|e|}}^{2^{|e|-2}} \sum_{\substack{J \\ J_p \subset I_p}} \sum_{I \in J_{e,1,k}} |\langle f, \psi_I\rangle| |\langle g, \psi_J\rangle| \sum_{R \in \{I, I_p\}} \left(\frac{\mu(R \cap J)}{\mu(R)}\right)^{\frac{1}{2}} \frac{F_\mu(I, J)}{k^\delta} \\ &\lesssim \sum_{|e| \geq e_M} \Big(\sum_{I \in \mathcal{D}_{2M}^c(Q)} |\langle f, \psi_I\rangle|^2 \sum_{k=2^{\theta|e|}}^{2^{|e|-2}} k^{-\delta} \sum_{R \in \{I, I_p\}} \sum_{J \in I_{-e,1,k}} \frac{\mu(R \cap J)}{\mu(R)}\Big)^{\frac{1}{2}} \\ &\qquad \times \Big(\sum_{J \in \mathcal{D}_M^c(Q)} |\langle g, \psi_J\rangle|^2 k_J^{-\delta}\Big)^{\frac{1}{2}} \\ &\lesssim \sum_{|e| \geq \log M} 2^{-|e|\theta\delta} \|f\|_{L^2(\mu)} \|g\|_{L^2(\mu)} \lesssim M^{-\frac{\alpha\delta}{\alpha+\delta/2}} \leq \epsilon, \end{aligned}$$

by the choice of θ and M.

10.6. Borderline cubes. Now we need to estimate the term B_6 and the terms Ad_1, and Ad_2, which we added to the paraproduct. All these terms contain cubes $I,J\in\mathcal{D}_M^{\mathrm{c}}(Q)$ such that $1\le k\le 2^{\theta|e|}+1$, that is, $\mathrm{inrdist}(I,J)-1\le \mathrm{ec}(I,J)^{-\theta}$. We show the work in detail only for B_6 since the same ideas can be used for the other two terms.

Recall the following notation used in Lemma 10.2: for $N\in\mathbb{N}$, which we chose in (44), we write $\mathcal{D}(Q)_{\ge N}=\{I\in\mathcal{D}_M^{\mathrm{c}}(Q):\ell(I)\ge 2^{-N}\ell(Q)\}$ and $\mathcal{D}(Q)_N=\{I\in\mathcal{D}_M^{\mathrm{c}}(Q):\ell(I)=2^{-N}\ell(Q)\}$. Moreover, for $I\in\mathcal{D}(Q)_{\ge N}$, let I_θ be the family of cubes $J\in\mathcal{D}(Q)_{\ge N}$ such that $1\le k\le 2^{\theta|e|}$.

Then we can rewrite B_6 as

$$B_6=\sum_{I\in\mathcal{D}(Q)_{\ge N}}\sum_{J\in I_\theta}\langle f,\psi_I\rangle\langle g,\psi_J\rangle\langle T\psi_I,\psi_J\rangle.$$

For each $I\in\mathcal{D}(Q)_{\ge N}$, let $I_{\max}$ be the family of cubes $J\in I_\theta$ that are maximal with respect to the inclusion. Then let I_{over} be the family of cubes $R\in\mathcal{D}(Q)_{\ge N}$ such that $J\subsetneq R$ for some $J\in I_{\max}$. We note that for all $I\in\mathcal{D}(Q)_{\ge N}$, either $Q\in I_{\max}$ (if $I\in Q_\theta$) or $Q\in I_{\mathrm{over}}$. So, we always have $Q\in I_\theta\cup I_{\mathrm{over}}$. We also note that all cubes in I_{over} satisfy that $k>2^{\theta|e|}$ with respect to I.

Now we include the cubes in I_{over}, and for each pair (I,J) in the sum defining B_6 we add the siblings either I or J that are not already contained in B_6. Then to the previous expression we add and subtract the term

$$\begin{aligned}A=&\sum_{I\in\mathcal{D}(Q)_{\ge N}}\sum_{J\in I_{\mathrm{over}}}\langle f,\psi_I\rangle\langle g,\psi_J\rangle\langle T\psi_I,\psi_J\rangle\\&+\sum_{I\in\mathcal{D}(Q)_{\ge N}}\sum_{J\in I_\theta}\sum_{\substack{J'\in\mathrm{ch}(J_p)\\J'\notin I_\theta}}\langle f,\psi_I\rangle\langle g,\psi_{J'}\rangle\langle T\psi_I,\psi_{J'}\rangle\\&+\sum_{J\in\mathcal{D}(Q)_{\ge N}}\sum_{I\in J_\theta}\sum_{\substack{I'\in\mathrm{ch}(I_p)\\I'\notin J_\theta}}\langle f,\psi_{I'}\rangle\langle g,\psi_J\rangle\langle T\psi_{I'},\psi_J\rangle.\end{aligned}$$

With this we obtain

$$|B_6|\lesssim\Big|\sum_{I\in\mathcal{D}(Q)_{\ge N}}\langle f,\psi_I\rangle\Big\langle T\psi_I,\sum_{J\in I_\theta\cup I_{\mathrm{over}}}\langle g,\psi_J\rangle\psi_J\Big\rangle\Big|+|A|.\tag{53}$$

In the last expression, the collections I_θ and I_{over} are not exactly the same as defined before. But we use the same notation for them because they consist of the same cubes as before plus their corresponding siblings if they were not initially in the expression defining B_6.

Since all pairs of cubes added satisfy that $k>2^{\theta|e|}$, we can apply the reasoning of any of the previous cases (adding and subtracting the corresponding part of a paraproduct when needed) to prove that the second term in (53) satisfies $|A|\lesssim\epsilon\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)}$. We note that in the case of Ad_i this is due to the fact that the expression $\langle T(\psi_I-\psi_{I,J}^{\mathrm{full}}),\psi_J\rangle$ for the described cubes satisfies the inequality given in Proposition 8.4. Then we only need to study the first term.

With this, for each $I\in\mathcal{D}(Q)_{\ge N}$, we have that $I_\theta\cup I_{\mathrm{over}}$ is a convex family of cubes that contains all siblings of each cube in the sum, has minimal cubes in $\mathcal{D}(Q)_N$ and maximal cube Q. Then, by summing a telescoping series, we have for each $I\in\mathcal{D}(Q)_{\ge N}$

$$\sum_{J\in I_\theta\cup I_{\mathrm{over}}}\langle g,\psi_J\rangle\psi_J=\sum_{J\in\mathcal{D}(I)_N\cap I_\theta}\langle g\rangle_J\chi_J-\langle g\rangle_Q\chi_Q.$$

We can easily estimate the part of the first term in (53) corresponding to the second term on the right-hand side of the previous expression. For this, we reason as follows:

$$\begin{aligned}\Big|\sum_{I\in\mathcal{D}(Q)_{\geq N}}\langle f,\psi_I\rangle\langle T\psi_I,\langle g\rangle_Q\chi_Q\rangle\Big| &= |\langle g\rangle_Q|\Big|\Big\langle\sum_{I\in\mathcal{D}(Q)_{\geq N}}\langle f,\psi_I\rangle\psi_I,T^*\chi_Q\Big\rangle\Big|\\ &\leq \mu(Q)^{-1}\|g\|_{L^2(\mu)}\mu(Q)^{\frac12}\Big\|\sum_{I\in\mathcal{D}(Q)_{\geq N}}\langle f,\psi_I\rangle\psi_I\Big\|_{L^2(\mu)}\|\chi_QT^*\chi_Q\|_{L^2(\mu)}.\end{aligned}$$

Let $h=\sum_{I\in\mathcal{D}(Q)_{\geq N}}\langle f,\psi_I\rangle\psi_I$. By Lemma 6.9,

$$\begin{aligned}\|h\|^2_{L^2(\mu)} &= \sum_{J\in\mathcal{D}(Q)_{\geq N}}\Big|\sum_{I\in\mathcal{D}(Q)_{\geq N}}\langle f,\psi_I\rangle\langle\psi_I,\psi_J\rangle\Big|^2+\|E_Qh\|_{L^2(\mu)}\\ &= \sum_{J\in\mathcal{D}(Q)_{\geq N}}\Big|\sum_{I\in\mathrm{ch}(J_p)}\langle f,\psi_I\rangle\langle\psi_I,\psi_J\rangle\Big|^2\\ &\lesssim \sum_{J\in\mathcal{D}(Q)_{\geq N}}\sum_{I\in\mathrm{ch}(J_p)}|\langle f,\psi_I\rangle|^2\lesssim\sum_{I\in\mathcal{D}(Q)_{\geq N-1}}|\langle f,\psi_I\rangle|^2.\end{aligned}$$

We have used that $E_Qh=0$, $|\langle\psi_I,\psi_J\rangle|\lesssim 1$, and that the cardinality of $\mathrm{ch}(J_p)$ is 2^n. Then by the testing condition (11) we get

$$\begin{aligned}\Big|\sum_{I\in\mathcal{D}(Q)_{\geq N}}\langle f,\psi_I\rangle\langle T\psi_I,\langle g\rangle_Q\chi_Q\rangle\Big| &\leq \mu(Q)^{-\frac12}\Big(\sum_{I\in\mathcal{D}(Q)_{\geq N-1}}|\langle f,\psi_I\rangle|^2\Big)^{\frac12}F_{T^*}(Q)\mu(Q)^{\frac12}\\ &\lesssim F_{T^*}(Q)\|f\|_{L^2(\mu)}<\epsilon.\end{aligned}$$

On the other hand, by Fubini's theorem, the remaining part of the first term on the right-hand side of (53) can be rewritten as

$$\begin{aligned}&\sum_{I\in\mathcal{D}(Q)_{\geq N}}\sum_{J\in\mathcal{D}(I)_N\cap I_\theta}\langle f,\psi_I\rangle\langle g\rangle_J\langle T\psi_I,\chi_J\rangle\\ &\qquad=\sum_{J\in\mathcal{D}(Q)_N}\langle g\rangle_J\Big\langle T\Big(\sum_{I\in\mathcal{D}(Q)_{\geq N}\cap J_\theta}\langle f,\psi_I\rangle\psi_I\Big),\chi_J\Big\rangle,\end{aligned}\tag{54}$$

where J_θ is defined as I_θ was defined before.

Now recall the following definition: for $J\in\mathcal{D}(Q)_N$, J^{fr} denotes the family of dyadic cubes $J'\in\mathcal{D}(Q)_N$ such that $\ell(J')=\ell(J)$ and $\mathrm{dist}(J',J)=0$. Then we note that, since J has minimal side length, the condition $I\in J_\theta$ implies that $J'\subset I$ for some $J'\in J^{\mathrm{fr}}$.

Moreover, the cardinality of J^{fr} is 3^n and so we can enumerate the cubes in J^{fr} as $\{J_j\}_{j=1}^{3^n}$ by their fixed position with respect to J. Then, for each $j\in\{1,\dots,3^n\}$ the cubes $I\in\mathcal{D}(Q)_{\geq N}\cap J_\theta$ such that $J_j\subset I$ form an increasing chain of cubes $I_j=I_{j,N}\subset I_{j,N-1}\subset\cdots\subset I_{j,k_j}$ parametrized by their side length $\ell(I_{j,k})=2^{-k}\ell(Q)$ with $k\in\{k_j,\dots,N\}\subset\{0,\dots,N\}$. Some chains may be empty. All these cubes depend on J, but we omit this dependence from the notation. Now, in each chain of cubes we also include the siblings of any cube already included in the chain. That way each chain is convex, and such that for each cube in the collection all its siblings are also in the collection. Then for each fixed $J \in \mathcal{D}(Q)_N$ we have

$$\sum_{I\in\mathcal{D}(Q)_{\geq N}\cap J_\theta} \langle f,\psi_I\rangle\psi_I = \sum_{j=1}^{3^n}\langle f\rangle_{I_j}\chi_{I_j} - \sum_{j=1}^{3^n}\langle f\rangle_{(I_{j,k_j})_p}\chi_{(I_{j,k_j})_p},$$

where $(I_{j,k_j})_p$ is the parent cube of I_{j,k_j}. To simplify notation, we will simply write I_{j,k_j}. With this, (54) can be written as

$$\sum_{J\in\mathcal{D}(Q)_N}\sum_{j=1}^{3^n}\langle f\rangle_{I_j}\langle g\rangle_J\langle T\chi_{I_j},\chi_J\rangle - \sum_{J\in\mathcal{D}(Q)_N}\sum_{j=1}^{3^n}\langle f\rangle_{I_{j,k_j}}\langle g\rangle_J\langle T\chi_{I_{j,k_j}},\chi_J\rangle = S_1 - S_2.$$

We now note that, since $t,x\in Q$, the kernel operator can be written as

$$K_{\gamma,Q}(t,x) = K(t,x)\bigg(1-\phi\Big(\frac{|t-x|}{\gamma}\Big)\bigg).$$

By the definition of the kernel we have $K_{\gamma,Q}(t,x)=0$ for $|t-x|<\gamma$. Then, if $\ell(I\vee J)<\gamma/3$ and $\operatorname{dist}(I,J)<\gamma/3$, we have

$$|t-x|\leq \ell(I)+\operatorname{dist}(I,J)+\ell(J)<\gamma$$

for all $t\in I$ and all $x\in J$, which implies $\langle T\chi_I,\chi_J\rangle = 0$.

Now, all cubes in the first sum S_1 satisfy $\ell(I_j)=\ell(J)$ and $\operatorname{dist}(I_j,J)=0$. Moreover, since $N>N_0\geq \log\frac{6\ell(Q)}{\gamma}$ and $\ell(I_j)=2^{-N}\ell(Q)$, we have $\ell(I_j)<\gamma/3$. Therefore, each term in the sum S_1 equals zero.

We now focus on S_2. Remember that the cubes I_{j,k_j} in that term satisfy $I_j\subset I_{j,k_j}$ with $\operatorname{dist}(I_j,J)=0$. Moreover, $1\leq k\leq 2^{\theta|e|}+1$, with $k=\operatorname{inrdist}(I_{j,k_j},I_j)=1+\frac{\operatorname{dist}(I_{j,k_j},I_j)}{\ell(I_j)}$, and $2^{-|e|\theta}=\frac{\ell(J)}{\ell(I_{j,k_j})}=\frac{\ell(I_j)}{\ell(I_{j,k_j})}$. Then

$$\begin{aligned}\operatorname{dist}(I_{j,k_j},I_j)\leq (k-1)\ell(I_j)\leq 2^{|e|\theta}\ell(I_j) &= \bigg(\frac{\ell(I_j)}{\ell(I_{j,k_j})}\bigg)^{1-\theta}\ell(I_{j,k_j})\\ &= 2^{-|e|(1-\theta)}\ell(I_{j,k_j})\leq \ell(I_{j,k_j}).\end{aligned}$$

Then, since $\ell(I_j)\leq\ell(I_{j,k_j})$, we get

$$\operatorname{dist}(I_{j,k_j},J)\leq \operatorname{dist}(I_j,J)+\ell(I_j)+\operatorname{dist}(I_{j,k_j},I_j)\leq 2\ell(I_{j,k_j}).$$

With this, when $\ell(I_{j,k_j})<\gamma/6$, we have $\operatorname{dist}(I_{j,k_j},J)<\gamma/3$ and so, as before, $\langle T\chi_{I_{j,k_j}},\chi_J\rangle=0$. This implies that the scales for which the dual pair is non-zero satisfy $\ell(I_{j,k_j})=2^{-k}\ell(Q)\geq\gamma/6$, that is, $k\leq\log\frac{6\ell(Q)}{\gamma}\leq N_0$. And since $k\in\{0,\dots,N\}$, that means that the non-zero terms in S_2 contain cubes I_{j,k_j} of at most N_0+1 different side lengths (in fact in the N_0+1 largest scales, all of them in $\{0,1,\dots,N_0\}$).

Now, to apply Fubini's theorem and change the order of summation, we need to rewrite the sum in

$$S_2 = \sum_{J\in\mathcal{D}(Q)_N}\sum_{j=1}^{3^n}\langle f\rangle_{J_{j,k_j}}\langle g\rangle_J\langle T\chi_{I_{j,k_j}},\chi_J\rangle \tag{55}$$

in terms of the cubes I_{j,k_j} instead of the cubes J.

Recall that in (55), for each $J\in\mathcal{D}(Q)_N$, each $I_j\in J^{\mathrm{fr}}$ with $j\in\{1,\dots,3^n\}$, and each scale $k\in\{0,\dots,N_0\}$, we have considered an associated cube $I_{j,k}\in J_\theta$ with side length $\ell(I_{j,k})=2^{-k}\ell(Q)$ and its siblings. Now we reparametrize the cubes we have up to now denoted by $I_{j,k}$ in the following way: for each scale $k\in\{0,\dots,N_0\}$ and each $i \in \{1, \dots, 2^{kn}\}$, we denote by $I^{i,k}$ the cubes such that $\ell(I^{i,k}) = 2^{-k}\ell(Q)$. We note that inside Q, for each $k \in \{0, \dots, N_0\}$ there are in total 2^{kn} of such cubes. Now, for each $I^{i,k}$ we define $\mathcal{J}^{i,k}$ as the family of cubes $J \in \mathcal{D}(Q)_N$ such that there exists $J' \in (I^{i,k})_\theta$. This implies $\operatorname{dist}(I_{i,k}, J') < 2^{|e|\theta}\ell(J')$. Finally, we denote $C^{i,k} = \bigcup_{J\in\mathcal{J}^{i,k}} 3J$. We note that $\mu(3J) \le \mu(C^{i,k})$. With this,

$$S_2 = \sum_{k=0}^{N_0-1}\sum_{i=0}^{2^{kn}}\sum_{J\in\mathcal{J}^{i,k}} \langle f\rangle_{I^{i,k}}\langle g\rangle_J\langle T\chi_{I^{i,k}}, \chi_J\rangle.$$

Now, for fixed $I_{i,k}$, let $\mathcal{J}_{0,0}$, $\mathcal{J}_{0,1}$, $\mathcal{J}_{1,0}$, and $\mathcal{J}_{1,1}$ be the collection of cubes in $J \in \mathcal{J}^{i,k}$ such that $\langle T\chi_{I^{i,k}}, \chi_J\rangle$ belongs to each quadrant of $\mathbb{C}$. More specifically, $J \in \mathcal{J}_{a,b}$ if and only if $(-1)^a \operatorname{Re}(\langle T\chi_{I^{i,k}}, \chi_J\rangle) \ge 0$ and $(-1)^b \operatorname{Im}(\langle T\chi_{I^{i,k}}, \chi_J\rangle) \ge 0$. Also let S_{l_1,l_2} be the union of the cubes in $\mathcal{J}_{l_1,l_2}$. Finally, we define

$$\tilde{S} = \bigcup_{k=0}^{N_0-1}\bigcup_{i=0}^{2^{kn}}\bigcup_{J\in\mathcal{J}^{i,k}} J,$$

that is, the union of all cubes $J \in \mathcal{D}(Q)_N$ such that $J \in (I^{i,k})_\theta$ for some i, k. We note that

$$\tilde{S} = \bigcup_{k=0}^{N_0-1}\bigcup_{i=0}^{2^{kn}}\bigcup_{l_1,l_2\in\{0,1\}}\bigcup_{J\in\mathcal{J}_{l_1,l_2}} J.$$

Before continuing, recall that in the decomposition obtained in (45) we first considered estimates for $\langle P_M^\perp T P_{2M}^\perp f, g_1\rangle$. In this case, the cubes $J \in \tilde{\mathcal{D}}$ and so they are open cubes. Therefore, $C^{i,k}$ and $\tilde{S}$ are open sets and they satisfy by the choice of N in (44) that $\mu(\tilde{S})$ is sufficiently small.

Then, since for $J \in \mathcal{J}_{a,b}$ we have

$$|\langle T\chi_{I^{i,k}}, \chi_J\rangle| \lesssim (-1)^a \operatorname{Re}\langle T\chi_{I^{i,k}}, \chi_J\rangle + (-1)^b \operatorname{Im}\langle T\chi_{I^{i,k}}, \chi_J\rangle,$$

we can write

$$\begin{aligned}
|S_2| &\le \sum_{k=0}^{N_0-1}\sum_{i=0}^{2^{kn}}\sum_{J\in\mathcal{J}^{i,k}} |\langle f\rangle_{I_{i,k}}||\langle g\rangle_J||\langle T\chi_{I^{i,k}}, \chi_J\rangle| \\
&\lesssim \|f\|_{L^\infty(\mu)}\|g\|_{L^\infty(\mu)} \sum_{k=0}^{N_0-1}\sum_{i=0}^{2^{kn}}\sum_{a,b=0}^{1}\Big(\sum_{J\in\mathcal{J}_{a,b}} (-1)^a \operatorname{Re}\langle T\chi_{I^{i,k}}, \chi_J\rangle \\
&\qquad\qquad + \sum_{J\in\mathcal{J}_{a,b}} (-1)^b \operatorname{Im}\langle T\chi_{I^{i,k}}, \chi_J\rangle\Big) \\
&= \|f\|_{L^\infty(\mu)}\|g\|_{L^\infty(\mu)} \sum_{k=0}^{N_0-1}\sum_{i=0}^{2^{kn}}\sum_{a,b=0}^{1}((-1)^a \operatorname{Re}\langle T\chi_{I^{i,k}}, \chi_{S_{a,b}}\rangle \\
&\qquad\qquad + (-1)^b \operatorname{Im}\langle T\chi_{I^{i,k}}, \chi_{S_{a,b}}\rangle) \\
&\lesssim \|f\|_{L^\infty(\mu)}\|g\|_{L^\infty(\mu)} \sum_{k=0}^{N_0-1}\sum_{i=0}^{2^{kn}}\sum_{a,b\in\{0,1\}} |\langle T\chi_{I^{i,k}}, \chi_{S_{a,b}}\rangle|.
\end{aligned}$$

Now we divide $S_{a,b}$ into $2n+1$ parts: $S_{a,b} = \bigcup_{j=0}^{2n} S_{j;a,b}$, where $S_{j;a,b}$ is the union of cubes $J \in \mathcal{D}(Q)_N$ such that $J \in (I^{i,k})_\theta$ for some i, k, and there is $I_j^{i,k} \in (I^{i,k})^{\mathrm{fr}}$ with $J \subset I_j^{i,k}$. This implies that $S_{j;a,b} \subset J_j^{i,k}$.

We work with $S_{j;a,b}$ for every $j \in \{0,1,\dots,2n\}$. By Lemma 5.3, the truncated operator $T_{\gamma,Q}$ is bounded on $L^2(\mu)$ with bounds $\|T_{\gamma,Q}\|_{2,2} \le \frac{\ell(Q)}{\gamma^\alpha} \le 2^{N_0}$. Then, since $S_{j;a,b} \subset \tilde{S}$, we have

$$|\langle T\chi_{I^{i,k}}, \chi_{S_{j;a,b}}\rangle| \le \|T\|_{2,2}\mu(I^{i,k})^{\frac{1}{2}}\mu(S_{j;a,b})^{\frac{1}{2}} \le 2^{N_0}\mu(I^{i,k})^{\frac{1}{2}}\mu(\tilde{S})^{\frac{1}{2}}.$$

With this,

$$\begin{aligned}|S_2| &\lesssim \mu(\tilde{S})^{\frac{1}{2}}\|f\|_{L^\infty(\mu)}\|g\|_{L^\infty(\mu)}2^{N_0}\sum_{k=0}^{N_0-1}\sum_{i=0}^{2^{kn}}\mu(I^{i,k})^{\frac{1}{2}}\\ &\lesssim \mu(\tilde{S})^{\frac{1}{2}}\|f\|_{L^\infty(\mu)}\|g\|_{L^\infty(\mu)}2^{N_0}\mu(Q)^{\frac{1}{2}}N_0 2^{N_0 n}\\ &\lesssim \mu(\tilde{S})^{\frac{1}{2}}\|f\|_{L^\infty(\mu)}\|g\|_{L^\infty(\mu)}2^{N_0(n+3)} < \epsilon.\end{aligned}$$

In the second last inequality we have used $\mu(I^{i,k}) \le \mu(Q) \le \rho(Q)\ell(Q)^\alpha \lesssim 2^{N_0}$ and so $\mu(Q)^{\frac{1}{2}} \le 2^{N_0}$. The last inequality holds because $\tilde{S} \subset C_N$ and from the choice of N in (44).

10.7. Estimates for $\langle TP_{2M}^{\perp}P_{M_N}f, g_1\rangle$ and $\langle Tf_1, g_{1,\partial}\rangle$. All previous work can be used verbatim to finally prove the right estimate for $\langle TP_{2M}^{\perp}P_{M_N}f, g_1\rangle$, the first term in (45).

To deal with the second term in (45), $\langle Tf_1, g_{1,\partial}\rangle$, we note first that the reasoning to estimate D_1, N_i, and P_i can be applied unchanged to this new case. For the term B_6, we implement a small change. Since B_6 is completely symmetrical with respect to the cubes I and J, we can switch the roles played by these cubes:

$$B_6 = \sum_{J\in\mathcal{D}(Q)_{\ge N}}\sum_{I\in I_\theta}\langle f,\psi_I\rangle\langle g,\psi_J\rangle\langle T\psi_I,\psi_J\rangle.$$

We now add and subtract the term

$$\begin{aligned}A' = &\sum_{J\in\mathcal{D}(Q)_{\ge N}}\sum_{I\in J_{\mathrm{over}}}\langle f,\psi_I\rangle\langle g,\psi_J\rangle\langle T\psi_I,\psi_J\rangle\\ &+\sum_{J\in\mathcal{D}(Q)_{\ge N}}\sum_{I\in J_\theta}\sum_{\substack{I'\in\mathrm{ch}(I_p)\\ I'\notin J_\theta}}\langle f,\psi_{I'}\rangle\langle g,\psi_J\rangle\langle T\psi_{I'},\psi_J\rangle\\ &+\sum_{I\in\mathcal{D}(Q)_{\ge N}}\sum_{J\in I_\theta}\sum_{\substack{J'\in\mathrm{ch}(J_p)\\ J'\notin I_\theta}}\langle f,\psi_I\rangle\langle g,\psi_{J'}\rangle\langle T\psi_I,\psi_{J'}\rangle,\end{aligned}$$

which, for the same reasons as before, satisfies $|A'| \lesssim \|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)}$. Then we rewrite the previous reasoning to obtain

$$\begin{aligned}|B_6 - A'| &\lesssim \sum_{I\in\mathcal{D}(Q)_N}\langle f\rangle_I\Big\langle T\chi_I, \sum_{J\in\mathcal{D}(Q)_{\ge N}\cap I_\theta}\langle g,\psi_J\rangle\psi_J\Big\rangle\\ &= \sum_{I\in\mathcal{D}(Q)_N}\sum_{i=1}^{3^n}\langle f\rangle_I\langle g\rangle_{I_i}\langle T\chi_I,\chi_{I_i}\rangle - \sum_{I\in\mathcal{D}(Q)_N}\sum_{i=1}^{3^n}\langle f\rangle_I\langle g\rangle_{I_{i,k_i}}\langle T\chi_I,\chi_{I_{i,k_i}}\rangle\\ &= S_1 - S_2.\end{aligned}$$

Again, we have that $S_1 = 0$, while we can reparametrize the sums in S_2 as we did before, to get

$$|S_2| \lesssim \|f\|_{L^\infty(\mu)}\|g\|_{L^\infty(\mu)}\sum_{k=0}^{N_0-1}\sum_{j=0}^{2^{kn}}\sum_{a,b=0}^{1}|\langle T\chi_{S_{a,b}},\chi_{J_{j,k}}\rangle|.$$

Now we note that the cubes $I \in \tilde{\mathcal{D}}$ are open and so N can be chosen large enough so that

$$|S_2| \lesssim \mu(\tilde{S})^{\frac{1}{2}} \|f\|_{L^\infty(\mu)} \|g\|_{L^\infty(\mu)} 2^{N_0(n+3)} < \epsilon.$$

This ends the estimate for $\langle Tf_1, g_{1,\partial} \rangle$.

10.8. Recursion process. For the last term in (45) $\langle Tf_{1,\partial}, g_{1,\partial} \rangle$ we iterate the previous argument. We first note that the supports of $f_{1,\partial}$ and $g_{1,\partial}$ are contained in the union of ∂I for all $I \in \mathcal{D}_1(Q)$ with $\ell(I) \geq 2^{-N}\ell(Q)$. This set, which we denote by $\partial\mathcal{D}_1(Q)$, consists of the intersection with Q of finitely many Euclidean affine spaces of dimension $n-1$, which are either pairwise parallel or pairwise perpendicular.

Now let $\mathcal{D} = \mathcal{T}_2\mathcal{D}$. We consider the families of cubes $\mathcal{T}_2\mathcal{D}(Q)$ and $\partial\mathcal{D}_2(Q)$ as the union of ∂I for all $I \in \mathcal{T}_2\mathcal{D}(Q)$ with $\ell(I) \geq 2^{-N}\ell(Q)$. We then decompose $f_{1,\partial} = f_2 + f_{2,\partial}$, where $f_{2,\partial} = f_{1,\partial}\chi_{\partial\mathcal{D}_2(Q)}$ and similarly for $g_{1,\partial}$. Now, using the Haar wavelet system $(\psi_I)_{I \in \mathcal{T}_2\mathcal{D}}$ we decompose as before:

$$\langle Tf_{1,\partial}, g_{1,\partial} \rangle = \langle Tf_{1,\partial}, g_2 \rangle + \langle Tf_2, g_{2,\partial} \rangle + \langle Tf_{2,\partial}, g_{2,\partial} \rangle.$$

Then we can apply all the previous work to estimate the first two terms.

For the third term, we note that the supports of $f_{2,\partial}$ and $g_{2,\partial}$ are now contained in $\partial\mathcal{D}_1(Q) \cap \partial\mathcal{D}_2(Q)$. We also note that $\partial\mathcal{D}_2(Q)$ consists of the intersection with Q of finitely many Euclidean affine spaces of dimension $n-1$, which are either pairwise parallel or pairwise perpendicular, and also either parallel or perpendicular to every affine space of dimension $n-1$ of $\partial\mathcal{D}_1(Q)$. Then $\partial\mathcal{D}_1(Q) \cap \partial\mathcal{D}_2(Q)$ is a set consisting of finitely many Euclidean affine spaces of dimension $n-2$.

Then, by repeating the same argument $k = n - \lfloor\alpha\rfloor + \delta(\alpha - \lfloor\alpha\rfloor)$ times in total, we obtain $P_{2M}^\perp P_{M_N} f = \sum_{i=1}^k f_i + f_{i,\partial}$ and likewise for $P_M^\perp P_{M_N} g$ such that the appropriate estimates hold for $|\langle Tf_i, \cdot \rangle|$ and $|\langle \cdot, T^* g_i \rangle|$ for all with $i \in \{1, \ldots, k\}$ and the functions $f_{k,\partial}$, $g_{k,\partial}$, are supported on $\bigcap_{i=1}^k \partial\mathcal{D}_i(Q)$. By repeating the previous reasoning on parallel and perpendicular affine spaces, we conclude that this set consists of finitely many Euclidean affine spaces of dimension $n-k = \lfloor\alpha\rfloor - \delta(\alpha - \lfloor\alpha\rfloor)$, which are either pairwise parallel or pairwise perpendicular.

But now we can show that $\bigcap_{i=1}^k \partial\mathcal{D}_i(Q)$ has measure zero with respect to μ. Let I be an arbitrary $n-k$-dimensional dyadic cube with side length $\ell(I)$. Let $(J_i)_{i=1}^m$ be a family of pairwise disjoint n-dimensional cubes J_i with fixed side length r such that $I \cap J_i \neq \emptyset$ and $I \subset \bigcup_{i=1}^m J_i$. This family has cardinality comparable to $m = \big(\frac{\ell(I)}{r}\big)^{n-k}$. Then

$$\mu(I) \leq \sum_{i=1}^m \mu(J_i) \lesssim \left(\frac{\ell(I)}{r}\right)^{n-k} r^\alpha = \ell(I)^{n-k} r^{\alpha-n+k}.$$

Since $\alpha - n + k = \alpha - \lfloor\alpha\rfloor + \delta(\alpha - \lfloor\alpha\rfloor) > 0$, we have

$$\mu(I) \lesssim \ell(I)^{n-k} \lim_{r\to 0} r^{\alpha-n+k} = 0$$

for all cubes I of dimension $n-k$. This shows that $\mu\big(\bigcap_{i=1}^k \partial\mathcal{D}_i(Q)\big) = 0$ and so $\langle Tf_{k,\partial}, g_{k,\partial} \rangle = 0$. This finishes the proof of the first part of the theorem, except for the last result in Proposition 10.3 below.

10.9. Similar estimates for cubes of different dimensions. The proof of the second part of the main Theorem 4.2 follows similar steps. As before, we first work with the classical n-dimensional dyadic grid $\mathcal{D}^n(Q) = \mathcal{D}_1(Q)$, and $\partial\mathcal{D}^n(Q)$ defined as the union of ∂I for all $I \in \mathcal{D}^n(Q)$ with $\ell(I) \geq 2^{-N}\ell(Q)$. Then we decompose $P_M^\perp P_{M_N} f = f_1 + f_{1,\partial}$, where $f_{1,\partial} = (P_M^\perp P_{M_N} f)\chi_{\partial D^n(Q)}$ and similarly for $P_M^\perp P_{M_N} g$. With this,

$$\langle TP_{2M}^\perp P_{M_N} f, P_M^\perp P_{M_N} g \rangle = \langle TP_{2M}^\perp P_{M_N} f, g_1 \rangle + \langle Tf_1, g_{1,\partial} \rangle + \langle TPf_{1,\partial}, g_{1,\partial} \rangle. \tag{56}$$

Then we use the previous reasoning to estimate the first two terms, namely, $\langle TP_{2M}^{\perp}P_{M_N}f, g_1\rangle$ and $\langle Tf_1, g_{1,\partial}\rangle$.

To control $\langle Tf_{1,\partial}, g_{1,\partial}\rangle$ we note that the supports of $f_{1,\partial}$ and $g_{1,\partial}$ are contained in $\partial\mathcal{D}^n(Q)$. We define:

- $\mathcal{D}^{n-1}(Q)$ as the family of dyadic cubes in $\partial\mathcal{D}^n(Q)$, which is a dyadic $n-1$-dimensional grid.
- $\partial\mathcal{D}^{n-1}(Q)$ as the union of ∂I for all $I \in \mathcal{D}^{n-1}(Q)$ with $\ell(I) \geq 2^{-N}\ell(Q)$.

We then decompose $f_{1,\partial} = f_2 + f_{2,\partial}$, with $f_{2,\partial} = f_{1,\partial}\chi_{\partial\mathcal{D}^{n-1}(Q)}$ and likewise for g. Similarly to before, we use the Haar wavelets $(\psi_I)_{I\in\mathcal{D}^{n-1}(Q)}$ to estimate the first two terms $\langle Tf_{1,\partial}, g_2\rangle$ and $\langle Tf_2, g_{2,\partial}\rangle$.

To control $\langle Tf_{2,\partial}, g_{2,\partial}\rangle$ we note that $f_{2,\partial}, g_{2,\partial}$ are supported on $\partial\mathcal{D}^{n-1}(Q)$ and we reiterate the process.

By repeating the same argument $k = n - \lfloor\alpha\rfloor + \delta(\alpha - \lfloor\alpha\rfloor)$ times, we obtain $P_{2M}^{\perp}P_{M_N}f = \sum_{i=1}^{k} f_i + f_{i,\partial}$ and similarly for $P_{M}^{\perp}P_{M_N}g$ such that the appropriate estimates hold for $|\langle Tf_i, \cdot\rangle|$ and $|\langle \cdot, T^*g_i\rangle|$ for all $i \in \{1, \dots, k\}$ and the functions $f_{k,\partial}$, $g_{k,\partial}$ are supported on $\partial\mathcal{D}^{n-k+1}(Q)$, for which we consider the $n-k$-dimensional grid $\mathcal{D}^{n-k}(Q)$. We prove as before that $\partial\mathcal{D}^{n-k+1}(Q)$ has measure zero with respect to μ.

Let $I \in \mathcal{D}^{n-k}(Q)$ be an arbitrary $n-k$-dimensional dyadic cube with side length $\ell(I)$. We cover I with a family of n-dimensional cubes $(J_i)_{i=1}^m$ with side length r and cardinality $m = \big(\frac{\ell(I)}{r}\big)^{n-k}$. Then again

$$\mu(I) \leq \sum_{i=1}^{m} \mu(J_i) \lesssim \left(\frac{\ell(I)}{r}\right)^{n-k} r^{\alpha} = \ell(I)^{n-k} r^{\alpha-n+k}.$$

As before, $\mu(I) \lesssim \ell(I)^{n-k}\lim_{r\to 0} r^{\alpha-n+k} = 0$ for every cube $I \in \mathcal{D}^{n-k}(Q)$ of dimension $n-k$. This shows that $\mu(\partial\mathcal{D}^{n-k+1}(Q)) = 0$ and so we have $\langle Tf_{k,\partial}, g_{k,\partial}\rangle = 0$, which finishes the proof of the second part of the theorem.

10.10. Domination by truncated operators. All work done since the start of the proof of Theorem 4.2 applies to the truncated operators $T_{\gamma,Q}$. To completely finish the result, we need Proposition 10.3, which proves that, when T is compact, the entire original operator T can be estimated above by the truncated operators. Its proof uses an argument developed in the appendix of the first chapter of [**23**], where the measure is doubling and the uncut operator T is assumed to be bounded.

Proposition 10.3. *Let $T_{\gamma,Q}$ be the uniformly bounded smooth truncated operators of Definition 5.2. Let $k = n - \lfloor\alpha\rfloor - \delta(\alpha - \lfloor\alpha\rfloor)$. Then for $f, g \in L^2(\mu)$ simple functions supported on a finite collection of cubes of $\mathcal{D}(Q)$, there exist functions $(f_i)_{i=1}^k$, $(g_i)_{i=1}^k$, $(f_{i,\partial})_{i=0}^k$, $(g_{i,\partial})_{i=1}^k$ with $\|f_i\|_{L^2(\mu)}, \|f_{i,\partial}\|_{L^2(\mu)} \leq \|f\|_{L^2(\mu)}$, and $\|g_i\|_{L^2(\mu)}, \|g_{i,\partial}\|_{L^2(\mu)} \leq \|g\|_{L^2(\mu)}$ such that for $\epsilon > 0$ there is $M_0 \in \mathbb{N}$ satisfying*

$$|\langle TP_{2M}^{\perp}P_{M_N}f, P_{M}^{\perp}P_{M_N}g\rangle| \lesssim \sup_{\gamma,Q} \sum_{i=1}^{k} |\langle T_{\gamma,Q}f_{i-1,\partial}, g_i\rangle| + |\langle T_{\gamma,Q}f_i, g_{i,\partial}\rangle| + \epsilon\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)} \tag{57}$$

for all $M > M_0$.

Furthermore, with the previous inequality and the fact that

$$|\langle T_{\gamma,Q}f_{i-1,\partial}, g_i\rangle| + |\langle T_{\gamma,Q}f_i, g_{i,\partial}\rangle| \lesssim \epsilon\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)},$$

for all $M > M_0$, we have

$$|\langle TP_{2M}^{\perp}P_{M_N}f, P_{M}^{\perp}P_{M_N}g\rangle| \lesssim \epsilon\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)}.$$

Proof: To prove (57), we fix $Q \in \mathcal{C}$ and $f, g \in L^2(\mu)$ simple functions compactly supported on $2^{-1}Q$ and with support on a finite collection of cubes of $\mathcal{D}(Q)$. Then f, g are constant on all cubes of $\mathcal{D}(Q)$ with side length smaller than or equal to a fixed arbitrary parameter. We further assume without loss of generality that f, g are supported on one quadrant of $\mathbb{R}^n$. Let $\epsilon > 0$ also be fixed.

Let $M > M_0 > 0$, $N > 0$, and M_N be the parameters fixed at the beginning of the proof of Theorem 4.2. Then we know that f, g are constant on all cubes of $\mathcal{D}(Q)$ with side length smaller than or equal to $2^{-N}\ell(Q)$. Recall that, by the way these parameters were chosen, $\|P_M^{\perp} f\|_{L^2(\mu)} < \epsilon \|f\|_{L^2(\mu)}$, similarly for g, and for $\ell(I) \le 2^{-N}\ell(Q)$ we have $F_\mu(I) + F_T(I) + F_{T^*}(I) < \epsilon$.

Since $T_{\gamma,Q}$ are uniformly bounded, by the uniform boundedness principle, there exists a sequence (γ_l) converging to zero and an operator T_0 bounded on $L^2(\mu)$ such that the operators $T_{\gamma_l,Q}$ weakly converge to T_0 in $L^2(\mu)$, namely,

$$\lim_{l\to\infty} \langle (T_{\gamma_l,Q} - T_0) f, g\rangle = 0,$$

and $\lim_{l\to\infty} \|(T_{\gamma_l,Q} - T_0) f\|_{L^2(\mu)} = 0$.

Consider the functions $P_{2M}^{\perp} P_{M_N} f$ and $P_M^{\perp} P_{M_N} g$, and recall the decompositions $P_{2M}^{\perp} P_{M_N} f = \sum_{i=1}^k f_i + f_{i,\partial}$, $P_M^{\perp} P_{M_N} g = \sum_{i=1}^k g_i + g_{i,\partial}$, given in (45) and (56) with cubes of different dimensions:

- First, $P_{2M}^{\perp} P_{M_N} f = f_1 + f_{1,\partial}$, where $f_{1,\partial} = (P_M^{\perp} P_{M_N} f)\chi_{\partial D^n(Q)}$, and $P_M^{\perp} P_{M_N} g = g_1 + g_{1,\partial}$, where $g_{1,\partial} = (P_M^{\perp} P_{M_N} g)\chi_{\partial D^n(Q)}$.
- For $i \in \{1, \dots, k\}$, $f_{i,\partial} = f_{i+1} + f_{i+1,\partial}$, where $f_{i+1,\partial} = f_{i,\partial}\chi_{\partial\mathcal{D}^{n-i+1}(Q)}$, and $g_{i,\partial} = g_{i+1} + g_{i+1,\partial}$, where $g_{i+1,\partial} = g_{i,\partial}\chi_{\partial\mathcal{D}^{n-i+1}(Q)}$.

With this, if we denote $f_{0,\partial} = P_{2M}^{\perp} P_{M_N} f$, by the recursive definitions just provided we can write

$$\langle T P_{2M}^{\perp} P_{M_N} f, P_M^{\perp} P_{M_N} g\rangle = \sum_{i=1}^{k} \langle T f_{i-1,\partial}, g_i\rangle + \langle T f_i, g_{i,\partial}\rangle + \langle T f_{k,\partial}, g_{k,\partial}\rangle, \tag{58}$$

where f_i, g_i are zero on $\partial\mathcal{D}^{n-i+1}(Q)$ and $\langle T f_{k,\partial}, g_{k,\partial}\rangle = 0$ as we saw before.

To control $\langle T f_{i-1,\partial}, g_i\rangle$, we note that by definition $f_{i-1,\partial}$ is a simple function supported on a finite collection of cubes of $\mathcal{D}^{n-i+1}(Q)$, while g_i is zero on $\partial\mathcal{D}^{n-i+1}(Q)$.

On the other hand, the control of the term $\langle T f_i, g_{i,\partial}\rangle$ requires a different consideration. We note that f_i is supported on an affine space of dimension $n-i+1$ and zero on $\partial\mathcal{D}^{n-i+1}(Q)$. Then, although $g_{i,\partial}$ is supported on an affine space of dimension $n-i$, we need to consider $g_{i,\partial}$ as a simple function supported on a finite collection of cubes of $\mathcal{D}^{n-i+1}(Q)$. This is possible since g is a simple function supported on a finite collection of cubes of $\mathcal{D}^n(Q)$.

From the previous explanation, we can write $f = f_{i-1,\partial} = \sum_{j=1}^m \langle f\rangle_{I_j}\chi_{I_j}$, with $I_j \in (\mathcal{D}^{n-i+1})^c_M(Q)$ such that $\ell(I_j) = 2^{-N}\ell(Q)$, $I_j \subset Q \subset 2^{-1}I'$, and small enough so that $F_T(I_j) < \epsilon$. Since $I_j \in \mathcal{D}^{n-i}(Q)$, the last inequality requires some considerations about the relative distance from the cube I_j to the origin of $\mathbb{R}^n$ (or its unit ball). But such considerations do not affect the property $F_T(I_j) < \epsilon$ since it can be obtained from the small side length $\ell(I_j) = 2^{-N}\ell(Q)$. We note that to apply similar ideas to $g_{i,\partial}$ we also use cubes in $(\mathcal{D}^{n-i+1})^c_M(Q)$.

For the next property, we denote each of the previous functions simply by f and g. We claim that for $i \in \{1, \dots, k\}$ and for f, g constant on all cubes of $\mathcal{D}(Q)$ with side length smaller than or equal to $2^{-N}\ell(Q)$ such that g is zero on $\partial\mathcal{D}^{n-i+1}(Q)$ there is a bounded function a_i such that

$$\langle Tf,g\rangle=\langle T_0 f,g\rangle+\langle a_i f,g\rangle, \quad \text{and} \quad |\langle a_i f,g\rangle|\lesssim \epsilon\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)}. \tag{59}$$

Similarly, $\langle f,T^*g\rangle=\langle f,T_0^*g\rangle+\langle f,b_i g\rangle$ for functions f that are zero on $\partial\mathcal{D}^{n-i+1}(Q)$, with $|\langle f,b_i g\rangle|\lesssim \epsilon\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)}$.

By assuming the claim, we can prove the statement. For each $i\in\{1,\dots,k\}$, let l_i be large enough so that

$$|\langle (T_0-T_{\gamma_{l_i},Q})f_{i-1,\partial},g_i\rangle|+|\langle (T_0-T_{\gamma_{l_i},Q})f_{i,\partial},g_{i,\partial}\rangle|<\epsilon\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)}.$$

Then

$$\begin{aligned}|\langle Tf_{i-1,\partial},g_i\rangle|&\leq |\langle T_{\gamma_{l_i},Q}f_{i-1,\partial},g_i\rangle|+|\langle (T_0-T_{\gamma_{l_i},Q})f_{i-1,\partial},g_i\rangle|+|\langle a_i f_{i-1,\partial},g_i\rangle|\\ &\leq \langle T_{\gamma_{l_i},Q}f_{i-1,\partial},g_i\rangle|+2\epsilon\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)},\end{aligned} \tag{60}$$

and likewise for $|\langle Tf_i,g_{i,\partial}\rangle|$. Then with (58) and (60), we get

$$\begin{aligned}|\langle TP_{2M}^{\perp}P_{M_N}f,P_M^{\perp}P_{M_N}g\rangle|\lesssim \sum_{i=1}^{k}\sup_{\gamma,Q}|\langle T_{\gamma,Q}f_{i-1,\partial},g_i\rangle|+|\langle T_{\gamma,Q}f_i,g_{i,\partial}\rangle|\\ +4k\epsilon\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)},\end{aligned}$$

which is comparable to the statement.

We now work to prove (59). If we denote $D=T-T_0$, we need to show that $\langle Df,g\rangle=\langle a_i f,g\rangle$. Let $I'\in\mathcal{C}$ such that $2Q\subset I'$.

(a) We first prove that, for $I\in\mathcal{D}^j(2^{-1}Q)$ with $j\in\{n,\dots,n-k+1\}$, and for all $g\in L^2(\mu)$ such that g is zero on $\partial\mathcal{D}^j(Q)$, we have

$$\langle D(\chi_I),g\rangle=\langle \chi_I D(\chi_{I'}),g\rangle. \tag{61}$$

(a.1) For this we first assume that g satisfies the additional condition $\operatorname{dist}(\sup g,I)>0$, and we prove that

$$\langle D(\chi_I),g\rangle=0=\langle \chi_I D(\chi_{I'}),g\rangle,$$

where the second equality is obvious. For this, for fixed $\epsilon'>0$, let $l_I\in\mathbb{N}$ be large enough so that $|\langle (T_{\gamma_{l_I},Q}-T_0)(\chi_I),g\rangle|<\epsilon'$ and $\operatorname{dist}(\sup g,I)>2\gamma_{l_I}$. Then

$$\langle D(\chi_I),g\rangle=\langle (T-T_{\gamma_{l_I},Q})(\chi_I),g\rangle+\langle (T_{\gamma_{l_I},Q}-T_0)(\chi_I),g\rangle.$$

Since $\sup g\cap I=\emptyset$, $x\in\sup g\subset 2^{-1}Q$, and $t\in I\subset 2^{-1}Q$, we have

$$\langle (T-T_{\gamma_{l_I},Q})(\chi_I),g\rangle=\iint_I K(t,x)\phi\Big(\frac{|t-x|}{\gamma_{l_I}}\Big)\,d\mu(t)g(x)\,d\mu(x)=0,$$

due to the facts that $\operatorname{supp}\phi\subset[-2,2]$ and $|t-x|\geq\operatorname{dist}(I,\sup g)>2\gamma_{l_I}$. Then

$$|\langle D(\chi_I),g\rangle|=|\langle (T_{\gamma_{l_I},Q}-T_0)(\chi_I),g\rangle|<\epsilon'.$$

Since the inequality holds for all $\epsilon'>0$, we conclude that $\langle D(\chi_I),g\rangle=0$ for all $g\in L^2(\mu)$ that are zero on $\partial\mathcal{D}^j(Q)$ and $\operatorname{dist}(\sup g,I)>0$.

(a.2) For the general case, we define $I^\lambda=\{x\in\mathbb{R}^n/\operatorname{dist}(x,I)<\lambda\ell(I)\}\subset(1+2\lambda)I$, and fix $\lambda\in(0,1/2)$ such that $I^\lambda\subset Q\subset I'$. Then

$$\begin{aligned}\langle D(\chi_I),g\rangle&=\langle (1-\chi_{I^\lambda})D(\chi_I),g\rangle+\langle \chi_{I^\lambda}D(\chi_I),g\rangle\\ &=\langle D(\chi_I),(1-\chi_{I^\lambda})g\rangle+\langle \chi_{I^\lambda}D(\chi_I),g\rangle.\end{aligned}$$

Since $\mathrm{dist}(\sup((1-\chi_{I^\lambda})g), I) \geq \mathrm{dist}(\sup(1-\chi_{I^\lambda}), I) \geq \lambda\ell(I) > 0$, by the previous case we have $\langle D(\chi_I), (1-\chi_{I^\lambda})g\rangle = 0$ for all $g \in L^2(\mu)$ that are zero in $\partial\mathcal{D}^j(Q)$. With this,

$$\langle D(\chi_I), g\rangle = \langle \chi_{I^\lambda} D(\chi_I), g\rangle. \tag{62}$$

Then, if we denote $I^c = I' \setminus I$, we can write

$$\langle D(\chi_I), g\rangle = \langle D(\chi_I), \chi_{I^\lambda} g\rangle = \langle D(\chi_{I'}), g\chi_{I^\lambda}\rangle - \langle D(\chi_{I^c}), g\chi_{I^\lambda}\rangle.$$

By the previous case again, we have that $\langle D(\chi_{I^c}), h\rangle = 0$ for all $h \in L^2(\mu)$ that are zero on $\partial\mathcal{D}^j(Q)$ and $\mathrm{dist}(\sup h, I^c) > 0$. Then, as in (62) applied to I^c and $g\chi_{I^\lambda}$, we have

$$\langle D(\chi_{I^c}), g\chi_{I^\lambda}\rangle = \langle \chi_{(I^c)^\lambda} D(\chi_{I^c}), g\chi_{I^\lambda}\rangle.$$

With this, for $\lambda > 0$ and $g \in L^2(\mu)$ that is zero on $\partial\mathcal{D}^j(Q)$, we get

$$\langle D(\chi_I), g\rangle = \langle \chi_{I^\lambda} D(\chi_{I'}), g\rangle - \langle \chi_{I^\lambda}\chi_{(I^c)^\lambda} D(\chi_I), g\rangle.$$

And since $\lambda > 0$ is arbitrary, we have

$$\langle D(\chi_I), g\rangle = \lim_{\lambda\to 0}\langle \chi_{I^\lambda} D(\chi_{I'}), g\rangle - \lim_{\lambda\to 0}\langle \chi_{I^\lambda}\chi_{(I^c)^\lambda} D(\chi_I), g\rangle. \tag{63}$$

We now work with each term separately.

(a.2.1) For the first term, we reason as follows. By the testing condition on T and the boundedness of T_0, we have $\|\chi_{I'} D(\chi_{I'})\|_{L^2(\mu)} \lesssim \mu(I')^{\frac{1}{2}}$. Then $D(\chi_{I'})g$ is integrable on I'. Moreover, $I^\lambda \subset I'$, $|\chi_{I^\lambda} D(\chi_{I'})g| \leq \chi_{I'}|D(\chi_{I'})g| \in L^1(\mu)$, and $\lim_{\lambda\to 0}\chi_{I^\lambda} = \chi_{\bar I}$, where $\bar I \in \mathcal{C}$ is the closed cube defined by the closure of I. Then, by Lebesgue's dominated convergence theorem,

$$\lim_{\lambda\to 0}\langle \chi_{I^\lambda} D(\chi_{I'}), g\rangle = \langle \chi_{\bar I} D(\chi_{I'}), g\rangle, = \langle \chi_I D(\chi_{I'}), g\rangle. \tag{64}$$

The last equality holds because g is zero on $\partial\mathcal{D}^j(Q)$ and so on ∂I.

(a.2.2) For the second term in (63), we work differently. Let $I_\lambda = (1-2\lambda)I$. We note that $\chi_{I^\lambda}\chi_{(I^c)^\lambda} = \chi_{I^\lambda\setminus \bar I} + \chi_{\bar I\setminus I_\lambda}$. Then

$$\langle \chi_{I_\lambda}\chi_{(I^c)_\lambda} D(\chi_I), g\rangle = \langle \chi_{I_\lambda\setminus \bar I} D(\chi_I), g\rangle + \langle \chi_{I\setminus I_\lambda} D(\chi_I), g\rangle, \tag{65}$$

where we can write I instead of $\bar I$ because g is zero on $\partial\mathcal{D}^j(Q)$.

- The second new term can be treated as before: $\|\chi_I D(\chi_I)\|_{L^2(\mu)} \lesssim \mu(I)^{\frac{1}{2}}$ and so $D(\chi_I)g$ is integrable on I. Moreover, $I \setminus I_\lambda \subset I$, $|\chi_{I\setminus I_\lambda} D(\chi_I)g| \leq \chi_I|D(\chi_I)g| \in L^1(\mu)$, and $\lim_{\lambda\to 0}\chi_{I\setminus I_\lambda} = \chi_A$, with $A \subseteq \partial I$. Then, by Lebesgue's dominated convergence theorem, we have

$$\lim_{\lambda\to 0}\langle \chi_{I\setminus I_\lambda} D(\chi_I), g\rangle = \langle \chi_A D(\chi_I), g\rangle = 0,$$

since g is zero on $A \subseteq \partial\mathcal{D}^j(Q)$.
- For the first term in (65), we proceed as follows. Let $S_r = \{x \in I^\lambda \setminus \bar I : 2^{-(r+1)}\ell(I^\lambda \setminus \bar I) < \mathrm{dist}(x, I) \leq 2^{-r}\ell(I^\lambda \setminus \bar I)\}$. Then since $I^\lambda \setminus \bar I = \bigcup_{r=0}^\infty S_r$,

$$\langle \chi_{I^\lambda\setminus \bar I} D(\chi_I), g\rangle = \Big\langle \lim_{R\to\infty}\sum_{r=0}^R \chi_{S_r} D(\chi_I), g\Big\rangle.$$

By Fatou's lemma,

$$
\begin{aligned}
|\langle \chi_{I^\lambda\setminus \bar{I}} D(\chi_I), g\rangle| &\leq \Big\langle \lim_{R\to\infty} \sum_{r=0}^{R} \chi_{S_r} |D(\chi_I)|, |g| \Big\rangle \\
&\leq \liminf_{R\to\infty} \sum_{r=0}^{R} \langle \chi_{S_r} |D(\chi_I)|, |g|\rangle.
\end{aligned} \tag{66}
$$

Given $\epsilon' > 0$, we choose l_r, which depends on r, λ, ϵ', I, and g, such that $2\gamma_{l_r} < 2^{-(r+1)}\ell(I_\lambda \setminus \bar{I})$ and $\|(T_{\gamma_{l_r},Q} - T_0)(\chi_I)\|_{L^2(\mu)}\|g\|_{L^2(\mu)} < \epsilon' 2^{-r}$. Then

$$\langle \chi_{S_r}|D(\chi_I)|, |g|\rangle \leq \langle \chi_{S_r}|(T - T_{\gamma_{l_r},Q})(\chi_I)|, |g|\rangle + \langle \chi_{S_r}|(T_{\gamma_{l_r},Q} - T_0)(\chi_I), |g|\rangle.$$

The second term can be bounded by

$$\|(T_{\gamma_{l_r},Q} - T_0)(\chi_I)\|_{L^2(\mu)}\|g\|_{L^2(\mu)} < \epsilon' 2^{-r}.$$

For the first term, since $S_r \cap I = \emptyset$,

$$\langle \chi_{S_r}|(T - T_{\gamma_{l_r},Q})(\chi_I)|, |g|\rangle = \int_{S_r} \left| \int_I K(t,x)\phi\Big(\frac{|t-x|}{\gamma_{l_r}}\Big)\, d\mu(t)\right| |g(x)|\, d\mu(x) = 0.$$

In the last equality we have used that $2\gamma_{l_r} < 2^{-(r+1)}\ell(I^\lambda \setminus \bar{I}) < \mathrm{dist}(x, I) \leq |t-x|$.

With both estimates we continue the estimate in (66) as

$$|\langle \chi_{I^\lambda\setminus\bar{I}} D(\chi_I), g\rangle| \leq \epsilon' \sum_{r=0}^{\infty} 2^{-r} \lesssim \epsilon'$$

for all $\epsilon' > 0$. Then $\langle \chi_{I^\lambda\setminus\bar{I}} D(\chi_I), g\rangle = 0$.

- By combining the decomposition (63), the equality (64), and the subsequent decomposition (65) with both terms zero, we get $\langle D(\chi_I), g\rangle = \langle \chi_I D(\chi_{I'}), g\rangle$, which is the equality claimed in (61).

(b) Now we use (61) to prove (59). By hypothesis, f, g are simple functions supported on a finite collection of cubes of $\mathcal{D}^n(Q)$ of side length $2^{-N}\ell(Q)$.

Let $f = \sum_{j=1}^{m} \langle f\rangle_{I_j}\chi_{I_j}$ with $I_j \in (\mathcal{D}^{n-i+1})^c_M(Q)$ such that $\ell(I_j) = 2^{-N}\ell(Q)$, $I_j \subset Q \subset 2^{-1}I'$ and small enough so that $F_T(I_j) < \epsilon$. Then, by (61),

$$\langle D(f), g\rangle = \sum_{j=1}^{m} \langle f\rangle_{I_j}\langle D(\chi_{I_j}), g\rangle = \sum_{j=1}^{m} \langle f\rangle_{I_j}\langle \chi_{I_j} D(\chi_{I'}), g\rangle = \langle f D(\chi_{I'}), g\rangle = \langle a_i f, g\rangle,$$

with $a_i = D(\chi_{I'})$. This proves equality (59).

Now we show that $|\langle a_i f, g\rangle| \lesssim \epsilon \|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)}$. Let $I, J \in \mathcal{D}^{n-i+1}(Q)$, $\tilde{J} \in \tilde{\mathcal{D}}^{n-i+1}(Q)$, with $\ell(I) = \ell(J) = 2^{-N}\ell(Q)$ and such that $\tilde{J}$ is the interior of J. By the definition of a_i and (61) again, we have

$$\langle a_i\chi_I, \chi_{\tilde{J}}\rangle = \langle \chi_{I\cap J} D(\chi_{I'}), \chi_{\tilde{J}}\rangle = \langle D(\chi_{I\cap J}), \chi_{\tilde{J}}\rangle. \tag{67}$$

If $I \cap J = \emptyset$, we get $\langle a_i\chi_I, \chi_{\tilde{J}}\rangle = 0$. Otherwise, since $\ell(I) = \ell(J)$, we have $I = J$. Then, by (67), the testing condition on T and the compactness of T_0,

$$
\begin{aligned}
|\langle a_i\chi_I, \chi_{\tilde{J}}\rangle| &= |\langle D\chi_J, \chi_{\tilde{J}}\rangle| \leq |\langle T\chi_J, \chi_{\tilde{J}}\rangle| + |\langle T_0\chi_J, \chi_{\tilde{J}}\rangle| \\
&\lesssim (\|\chi_J T\chi_J\|_{L^2(\mu)} + \|T_0\chi_J\|_{L^2(\mu)})\mu(\tilde{J})^{\frac{1}{2}} \\
&\lesssim F_T(J)\mu(J)^{\frac{1}{2}}\mu(\tilde{J})^{\frac{1}{2}} < \epsilon\mu(J)^{\frac{1}{2}}\mu(\tilde{J})^{\frac{1}{2}}.
\end{aligned}
$$

We now write $f = \sum_{I\in\mathcal{D}^{n-i+1}(Q)_N}\langle f\rangle_I\chi_I$ and $g = \sum_{\tilde{I}\in\tilde{\mathcal{D}}^{n-i+1}(Q)_N}\langle g\rangle_{\tilde{I}}\chi_{\tilde{I}}$, where $F_T(I) < \epsilon$, and $\tilde{I}$ is the interior of I so that g is zero on ∂I. We have just seen that $\langle a_i\chi_I,\chi_{\tilde{J}}\rangle = 0$ for $I \neq J$. Then

$$\begin{aligned}
|\langle a_i f, g\rangle| &= \Bigg|\sum_{I\in\mathcal{D}^{n-i+1}(Q)_N}\ \sum_{\tilde{J}\in\tilde{\mathcal{D}}^{n-i+1}(Q)_N}\langle f\rangle_I\langle g\rangle_{\tilde{J}}\langle a_i\chi_I,\chi_{\tilde{J}}\rangle\Bigg| \\
&\le \sum_{I\in\mathcal{D}^{n-i+1}(Q)_N}|\langle f\rangle_I||\langle g\rangle_{\tilde{I}}||\langle a_i\chi_I,\chi_{\tilde{I}}\rangle| \\
&\lesssim \epsilon\sum_{I\in\mathcal{D}^{n-i+1}(Q)_N}|\langle f\rangle_I||\langle g\rangle_{\tilde{I}}|\mu(I)^{\frac{1}{2}}\mu(\tilde{I})^{\frac{1}{2}} \\
&\lesssim \epsilon\Bigg(\sum_{I\in\mathcal{D}^{n-i+1}(Q)_N}|\langle f\rangle_I|^2\mu(I)\Bigg)^{\frac{1}{2}}\Bigg(\sum_{\tilde{I}\in\tilde{\mathcal{D}}^{n-i+1}(Q)_N}|\langle g\rangle_{\tilde{I}}|^2\mu(\tilde{I})\Bigg)^{\frac{1}{2}} \\
&\lesssim \epsilon\|f\|_{L^2(\mu)}\|g\|_{L^2(\mu)}.
\end{aligned}$$

□

References

[1] P. AUSCHER, S. HOFMANN, C. MUSCALU, T. TAO, AND C. THIELE, Carleson measures, trees, extrapolation, and T(b) theorems, *Publ. Mat.* **46(2)** (2002), 257–325. DOI: 10.5565/PUBLMAT_46202_01.

[2] P. AUSCHER AND E. ROUTIN, Local Tb theorems and Hardy inequalities, *J. Geom. Anal.* **23(1)** (2013), 303–374. DOI: 10.1007/s12220-011-9249-1.

[3] M. CHRIST, A $T(b)$ theorem with remarks on analytic capacity and the Cauchy integral, *Colloq. Math.* **60/61(2)** (1990), 601–628. DOI: 10.4064/cm-60-61-2-601-628.

[4] M. CHRIST, *Lectures on Singular Integral Operators*, CBMS Regional Conf. Ser. in Math. **77**, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1990. DOI: 10.1090/CBMS/077.

[5] G. DAVID, *Wavelets and Singular Integrals on Curves and Surfaces*, Lecture Notes in Math. **1465**, Springer-Verlag, Berlin, 1991. DOI: 10.1007/BFb0091544.

[6] E. B. FABES, M. JODEIT, JR., AND N. M. RIVIÈRE, Potential techniques for boundary value problems on C^1-domains, *Acta Math.* **141(3-4)** (1978), 165–186. DOI: 10.1007/BF02545747.

[7] S. HOFMANN, On singular integrals of Calderón-type in $\mathbb{R}^n$, and BMO, *Rev. Mat. Iberoamericana* **10(3)** (1994), 467–505. DOI: 10.4171/RMI/159.

[8] S. HOFMANN, A proof of the local Tb Theorem for standard Calderón–Zygmund operators, Preprint (2007). arXiv:0705.0840.

[9] S. HOFMANN, Local $T(b)$ theorems and applications in PDE, in: *Harmonic Analysis and Partial Differential Equations*, Contemp. Math. **505**, American Mathematical Society, Providence, RI, 2010, pp. 29–52. DOI: 10.1090/conm/505/09914.

[10] T. P. HYTÖNEN, An operator-valued Tb theorem, *J. Funct. Anal.* **234(2)** (2006), 420–463. DOI: 10.1016/j.jfa.2005.11.001.

[11] T. P. HYTÖNEN, The sharp weighted bound for general Calderón–Zygmund operators, *Ann. of Math. (2)* **175(3)** (2012), 1473–1506. DOI: 10.4007/annals.2012.175.3.9.

[12] T. HYTÖNEN AND H. MARTIKAINEN, On general local Tb theorems, *Trans. Amer. Math. Soc.* **364(9)** (2012), 4819–4846. DOI: 10.1090/S0002-9947-2012-05599-1.

[13] T. HYTÖNEN AND F. NAZAROV, The local Tb theorem with rough test functions, *Adv. Math.* **372** (2020), 107306, 36 pp. DOI: 10.1016/j.aim.2020.107306.

[14] T. P. HYTÖNEN AND A. V. VÄHÄKANGAS, The local non-homogeneous Tb theorem for vector-valued functions, *Glasg. Math. J.* **57(1)** (2015), 17–82. DOI: 10.1017/S0017089514000123.

[15] C. E. KENIG, *Harmonic Analysis Techniques for Second Order Elliptic Boundary Value Problems*, CBMS Regional Conf. Ser. in Math. **83**, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1994. DOI: 10.1090/cbms/083.

[16] M. T. Lacey and A. V. Vähäkangas, The perfect local Tb theorem and twisted martingale transforms, *Proc. Amer. Math. Soc.* **142(5)** (2014), 1689–1700. DOI: 10.1090/S0002-9939-2014-11930-7.

[17] H. Martikainen, M. Mourgoglou, and E. Vuorinen, A new approach to non-homogeneous local Tb theorems, *J. Anal. Math.* **143(1)** (2021), 95–121. DOI: 10.1007/s11854-021-0147-6.

[18] F. Nazarov, S. Treil, and A. Volberg, Cauchy integral and Calderón–Zygmund operators on nonhomogeneous spaces, *Internat. Math. Res. Notices* **1997(15)** (1997), 703–726. DOI: 10.1155/S1073792897000469.

[19] F. Nazarov, S. Treil, and A. Volberg, Weak type estimates and Cotlar inequalities for Calderón–Zygmund operators on nonhomogeneous spaces, *Internat. Math. Res. Notices* **1998(9)** (1998), 463–487. DOI: 10.1155/S1073792898000312.

[20] F. Nazarov, S. Treil, and A. Volberg, Accretive system Tb-theorems on nonhomogeneous spaces, *Duke Math. J.* **113(2)** (2002), 259–312. DOI: 10.1215/S0012-7094-02-11323-4.

[21] F. Nazarov, S. Treil, and A. Volberg, The Tb-theorem on non-homogeneous spaces, *Acta Math.* **190(2)** (2003), 151–239. DOI: 10.1007/BF02392690.

[22] F. Nazarov, S. Treil, and A. Volberg, The Tb-theorem on non-homogeneous spaces that proves a conjecture of Vitushkin, Preprint (2014). arXiv:1401.2479v2.

[23] E. M. Stein, *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, With the assistance of Timothy S. Murphy, Princeton Math. Ser. **43**, Monogr. Harmon. Anal. III, Princeton University Press, Princeton, NJ, 1993. DOI: 10.1515/9781400883929.

[24] X. Tolsa, *Analytic Capacity, the Cauchy Transform, and Non-homogeneous Calderón–Zygmund Theory*, Progr. Math. **307**, Birkhäuser/Springer, Cham, 2014. DOI: 10.1007/978-3-319-00596-6.

[25] P. Villarroya, A characterization of compactness for singular integrals, *J. Math. Pures Appl. (9)* **104(3)** (2015), 485–532. DOI: 10.1016/j.matpur.2015.03.006.

[26] P. Villarroya, A global Tb theorem for compactness and boundedness of Calderón–Zygmund operators, *J. Math. Anal. Appl.* **480(1)** (2019), 123323, 41 pp. DOI: 10.1016/j.jmaa.2019.07.013.

Department of Applied Mathematics, Santa Clara University, Santa Clara, CA 95053
E-mail address: paco.villarroya@scu.edu

Received on October 19, 2022.
Accepted on January 29, 2024.