PSEUDO-LOCALIZATION OF SINGULAR INTEGRALS AND NONCOMMUTATIVE CALDERÓN-ZYGMUND THEORY

JAVIER PARCET

Contents

Introduction

After the pioneer work of Calderón and Zygmund in the 50's, the systematic study of singular integrals has become a corner stone in harmonic analysis with deep implications in mathematical physics, partial differential equations and other mathematical disciplines. Subsequent generalizations of Calderón-Zygmund theory have essentially pursued two lines. We may either consider more general domains or ranges for the functions considered. In the first case, the Euclidean space is replaced by metric spaces equipped with a doubling or non-doubling measure of polynomial growth. In the second case, the real or complex fields are replaced by a Banach space in which martingale differences are unconditional. Historically, the study of singular integrals acting on matrix or operator valued functions has been considered part of the vector-valued theory. This is however a limited approach in the noncommutative setting and we propose to regard these functions as operators in a suitable von Neumann algebra, generalizing so the domain and not the range of classical functions. A far reaching aspect of our approach is the stability of the product fg and the absolute value $|f| = \sqrt{f^*f}$ for operator-valued functions, a fundamental property not exploited in the vector theory. In this paper we follow the original Calderón-Zygmund program and present a non-commutative scalar-valued Calderón-Zygmund theory, emancipated from the vector theory.

Partially supported by 'Programa Ramón y Cajal, 2005' and
also by Grants MTM2004-00678 and CCG06-UAM/ESP-0286, Spain.
2000 Mathematics Subject Classification: 42B20, 42B25, 46L51, 46L52, 46L53.
Key words: Calderón-Zygmund operator, almost orthogonality, noncommutative martingale.

Noncommutative harmonic analysis (understood in a wide sense) has received much attention in recent years. The functional analytic approach given by operator space theory and the new methods from quantum/free probability have allowed to study a great variety of topics. We find in the recent literature noncommutative analogs of Khintchine and Rosenthal inequalities, a settled noncommutative theory of martingale inequalities, new results on Fourier/Schur multipliers, matrix A_p weights and a sharpened Carleson embedding theorem, see [20, 29, 31, 43, 47, 57] and the references therein. However, no essential progress has been made in the context of singular integral operators.

Our original motivation was the weak type boundedness of Calderón-Zygmund operators acting on operator-valued functions, a well-known problem which has remained open since the beginning of the vector-valued theory in the 80's. This fits in the context of Mei's recent paper [36]. Our main tools for its solution are two. On one hand, the failure of some classical estimates in the noncommutative setting forces us to have a deep understanding of how the L_2-mass of a singular integral is concentrated around the support of the function on which it acts. To that aim, we have developed a *pseudo-localization principle* for singular integrals which is of independent interest, even in the classical theory. This is used in conjunction with a *noncommutative form of Calderón-Zygmund decomposition* which we have constructed using the theory of noncommutative martingales. As a byproduct of our weak type inequality, we obtain the sharp asymptotic behavior of the constants for the strong L_p inequalities as $p \to 1$ and $p \to \infty$, which are not known. At the end of the paper we generalize our results to certain singular integrals including operator-valued kernels and functions at the same time. A deep knowledge of this kind of *fully noncommutative* operators is a central aim in noncommutative harmonic analysis. Our methods in this paper open a door to work in the future with more general classes of operators.

1. Terminology. Let us set some notation that will remain fixed all through out the paper. Let $\mathcal{M}$ be a semifinite von Neumann algebra equipped with a normal semifinite faithful trace τ. Let us consider the algebra $\mathcal{A}_B$ of essentially bounded $\mathcal{M}$-valued functions

$$\mathcal{A}_B = \Big\{ f : \mathbb{R}^n \to \mathcal{M} \, \Big| \, f \text{ strongly measurable s.t. } \operatorname*{ess\,sup}_{x \in \mathbb{R}^n} \|f(x)\|_{\mathcal{M}} < \infty \Big\},$$

equipped with the *n.s.f.* trace

$$\varphi(f) = \int_{\mathbb{R}^n} \tau(f(x))\, dx.$$

The weak-operator closure $\mathcal{A}$ of $\mathcal{A}_B$ is a von Neumann algebra. If $1 \le p \le \infty$, we write $L_p(\mathcal{M})$ and $L_p(\mathcal{A})$ for the noncommutative L_p spaces associated to the pairs $(\mathcal{M}, \tau)$ and $(\mathcal{A}, \varphi)$. The lattices of projections are written $\mathcal{M}_\pi$ and $\mathcal{A}_\pi$, while $\mathbf{1}_\mathcal{M}$ and $\mathbf{1}_\mathcal{A}$ stand for the unit elements.

The set of dyadic cubes in $\mathbb{R}^n$ is denoted by $\mathcal{Q}$. The size of any cube Q in $\mathbb{R}^n$ is defined as the length $\ell(Q)$ of one of its edges. Given an integer $k \in \mathbb{Z}$, we use $\mathcal{Q}_k$ for the subset of $\mathcal{Q}$ formed by cubes Q of the k-th generation, i.e. those of size $1/2^k$. If Q is a dyadic cube and $f : \mathbb{R}^n \to \mathcal{M}$ is integrable on Q, we set the average

$$f_Q = \frac{1}{|Q|} \int_Q f(y)\, dy.$$

Let us write $(\mathsf{E}_k)_{k\in\mathbb{Z}}$ for the family of conditional expectations associated to the classical dyadic filtration on $\mathbb{R}^n$. E_k will also stand for the tensor product $\mathsf{E}_k \otimes id_{\mathcal{M}}$ acting on $\mathcal{A}$. If $1 \le p < \infty$ and $f \in L_p(\mathcal{A})$

$$\mathsf{E}_k(f) = f_k = \sum_{Q\in\mathcal{Q}_k} f_Q 1_Q.$$

We shall denote by $(\mathcal{A}_k)_{k\in\mathbb{Z}}$ the corresponding filtration $\mathcal{A}_k = \mathsf{E}_k(\mathcal{A})$.

If $Q \in \mathcal{Q}$, its dyadic father $\widehat{Q}$ is the only dyadic cube containing Q with double size. Given $\delta > 1$, the δ-concentric father of Q is the only cube δQ concentric with the cube Q and such that $\ell(\delta Q) = \delta \ell(Q)$. In this paper we will mainly work with dyadic and 9-concentric fathers. Note that in the classical theory 2-concentric fathers are typically enough. We shall write just L_p to refer to the commutative L_p space on $\mathbb{R}^n$ equipped with the Lebesgue measure dx.

2. Statement of the problem. Just to motivate our problem and for the sake of simplicity, the reader may think for the moment that $(\mathcal{M},\tau)$ is given by the pair (M_m, tr) formed by the algebra of $m \times m$ square matrices equipped with the standard trace. In this particular case, the von Neumann algebra $\mathcal{A} = \mathcal{A}_B$ becomes the space of essentially bounded matrix-valued functions. Let us consider a Calderón-Zygmund operator formally given by

$$Tf(x) = \int_{\mathbb{R}^n} k(x,y) f(y)\, dy.$$

As above, let $L_p(\mathcal{M})$ be the noncommutative L_p space associated to $(\mathcal{M},\tau)$. If $\mathcal{M}$ is the algebra of $m \times m$ matrices we recover the Schatten p-class over M_m, for a general definition see below. The first question which arises is whether or not the singular integral T is bounded on $L_p(\mathcal{A})$ for $1 < p < \infty$. The space $L_p(\mathcal{A})$ is defined as the closure of $\mathcal{A}_B$ with respect to the norm

$$\|f\|_p = \Big(\int_{\mathbb{R}^n} \tau\left(|f(x)|^p\right) dx \Big)^{\frac{1}{p}}.$$

In other words, $L_p(\mathcal{A})$ is isometric to the Bochner L_p space with values in $L_p(\mathcal{M})$. In particular, when dealing with the Hilbert transform and by a well-known result of Burkholder [5, 6], the boundedness on $L_p(\mathcal{A})$ reduces to the fact that $L_p(\mathcal{M})$ is a UMD Banach space for $1 < p < \infty$, see also [2, 4]. After some partial results of Bourgain [3], it was finally Figiel [15] who showed in 1989 (using an ingenious martingale approach) that the UMD property implies the L_p boundedness of the corresponding vector-valued singular integrals associated to generalized kernels.

The second natural question has to do with a suitable weak type inequality for $p = 1$. Namely, such inequality is typically combined in the classical theory with the real interpolation method to produce extrapolation results on the L_p boundedness of Calderón-Zygmund and other related operators. The problem of finding the right weak type inequality is subtler since arguments from the vector-valued theory are no longer at our disposal. Indeed, in terms of Bochner spaces we may generalize the previous situation by considering the mapping T from $L_1(\mathbb{R}^n; \mathrm{X})$ to $L_{1,\infty}(\mathbb{R}^n; \mathrm{X})$ with $\mathrm{X} = L_1(\mathcal{M})$. However, $L_1(\mathcal{M})$ is not UMD and the resulting operator is not bounded. On the contrary, using operators rather than vectors (i.e. working directly on the algebra $\mathcal{A}$) we may consider the operator $T : L_1(\mathcal{A}) \to L_{1,\infty}(\mathcal{A})$ where $L_{1,\infty}(\mathcal{A})$ denotes the corresponding noncommutative Lorentz space, to be defined below. The only result on this line is the weak type $(1,1)$ boundedness of the Hilbert transform for operator-valued functions, proved by Randrianantoanina in [49]. He followed Kolmogorov's approach, exploiting the conjugation nature of the Hilbert transform (defined in a very wide setting via Arveson's [1] maximal subdiagonal algebras) and applying complex variable methods. As is well-known this is no longer valid for other Calderón-Zygmund operators and new real variable methods are needed. In the classical case, these methods live around the celebrated Calderón-Zygmund decomposition. One of the main purposes of this paper is to supply the right real variable methods in the noncommutative context. As we will see, there are significant differences.

Using real interpolation, our main result gives an extrapolation method which produces the L_p boundedness results discussed in the paragraph above and provides the sharp asymptotic behavior of the constants, for which the UMD approach is inefficient. Moreover, when working with operator-valued kernels we obtain new strong L_p inequalities. We should warn the reader not to confuse this setting with that of Rubio de Francia, Ruiz and Torrea [53], Hytönen [21] and Hytönen/Weis [22, 23], where the mentioned limitations of the vector-valued theory appear.

3. Calderón-Zygmund decomposition. Let us recall the formulation of the classical decomposition for scalar-valued integrable functions. If $f \in L_1$ is positive and $\lambda \in \mathbb{R}_+$, we consider the level set

$$\mathrm{E}_\lambda = \Big\{ x \in \mathbb{R}^n \, \big| \, M_d f(x) > \lambda \Big\},$$

where the dyadic Hardy-Littlewood maximal function $M_d f$ is greater than λ. If we write $\mathrm{E}_\lambda = \bigcup_j Q_j$ as a disjoint union of maximal dyadic cubes, we may decompose $f = g + b$ where the good and bad parts are given by

$$g = f 1_{\mathrm{E}_\lambda^c} + \sum\nolimits_j f_{Q_j} 1_{Q_j} \quad \text{and} \quad b = \sum\nolimits_j (f - f_{Q_j}) 1_{Q_j}$$

Letting $b_j = (f - f_{Q_j}) 1_{Q_j}$, we have

i) $\|g\|_1 \le \|f\|_1$ and $\|g\|_\infty \le 2^n \lambda$.
ii) $\operatorname{supp} b_j \subset Q_j$, $\int_{Q_j} b_j = 0$ and $\sum_j \|b_j\|_1 \le 2\|f\|_1$.

These properties are crucial for the analysis of singular integral operators.

In this paper we use the so-called Cuculescu's construction [9] to produce a sequence $(p_k)_{k \in \mathbb{Z}}$ of disjoint projections in $\mathcal{A}$ which constitute the noncommutative counterpart of the characteristic functions supported by the sets

$$\mathrm{E}_\lambda(k) = \bigcup_{\substack{Q_j \subset \mathrm{E}_\lambda \\ \ell(Q_j) = 1/2^k}} Q_j.$$

Cuculescu's construction will be properly introduced in the text. It has proved to be the right tool from the theory of noncommutative martingales to deal with inequalities of weak type. Indeed, Cuculescu proved in [9] the noncommutative Doob's maximal weak type inequality. Moreover, these techniques were used by Randrianantoanina to prove several weak type inequalities for noncommutative martingales [50, 51, 52] and by Junge and Xu in their remarkable paper [31]. In fact, a strong motivation for this paper relies on [44], where similar methods were applied to obtain Gundy's decomposition for noncommutative martingales. It is well-known that the probabilistic analog of Calderón-Zygmund decomposition is precisely Gundy's decomposition. However, in contrast to the classical theory, the noncommutative analogue of Calderón-Zygmund decomposition turns out to be much harder than Gundy's decomposition. Although we shall justify this below in further detail, the main reason is that singular integral operators do not localize the support of the function on which it acts, something that happens for instance with martingale transforms or martingale square functions.

Let us now formulate the noncommutative Calderón-Zygmund decomposition. If $f \in L_1(\mathcal{A})_+$ and $\lambda \in \mathbb{R}_+$, we consider the disjoint projections $(p_k)_{k \in \mathbb{Z}}$ given by Cuculescu's construction. Let p_∞ denote the projection onto the ortho-complement of the range of $\sum_k p_k$. In particular, using the terminology $\widehat{\mathbb{Z}} = \mathbb{Z} \cup \{\infty\}$ we find the relation

$$\sum_{k \in \widehat{\mathbb{Z}}} p_k = \mathbf{1}_{\mathcal{A}}.$$

Then, the *good* and *bad* parts are given by

$$g = \sum_{i,j \in \widehat{\mathbb{Z}}} p_i f_{i \vee j} p_j \quad \text{and} \quad b = \sum_{i,j \in \widehat{\mathbb{Z}}} p_i (f - f_{i \vee j}) p_j,$$

with $i \vee j = \max(i,j)$. We will show how this generalizes the classical decomposition.

4. Main weak type inequality. Let Δ denote the diagonal of $\mathbb{R}^n \times \mathbb{R}^n$. We will write in what follows T to denote a linear map $\mathcal{S} \to \mathcal{S}'$ from test functions to distributions which is associated to a given kernel $k : \mathbb{R}^{2n} \setminus \Delta \to \mathbb{C}$. This means that for any smooth test function f with compact support, we have

$$Tf(x) = \int_{\mathbb{R}^n} k(x,y) f(y) \, dy \quad \text{for all} \quad x \notin \operatorname{supp} f.$$

Given two points $x, y \in \mathbb{R}^n$, the distance $|x-y|$ between x and y will be taken for convenience with respect to the $\ell_\infty(n)$ metric. As usual, we impose size and smoothness conditions on the kernel:

a) If $x, y \in \mathbb{R}^n$, we have

$$|k(x,y)| \lesssim \frac{1}{|x-y|^n}.$$

b) There exists $0 < \gamma \le 1$ such that

$$\begin{aligned} \big|k(x,y) - k(x',y)\big| &\lesssim \frac{|x-x'|^\gamma}{|x-y|^{n+\gamma}} \quad \text{if} \quad |x-x'| \le \frac{1}{2}|x-y|, \\ \big|k(x,y) - k(x,y')\big| &\lesssim \frac{|y-y'|^\gamma}{|x-y|^{n+\gamma}} \quad \text{if} \quad |y-y'| \le \frac{1}{2}|x-y|. \end{aligned}$$

We will refer to this γ as the Lipschitz smoothness parameter of the kernel.

Theorem A. *Let T be a generalized Calderón-Zygmund operator associated to a kernel satisfying the size and smoothness estimates above. Assume that T is bounded on L_q for some $1 < q < \infty$. Then, given any $f \in L_1(\mathcal{A})$, the estimate below holds for some constant $\mathrm{c}_{n,\gamma}$ depending only on the dimension n and the Lipschitz smoothness parameter γ*

$$\sup_{\lambda > 0} \lambda \, \varphi \Big\{ |Tf| > \lambda \Big\} \le \mathrm{c}_{n,\gamma} \, \|f\|_1.$$

In particular, given $1 < p < \infty$ *and* $f \in L_p(\mathcal{A})$*, we find*

$$\|Tf\|_p \le \mathrm{c}_{n,\gamma} \, \frac{p^2}{p-1} \, \|f\|_p.$$

The expression $\sup_{\lambda > 0} \lambda \, \varphi \big\{ |Tf| > \lambda \big\}$ is just a slight abuse of notation to denote the noncommutative weak L_1 norm, to be rigorously defined below. We find it though more intuitive, since it is reminiscent of the classical terminology. Theorem A provides a positive answer to our problem for any singular integral associated to a generalized Calderón-Zygmund kernel satisfying the size/smoothness conditions imposed above. Moreover, the asymptotic behavior of the constants as $p \to 1$ and $p \to \infty$ is optimal. Independently, Tao Mei has recently obtained another argument for this which does not include the weak type inequality [38]. We shall present it at the end of the paper, since we shall use it indirectly to obtain weak type inequalities for singular integrals associated to operator-valued kernels. In the language of operator space theory and following Pisier's characterization [45, 46] of complete boundedness we immediately obtain:

Corollary. *Let* T *be a generalized* L_q*-bounded and* γ*-Lipschitz Calderón-Zygmund operator. Let us equip* L_p *with its natural operator space structure. Then, the cb-norm of* $T : L_p \to L_p$ *is controlled by*

$$\mathrm{c}_{n,\gamma} \, \frac{p^2}{p-1}.$$

Thus, the growth rate as $p \to 1$ *or* $p \to \infty$ *coincides with the Banach space case.*

Before going on, a few remarks are in order:

a) It is standard to reduce the proof of Theorem A to the case $q = 2$.

b) The reader might think that our hypothesis on Lipschitz smoothness for the first variable is unnecessary to obtain the weak type inequality and that only smoothness with respect to the second variable is needed. Namely, this is the case in the classical theory. It is however not the case in this paper because the use of certain almost orthogonality arguments (see below) forces us to apply both kinds of smoothness. We refer to Remark 2.11 for the specific point where the x-Lipschitz smoothness is applied and to Remark 5.5 for more in depth discussion on the conditions imposed on the kernel.

c) In the classical case E_λ is a perfectly delimited region of $\mathbb{R}^n$. In particular, we may construct the dilation $9\,\mathrm{E}_\lambda = \bigcup_j 9Q_j$. This set is useful to estimate the bad part b since it has two crucial properties. First, it is small because $|9\,\mathrm{E}_\lambda| \sim |\mathrm{E}_\lambda|$ and E_λ satisfies the Hardy-Littlewood weak maximal inequality. Second, its complement is far away from E_λ (the support of b) so that Tb restricted to $\mathbb{R}^n \setminus 9\,\mathrm{E}_\lambda$ avoids the singularity of the kernel. The problem that we find in the noncommutative case is that E_λ is no longer a region in $\mathbb{R}^n$. Indeed, given a dyadic cube Q and a positive $f \in L_1$, we have either $f_Q > \lambda$ or not and this dichotomy completely determines the set E_λ. However, for $f \in L_1(\mathcal{A})_+$ the average f_Q is a positive operator (not a positive number) and the dichotomy disappears since the condition $f_Q > \lambda$ is only satisfied in part of the spectrum of f_Q. This difficulty is inherent to the noncommutativity and is motivated by the lack of a total order in the positive cone of $\mathcal{M}$. It also produces difficulties to define noncommutative maximal functions

[9, 24], a problem that required the recent theory of operator spaces for its solution and is in the heart of the matter. Our construction of the right noncommutative analog ζ of $\mathbb{R}^n \setminus 9\mathrm{E}_\lambda$ is a key step in this paper, see Lemma 4.2 below. Here it is relevant to recall that, quite unexpectedly (in contrast with the classical case) we shall need the projection ζ to deal with *both* the good and the bad parts.

d) Another crucial difference with the classical setting and maybe the hardest point to overcome is the lack of estimates i) and ii) above in the noncommutative framework. Indeed, given $f \in L_1(\mathcal{A})_+$ we only have such estimates for the diagonal terms

$$\sum\nolimits_k p_k f_k p_k \quad \text{and} \quad \sum\nolimits_k p_k (f - f_k) p_k.$$

A more detailed discussion on this topic is given in Appendix B below. Let us now explain how we face the lack of the classical inequalities. Since $\mathbf{1}_\mathcal{A} - \zeta$ is the noncommutative analog of $9\mathrm{E}_\lambda$ which is *small* as explained above, we can use the noncommutative Hardy-Littlewood weak maximal inequality to reduce our problem to estimate the terms $\zeta T(g) \zeta$ and $\zeta T(b) \zeta$. A very naive and formally incorrect way to explain what to do here is the following. Given a fixed positive integer s, we find *something like*

$$\begin{aligned} \Big\| \zeta T \Big(\sum_{|i-j|=s} p_i f_{i \vee j} p_j \Big) \zeta \Big\|_2 & \lesssim s 2^{-\gamma s} \Big\| \sum\nolimits_k p_k f_k p_k \Big\|_2, \\ \Big\| \zeta T \Big(\sum_{|i-j|=s} p_i (f - f_{i \vee j}) p_j \Big) \zeta \Big\|_1 & \lesssim s 2^{-\gamma s} \Big\| \sum\nolimits_k p_k (f - f_k) p_k \Big\|_1, \end{aligned}$$

where γ is the Lipschitz smoothness parameter of the kernel. In other words, we may estimate the action of $\zeta T(\cdot)\zeta$ on the terms in the s-th upper and lower diagonals by $s 2^{-\gamma s}$ times the *corresponding size* of the main diagonal. Then, recalling that i) and ii) hold on the diagonal, it is standard to complete the argument. We urge however the reader to understand this just as a motivation (not as a claim) since the argument is quite more involved than this. For instance, we will need to replace the off-diagonal terms of g by other $g_{k,s}$'s satisfying

$$\sum_{k,s} g_{k,s} = \sum_{i \neq j} p_i f_{i \vee j} p_j.$$

A rough way of rephrasing this phenomenon is to say that Calderón-Zygmund operators are *almost diagonal* when acting on operator-valued functions. In other contexts, this almost diagonal nature has already appeared in the literature. The wavelet proof of the $T1$ theorem [39] exhibits this property of singular integrals with respect to the Haar system in $\mathbb{R}^n$. This also applies in the context of Clifford analysis [40]. Moreover, some deep results in [7, 13] (which we will comment below) use this almost diagonal nature as a key idea. It is also worthy of mention that these difficulties do not appear in [44]. The reason is that the operators for which Gundy's decomposition is typically applied (martingale transforms or martingale square functions) do not move the support of the original function/operator. This means that the action of T over the off-diagonal terms is essentially supported by $\mathbf{1}_\mathcal{A} - \zeta$. Consequently, these terms are controlled by means of the noncommutative analog of Doob's maximal weak type inequality, see [44] for further details. As a byproduct, we observe that the pseudo-localization principle which we present below is not needed in [44].

5. Pseudo-localization. A key point in our argument is the behavior of singular integrals acting on the off-diagonal terms $p_i f_{i \vee j} p_j$ and $p_i(f - f_{i\vee j})p_j$, as a function of the parameter $s = |i-j|$ in a region $\zeta \approx \mathbb{R}^n \setminus 9\mathrm{E}_\lambda$ which is in some sense far away from their (left and right) support. The idea we need to exploit relies on the following principle: more regularity of the kernel of T implies a faster decay of Tf far away from the support of f. That is why the b-terms $p_i(f - f_{i\vee j})p_j$ are better than the g-terms $p_i f_{i\vee j} p_j$. Indeed, the cancellation of $f - f_{i\vee j}$ allows to subtract a piecewise constant function from the kernel (in the standard way) to apply the smoothness properties of it and obtain suitable L_1 estimates. However, the off-diagonal g-terms are not mean-zero (at least at first sight) with respect to $\int_{\mathbb{R}^n}$ and we need more involved tools to prove this pseudo-localization property in the L_2 metric. For the sake of clarity and since our result might be of independent interest even in the classical theory, we state it for scalar-valued functions. The way we apply it in our noncommutative setting will be clarified along the text, see Theorem 5.2 for the noncommutative form of this principle.

Since we are assuming that T is bounded on L_2, we may further assume by homogeneity that it is of norm 1. In the sequel, we will only consider L_2-normalized Calderón-Zygmund operators. Our result is related to the following problem.

An L_2-localization problem. *Given $f : \mathbb{R}^n \to \mathbb{C}$ in L_2 and $0 < \delta < 1$, find the sets $\Sigma_{f,\delta}$ such that the inequality below holds for all normalized Calderón-Zygmund operator satisfying the imposed size/smoothness conditions*

$$\Big(\int_{\mathbb{R}^n \setminus \Sigma_{f,\delta}} |Tf(x)|^2 dx \Big)^{\frac12} \le \delta \Big(\int_{\mathbb{R}^n} |f(x)|^2 dx \Big)^{\frac12}.$$

Given $f : \mathbb{R}^n \to \mathbb{C}$ in L_2, let f_k and df_k denote the k-th condition expectation of f with respect to the standard dyadic filtration and its corresponding k-th martingale difference. That is, we have $df_k = \sum_{Q \in \mathcal{Q}_k} \big(f_Q - f_{\widehat{Q}} \big) 1_Q$. Let $\mathcal{R}_k$ be the class of sets in $\mathbb{R}^n$ being the union of a family of cubes in $\mathcal{Q}_k$. Given such an $\mathcal{R}_k$-set $\Omega = \bigcup_j Q_j$, we shall work with the dilations $9\Omega = \bigcup_j 9Q_j$, where $9Q$ denotes the 9-concentric father of Q. We shall prove the following result.

A pseudo-localization principle. *Let us fix a positive integer s. Given a function f in L_2 and any integer k, we define Ω_k to be the smallest $\mathcal{R}_k$-set containing the support of df_{k+s}. If we further consider the set*

$$\Sigma_{f,s} = \bigcup_{k \in \mathbb{Z}} 9\Omega_k,$$

then we have the localization estimate

$$\Big(\int_{\mathbb{R}^n \setminus \Sigma_{f,s}} |Tf(x)|^2 \, dx \Big)^{\frac12} \ \le \ \mathrm{c}_{n,\gamma} \, s \, 2^{-\gamma s/4} \Big(\int_{\mathbb{R}^n} |f(x)|^2 dx \Big)^{\frac12},$$

for any L_2-normalized Calderón-Zygmund operator with Lipschitz parameter γ.

Given any integer $k \in \mathbb{Z}$, we are considering the smallest set Ω_k containing $\operatorname{supp} df_{k+s}$ and belonging to an s times *coarser* topology. This procedure gives rise to an apparently artificial *shift condition*

$$\operatorname{supp} df_{k+s} \subset \Omega_k$$

which gives a measure of how much we should enlarge $\Sigma_f = \bigcup_k \mathrm{supp}\, df_k$ at every scale. However, this condition (or its noncommutative analog) is quite natural in our setting since it is satisfied by the off-diagonal terms of g, precisely those for which our previous tools did not work. In the classical/commutative setting there are some natural situations for which our result applies and some others which limit the applicability of it. For instance, at first sight our result is only applicable for functions f satisfying $f_m = 0$ for some integer m. There are also some other natural questions such as an L_p analog of our result or an equivalent formulation using a Littlewood-Paley decomposition, instead of martingale differences. For the sake of clarity in our exposition, we prove the result in the body of the paper and we postpone these further comments to Appendix A below.

The proof of this result reduces to a *shifted form of the* $T1$ *theorem* in a sense to be explained below. In particular, almost orthogonality methods are essential in our approach. Compared to the standard proofs of the $T1$ theorem, with wavelets [39] or more generally with approximations of the identity [54], we need to work in a dyadic/martingale setting forced by the role of Cuculescu's construction in this paper. This produces a lack of smoothness in the functions we work with, requiring quite involved estimates to obtain almost orthogonality results. An apparently new aspect of our estimates is the asymmetry of our bounds when applying Schur lemma, see Remark 2.1 for more details.

Let us briefly comment the relation of our result with two papers by Christ [7] and Duoandikoetxea/Rubio de Francia [13]. Although both papers already exploited the almost diagonal nature of Calderón-Zygmund operators, only convolution-type singular integrals are considered and no localization result is pursued there. Being more specific, a factor $2^{-\gamma s}$ is obtained in [7] for the bad part of Calderón-Zygmund decomposition. As explained above, we need to produce this factor for the good part. This is very unusual (or even new) in the literature. Nevertheless, the way we have stated our pseudo-localization result shows that the key property is the shift condition $\mathrm{supp}\, df_{k+s} \subset \Omega_k$, regardless we work with good or bad parts. On the other hand, in [13] Littlewood-Paley theory and the commutativity produced by the use of convolution operators is used to obtain related estimates in L_p with $p \neq 2$. In particular, almost orthogonality does not play any role there. The lack of a suitable noncommutative Littlewood-Paley theory and our use of generalized Calderón-Zygmund operators make their argument not applicable here.

6. Operator-valued kernels. At the end of the paper we extend our main results to certain Calderón-Zygmund operators associated to kernels $k : \mathbb{R}^{2n} \setminus \Delta \to \mathcal{M}$ satisfying the canonical size/smoothness conditions. In other words, we replace the absolute value by the norm in $\mathcal{M}$:

a) If $x, y \in \mathbb{R}^n$, we have
$$\|k(x,y)\|_{\mathcal{M}} \lesssim \frac{1}{|x-y|^n}.$$

b) There exists $0 < \gamma \leq 1$ such that
$$\begin{aligned} \big\|k(x,y) - k(x',y)\big\|_{\mathcal{M}} &\lesssim \frac{|x-x'|^\gamma}{|x-y|^{n+\gamma}} \quad \text{if} \quad |x-x'| \leq \frac{1}{2}|x-y|, \\ \big\|k(x,y) - k(x,y')\big\|_{\mathcal{M}} &\lesssim \frac{|y-y'|^\gamma}{|x-y|^{n+\gamma}} \quad \text{if} \quad |y-y'| \leq \frac{1}{2}|x-y|. \end{aligned}$$

Unfortunately, not every such kernel satisfies the analog of Theorem A. Namely, we shall construct (using classical Littlewood-Paley methods) a simple kernel satisfying the size and smoothness conditions above and giving rise to a Calderón-Zygmund operator bounded on $L_2(\mathcal{A})$ but not on $L_p(\mathcal{A})$ for $1 < p < 2$. However, a detailed inspection of our proof of Theorem A and a few auxiliary results will show that the key condition (together with the size/smoothness hypotheses on the kernel) for the operator T is to be an $\mathcal{M}$-bimodule map. Of course, this always holds in the context of Theorem A. When dealing with operator-valued kernels this is false in general, but it holds for instance when dealing with *standard* Calderón-Zygmund operators

$$Tf(x) \,=\, \xi f(x) \,+\, \lim_{\varepsilon \to 0} \int_{|x-y|>\varepsilon} k(x,y)\, f(y)\, dy$$

associated to a commuting kernel $k : \mathbb{R}^{2n} \setminus \Delta \to \mathcal{Z}_\mathcal{M}$, with $\mathcal{Z}_\mathcal{M} = \mathcal{M} \cap \mathcal{M}'$ standing for the center of $\mathcal{M}$. Note that we are only requiring the $\mathcal{M}$-bimodule property to hold on the singular integral part, since the multiplier part is always well-behaved as far as $\xi \in \mathcal{A}$. Note also that when $\mathcal{M}$ is a factor, any commuting kernel must be scalar-valued and we go back to Theorem A.

Theorem B. *Let T be a generalized Calderón-Zygmund operator associated to an operator-valued kernel $k : \mathbb{R}^{2n} \setminus \Delta \to \mathcal{M}$ satisfying the imposed size/smoothness conditions. Assume that T is an $\mathcal{M}$-bimodule map bounded on $L_q(\mathcal{A})$ for some $1 < q < \infty$. Then, the following weak type inequality holds for some constant $\mathrm{c}_{n,\gamma}$ depending only on the dimension n and the Lipschitz smoothness parameter γ*

$$\sup_{\lambda>0} \lambda\, \varphi \Big\{ |Tf| > \lambda \Big\} \le \mathrm{c}_{n,\gamma} \, \|f\|_1.$$

In particular, given $1 < p < \infty$ and $f \in L_p(\mathcal{A})$, we find

$$\|Tf\|_p \le \mathrm{c}_{n,\gamma} \, \frac{p^2}{p-1} \, \|f\|_p.$$

The strong L_p inequalities stated in Theorem B do not follow from a UMD-type argument as it happened with Theorem A. In particular, these L_p estimates seem to be new and independently obtained by Tao Mei as pointed above.

7. Appendices. We conclude the paper with two appendices. A further analysis on pseudo-localization is given in Appendix A. This mainly includes remarks related to our result, some conjectures on possible generalizations and a corollary on the rate of decreasing of the L_2 mass of a singular integral far away from the support of the function on which it acts. In Appendix B we study the noncommutative form of Calderón-Zygmund decomposition in further detail. In particular, we give some weighted inequalities for the good and bad parts which generalize the classical L_1 and L_2 estimates satisfied by these functions. The sharpness of our estimates remains as an open interesting question.

Remark. The value of the constant $\mathrm{c}_{n,\gamma}$ will change from one instance to another.

Acknowledgement. I would like to thank J.M. Martell, F. Soria and Q. Xu for discussions related to the content of this paper and specially to Tao Mei for keeping me up to date on his related work.

1. Noncommutative integration

We begin with a quick survey of definitions and results on noncommutative L_p spaces and related topics that will be used along the paper. All or most of it will be well-known to experts in the field. The right framework for a noncommutative analog of measure theory and integration is von Neumann algebra theory. We refer to [32, 55] for a systematic study of von Neumann algebras and to the recent survey by Pisier/Xu [48] for a detailed exposition of noncommutative L_p spaces.

1.1. **Noncommutative L_p.** A *von Neumann algebra* is a weak-operator closed C*-algebra. By the Gelfand-Naimark-Segal theorem, any von Neumann algebra $\mathcal{M}$ can be embedded in the algebra $\mathcal{B}(\mathcal{H})$ of bounded linear operators on some Hilbert space $\mathcal{H}$. In what follows we will identify $\mathcal{M}$ with a subalgebra of $\mathcal{B}(\mathcal{H})$. The positive cone $\mathcal{M}_+$ is the set of positive operators in $\mathcal{M}$. A *trace* $\tau : \mathcal{M}_+ \to [0,\infty]$ on $\mathcal{M}$ is a linear map satisfying the tracial property $\tau(a^*a) = \tau(aa^*)$. It is said to be *normal* if $\sup_\alpha \tau(a_\alpha) = \tau(\sup_\alpha a_\alpha)$ for any bounded increasing net (a_α) in $\mathcal{M}_+$; it is *semifinite* if for any non-zero $a \in \mathcal{M}_+$, there exists $0 < a' \le a$ such that $\tau(a') < \infty$ and it is *faithful* if $\tau(a) = 0$ implies $a = 0$. Taking into account that τ plays the role of the integral in measure theory, all these properties are quite familiar. A von Neumann algebra $\mathcal{M}$ is called *semifinite* whenever it admits a normal semifinite faithful (*n.s.f.* in short) trace τ. Except for a brief comment in Remark 5.4 below we shall always work with semifinite von Neumann algebras. Recalling that any operator a can be written as a linear combination $a_1 - a_2 + ia_3 - ia_4$ of four positive operators, we can extend τ to the whole algebra $\mathcal{M}$. Then, the tracial property can be restated in the familiar way $\tau(ab) = \tau(ba)$ for all $a, b \in \mathcal{M}$.

According to the GNS construction, it is easily seen that the noncommutative analogs of measurable sets (or equivalently characteristic functions of those sets) are orthogonal projections. Given $a \in \mathcal{M}_+$, the support projection of a is defined as the least projection q in $\mathcal{M}$ such that $qa = a = aq$ and will be denoted by $\mathrm{supp}\, a$. Let $\mathcal{S}_+$ be the set of all $a \in \mathcal{M}_+$ such that $\tau(\mathrm{supp}\, a) < \infty$ and set $\mathcal{S}$ to be the linear span of $\mathcal{S}_+$. If we write $|x|$ for the operator $(x^*x)^{\frac{1}{2}}$, we can use the spectral measure $\gamma_{|x|} : \mathbb{R}_+ \to \mathcal{B}(\mathcal{H})$ of the operator $|x|$ to define

$$|x|^p = \int_{\mathbb{R}_+} s^p \, d\gamma_{|x|}(s) \quad \text{for} \quad 0 < p < \infty.$$

We have $x \in \mathcal{S} \Rightarrow |x|^p \in \mathcal{S}_+ \Rightarrow \tau(|x|^p) < \infty$. If we set $\|x\|_p = \tau(|x|^p)^{\frac{1}{p}}$, it turns out that $\| \ \|_p$ is a norm in $\mathcal{S}$ for $1 \le p < \infty$ and a p-norm for $0 < p < 1$. Using that $\mathcal{S}$ is a w^*-dense $*$-subalgebra of $\mathcal{M}$, we define the *noncommutative L_p space* $L_p(\mathcal{M})$ associated to the pair $(\mathcal{M}, \tau)$ as the completion of $(\mathcal{S}, \| \ \|_p)$. On the other hand, we set $L_\infty(\mathcal{M}) = \mathcal{M}$ equipped with the operator norm. Many of the fundamental properties of classical L_p spaces like duality, real and complex interpolation... can be transferred to this setting. The most important properties for our purposes are the following:

- Hölder inequality. If $1/r = 1/p + 1/q$, we have $\|ab\|_r \le \|a\|_p \|b\|_q$.
- The trace τ extends to a continuous functional on $L_1(\mathcal{M})$: $|\tau(x)| \le \|x\|_1$.

We refer to [48] for a definition of L_p over non-semifinite von Neumann algebras.

1.2. **Noncommutative symmetric spaces.** Let

$$\mathcal{M}' = \Big\{ b \in \mathcal{B}(\mathcal{H}) \, \big| \, ab = ba \text{ for all } a \in \mathcal{M} \Big\}$$

be the commutant of $\mathcal{M}$. A closed densely-defined operator on $\mathcal{H}$ is *affiliated* with $\mathcal{M}$ when it commutes with every unitary u in the commutant $\mathcal{M}'$. Recall that $\mathcal{M} = \mathcal{M}''$ and this implies that every $a \in \mathcal{M}$ is affiliated with $\mathcal{M}$. The converse fails in general since we may find unbounded operators. If a is a densely defined self-adjoint operator on $\mathcal{H}$ and $a = \int_\mathbb{R} s \, d\gamma_a(s)$ is its spectral decomposition, the spectral projection $\int_\mathcal{R} d\gamma_a(s)$ will be denoted by $\chi_\mathcal{R}(a)$. An operator a affiliated with $\mathcal{M}$ is *τ-measurable* if there exists $s > 0$ such that

$$\tau\big(\chi_{(s,\infty)}(|a|)\big) = \tau\big\{|a| > s\big\} < \infty.$$

The *generalized singular-value* $\mu(a) : \mathbb{R}_+ \to \mathbb{R}_+$ is defined by

$$\mu_t(a) = \inf \Big\{ s > 0 \, \big| \, \tau\big\{|x| > s\big\} \le t \Big\}.$$

This provides us with a noncommutative analogue of the so-called non-increasing rearrangement of a given function. We refer to [14] for a detailed exposition of the function $\mu(a)$ and the corresponding notion of convergence in measure.

If $L_0(\mathcal{M})$ denotes the $*$-algebra of τ-measurable operators, we have the following equivalent definition of L_p

$$L_p(\mathcal{M}) = \Big\{ a \in L_0(\mathcal{M}) \, \big| \, \Big(\int_{\mathbb{R}_+} \mu_t(a)^p \, dt \Big)^{\frac{1}{p}} < \infty \Big\}.$$

The same procedure applies to symmetric spaces. Given the pair $(\mathcal{M}, \tau)$, let X be a rearrangement invariant quasi-Banach function space on the interval $(0, \tau(\mathbf{1}_\mathcal{M}))$. The *noncommutative symmetric space* $\mathrm{X}(\mathcal{M})$ is defined by

$$\mathrm{X}(\mathcal{M}) = \Big\{ a \in L_0(\mathcal{M}) \, \big| \, \mu(a) \in \mathrm{X} \Big\} \quad \text{with} \quad \|a\|_{\mathrm{X}(\mathcal{M})} = \|\mu(a)\|_\mathrm{X}.$$

It is known that $\mathrm{X}(\mathcal{M})$ is a Banach (resp. quasi-Banach) space whenever X is a Banach (resp. quasi-Banach) function space. We refer the reader to [11, 58] for more in depth discussion of this construction. Our interest in this paper is restricted to noncommutative L_p-spaces and *noncommutative weak L_1-spaces*. Following the construction of symmetric spaces of measurable operators, the noncommutative weak L_1-space $L_{1,\infty}(\mathcal{M})$, is defined as the set of all a in $L_0(\mathcal{M})$ for which the quasi-norm

$$\|a\|_{1,\infty} = \sup_{t>0} t \, \mu_t(x) = \sup_{\lambda>0} \lambda \tau \Big\{ |x| > \lambda \Big\}$$

is finite. As in the commutative case, the noncommutative weak L_1 space satisfies a quasi-triangle inequality that will be used below with no further reference. Indeed, the following inequality holds for $a_1, a_2 \in L_{1,\infty}(\mathcal{M})$

$$\lambda \, \tau \Big\{ |a_1 + a_2| > \lambda \Big\} \le \lambda \, \tau \Big\{ |a_1| > \lambda/2 \Big\} + \lambda \, \tau \Big\{ |a_2| > \lambda/2 \Big\}.$$

1.3. **Noncommutative martingales.** Consider a von Neumann subalgebra (a weak* closed $*$-subalgebra) $\mathcal{N}$ of $\mathcal{M}$. A *conditional expectation* $\mathcal{E} : \mathcal{M} \to \mathcal{N}$ from $\mathcal{M}$ onto $\mathcal{N}$ is a positive contractive projection. The conditional expectation $\mathcal{E}$ is called *normal* if the adjoint map $\mathcal{E}^*$ satisfies $\mathcal{E}^*(\mathcal{M}_*) \subset \mathcal{N}_*$. In this case, there is a map $\mathcal{E}_* : \mathcal{M}_* \to \mathcal{N}_*$ whose adjoint is $\mathcal{E}$. Note that such normal conditional expectation exists if and only if the restriction of τ to the von Neumann subalgebra $\mathcal{N}$ remains semifinite, see e.g. Theorem 3.4 in [55]. Any such conditional expectation is trace preserving (i.e. $\tau \circ \mathcal{E} = \tau$) and satisfies the bimodule property

$$\mathcal{E}(a_1 b a_2) = a_1 \mathcal{E}(b) a_2 \quad \text{for all} \quad a_1, a_2 \in \mathcal{N} \text{ and } b \in \mathcal{M}.$$

Let $(\mathcal{M}_k)_{k \geq 1}$ be an increasing sequence of von Neumann subalgebras of $\mathcal{M}$ such that the union of the $\mathcal{M}_k$'s is weak* dense in $\mathcal{M}$. Assume that for every $k \geq 1$, there is a normal conditional expectation $\mathcal{E}_k : \mathcal{M} \to \mathcal{M}_k$. Note that for every $1 \leq p < \infty$ and $k \geq 1$, $\mathcal{E}_k$ extends to a positive contraction $\mathcal{E}_k : L_p(\mathcal{M}) \to L_p(\mathcal{M}_k)$. A *noncommutative martingale* with respect to the filtration $(\mathcal{M}_k)_{k \geq 1}$ is a sequence $a = (a_k)_{k \geq 1}$ in $L_1(\mathcal{M})$ such that

$$\mathcal{E}_j(a_k) = a_j \quad \text{for all} \quad 1 \leq j \leq k < \infty.$$

If additionally $a \subset L_p(\mathcal{M})$ for some $1 \leq p \leq \infty$ and $\|a\|_p = \sup_{k \geq 1} \|a_k\|_p < \infty$, then a is called an *L_p-bounded martingale*. Given a martingale $a = (a_k)_{k \geq 1}$, we assume the convention that $a_0 = 0$. Then, the martingale difference sequence $da = (da_k)_{k \geq 1}$ associated to x is defined by $da_k = a_k - a_{k-1}$.

The next result due to Cuculescu [9] was the first known result in the theory and will be crucial in this paper. It can be viewed as a noncommutative analogue of the classical weak type $(1,1)$ boundedness of Doob's maximal function.

Cuculescu's construction. *Suppose $a = (a_1, a_2, \ldots)$ is a positive L_1 martingale relative to the filtration $(\mathcal{M}_k)_{k \geq 1}$ and let λ be a positive number. Then there exists a decreasing sequence of projections*

$$q(\lambda)_1, q(\lambda)_2, q(\lambda)_3, \ldots$$

in $\mathcal{M}$ satisfying the following properties

i) *$q(\lambda)_k$ commutes with $q(\lambda)_{k-1} a_k q(\lambda)_{k-1}$ for each $k \geq 1$.*
ii) *$q(\lambda)_k$ belongs to $\mathcal{M}_k$ for each $k \geq 1$ and $q(\lambda)_k a_k q(\lambda)_k \leq \lambda q(\lambda)_k$.*
iii) *The following estimate holds*

$$\tau\Big(\mathbf{1}_\mathcal{M} - \bigwedge_{k \geq 1} q(\lambda)_k\Big) \leq \frac{1}{\lambda} \sup_{k \geq 1} \|a_k\|_1.$$

Explicitly, we set $q(\lambda)_0 = \mathbf{1}_\mathcal{M}$ and define $q(\lambda)_k = \chi_{(0,\lambda]}(q(\lambda)_{k-1} a_k q(\lambda)_{k-1})$.

The theory of noncommutative martingales has achieved considerable progress in recent years. The renewed interest on this topic started from the fundamental paper of Pisier and Xu [47], where they introduced a new functional analytic approach to study Hardy spaces and the Burkholder-Gundy inequalities for noncommutative martingales. Shortly after, many classical inequalities have been transferred to the noncommutative setting. A noncommutative analogue of Doob's maximal function [24], the noncommutative John-Nirenberg theorem [26], extensions of Burkholder inequalities for conditioned square functions [30] and related weak type inequalities [50, 51, 52]; see [44] for a simpler approach to some of them.

2. A pseudo-localization principle

Let us now proceed with the proof of the pseudo-localization principle stated in the Introduction. In the course of it we will see the link with a shifted form of the $T1$ theorem, which is formulated in a dyadic martingale setting. Since we are concerned with its applications to our noncommutative problem, we leave a more in depth analysis of our result to Appendix A below.

2.1. **Three auxiliary results.** We need some well-known results that live around David-Journé's $T1$ theorem. Cotlar lemma is very well-known and its proof can be found in [12, 54]. We include the proof of Schur lemma, since our statement and proof is non-standard, see Remark 2.1 below for details. The localization estimate at the end follows from [39]. We give the proof for completeness.

Cotlar lemma. *Let $\mathcal{H}$ be a Hilbert space and let us consider a family $(T_k)_{k\in\mathbb{Z}}$ of bounded operators on $\mathcal{H}$ with finitely many non-zero T_k's. Assume that there exists a summable sequence $(\alpha_k)_{k\in\mathbb{Z}}$ such that*

$$\max \Big\{ \big\| T_i^* T_j \big\|_{\mathcal{B}(\mathcal{H})}, \big\| T_i T_j^* \big\|_{\mathcal{B}(\mathcal{H})} \Big\} \ \le \ \alpha_{i-j}^2$$

for all $i,j \in \mathbb{Z}$. Then we automatically have

$$\Big\| \sum\nolimits_k T_k \Big\|_{\mathcal{B}(\mathcal{H})} \ \le \ \sum\nolimits_k \alpha_k.$$

Schur lemma. *Let T be given by*

$$Tf(x) = \int_{\mathbb{R}^n} k(x,y) \, f(y) \, dy.$$

Let us define the Schur integrals associated to k

$$\begin{array}{rcl} \mathcal{S}_1(x) & = & \displaystyle\int_{\mathbb{R}^n} \big| k(x,y) \big| \, dy, \\ \mathcal{S}_2(y) & = & \displaystyle\int_{\mathbb{R}^n} \big| k(x,y) \big| \, dx. \end{array}$$

Assume that both $\mathcal{S}_1$ and $\mathcal{S}_2$ belong to L_∞. Then, T is bounded on L_2 and

$$\|T\|_{\mathcal{B}(L_2)} \le \sqrt{\big\|\mathcal{S}_1\big\|_\infty \big\|\mathcal{S}_2\big\|_\infty}.$$

Proof. By the Cauchy-Schwarz inequality we obtain

$$\begin{array}{l} \displaystyle\Big(\int_{\mathbb{R}^n} \Big| \int_{\mathbb{R}^n} k(x,y) \, f(y) \, dy \Big|^2 dx \Big)^{\frac12} \\ \quad \le \ \displaystyle\Big(\int_{\mathbb{R}^n} \Big[\int_{\mathbb{R}^n} |k(x,y)| \, |f(y)| \, dy \Big]^2 dx \Big)^{\frac12} \\ \quad \le \ \displaystyle\Big(\int_{\mathbb{R}^n} \Big[\int_{\mathbb{R}^n} |k(x,y)| \, dy \Big] \Big[\int_{\mathbb{R}^n} |k(x,y)| \, |f(y)|^2 \, dy \Big] \, dx \Big)^{\frac12} \\ \quad \le \ \displaystyle\sqrt{\big\|\mathcal{S}_1\big\|_\infty} \Big(\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |k(x,y)| \, |f(y)|^2 \, dy \, dx \Big)^{\frac12} \\ \quad \le \ \displaystyle\sqrt{\big\|\mathcal{S}_1\big\|_\infty \big\|\mathcal{S}_2\big\|_\infty} \Big(\int_{\mathbb{R}^n} |f(y)|^2 \, dy \Big)^{\frac12}. \end{array}$$

□

Remark 2.1. Typically, Schur lemma is formulated as

$$\|T\|_{\mathcal{B}(L_2)} \le \frac{1}{2}\Big(\big\|\mathcal{S}_1\big\|_\infty + \big\|\mathcal{S}_2\big\|_\infty\Big),$$

see e.g. [39, 54]. This might happen because we usually have $\|\mathcal{S}_1\|_\infty \sim \|\mathcal{S}_2\|_\infty$, by certain symmetry in the estimates. In particular, the cases for which the arithmetic mean does not help but the geometric mean does are very rare in the literature, or even (as far as we know) not existent! However, motivated by a lack of symmetry in our estimates, this is exactly the case in this paper.

A localization estimate. *Assume that*

$$|k(x,y)| \lesssim \frac{1}{|x-y|^n} \quad \text{for all} \quad x,y \in \mathbb{R}^n.$$

Let T be a Calderón-Zygmund operator associated to the kernel k and assume that T is L_2-normalized. Then, given $x_0 \in \mathbb{R}^n$ and $r_1, r_2 \in \mathbb{R}_+$ with $r_2 > 2r_1$, the estimate below holds for any pair f, g of bounded scalar-valued functions respectively supported by $\mathsf{B}_{r_1}(x_0)$ and $\mathsf{B}_{r_2}(x_0)$

$$\big|\big\langle Tf, g\big\rangle\big| \le \mathrm{c}_n\, r_1^n \log(r_2/r_1)\, \|f\|_\infty \|g\|_\infty.$$

Proof. Let us write B for the ball $\mathsf{B}_{3r_1/2}(x_0)$ and let us consider a smooth function ρ which is identically 1 on B and identically 0 outside $\mathsf{B}_{2r_1}(x_0)$. Set $\eta = 1-\rho$ so that we may decompose

$$\big\langle Tf, g\big\rangle = \big\langle Tf, \rho g\big\rangle + \big\langle Tf, \eta g\big\rangle.$$

For the first term we have

$$\begin{aligned} \big|\big\langle Tf, \rho g\big\rangle\big| &\le \|Tf\|_2 \|\rho g\|_2 \ \le\ \|f\|_2 \|\rho g\|_2 \\ &\le \|f\|_\infty \|g\|_\infty \sqrt{\big|\mathrm{supp} f\big|\big|\mathrm{supp}(\rho g)\big|} \ \le\ \mathrm{c}_n\, r_1^n\, \|f\|_\infty \|g\|_\infty. \end{aligned}$$

On the other hand, for the second term we have

$$|\langle Tf, \eta g\rangle| = \Big| \int_{\mathsf{B}_{r_2}(x_0)\setminus \mathsf{B}} \Big(\int_{\mathsf{B}_{r_1}(x_0)} k(x,y)\, f(y)\, dy \Big) \, \overline{\eta g}(x)\, dx \Big|.$$

The latter integral is clearly bounded by

$$\|f\|_\infty \|g\|_\infty \int_\Omega \frac{dx\, dy}{|x-y|^n}$$

with $\Omega = (\mathsf{B}_{r_2}(x_0) \setminus \mathsf{B}) \times \mathsf{B}_{r_1}(x_0)$. However, it is easily checked that an upper bound for the double integral given above is provided by $\mathrm{c}_n\, r_1^n \log(r_2/r_1)$, where c_n is a constant depending only on n. This completes the proof. □

2.2. **Shifted $T1$ theorem.** By the conditions imposed on T in the Introduction, it is clear that its adjoint T^* is an L_2-normalized Calderón-Zygmund operator with kernel $k^*(x,y) = \overline{k(y,x)}$ satisfying the same size and smoothness estimates. This implies that T^*1 (understood in a weak sense, see e.g. [54] for details) belongs to BMO, the space of functions with bounded mean oscillation. In addition, if $\Delta_j = \mathsf{E}_j - \mathsf{E}_{j-1}$ denotes the dyadic martingale difference operator, it is also well known that for any $\rho \in \mathrm{BMO}$ the dyadic paraproduct against ρ

$$\Pi_\rho(f) = \sum_{j=-\infty}^{\infty} \Delta_j(\rho) \mathsf{E}_{j-1}(f)$$

is bounded on L_2. Here it is necessary to know how BMO is related to its dyadic version BMO_d, see [16] and [35] for details. It is clear that $\Pi_\rho(1) = \rho$ and the adjoint of Π_ρ is given by the operator

$$\Pi_\rho^*(f) = \sum_{j=-\infty}^{\infty} \mathsf{E}_{j-1}\big(\overline{\Delta_j(\rho)} f\big).$$

Thus, since $T^*1 \in \mathrm{BMO}$ we may write

$$T = T_0 + \Pi_{T^*1}^*.$$

According to our previous considerations, both T_0 and $\Pi_{T^*1}^*$ are Calderón-Zygmund operators bounded on L_2 and their kernels satisfy the standard size and smoothness conditions imposed on T with the same Lipschitz smoothness parameter γ, see [54] for the latter assertion. Moreover, the operator T_0 now satisfies $T_0^*1 = 0$. Now we use that T_0^*1 is the weak* limit of a sequence $(T_0^* \rho_k)_{k \ge 1}$ in BMO, where the ρ_k's are increasing bump functions which converge to 1. In particular, the relation below holds for any $f \in H_1$

$$\int_{\mathbb{R}^n} T_0 f(x)\, dx = 0. \tag{2.1}$$

Indeed, we have $\langle T_0 f, 1\rangle = \langle f, T_0^*1\rangle = 0$. The use of paraproducts is exploited in the $T1$ theorem to produce the cancellation condition (2.1), which is a key assumption to make Cotlar lemma effective in this setting. The paraproduct term is typically estimated using Carleson's lemma, although we will not need it here. What we shall do is to prove that our theorem for T_0 and $\Pi_{T^*1}^*$ reduces to prove a shifted form of the $T1$ theorem. In this paragraph we only deal with T_0.

Let T be a generalized Calderón-Zygmund operator as in the statement of our result and assume that T satisfies the cancellation condition (2.1), so that there is no need to use the notation T_0 in what follows. Let us write

$$\mathbb{R}^n \setminus \Sigma_{f,s} = \bigcap_{k \in \mathbb{Z}} \Theta_k \qquad \text{with} \qquad \mathbb{R}^n \setminus \Theta_k = 9\Omega_k.$$

Denote by E_k the k-th dyadic conditional expectation and by Δ_k the martingale difference operator $\mathsf{E}_k - \mathsf{E}_{k-1}$, so that $\mathsf{E}_k(f) = f_k$ and $\Delta_k(f) = df_k$. Recall that Ω_k and Θ_k are $\mathcal{R}_k$-sets. In particular, the action of multiplying by the characteristic functions 1_{Ω_k} or 1_{Θ_k} commutes with E_j for all $j \ge k$. Then we consider the following decomposition

$$1_{\mathbb{R}^n \setminus \Sigma_{f,s}} Tf = 1_{\mathbb{R}^n \setminus \Sigma_{f,s}} \Big(\sum\nolimits_k \mathsf{E}_k T \Delta_{k+s} 1_{\Omega_k} + \sum\nolimits_k (id - \mathsf{E}_k) 1_{\Theta_k} T 1_{\Omega_k} \Delta_{k+s} \Big)(f).$$

Note that we have used here the shift condition $\mathrm{supp}\, df_{k+s} \subset \Omega_k$ as well as the commutation relations mentioned above in conjunction with $\mathbb{R}^n \setminus \Sigma_{f,s} \subset \Theta_k$. Next we observe that $1_{\Theta_k} T 1_{\Omega_k} = 1_{\Theta_k} T_{4\cdot 2^{-k}} 1_{\Omega_k}$, where T_ε denotes the truncated singular integral formally given by

$$T_\varepsilon f(x) = \int_{|x-y|>\varepsilon} k(x,y) f(y)\, dy.$$

Indeed, we have

$$1_{\Theta_k} T 1_{\Omega_k} f(x) \ = \ 1_{\Theta_k}(x) \sum_{\substack{Q \in \mathcal{Q}_k \\ Q \cap \Omega_k \neq \emptyset}} 1_{\mathbb{R}^n \setminus 9Q}(x) \int_Q k(x,y) f(y) \, dy,$$

from where the claimed identity follows, since we have

$$\mathrm{dist}(Q, \mathbb{R}^n \setminus 9Q) = 4 \cdot 2^{-k}$$

for all $Q \in \mathcal{Q}_k$. Taking all these considerations into account, we deduce

$$1_{\mathbb{R}^n \setminus \Sigma_{f,s}} Tf = 1_{\mathbb{R}^n \setminus \Sigma_{f,s}} \Big(\sum\nolimits_k \mathsf{E}_k T \Delta_{k+s} + \sum\nolimits_k (id - \mathsf{E}_k) T_{4 \cdot 2^{-k}} \Delta_{k+s} \Big)(f).$$

In particular, our problem reduces to estimate the norm in $\mathcal{B}(L_2)$ of

$$\Phi_s = \sum\nolimits_k \mathsf{E}_k T \Delta_{k+s} \quad \text{and} \quad \Psi_s = \sum\nolimits_k (id - \mathsf{E}_k) T_{4 \cdot 2^{-k}} \Delta_{k+s}.$$

Both Φ_s and Ψ_s are reminiscent of well-known operators (in a sense $\mathsf{E}_k T$ and $T_{4 \cdot 2^{-k}}$ behave here in the same way) appearing in the proof of the $T1$ theorem by David and Journé [10]. Indeed, what we find (in the context of dyadic martingales) is exactly the s-shifted analogs meaning that we replace Δ_k by Δ_{k+s}. In summary, we have proved that under the assumption that cancellation condition (2.1) holds our main result reduces to the proof of the theorem below.

Shifted $T1$ theorem. *Let T be an L_2-normalized Calderón-Zygmund operator with Lipschitz parameter γ. Assume that $T^*1 = 0$ or, in other words, that we have $\int_{\mathbb{R}^n} Tf(x) \, dx = 0$ for any $f \in H_1$. Then, we have*

$$\|\Phi_s\|_{\mathcal{B}(L_2)} = \Big\| \sum\nolimits_k \mathsf{E}_k T \Delta_{k+s} \Big\|_{\mathcal{B}(L_2)} \le \mathrm{c}_{n,\gamma} \, s \, 2^{-\gamma s/4}.$$

Moreover, regardless the value of T^*1 we also have

$$\|\Psi_s\|_{\mathcal{B}(L_2)} = \Big\| \sum\nolimits_k (id - \mathsf{E}_k) T_{4 \cdot 2^{-k}} \Delta_{k+s} \Big\|_{\mathcal{B}(L_2)} \le \mathrm{c}_{n,\gamma} \, 2^{-\gamma s/2}.$$

Remark 2.2. For some time, our hope was to estimate

$$\Big\| \sum\nolimits_k T_{4 \cdot 2^{-k}} \Delta_{k+s} \Big\|_{\mathcal{B}(L_2)}$$

since we believed that the truncation of order 2^{-k} in conjunction with the action of Δ_{k+s} was enough to produce the right decay. Note that our pseudo-localization result could also be deduced from this estimate. However, the cancellation produced by the paraproduct decomposition in Φ_s and by the presence of the term $id - \mathsf{E}_k$ in Ψ_s play an essential role in the argument.

2.3. **Paraproduct argument.** Now we show how the estimate of the paraproduct term also reduces to the shifted $T1$ theorem stated above. Indeed, let us write Π instead of $\Pi_{T^*1}^*$ to simplify the notation. Then, as we did above, it is straightforward to see that

$$1_{\mathbb{R}^n \setminus \Sigma_{f,s}} \Pi f = 1_{\mathbb{R}^n \setminus \Sigma_{f,s}} \Big(\sum\nolimits_k \mathsf{E}_k \Pi \Delta_{k+s} 1_{\Omega_k} + \sum\nolimits_k (id - \mathsf{E}_k) \Pi_{4 \cdot 2^{-k}} \Delta_{k+s} \Big)(f).$$

Recalling one more time that Π is an L_2-bounded generalized Calderón-Zygmund operator satisfying the same size and smoothness conditions as T, the estimate for the second operator

$$\Big\| \sum\nolimits_k (id - \mathsf{E}_k)\Pi_{4\cdot 2^{-k}} \Delta_{k+s} \Big\|_{\mathcal{B}(L_2)} \le c_{n,\gamma}\, 2^{-\gamma s/2}$$

follows from the second assertion of the shifted $T1$ theorem. Here it is essential to note that the hypothesis $T^*1 = 0$ is not needed for Ψ_s. Therefore, it only remains to estimate the first operator. However, we claim that $1_{\mathbb{R}^n \setminus \Sigma_{f,s}} \sum_k \mathsf{E}_k \Pi \Delta_{k+s} 1_{\Omega_k} f$ is identically zero. Let us prove this assertion. We have

$$1_{\mathbb{R}^n \setminus \Sigma_{f,s}} \sum\nolimits_k \mathsf{E}_k \Pi \Delta_{k+s} 1_{\Omega_k} f \, = \, 1_{\mathbb{R}^n \setminus \Sigma_{f,s}} \sum\nolimits_k \mathsf{E}_k \sum\nolimits_j \mathsf{E}_{j-1}\Big(\overline{\Delta_j(T^*1)}\, 1_{\Omega_k}\, df_{k+s} \Big).$$

If we fix the integer k, all the j-terms on the second sum above vanish except for the term associated to $j = k+s$. Indeed, if $j < k+s$ we use $\mathsf{E}_{j-1} = \mathsf{E}_{j-1}\mathsf{E}_{k+s-1}$ and obtain

$$\begin{aligned} \mathsf{E}_{j-1}\Big(\overline{\Delta_j(T^*1)}\, 1_{\Omega_k}\, df_{k+s} \Big) &= \mathsf{E}_{j-1}\Big(\mathsf{E}_{k+s-1}\big(\overline{\Delta_j(T^*1)}\, 1_{\Omega_k}\, df_{k+s} \big) \Big) \\ &= \mathsf{E}_{j-1}\Big(\overline{\Delta_j(T^*1)}\, \mathsf{E}_{k+s-1}\big(1_{\Omega_k}\, df_{k+s} \big) \Big) \, = \, 0. \end{aligned}$$

If $j > k+s$ we have

$$\mathsf{E}_{j-1}\Big(\overline{\Delta_j(T^*1)}\, 1_{\Omega_k}\, df_{k+s} \Big) \, = \, \mathsf{E}_{j-1}\big(\overline{\Delta_j(T^*1)} \big)\, 1_{\Omega_k}\, df_{k+s} \, = \, 0.$$

In particular, we obtain the following identity

$$\begin{aligned} 1_{\mathbb{R}^n \setminus \Sigma_{f,s}} \sum\nolimits_k \mathsf{E}_k \Pi \Delta_{k+s} 1_{\Omega_k} f &= 1_{\mathbb{R}^n \setminus \Sigma_{f,s}} \sum\nolimits_k \mathsf{E}_k \Big(\overline{\Delta_{k+s}(T^*1)}\, 1_{\Omega_k}\, df_{k+s} \Big) \\ &= 1_{\mathbb{R}^n \setminus \Sigma_{f,s}} \sum\nolimits_k 1_{\Omega_k}\, \mathsf{E}_k \Big(\overline{\Delta_{k+s}(T^*1)}\, df_{k+s} \Big) = 0. \end{aligned}$$

The last identity follows from the fact that $\Omega_k \subset \Sigma_{f,s}$ and $\mathbb{R}^n \setminus \Sigma_{f,s}$ are disjoint.

2.4. **Estimating the norm of Φ_s.** Now we estimate the operator norm of the sum Φ_s under the assumption that the cancellation condition (2.1) holds for T. We begin by identifying the kernel of the operators appearing in Φ_s. Let us denote by $k_{\mathsf{e},k}$ and $k_{\delta,k+s}$ the kernels of E_k and Δ_{k+s} respectively. The kernel of the operator $\mathsf{E}_k T \Delta_{k+s}$ is then given by

$$k_{s,k}(x,y) \, = \, \int_{\mathbb{R}^n \times \mathbb{R}^n} k_{\mathsf{e},k}(x,w)\, k(w,z)\, k_{\delta,k+s}(z,y)\, dw\, dz.$$

It is straightforward to verify that

$$\begin{aligned} k_{\mathsf{e},k}(x,w) &= 2^{nk} \sum_{R \in \mathcal{Q}_k} 1_{R \times R}(x,w), \\ k_{\delta,k+s}(z,y) &= 2^{n(k+s)} \sum_{Q \in \mathcal{Q}_{k+s}} \Big(1_{Q \times Q}(z,y) - \frac{1}{2^n}\, 1_{Q \times \widehat{Q}}(z,y) \Big). \end{aligned}$$

Given $x, y \in \mathbb{R}^n$, define R_x to be the only cube in $\mathcal{Q}_k$ containing x, while Q_y will stand for the only cube in $\mathcal{Q}_{k+s}$ containing y. Moreover, let $Q_2, Q_3, \ldots, Q_{2^n}$ be the remaining cubes in $\mathcal{Q}_{k+s}$ sharing dyadic father with Q_y. Let us introduce the following functions

$$\phi_{R_x}(w) \;\; = \;\; \frac{1}{|R_x|}\, 1_{R_x}(w),$$

$$\psi_{\widehat{Q}_y}(z) \quad = \quad \frac{1}{|\widehat{Q}_y|} \sum_{j=2}^{2^n} 1_{Q_y}(z) - 1_{Q_j}(z).$$

Then the kernel $k_{s,k}(x,y)$ can be written as follows

$$k_{s,k}(x,y) = \Big\langle T(\psi_{\widehat{Q}_y}), \phi_{R_x} \Big\rangle. \tag{2.2}$$

Notice that $\psi_{\widehat{Q}_y} \in H_1$ since it is a linear combination of atoms.

2.4.1. *Schur type estimates.* In this paragraph we give pointwise estimates for the kernels $k_{s,k}$ and use them to obtain upper bounds of the Schur integrals associated to them. Both will be used below to produce Cotlar type estimates.

Lemma 2.3. *The following estimates hold*:

a) *If $y \in \mathbb{R}^n \setminus 3R_x$, we have*

$$\big|k_{s,k}(x,y)\big| \le \mathrm{c}_n \, 2^{-\gamma(k+s)} \frac{1}{|x-y|^{n+\gamma}}.$$

b) *If $y \in 3R_x \setminus R_x$, we have*

$$\big|k_{s,k}(x,y)\big| \le \mathrm{c}_{n,\gamma} \, 2^{-\gamma(k+s)} 2^{nk} \min \bigg\{ \int_{R_x} \frac{dw}{|w-\mathrm{c}_y|^{n+\gamma}}, s2^{\gamma(k+s)} \bigg\}.$$

c) *Similarly, if $y \in R_x$ we have*

$$\big|k_{s,k}(x,y)\big| \le \mathrm{c}_{n,\gamma} \, 2^{-\gamma(k+s)} 2^{nk} \min \bigg\{ \int_{\mathbb{R}^n \setminus R_x} \frac{dw}{|w-\mathrm{c}_y|^{n+\gamma}}, s2^{\gamma(k+s)} \bigg\}.$$

The constant $\mathrm{c}_{n,\gamma}$ only depends on n and γ; c_y denotes the center of the cube $\widehat{Q}_y$.

Proof. We proceed in several steps.

The first estimate. Using

$$\int_{\mathbb{R}^n} \psi_{\widehat{Q}_y}(z)\, dz = 0,$$

we obtain the following identity where c_y denotes the center of $\widehat{Q}_y$

$$\big|k_{s,k}(x,y)\big| \; = \; \Big| \int_{R_x \times \widehat{Q}_y} \phi_{R_x}(w) \big[k(w,z) - k(w,\mathrm{c}_y)\big] \psi_{\widehat{Q}_y}(z)\, dw\, dz \Big|.$$

Since $|z-\mathrm{c}_y| \le \frac{1}{2}|w-\mathrm{c}_y|$ for $(w,z) \in R_x \times \widehat{Q}_y$, Lipschitz smoothness gives

$$\big|k_{s,k}(x,y)\big| \; \le \; \int_{R_x \times \widehat{Q}_y} \phi_{R_x}(w) \frac{|z-\mathrm{c}_y|^\gamma}{|w-\mathrm{c}_y|^{n+\gamma}} |\psi_{\widehat{Q}_y}(z)|\, dw\, dz.$$

Then, we use $|w-\mathrm{c}_y| \ge \frac{1}{3}|x-y|$ and $|z-\mathrm{c}_y| \le 2^{-(k+s)}$ for $(w,z) \in R_x \times \widehat{Q}_y$

$$\big|k_{s,k}(x,y)\big| \; \le \; \mathrm{c}_n \frac{2^{-\gamma(k+s)}}{|x-y|^{n+\gamma}} \int_{R_x \times \widehat{Q}_y} \phi_{R_x}(w) |\psi_{\widehat{Q}_y}(z)|\, dw\, dz \le \mathrm{c}_n \frac{2^{-\gamma(k+s)}}{|x-y|^{n+\gamma}}.$$

The second estimate. By (2.1), we have

$$\int_{\mathbb{R}^n} T(\psi_{\widehat{Q}_y})(w)\, dw = 0.$$

Using this cancellation, we shall use the following relations:

- If $y \notin R_x \Rightarrow \widehat{Q}_y \not\subset R_x$ and $\big|k_{s,k}(x,y)\big| \;=\; \dfrac{1}{|R_x|}\Big|\displaystyle\int_{R_x} T(\psi_{\widehat{Q}_y})(w)\,dw\Big|.$
- If $y \in R_x \Rightarrow \widehat{Q}_y \subset R_x$ and $\big|k_{s,k}(x,y)\big| \;=\; \dfrac{1}{|R_x|}\Big|\displaystyle\int_{\mathbb{R}^n\setminus R_x} T(\psi_{\widehat{Q}_y})(w)\,dw\Big|.$

In the first case, we may have

b1) $3\widehat{Q}_y \cap R_x = \emptyset$,
b2) $3\widehat{Q}_y \cap R_x \neq \emptyset$.

If $3\widehat{Q}_y \cap R_x = \emptyset$, we may use Lipschitz smoothness as above to obtain

$$\begin{aligned}\big|k_{s,k}(x,y)\big| &\le \frac{1}{|R_x|}\int_{R_x\times\widehat{Q}_y} \frac{|z-\mathrm{c}_y|^\gamma}{|w-\mathrm{c}_y|^{n+\gamma}}|\psi_{\widehat{Q}_y}(z)|\,dw\,dz,\\ &\le \mathrm{c}_n 2^{-\gamma(k+s)}\,2^{nk}\int_{R_x}\frac{dw}{|w-\mathrm{c}_y|^{n+\gamma}}.\end{aligned}$$

On the other hand, if $3\widehat{Q}_y \cap R_x \neq \emptyset$ we use the latter estimate on $R_x \setminus 3\widehat{Q}_y$

$$\begin{aligned}\big|k_{s,k}(x,y)\big| &\le \mathrm{c}_n\,2^{-\gamma(k+s)}\,2^{nk}\int_{R_x\setminus 3\widehat{Q}_y}\frac{dw}{|w-\mathrm{c}_y|^{n+\gamma}}\\ &+ \frac{\mathrm{c}_n}{|R_x|\,|\widehat{Q}_y|}\int_{(R_x\cap 3\widehat{Q}_y)\times\widehat{Q}_y}|k(w,z)|\,dw\,dz.\end{aligned}$$

We claim that the second term on the right is dominated by the first one, up to a constant $\mathrm{c}_{n,\gamma}$ depending only on n and γ. Indeed, let us write $\delta_z = \mathrm{dist}(z,\partial\widehat{Q}_y)$ with $\partial\Omega$ denoting the boundary of Ω. The size estimate for the kernel gives

$$\frac{1}{|R_x|\,|\widehat{Q}_y|}\int_{(R_x\cap 3\widehat{Q}_y)\times\widehat{Q}_y}\big|k(w,z)\big|\,dw\,dz \le 2^{nk}\,2^{n(k+s)}\int_{\widehat{Q}_y}\int_{R_x\cap 3\widehat{Q}_y}\frac{dw}{|w-z|^n}\,dz.$$

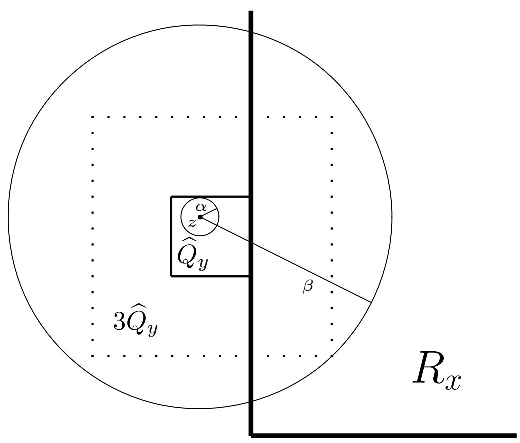


FIGURE I
We have $\alpha = \delta_z$ and $\beta \le 2\sqrt{n}2^{-(k+s-1)}$

According to Figure I, we easily see that

$$\begin{aligned}
& 2^{nk}\, 2^{n(k+s)} \int_{\widehat{Q}_y} \int_{R_x \cap 3\widehat{Q}_y} \frac{dw}{|w-z|^n}\, dz \\
\le \ & \mathrm{c}_n 2^{nk}\, 2^{n(k+s)} \int_{\widehat{Q}_y} \Big[\int_{\mathrm{S}_{n-1}} \Big(\int_{\delta_z}^{2\sqrt{n}\, 2^{-(k+s-1)}} \frac{dr}{r} \Big)\, d\sigma \Big]\, dz \\
\le \ & \mathrm{c}_n 2^{nk}\, 2^{n(k+s)} \int_{\widehat{Q}_y} \log \Big(\frac{2\sqrt{n}\, 2^{-(k+s-1)}}{\delta_z} \Big)\, dz \cdot \sigma(\mathrm{S}_{n-1}) \\
\sim \ & \mathrm{c}_n 2^{nk}\, 2^{n(k+s)} \int_0^{\sqrt{n}/2^{k+s}} \log \Big(\frac{4\sqrt{n}\, 2^{-(k+s)}}{\sqrt{n}\, 2^{-(k+s)} - r} \Big)\, r^{n-1}\, dr \ \le \ \mathrm{c}_n 2^{nk}.
\end{aligned}$$

This gives rise to

$$\big| k_{s,k}(x,y) \big| \le \mathrm{c}_n\, 2^{-\gamma(k+s)}\, 2^{nk} \int_{R_x} \frac{dw}{|w - \mathrm{c}_y|^{n+\gamma}} \ + \ \mathrm{c}_n 2^{nk}\, 1_{\mathcal{U}_{s,k}^x}(y),$$

where the set $\mathcal{U}_{s,k}^x$ is defined by

$$\mathcal{U}_{s,k}^x = \Big\{ y \in \mathbb{R}^n \setminus R_x \ \Big| \ \mathrm{dist}(y, \partial R_x) < 2^{-(k+s-1)} \Big\}.$$

However, it is easily seen that for $y \in \mathcal{U}_{s,k}^x$ we have

$$2^{-\gamma(k+s)} \int_{R_x} \frac{dw}{|w - \mathrm{c}_y|^{n+\gamma}} \ \ge \ \mathrm{c}_n 2^{-\gamma(k+s)} \int_{\mathrm{S}_{n-1}} \Big(\int_{2^{-(k+s)}}^{2^{-(k+s)} + 2^{-k}} \frac{dr}{r^{1+\gamma}} \Big)\, d\sigma \ge \mathrm{c}_{n,\gamma}.$$

In particular, we deduce our claim and so

$$\big| k_{s,k}(x,y) \big| \le \mathrm{c}_{n,\gamma}\, 2^{-\gamma(k+s)}\, 2^{nk} \int_{R_x} \frac{dw}{|w - \mathrm{c}_y|^{n+\gamma}}.$$

In the second case $\widehat{Q}_y \subset R_x$, we may have

- c1) $3\widehat{Q}_y \cap (\mathbb{R}^n \setminus R_x) = \emptyset$,
- c2) $3\widehat{Q}_y \cap (\mathbb{R}^n \setminus R_x) \neq \emptyset$.

The argument in this case is entirely similar. Indeed, if the intersection is empty we use Lipschitz smoothness one more time and the same argument as above gives

$$\big| k_{s,k}(x,y) \big| \ \le \ 2^{-\gamma(k+s)}\, 2^{nk} \int_{\mathbb{R}^n \setminus R_x} \frac{dw}{|w - \mathrm{c}_y|^{n+\gamma}}.$$

If the intersection is not empty, the inequality

$$\frac{1}{|R_x|\, |\widehat{Q}_y|} \int_{((\mathbb{R}^n \setminus R_x) \cap 3\widehat{Q}_y) \times \widehat{Q}_y} |k(w,z)|\, dw\, dz \ \le \ \mathrm{c}_n 2^{nk}$$

can be proved as above. This gives rise to the estimate

$$\big| k_{s,k}(x,y) \big| \le \mathrm{c}_n\, 2^{-\gamma(k+s)}\, 2^{nk} \int_{\mathbb{R}^n \setminus R_x} \frac{dw}{|w - \mathrm{c}_y|^{n+\gamma}} \ + \ \mathrm{c}_n 2^{nk}\, 1_{\mathcal{V}_{s,k}^x}(y),$$

where the set $\mathcal{V}_{s,k}^x$ is defined by

$$\mathcal{V}_{s,k}^x = \Big\{ y \in R_x \ \Big| \ \mathrm{dist}(y, \partial R_x) < 2^{-(k+s-1)} \Big\}.$$

Now we use that for $y \in \mathcal{V}_{s,k}^x$ we have

$$2^{-\gamma(k+s)} \int_{\mathbb{R}^n \setminus R_x} \frac{dw}{|w - \mathrm{c}_y|^{n+\gamma}} \ \geq \ \mathrm{c}_n 2^{-\gamma(k+s)} \int_{\mathrm{S}_{n-1}} \Big(\int_{2^{-(k+s)}}^{\infty} \frac{dr}{r^{1+\gamma}} \Big) \, d\sigma \geq \mathrm{c}_{n,\gamma}.$$

Our estimates prove the first halves of inequalities b) and c) above.

The third estimate. It remains to prove that

$$\big| k_{s,k}(x,y) \big| = \Big| \Big\langle T(\psi_{\widehat{Q}_y}), \phi_{R_x} \Big\rangle \Big| \leq \mathrm{c}_n s \, 2^{nk}.$$

Since $y \in 3R_x$, the localization estimate in Paragraph 2.1 gives

$$\begin{aligned} \big| k_{s,k}(x,y) \big| & \leq & \mathrm{c}_n \, \ell(\widehat{Q}_y)^n \log \Big(\frac{\ell(3R_x)}{\ell(\widehat{Q}_y)} \Big) \, \|\phi_{R_x}\|_\infty \|\psi_{\widehat{Q}_y}\|_\infty \\ & = & \mathrm{c}_n \, |\widehat{Q}_y| \, \log \big(3 \, 2^{s-1} \big) \, \frac{1}{|R_x|} \, \frac{2^n - 1}{|\widehat{Q}_y|} \ \leq \ \mathrm{c}_n \, s \, 2^{nk}. \end{aligned}$$

We have used that T is assumed to be L_2-normalized. The proof is complete. □

Lemma 2.4. *Let us define*

$$\begin{aligned} \mathcal{S}_{s,k}^1(x) & = & \int_{\mathbb{R}^n} \big| k_{s,k}(x,y) \big| \, dy, \\ \mathcal{S}_{s,k}^2(y) & = & \int_{\mathbb{R}^n} \big| k_{s,k}(x,y) \big| \, dx. \end{aligned}$$

Then, there exists a constant $\mathrm{c}_{n,\gamma}$ *depending only on* n, γ *such that*

$$\begin{aligned} \mathcal{S}_{s,k}^1(x) & \leq & \frac{\mathrm{c}_{n,\gamma} \, s}{2^{\gamma s}} \quad \textit{for all} \quad (x,k) \in \mathbb{R}^n \times \mathbb{Z}, \\ \mathcal{S}_{s,k}^2(y) & \leq & \mathrm{c}_{n,\gamma} \, s \quad \textit{for all} \quad (y,k) \in \mathbb{R}^n \times \mathbb{Z}. \end{aligned}$$

Proof. We estimate $\mathcal{S}_{s,k}^1$ and $\mathcal{S}_{s,k}^2$ in turn.

Estimate of $\mathcal{S}_{s,k}^1(x)$. Given $x \in \mathbb{R}^n$, define the cube R_x as above. Then we decompose the integral defining $\mathcal{S}_{s,k}^1(x)$ into three regions according to Lemma 2.3 and estimate each one independently. Using Lemma 2.3 a) we find

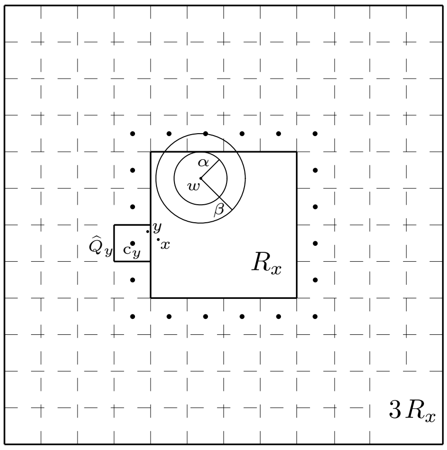
$$\int_{\mathbb{R}^n \setminus 3R_x} \big| k_{s,k}(x,y) \big| \, dy \ \leq \ \mathrm{c}_n \, 2^{-\gamma(k+s)} \int_{\mathbb{R}^n \setminus 3R_x} \frac{dy}{|x-y|^{n+\gamma}} \ \leq \ \mathrm{c}_n \, 2^{-\gamma s}. \tag{2.3}$$

On the other hand, the first estimate in Lemma 2.3 b) gives

$$\int_{3R_x \setminus R_x} \big| k_{s,k}(x,y) \big| \, dy \ \leq \ \mathrm{c}_{n,\gamma} \, 2^{-\gamma(k+s)} \, 2^{nk} \int_{3R_x \setminus R_x} \int_{R_x} \frac{1}{|w - \mathrm{c}_y|^{n+\gamma}} \, dw \, dy.$$

Now we set $\delta_w = \mathrm{dist}(w, \partial R_x)$ for $w \in R_x$. Then we clearly have

$$\widehat{\delta}_w \equiv \delta_w + 2^{-(k+s)} \leq \delta_w + \mathrm{dist}(\mathrm{c}_y, \partial R_x) \leq |w - \mathrm{c}_y|.$$



$(\alpha, \beta) = (\delta_w, \widehat{\delta}_w)$

FIGURE II
Even if x and y are close, we have a $(\widehat{\delta}_w - \delta_w)$ – margin

In particular, we find (see Figure II above)

$$\int_{3R_x \setminus R_x} \frac{dy}{|w - \mathrm{c}_y|^{n+\gamma}} \lesssim \frac{|\mathsf{B}_{\widehat{\delta}_w}(w)|}{\widehat{\delta}_w^{n+\gamma}} + \int_{\mathbb{R}^n \setminus \mathsf{B}_{\widehat{\delta}_w}(w)} \frac{dy}{|w-y|^{n+\gamma}} \sim 1/\widehat{\delta}_w^{\gamma}.$$

This provides us with the estimate

$$\int_{3R_x \setminus R_x} \int_{R_x} \frac{1}{|w - \mathrm{c}_y|^{n+\gamma}}\, dw\, dy \quad \le \quad \mathrm{c}_n \int_{\mathrm{S}_{n-1}} \int_0^{2^{-k}} \frac{r^{n-1}}{(2^{-(k+s)} + 2^{-k} - r)^{\gamma}}\, dr\, d\sigma.$$

Using $t = 2^{-k} + 2^{-(k+s)} - r$ and the bound $r \le 2^{-k}$

$$\begin{aligned} \int_{3R_x \setminus R_x} \int_{R_x} \frac{1}{|w - \mathrm{c}_y|^{n+\gamma}}\, dw\, dy \quad & \le \quad \mathrm{c}_n 2^{-(n-1)k} \int_{2^{-(k+s)}}^{2^{-k}} \frac{dt}{t^{\gamma}} \\ & \le \quad \mathrm{c}_n \begin{cases} s\, 2^{-(n-1)k} & \text{if } \gamma = 1, \\ \mathrm{c}_\gamma\, 2^{-nk}\, 2^{\gamma k} & \text{if } 0 < \gamma < 1. \end{cases} \end{aligned}$$

In summary, combining our estimates we have obtained

$$\int_{3R_x \setminus R_x} \big| k_{s,k}(x,y) \big|\, dy \ \le\ \mathrm{c}_{n,\gamma}\, s\, 2^{-\gamma s}. \tag{2.4}$$

It remains to control the integral over R_x. By Lemma 2.3 c)

$$\int_{R_x} \big| k_{s,k}(x,y) \big|\, dy \ \le\ \mathrm{c}_{n,\gamma}\, 2^{-\gamma(k+s)}\, 2^{nk} \int_{R_x} \int_{\mathbb{R}^n \setminus R_x} \frac{1}{|w - \mathrm{c}_y|^{n+\gamma}}\, dw\, dy.$$

For any given $y \in R_x$, we set again

$$\delta_{\mathrm{c}_y} = \mathrm{dist}(\mathrm{c}_y, \partial R_x) \ge 2^{-(k+s)}.$$

Arguing as above, we may use polar coordinates to obtain

$$\int_{R_x} \int_{\mathbb{R}^n \setminus R_x} \frac{1}{|w - \mathrm{c}_y|^{n+\gamma}} \, dw \, dy \ \leq \ \int_{R_x} \Big(\int_{\mathrm{S}_{n-1}} \int_{\delta_{\mathrm{c}_y}}^{\infty} \frac{r^{n-1}}{r^{n+\gamma}} \, dr \, d\sigma \Big) \, dy \ \sim \ \int_{R_x} \frac{dy}{\delta_{\mathrm{c}_y}^{\gamma}}$$

$$\sim \ \int_{\mathrm{S}_{n-1}} \int_{0}^{2^{-k} - 2^{-(k+s)}} \frac{r^{n-1}}{(2^{-k} - r)^{\gamma}} \, dr \, d\sigma$$

$$+ \ \int_{\mathrm{S}_{n-1}} \int_{2^{-k} - 2^{-(k+s)}}^{2^{-k}} \frac{r^{n-1}}{2^{-\gamma(k+s)}} \, dr \, d\sigma.$$

The first integral is estimated as above

$$\int_{\mathrm{S}_{n-1}} \int_{0}^{2^{-k} - 2^{-(k+s)}} \frac{r^{n-1}}{(2^{-k} - r)^{\gamma}} \, dr \, d\sigma \ \leq \ \mathrm{c}_n \begin{cases} s \, 2^{-(n-1)k} & \text{if } \gamma = 1, \\ \mathrm{c}_\gamma \, 2^{-nk} \, 2^{\gamma k} & \text{if } 0 < \gamma < 1, \end{cases}$$

as for the second we obtain an even better bound. Indeed, we have

$$\begin{aligned} \int_{\mathrm{S}_{n-1}} \int_{2^{-k} - 2^{-(k+s)}}^{2^{-k}} \frac{r^{n-1}}{2^{-\gamma(k+s)}} \, dr \, d\sigma \ &\sim \ 2^{\gamma(k+s)} \Big(2^{-nk} - \big[2^{-k} - 2^{-(k+s)} \big]^n \Big) \\ &= \ 2^{\gamma(k+s)} \, 2^{-nk} \Big(1 - \big[1 - 2^{-s} \big]^n \Big) \\ &\leq \ 2^{\gamma(k+s)} \, 2^{-nk} \sum_{j=1}^{n} \binom{n}{j} 2^{-sj} \\ &\leq \ \mathrm{c}_n \, 2^{-nk} \, 2^{\gamma k}. \end{aligned}$$

Writing all together we finally get

$$\int_{R_x} \big| k_{s,k}(x,y) \big| \, dy \ \leq \ \mathrm{c}_{n,\gamma} \, s \, 2^{-\gamma s}. \tag{2.5}$$

According to (2.3), (2.4) and (2.5) we obtain the upper bound $\mathcal{S}_{s,k}^1(x) \leq \mathrm{c}_{n,\gamma} \, s \, 2^{-\gamma s}$.

Estimate of $\mathcal{S}_{s,k}^2(y)$. Given a fixed point y, we consider a partition $\mathbb{R}^n = \Omega_1 \cup \Omega_2$ where Ω_1 is the set of points x such that $y \notin 3R_x$ and $\Omega_2 = \mathbb{R}^n \setminus \Omega_1$. In the region Ω_1 we may proceed as in (2.3). On the other hand, inside Ω_2 and according to Lemma 2.3 we know that $|k_{s,k}(x,y)| \leq \mathrm{c}_{n,\gamma} \, s \, 2^{nk}$. This means that we have

$$\mathcal{S}_{s,k}^2(y) \leq \mathrm{c}_{n,\gamma} \Big(2^{-\gamma s} + |\Omega_2| \, s \, 2^{nk} \Big) = \mathrm{c}_{n,\gamma} \Big(2^{-\gamma s} + |3R_y| \, s \, 2^{nk} \Big) \leq \mathrm{c}_{n,\gamma} \, s.$$

This upper bound holds for all $(y,k) \in \mathbb{R}^n \times \mathbb{Z}$. Hence, the proof is complete. □

2.4.2. *Cotlar type estimates.* Let us write $\Lambda_{s,k}$ for $\mathsf{E}_k T \Delta_{k+s}$. According to the pairwise orthogonality of martingale differences, we have $\Lambda_{s,i} \Lambda_{s,j}^* = 0$ whenever $i \neq j$. In particular, it follows from Cotlar lemma that it suffices to control the norm of the operators $\Lambda_{s,i}^* \Lambda_{s,j}$. Explicitly, our estimate for Φ_s stated in the shifted $T1$ theorem will be deduced from

$$\big\| \Lambda_{s,i}^* \Lambda_{s,j} \big\|_{\mathcal{B}(L_2)} \ \leq \ \mathrm{c}_{n,\gamma} \, s^2 \, 2^{-\gamma s/2} \alpha_{i-j}^2$$

for some summable sequence $(\alpha_k)_{k \in \mathbb{Z}}$. The kernel of $\Lambda_{s,i}^* \Lambda_{s,j}$ is given by

$$k_{i,j}^s(x,y) \ = \ \int_{\mathbb{R}^n} \overline{k_{s,i}(z,x)} \, k_{s,j}(z,y) \, dz.$$

Before proceeding with our estimates we need to point out another cancellation property which easily follows from (2.1). Given $r > 0$ and a point $y \in \mathbb{R}^n$, let $f(z) = 1_{\mathsf{B}_r(y)}(z)/|\mathsf{B}_r(y)|$. Then it is clear that $\Delta_{k+s} f = df_{k+s}$ is in H_1 since it can be written as a linear combination of atoms. According to our cancellation condition (2.1) we find

$$\int_{\mathbb{R}^n} \mathsf{E}_k T \Delta_{k+s} f(x) \, dx = \int_{\mathbb{R}^n} T df_{k+s}(x) \, dx = 0.$$

In terms of the kernels, this identity is written as

$$\int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} k_{s,k}(x,z) f(z) \, dz \Big) \, dx = 0.$$

Using Fubini theorem (our estimates in Lemma 2.3 ensure the integrability) and taking the limit as $r \to 0$, the Lebesgue differentiation theorem implies the following identity, which holds for almost every point y

$$\int_{\mathbb{R}^n} k_{s,k}(x,y) \, dx = 0. \tag{2.6}$$

This holds for all $k \in \mathbb{Z}$ and we deduce

$$\begin{aligned} k_{i,j}^s(x,y) &= \int_{\mathbb{R}^n} \overline{k_{s,i}(z,x)} \Big(k_{s,j}(z,y) - k_{s,j}(x,y) \Big) \, dz \\ &= \int_{\mathbb{R}^n} \Big(\overline{k_{s,i}(z,x)} - \overline{k_{s,i}(y,x)} \Big) \, k_{s,j}(z,y) \; dz. \end{aligned}$$

In order to estimate the kernels $k_{i,j}^s$, we use the first or the second expression above according to whether $i \ge j$ or not. Since the estimates are entirely similar we shall assume in what follows that $i \ge j$ and work in the sequel with the first expression above. Moreover, given $w \in \mathbb{R}^n$ we shall write all through out this paragraph R_w for the only cube in $\mathcal{Q}_j$ containing w. Then, since $R_z = R_x$ whenever $z \in R_x$, it follows from (2.2) that

$$\begin{aligned} k_{i,j}^s(x,y) &= \int_{\mathbb{R}^n \setminus 3R_x} \overline{k_{s,i}(z,x)} \Big(k_{s,j}(z,y) - k_{s,j}(x,y) \Big) \, dz \\ &+ \int_{3R_x \setminus R_x} \overline{k_{s,i}(z,x)} \Big(k_{s,j}(z,y) - k_{s,j}(x,y) \Big) \, dz. \end{aligned}$$

If $\alpha_{i,j}^s(x,y)$ and $\beta_{i,j}^s(x,y)$ are the first and second terms above, let

$$\begin{aligned} \mathcal{S}_{i,j,s}^{1,\alpha}(x) &= \int_{\mathbb{R}^n} \big| \alpha_{i,j}^s(x,y) \big| \, dy, \\ \mathcal{S}_{i,j,s}^{2,\alpha}(y) &= \int_{\mathbb{R}^n} \big| \alpha_{i,j}^s(x,y) \big| \, dx, \\ \mathcal{S}_{i,j,s}^{1,\beta}(x) &= \int_{\mathbb{R}^n} \big| \beta_{i,j}^s(x,y) \big| \, dy, \\ \mathcal{S}_{i,j,s}^{2,\beta}(y) &= \int_{\mathbb{R}^n} \big| \beta_{i,j}^s(x,y) \big| \, dx. \end{aligned}$$

According to Schur lemma from Paragraph 2.1, we obtain the upper bound

$$\big\| \Lambda_{s,i}^* \Lambda_{s,j} \big\|_{\mathcal{B}(L_2)} \le \sqrt{\Big(\big\| \mathcal{S}_{i,j,s}^{1,\alpha} \big\|_\infty + \big\| \mathcal{S}_{i,j,s}^{1,\beta} \big\|_\infty \Big) \Big(\big\| \mathcal{S}_{i,j,s}^{2,\alpha} \big\|_\infty + \big\| \mathcal{S}_{i,j,s}^{2,\beta} \big\|_\infty \Big)}. \tag{2.7}$$

Lemma 2.5. *We have*

$$\max\Big\{\big\|\mathcal{S}_{i,j,s}^{1,\alpha}\big\|_\infty, \big\|\mathcal{S}_{i,j,s}^{1,\beta}\big\|_\infty\Big\} \, \le \, \mathrm{c}_{n,\gamma}\, s\,(s+|i-j|)\, 2^{-\gamma s}.$$

Proof. According to Lemma 2.3, we know that

$$\big|k_{s,i}(z,x)\big| \le \mathrm{c}_n\, 2^{-\gamma(i+s)} \frac{1}{|x-z|^{n+\gamma}}$$

whenever $z \notin 3R_x$. Moreover, Lemma 2.4 gives

$$\int_{\mathbb{R}^n} \big|k_{s,j}(z,y) - k_{s,j}(x,y)\big|\, dy \, \le \, \mathrm{c}_{n,\gamma}\, s\, 2^{-\gamma s}.$$

If we combine the two estimates above, we obtain

$$\begin{aligned} \mathcal{S}_{i,j,s}^{1,\alpha}(x) & \le \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n \setminus 3R_x} \big|k_{s,i}(z,x)\big| \, \big|k_{s,j}(z,y) - k_{s,j}(x,y)\big|\, dz \Big)\, dy \\ & = \int_{\mathbb{R}^n \setminus 3R_x} \big|k_{s,i}(z,x)\big| \Big(\int_{\mathbb{R}^n} \big|k_{s,j}(z,y) - k_{s,j}(x,y)\big|\, dy \Big)\, dz \\ & \le \mathrm{c}_{n,\gamma}\, s\, 2^{-\gamma s} 2^{-\gamma(i+s)} \int_{\mathbb{R}^n \setminus 3R_x} \frac{dz}{|x-z|^{n+\gamma}} \, \le \, \mathrm{c}_{n,\gamma}\, s\, 2^{-2\gamma s} 2^{-\gamma |i-j|}. \end{aligned}$$

The last inequality uses the assumption $i \ge j$, so that $i-j = |i-j|$. The estimate above holds for all $x \in \mathbb{R}^n$. Hence, since $\mathrm{c}_{n,\gamma} s\, 2^{-2\gamma s} 2^{-\gamma|i-j|}$ is much smaller than $\mathrm{c}_{n,\gamma} s\,(s+|i-j|)\, 2^{-\gamma s}$, it is clear that the first function satisfies the thesis. Let us now proceed with the second function. To that aim we observe that $k_{s,j}(z,y)$ is j-measurable as a function in z, meaning that $\mathsf{E}_j(k_{s,j}(\,\cdot\,,y))(z) = k_{s,j}(z,y)$. This follows from (2.2). In particular, the same holds for the function

$$1_{3R_x \setminus R_x}(z) \Big(k_{s,j}(z,y) - k_{s,j}(x,y) \Big).$$

Therefore, using the integral invariance of conditional expectations

$$\begin{aligned} \beta_{i,j}^s(x,y) & = \int_{3R_x \setminus R_x} \overline{\mathsf{E}_j(k_{s,i}(\,\cdot\,,x))(z)} \Big(k_{s,j}(z,y) - k_{s,j}(x,y) \Big)\, dz \\ & = \sum_{R \sim R_x} \int_R \overline{\frac{1}{|R|} \int_R k_{s,i}(w,x)\, dw} \Big(k_{s,j}(z,y) - k_{s,j}(x,y) \Big)\, dz, \end{aligned}$$

where $R \sim R_x$ is used to denote that R is a neighbor of R_x in $\mathcal{Q}_j$. That is, the neighbors of R_x form a partition of $3R_x \setminus R_x$ formed by $3^n - 1$ cubes in $\mathcal{Q}_j$. If c_R denotes the center of R, we use that $k_{s,j}(z,y) = k_{s,j}(\mathrm{c}_R, y)$ for $z \in R$ and obtain the estimate

$$\big|\beta_{i,j}^s(x,y)\big| \le \sum_{R \sim R_x} \Big| \int_R k_{s,i}(w,x)\, dw \Big| \, \big|k_{s,j}(\mathrm{c}_R, y) - k_{s,j}(x,y)\big|. \tag{2.8}$$

This, combined with Lemma 2.4, produces

$$\mathcal{S}_{i,j,s}^{1,\beta}(x) \, \le \, \mathrm{c}_{n,\gamma}\, s\, 2^{-\gamma s} \sum_{R \sim R_x} \Big| \int_R k_{s,i}(w,x)\, dw \Big|. \tag{2.9}$$

Let us now estimate the integral. If $w \in S_w \in \mathcal{Q}_i$ and $x \in O_x \in \mathcal{Q}_{i+s}$

$$\int_R k_{s,i}(w,x)\, dw \;\; = \;\; \int_R \Big\langle T\psi_{\widehat{O}_x}, \phi_{S_w} \Big\rangle\, dw$$

$$= \sum_{S \subset R, S \in \mathcal{Q}_i} \int_S T\psi_{\widehat{O}_x}(z)\, dz \;=\; \Big\langle T\psi_{\widehat{O}_x}, 1_R \Big\rangle.$$

Now we use the localization estimate from Paragraph 2.1 to obtain

$$\Big| \Big\langle T\psi_{\widehat{O}_x}, 1_R \Big\rangle \Big| \;\le\; \mathrm{c}_n\, \ell(\widehat{O}_x)^n \log \Big(\frac{\ell(3R)}{\ell(\widehat{O}_x)} \Big) \|\psi_{\widehat{O}_x}\|_\infty \|1_R\|_\infty \;\le\; \mathrm{c}_n \big(s + |i-j| \big).$$

Since there are $3^n - 1$ neighbors, this estimate completes the proof with (2.9). □

Lemma 2.6. *We have*

$$\max \Big\{ \big\| \mathcal{S}^{2,\alpha}_{i,j,s} \big\|_\infty, \big\| \mathcal{S}^{2,\beta}_{i,j,s} \big\|_\infty \Big\} \;\le\; \mathrm{c}_n s^2 \big(1 + |i-j|\big) \, 2^{-\gamma|i-j|}.$$

Proof. Once again, Lemma 2.3 gives

$$\big| k_{s,i}(z,x) \big| \le \mathrm{c}_n\, 2^{-\gamma(i+s)} \frac{1}{|x-z|^{n+\gamma}} \qquad \text{for} \qquad z \notin 3\,R_x.$$

This, together with Fubini theorem produces

$$\begin{aligned}
\mathcal{S}^{2,\alpha}_{i,j,s}(y) \;&\le\; \mathrm{c}_n \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n \setminus 3R_x} \frac{2^{-\gamma(i+s)}}{|x-z|^{n+\gamma}} \big| k_{s,j}(z,y) - k_{s,j}(x,y) \big|\, dz \Big)\, dx \\
&\le\; \mathrm{c}_n \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n \setminus \mathsf{B}_{2^{-j}}(z)} \frac{2^{-\gamma(i+s)}}{|x-z|^{n+\gamma}}\, dx \Big) \big| k_{s,j}(z,y) \big|\, dz \\
&+\; \mathrm{c}_n \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n \setminus \mathsf{B}_{2^{-j}}(x)} \frac{2^{-\gamma(i+s)}}{|x-z|^{n+\gamma}}\, dz \Big) \big| k_{s,j}(x,y) \big|\, dx \\
&=\; \mathrm{c}_n\, 2^{-\gamma s} 2^{-\gamma|i-j|} \int_{\mathbb{R}^n} \big| k_{s,j}(w,y) \big|\, dw.
\end{aligned}$$

Now, according to Lemma 2.4 we now that the integral on the right is bounded by $\mathrm{c}_{n,\gamma} s$ for all y in $\mathbb{R}^n$. Therefore, the L_∞ norm of the first function is much smaller than our upper bound. Let us now estimate the second function. If we proceed as in Lemma 2.5 and use (2.8), we find

$$\mathcal{S}^{2,\beta}_{i,j,s}(y) \;\le\; \int_{\mathbb{R}^n} \sum_{R \sim R_x} \Big| \int_R k_{s,i}(w,x)\, dw \Big| \big| k_{s,j}(\mathrm{c}_R, y) - k_{s,j}(x,y) \big|\, dx.$$

Now we need a different estimate for the integral of $k_{s,i}(\,\cdot\,, x)$ over the neighbor cubes R of R_x. Indeed, combining the pointwise estimates obtained in Lemma 2.3 it easily follows that

$$\big| k_{s,i}(w,x) \big| \;\le\; \mathrm{c}_{n,\gamma} \frac{s\, 2^{ni}}{\big(1 + 2^i |x-w|\big)^{n+\gamma}} \qquad \text{for all} \qquad (w,x) \in \mathbb{R}^n \times \mathbb{R}^n. \tag{2.10}$$

If we set $\delta_x = \mathrm{dist}(x, \partial R_x) \le \mathrm{dist}(x, \partial R)$, we get

$$\begin{aligned}
\Big| \int_R k_{s,i}(w,x)\, dw \Big| \;&\le\; \mathrm{c}_{n,\gamma}\, s \int_R \frac{2^{ni}}{\big(1 + 2^i |x-w|\big)^{n+\gamma}}\, dw \\
&\le\; \mathrm{c}_{n,\gamma}\, s \int_{\mathrm{S}_{n-1}} \Big(\int_{\delta_x}^\infty \frac{2^{ni} r^{n-1}}{\big(1 + 2^i r\big)^{n+\gamma}}\, dr \Big)\, d\sigma \\
&=\; \mathrm{c}_{n,\gamma}\, s \int_{2^i \delta_x}^\infty \frac{z^{n-1}}{\big(1+z\big)^{n+\gamma}}\, dz \;\le\; \mathrm{c}_{n,\gamma}\, s\, \frac{1}{\big(1 + 2^i \delta_x\big)^\gamma}.
\end{aligned}$$

Using (2.10) for $k_{s,j}$, we have

$$\mathcal{S}_{i,j,s}^{2,\beta}(y) \le \mathrm{c}_{n,\gamma}\, s^2\, 2^{nj}\, \Upsilon(i,j,\gamma)$$

where the term $\Upsilon(i,j,\gamma)$ is given by

$$\int_{\mathbb{R}^n} \sum_{R\sim R_x} \frac{1}{\big(1+2^i\delta_x\big)^\gamma} \left(\frac{1}{\big(1+2^j|\mathrm{c}_R - y|\big)^{n+\gamma}} + \frac{1}{\big(1+2^j|x-y|\big)^{n+\gamma}} \right) dx.$$

It is straightforward to see that it suffices to estimate the integral

$$\int_{\mathbb{R}^n} \frac{1}{\big(1+2^i\delta_x\big)^\gamma} \frac{1}{\big(1+2^j|x-y|\big)^{n+\gamma}}\, dx. \tag{2.11}$$

Indeed, both functions inside the big bracket above are comparable and the sum $\sum_{R\sim R_x}$ can be deleted since it only provides an extra factor of 3^n-1. Now, the main idea to estimate (2.11) is to observe that the two functions in the integrand are nearly independent inside any dyadic cube of $\mathcal{Q}_j$. Let us be more explicit, we have

$$\begin{aligned}
&\int_{\mathbb{R}^n} \frac{1}{\big(1+2^i\delta_x\big)^\gamma} \frac{1}{\big(1+2^j|x-y|\big)^{n+\gamma}}\, dx \\
&\quad\le \sum_{R\in\mathcal{Q}_j} \int_R \frac{1}{\big(1+2^i\delta_x\big)^\gamma}\, dx \quad \frac{1}{\big(1+2^j\,\mathrm{dist}(R,R_y)\big)^{n+\gamma}} \\
&\quad\le \sup_{R\in\mathcal{Q}_j} \Big(\int_R \frac{1}{\big(1+2^i\delta_x\big)^\gamma}\, dx \Big) \sum_{R\in\mathcal{Q}_j} \frac{1}{\big(1+2^j\,\mathrm{dist}(R,R_y)\big)^{n+\gamma}} \\
&\quad\sim \sup_{R\in\mathcal{Q}_j} \Big(\int_R \frac{1}{\big(1+2^i\delta_x\big)^\gamma}\, dx \Big) \int_{\mathbb{R}^n} \frac{2^{nj}}{\big(1+2^j\,|x-y|\big)^{n+\gamma}}\, dx.
\end{aligned}$$

The integral on the right is majorized by an absolute constant. Moreover, recalling that δ_x stands for $\mathrm{dist}(x,\partial R_x)$ and that $R_x = R$ for any $x\in R\in\mathcal{Q}_j$, it is clear that the integral on the left does not depend on the chosen cube R, so that the supremum is unnecessary. To estimate this integral we set $\lambda = 1-2^{j-i}$

$$\begin{aligned}
\int_R \frac{dx}{\big(1+2^i\delta_x\big)^\gamma} &\sim \int_{\mathrm{S}_{n-1}} \Big(\int_0^{2^{-j}} \frac{r^{n-1}}{\big(1+2^i(2^{-j}-r)\big)^\gamma}\, dr \Big)\, d\sigma \\
&\sim \int_0^{\lambda 2^{-j}} \frac{r^{n-1}}{\big(1+2^i(2^{-j}-r)\big)^\gamma}\, dr + \int_{\lambda 2^{-j}}^{2^{-j}} \frac{r^{n-1}}{\big(1+2^i(2^{-j}-r)\big)^\gamma}\, dr.
\end{aligned}$$

The first integral is majorized by

$$\begin{aligned}
2^{-\gamma i} \int_0^{\lambda 2^{-j}} \frac{r^{n-1}}{(2^{-j}-r)^\gamma}\, dr &\le 2^{-\gamma i}\lambda^{n-1} 2^{-nj} 2^j \int_0^{\lambda 2^{-j}} \frac{dr}{(2^{-j}-r)^\gamma} \\
&\le 2^{-nj} \begin{cases} |i-j| 2^{-|i-j|} & \text{if } \gamma = 1, \\ \mathrm{c}_\gamma 2^{-\gamma|i-j|} & \text{if } 0<\gamma<1. \end{cases}
\end{aligned}$$

The second integral is majorized by

$$\int_{\lambda 2^{-j}}^{2^{-j}} r^{n-1}\, dr \le 2^{-nj} 2^j (2^{-j} - \lambda 2^{-j}) = 2^{-nj} 2^{-|i-j|}.$$

Combining our estimates we finally get

$$\mathcal{S}^{2,\beta}_{i,j,s}(y) \le \mathrm{c}_{n,\gamma}\, s^2 \big(1+|i-j|\big)\, 2^{-\gamma|i-j|}.$$

Since the last estimate holds for all $y \in \mathbb{R}^n$, the proof is complete. □

Conclusion. According to (2.7), Lemmas 2.5 and 2.6 give

$$\big\|\Lambda^*_{s,i}\Lambda_{s,j}\big\|_{\mathcal{B}(L_2)} \le \mathrm{c}_{n,\gamma}\sqrt{s^3\,(s+|i-j|)^2\, 2^{-\gamma s}\, 2^{-\gamma|i-j|}} \le \mathrm{c}_{n,\gamma}\, s^2\, 2^{-\gamma s/2}\, \alpha^2_{i-j},$$

where $\alpha_k = (1+|k|)^{\frac12}\, 2^{-\gamma|k|/4}$. In particular, Cotlar lemma provides the estimate

$$\|\Phi_s\|_{\mathcal{B}(L_2)} \; = \; \Big\|\sum\nolimits_k \mathsf{E}_k T \Delta_{k+s}\Big\|_{\mathcal{B}(L_2)} \; \le \; \mathrm{c}_{n,\gamma}\, s\, 2^{-\gamma s/4} \sum\nolimits_k \alpha_k \; = \; \mathrm{c}_{n,\gamma}\, s\, 2^{-\gamma s/4}.$$

2.5. **Estimating the norm of Ψ_s.** We finally estimate the operator norm of Ψ_s. This will complete the proof of our pseudo-localization principle. We shall adapt some of the notation introduced in the previous paragraph. Namely, we shall now write $\Lambda_{s,k}$ when referring to the operator $(id-\mathsf{E}_k)T_{4\cdot 2^{-k}}\Delta_{k+s}$ and $k_{s,k}(x,y)$ will be reserved for its kernel. Arguing as above it is simple to check that we have

$$k_{s,k}(x,y) = T_{4\cdot 2^{-k}}\psi_{\widehat{Q}_y}(x) - \Big\langle T_{4\cdot 2^{-k}}\psi_{\widehat{Q}_y}, \phi_{R_x}\Big\rangle.$$

We shall use the terminology

$$\begin{aligned} k^1_{s,k}(x,y) &= T_{4\cdot 2^{-k}}\psi_{\widehat{Q}_y}(x), \\ k^2_{s,k}(x,y) &= \Big\langle T_{4\cdot 2^{-k}}\psi_{\widehat{Q}_y}, \phi_{R_x}\Big\rangle. \end{aligned}$$

2.5.1. *Schur type estimates.*

Lemma 2.7. *Let us consider the sets*

$$\mathcal{W}^x_{s,k} \; = \; \Big\{ w \in \mathbb{R}^n \;\big|\; 4\cdot 2^{-k} - 2^{-(k+s-1)} \le |x-w| < 4\cdot 2^{-k} + 2^{-(k+s-1)} \Big\}.$$

Then, the following pointwise estimate holds

$$\big|k_{s,k}(x,y)\big| \; \le \; \mathrm{c}_n 1_{\mathbb{R}^n\setminus \mathsf{B}_{2\cdot 2^{-k}}(x)}(y)\Big(\frac{2^{-\gamma(k+s)}}{|x-y|^{n+\gamma}} + 2^{nk}\, 1_{\mathcal{W}^x_{s,k}}(y)\Big).$$

Proof. We have

$$k^1_{s,k}(x,y) \; = \; \int_{\widehat{Q}_y} 1_{\mathbb{R}^n\setminus \mathsf{B}_{4\cdot 2^{-k}}(x)}(z)\, k(x,z)\, \psi_{\widehat{Q}_y}(z)\, dz.$$

If $|x-y| \le 3\cdot 2^{-k}$ we have

$$|x-z| \le |x-y| + |y-z| \le |x-y| + 2^{-(k+s-1)} \le 4\cdot 2^{-k}$$

since $z \in \widehat{Q}_y$. In particular, we obtain

$$k^1_{s,k}(x,y) = 0 \qquad \text{whenever} \qquad |x-y| \le 3\cdot 2^{-k}.$$

If $|x-y| > 5\cdot 2^{-k}$, then we have for $z \in \widehat{Q}_y$

$$|x-z| \ge |x-y| - |z-y| \ge |x-y| - 2^{-(k+s-1)} > 4\cdot 2^{-k}.$$

Thus, we can argue in the usual way and obtain

$$\big|k^1_{s,k}(x,y)\big| \quad = \quad \Big|\int_{\widehat{Q}_y}\Big(k(x,z) - k(x,\mathrm{c}_y)\Big)\,\psi_{\widehat{Q}_y}(z)\, dz\Big|$$

$$\leq \ c_n \frac{2^{-\gamma(k+s)}}{|x-y|^{n+\gamma}} \int_{\widehat{Q}_y} \big|\psi_{\widehat{Q}_y}(z)\big| \, dz \ \leq \ c_n \frac{2^{-\gamma(k+s)}}{|x-y|^{n+\gamma}}.$$

If $3 \cdot 2^{-k} < |x-y| \leq 5 \cdot 2^{-k}$, we write $k_{s,k}^1(x,y)$ as a sum of two integrals

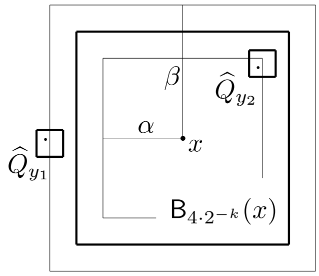
$$\begin{aligned} k_{s,k}^1(x,y) &= \int_{\mathbb{R}^n} 1_{\mathbb{R}^n \setminus \mathsf{B}_{4 \cdot 2^{-k}}(x)}(z) \Big(k(x,z) - k(x,c_y) \Big) \psi_{\widehat{Q}_y}(z) \, dz \\ &+ \int_{\mathbb{R}^n} k(x,c_y) \Big(1_{\mathbb{R}^n \setminus \mathsf{B}_{4 \cdot 2^{-k}}(x)}(z) - 1_{\mathbb{R}^n \setminus \mathsf{B}_{4 \cdot 2^{-k}}(x)}(c_y) \Big) \psi_{\widehat{Q}_y}(z) \, dz \\ &= \mathsf{A}_1 + \mathsf{B}_1. \end{aligned}$$

Here we have used that $\psi_{\widehat{Q}_y}$ is mean-zero. Lipschitz smoothness gives once more

$$|\mathsf{A}_1| \ \leq \ c_n \frac{2^{-\gamma(k+s)}}{|x-y|^{n+\gamma}}.$$

To estimate B_1 we use the size condition on the kernel

$$|\mathsf{B}_1| \ \leq \ \frac{c_n}{|x-c_y|^n} \frac{1}{|\widehat{Q}_y|} \int_{\widehat{Q}_y} \big| 1_{\mathbb{R}^n \setminus \mathsf{B}_{4 \cdot 2^{-k}}(x)}(z) - 1_{\mathbb{R}^n \setminus \mathsf{B}_{4 \cdot 2^{-k}}(x)}(c_y) \big| \, dz.$$



$$(\alpha, \beta) = \big(4 \cdot 2^{-k} - 2^{-(k+s-1)}, 4 \cdot 2^{-k} + 2^{-(k+s-1)}\big)$$

FIGURE III
If $y \notin \mathcal{W}_{s,k}^x = \mathsf{B}_\beta(x) \setminus \mathsf{B}_\alpha(x)$, we have $\widehat{Q}_y \cap \partial \mathsf{B}_{4 \cdot 2^{-k}}(x) = \emptyset$

Since $3 \cdot 2^{-k} < |x-y| \leq 5 \cdot 2^{-k}$, we have $c_n |x-c_y|^{-n} \sim c_n 2^{nk}$. Moreover, the only z's for which the integrand above is not zero are those with (z, c_y) lying at different sides of $\partial \mathsf{B}_{4 \cdot 2^{-k}}(x)$. This can only happen when $y \in \mathcal{W}_{s,k}^x$ and we get

$$|\mathsf{B}_1| \leq c_n 2^{nk} \frac{1_{\mathcal{W}_{s,k}^x}(y)}{|\widehat{Q}_y|} \int_{\widehat{Q}_y} \big| 1_{\mathbb{R}^n \setminus \mathsf{B}_{4 \cdot 2^{-k}}(x)}(z) - 1_{\mathbb{R}^n \setminus \mathsf{B}_{4 \cdot 2^{-k}}(x)}(c_y) \big| \, dz \leq c_n 2^{nk} 1_{\mathcal{W}_{s,k}^x}(y).$$

Combining our estimates obtained so far we get

$$(2.12) \qquad \big| k_{s,k}^1(x,y) \big| \ \leq \ c_n 1_{\mathbb{R}^n \setminus \mathsf{B}_{3 \cdot 2^{-k}}(x)}(y) \Big(\frac{2^{-\gamma(k+s)}}{|x-y|^{n+\gamma}} + 2^{nk} 1_{\mathcal{W}_{s,k}^x}(y) \Big).$$

Let us now study pointwise estimates for the kernel

$$k_{s,k}^2(x,y) \ = \ \frac{1}{|R_x|} \int_{R_x} \Big(\int_{\widehat{Q}_y} 1_{\mathbb{R}^n \setminus \mathsf{B}_{4 \cdot 2^{-k}}(w)}(z) \, k(w,z) \, \psi_{\widehat{Q}_y}(z) \, dz \Big) \, dw.$$

If $|x-y| \le 2 \cdot 2^{-k}$ we have

$$|w-z| \le |w-x| + |x-y| + |y-z| \le 4 \cdot 2^{-k}$$

for all $(w,z) \in R_x \times \widehat{Q}_y$. This gives

$$k_{s,k}^2(x,y) = 0 \qquad \text{whenever} \qquad |x-y| \le 2 \cdot 2^{-k}.$$

If $|x-y| > 6 \cdot 2^{-k}$, then we have for $(w,z) \in R_x \times \widehat{Q}_y$

$$|w-z| \ge |x-y| - |x-w| - |z-y| > 4 \cdot 2^{-k}.$$

Therefore, we obtain as usual the estimate

$$\begin{aligned} \big|k_{s,k}^2(x,y)\big| &\le \frac{1}{|R_x|} \int_{R_x} \Big(\int_{\widehat{Q}_y} \big|k(w,z) - k(w,\mathrm{c}_y)\big| \, \big|\psi_{\widehat{Q}_y}(z)\big| \, dz \Big) \, dw \\ &\le \frac{2^{-\gamma(k+s)}}{|R_x|} \int_{R_x} \frac{1}{|w-\mathrm{c}_y|^{n+\gamma}} \Big(\int_{\widehat{Q}_y} \big|\psi_{\widehat{Q}_y}(z)\big| \, dz \Big) \, dw \\ &\le \mathrm{c}_n \frac{2^{-\gamma(k+s)}}{|R_x|} \int_{R_x} \frac{dw}{|w-\mathrm{c}_y|^{n+\gamma}} \ = \ \mathrm{c}_n \frac{2^{-\gamma(k+s)}}{|x-y|^{n+\gamma}}. \end{aligned}$$

When $2 \cdot 2^{-k} < |x-y| \le 6 \cdot 2^{-k}$ we have the two integrals

$$\begin{aligned} & k_{s,k}^2(x,y) \\ &= \frac{1}{|R_x|} \int_{R_x \times \widehat{Q}_y} 1_{\mathbb{R}^n \setminus \mathsf{B}_{4 \cdot 2^{-k}}(w)}(z) \Big(k(w,z) - k(w,\mathrm{c}_y) \Big) \psi_{\widehat{Q}_y}(z) \, dw \, dz \\ &+ \frac{1}{|R_x|} \int_{R_x \times \widehat{Q}_y} k(w,\mathrm{c}_y) \Big(1_{\mathbb{R}^n \setminus \mathsf{B}_{4 \cdot 2^{-k}}(w)}(z) - 1_{\mathbb{R}^n \setminus \mathsf{B}_{4 \cdot 2^{-k}}(w)}(\mathrm{c}_y) \Big) \psi_{\widehat{Q}_y}(z) \, dw \, dz \\ &= \mathsf{A}_2 + \mathsf{B}_2. \end{aligned}$$

By Lipschitz smoothness, we may estimate A_2 by

$$|\mathsf{A}_2| \le \mathrm{c}_n \frac{2^{-\gamma(k+s)}}{|R_x| \, |\widehat{Q}_y|} \int_{R_x \times \widehat{Q}_y} \frac{1_{\mathbb{R}^n \setminus \mathsf{B}_{4 \cdot 2^{-k}}(w)}(z)}{|w-z|^{n+\gamma}} \, dw \, dz \le \mathrm{c}_n 2^{nk} \, 2^{-\gamma s}$$

since $4 \cdot 2^{-k} \le |w-z| \le |w-x| + |x-y| + |y-z| \le 8 \cdot 2^{-k}$. That is

$$|\mathsf{A}_2| \le \mathrm{c}_n \frac{2^{-\gamma(k+s)}}{|x-y|^{n+\gamma}} \qquad \big(2 \cdot 2^{-k} < |x-y| \le 6 \cdot 2^{-k} \big).$$

To estimate B_2 we first observe that

$$|w-\mathrm{c}_y| \ge |x-y| - |x-w| - |\mathrm{c}_y - y| \ge (2 - 1 - \tfrac{1}{2}) 2^{-k} = \tfrac{1}{2} 2^{-k}.$$

Then we apply the size estimate for the kernel and Fubini theorem

$$\begin{aligned} |\mathsf{B}_2| &\le \frac{1}{|R_x| \, |\widehat{Q}_y|} \int_{R_x \times \widehat{Q}_y} \frac{\big| 1_{\mathbb{R}^n \setminus \mathsf{B}_{4 \cdot 2^{-k}}(w)}(z) - 1_{\mathbb{R}^n \setminus \mathsf{B}_{4 \cdot 2^{-k}}(w)}(\mathrm{c}_y) \big|}{|w-\mathrm{c}_y|^n} \, dw \, dz \\ &\le \frac{\mathrm{c}_n 2^{nk}}{|R_x| \, |\widehat{Q}_y|} \int_{\widehat{Q}_y} \Big(\int_{R_x} \big| 1_{\mathbb{R}^n \setminus \mathsf{B}_{4 \cdot 2^{-k}}(w)}(z) - 1_{\mathbb{R}^n \setminus \mathsf{B}_{4 \cdot 2^{-k}}(w)}(\mathrm{c}_y) \big| \, dw \Big) \, dz. \end{aligned}$$

In the integral inside the brackets, the points z and c_y are fixed. Moreover, since $z \in \widehat{Q}_y$ we know that $|z - \mathrm{c}_y| \le 2^{-(k+s)}$. Therefore, we find that the only w's for which the integrand of the inner integral is not zero live in

$$\mathcal{W}_{s+1,k}^{\mathrm{c}_y} = \Big\{ w \in \mathbb{R}^n \ \big| \ 4 \cdot 2^{-k} - 2^{-(k+s)} \le |w - \mathrm{c}_y| < 4 \cdot 2^{-k} + 2^{-(k+s)} \Big\}.$$

This automatically gives the estimate

$$|\mathsf{B}_2| \le \frac{c_n 2^{nk}}{|R_x|} |\mathcal{W}_{s+1,k}^{c_y}| \le c_n 2^{nk} 2^{-s} \le c_n \frac{2^{-\gamma(k+s)}}{|x-y|^{n+\gamma}}.$$

Our partial estimates so far produce the global estimate

$$\big|k_{s,k}^2(x,y)\big| \le c_n 1_{\mathbb{R}^n \setminus \mathsf{B}_{2\cdot 2^{-k}}(x)}(y) \frac{2^{-\gamma(k+s)}}{|x-y|^{n+\gamma}}. \tag{2.13}$$

The assertion then follows from a combination of inequalities (2.12) and (2.13). □

Lemma 2.8. *Let us define*

$$\begin{aligned} \mathcal{S}_{s,k}^1(x) &= \int_{\mathbb{R}^n} \big|k_{s,k}(x,y)\big| \, dy, \\ \mathcal{S}_{s,k}^2(y) &= \int_{\mathbb{R}^n} \big|k_{s,k}(x,y)\big| \, dx. \end{aligned}$$

Then there exists a constant c_n *such that*

$$\max \Big\{ \mathcal{S}_{s,k}^1(x), \mathcal{S}_{s,k}^2(y) \Big\} \le c_n 2^{-\gamma s}.$$

Proof. According to Lemma 2.7 and $|\mathcal{W}_{s,k}^x| \le c_n 2^{-nk} 2^{-s}$

$$\mathcal{S}_{s,k}^1(x) \le c_n \int_{\mathbb{R}^n \setminus \mathsf{B}_{2\cdot 2^{-k}}(x)} \Big(\frac{2^{-\gamma(k+s)}}{|x-y|^{n+\gamma}} + 2^{nk} 1_{\mathcal{W}_{s,k}^x}(y) \Big) \, dy \le c_n 2^{-\gamma s}.$$

The same argument applies for $\mathcal{S}_{s,k}^2(y)$, since we have $1_{\mathcal{W}_{s,k}^x}(y) = 1_{\mathcal{W}_{s,k}^y}(x)$. □

2.5.2. *Cotlar type estimates.* We have again $\Lambda_{s,i}\Lambda_{s,j}^* = 0$ for $i \neq j$, so that we are reduced (by Cotlar lemma) to estimate the norms of $\Lambda_{s,i}^*\Lambda_{s,j}$ in $\mathcal{B}(L_2)$. The kernel of $\Lambda_{s,i}^*\Lambda_{s,j}$ is given by

$$k_{i,j}^s(x,y) = \int_{\mathbb{R}^n} \overline{k_{s,i}(z,x)} k_{s,j}(z,y) \, dz.$$

Taking $f(z) = 1_{\mathsf{B}_r(y)}(z)/|\mathsf{B}_r(y)|$, we note

$$\frac{1}{|\mathsf{B}_r(y)|} \int_{\mathsf{B}_r(y)} \Big(\int_{\mathbb{R}^n} k_{s,k}(x,z) \, dx \Big) \, dz = \int_{\mathbb{R}^n} (id - \mathsf{E}_k) T_{4\cdot 2^{-k}} \Delta_{k+s} f(x) \, dx = 0,$$

due to the integral invariance of conditional expectations. Taking the limit as $r \to 0$, we deduce from Lebesgue differentiation theorem that the cancellation condition (2.6) also holds for our new kernels $k_{s,k}(x,y)$ and for a.e. $y \in \mathbb{R}^n$. In particular, the same discussion as above leads us to use (2.6) in one way or another according to $i \ge j$ or viceversa. Both cases can be estimated in the same way. Thus we assume in what follows that $i \ge j$ and use the expression

$$\begin{aligned} k_{i,j}^s(x,y) &= \int_{\mathbb{R}^n} \overline{k_{s,i}(z,x)} \Big(k_{s,j}(z,y) - k_{s,j}(x,y) \Big) \, dz \\ &= \int_{\mathbb{R}^n \setminus \mathsf{B}_{2\cdot 2^{-j}}(x)} \overline{k_{s,i}(z,x)} \Big(k_{s,j}(z,y) - k_{s,j}(x,y) \Big) \, dz \\ &+ \int_{\mathsf{B}_{2\cdot 2^{-j}}(x) \setminus \mathsf{B}_{2\cdot 2^{-i}}(x)} \overline{k_{s,i}(z,x)} \Big(k_{s,j}(z,y) - k_{s,j}(x,y) \Big) \, dz. \end{aligned}$$

Observe that the integrand vanishes for z in $\mathsf{B}_{2\cdot 2^{-i}}(x)$ since $k_{s,i}(z,x)$ does, according to Lemma 2.7. Let us write $\alpha_{i,j}^s(x,y)$ and $\beta_{i,j}^s(x,y)$ for the first and second terms on the right. Then (as before) we need to estimate the quantity

$$\sqrt{\Big(\big\|\mathcal{S}_{i,j,s}^{1,\alpha}\big\|_\infty + \big\|\mathcal{S}_{i,j,s}^{1,\beta}\big\|_\infty\Big)\Big(\big\|\mathcal{S}_{i,j,s}^{2,\alpha}\big\|_\infty + \big\|\mathcal{S}_{i,j,s}^{2,\beta}\big\|_\infty\Big)},$$

where the $\mathcal{S}$ functions are given by

$$\begin{aligned}
\mathcal{S}_{i,j,s}^{1,\alpha}(x) &= \int_{\mathbb{R}^n} \big|\alpha_{i,j}^s(x,y)\big|\,dy,\\
\mathcal{S}_{i,j,s}^{2,\alpha}(y) &= \int_{\mathbb{R}^n} \big|\alpha_{i,j}^s(x,y)\big|\,dx,\\
\mathcal{S}_{i,j,s}^{1,\beta}(x) &= \int_{\mathbb{R}^n} \big|\beta_{i,j}^s(x,y)\big|\,dy,\\
\mathcal{S}_{i,j,s}^{2,\beta}(y) &= \int_{\mathbb{R}^n} \big|\beta_{i,j}^s(x,y)\big|\,dx.
\end{aligned}$$

Lemma 2.9. *We have*

$$\max\Big\{\big\|\mathcal{S}_{i,j,s}^{1,\alpha}\big\|_\infty, \big\|\mathcal{S}_{i,j,s}^{1,\beta}\big\|_\infty\Big\} \le \mathrm{c}_n 2^{-2\gamma s}.$$

Proof. According to Lemma 2.7, we know that

$$\big|k_{s,i}(z,x)\big| \le \mathrm{c}_n\Big(\frac{2^{-\gamma(i+s)}}{|x-z|^{n+\gamma}} + 2^{ni}\,1_{\mathcal{W}_{s,i}^x}(z)\Big)$$

for all $z \in \mathbb{R}^n \setminus \mathsf{B}_{2\cdot 2^{-j}}(x)$. Moreover, Lemma 2.8 gives

$$\int_{\mathbb{R}^n} \big|k_{s,j}(z,y) - k_{s,j}(x,y)\big|\,dy \le \mathrm{c}_n 2^{-\gamma s}.$$

Combining these estimates we find an L_∞ bound for $\mathcal{S}_{i,j,s}^{1,\alpha}$

$$\begin{aligned}
\mathcal{S}_{i,j,s}^{1,\alpha}(x) &\le \mathrm{c}_n 2^{-\gamma s}\int_{\mathbb{R}^n\setminus \mathsf{B}_{2\cdot 2^{-j}}(x)} \Big(\frac{2^{-\gamma(i+s)}}{|x-z|^{n+\gamma}} + 2^{ni}\,1_{\mathcal{W}_{s,i}^x}(z)\Big)\,dz\\
&\le \mathrm{c}_n 2^{-2\gamma s}2^{-\gamma|i-j|} + \mathrm{c}_n 2^{-\gamma s}2^{ni}\big|(\mathbb{R}^n\setminus \mathsf{B}_{2\cdot 2^{-j}}(x)) \cap \mathcal{W}_{s,i}^x\big|.
\end{aligned}$$

We claim that $\mathcal{S}_{i,j,s}^{1,\alpha}(x) \le \mathrm{c}_n 2^{-2\gamma s}2^{-\gamma|i-j|}$. Indeed, if the intersection above is empty there is nothing to prove. If it is not empty, the following inequality must hold

$$4\cdot 2^{-i} + 2^{-(i+s-1)} > 2\cdot 2^{-j}.$$

This implies that we can only have $i=j$ or $i=j+1$ and hence

$$2^{-\gamma s}2^{ni}\big|(\mathbb{R}^n\setminus \mathsf{B}_{2\cdot 2^{-j}}(x))\cap\mathcal{W}_{s,i}^x\big| \le 2^{-\gamma s}2^{ni}\big|\mathcal{W}_{s,i}^x\big| \le 2^{-2\gamma s} \sim 2^{-2\gamma s}2^{-\gamma|i-j|}.$$

Therefore, the first function clearly satisfies the thesis. Let us now analyze the second function. To that aim we proceed exactly as above in $\mathsf{B}_{2\cdot 2^{-j}}(x)\setminus \mathsf{B}_{2\cdot 2^{-i}}(x)$ and obtain

$$\begin{aligned}
\mathcal{S}_{i,j,s}^{1,\beta}(x) &\le \mathrm{c}_n 2^{-\gamma s}\int_{\mathsf{B}_{2\cdot 2^{-j}}(x)\setminus \mathsf{B}_{2\cdot 2^{-i}}(x)} \Big(\frac{2^{-\gamma(i+s)}}{|x-z|^{n+\gamma}} + 2^{ni}\,1_{\mathcal{W}_{s,i}^x}(z)\Big)\,dz\\
&\le \mathrm{c}_n 2^{-2\gamma s} + \mathrm{c}_n 2^{-\gamma s}2^{ni}\big|(\mathbb{R}^n\setminus \mathsf{B}_{2\cdot 2^{-j}}(x))\cap\mathcal{W}_{s,i}^x\big| \le \mathrm{c}_n 2^{-2\gamma s}.
\end{aligned}$$

□

Lemma 2.10. *We have*

$$\max\Big\{\big\|\mathcal{S}_{i,j,s}^{2,\alpha}\big\|_\infty, \big\|\mathcal{S}_{i,j,s}^{2,\beta}\big\|_\infty\Big\} \le \mathrm{c}_{n,\gamma}\,\big(1+|i-j|\big)\,2^{-\gamma|i-j|}.$$

Proof. For the first function we have

$$\begin{aligned}
\mathcal{S}_{i,j,s}^{2,\alpha}(y) & \le \mathrm{c}_n \int_{\mathbb{R}^n} \Big[\int_{\mathbb{R}^n\setminus\mathsf{B}_{2\cdot 2^{-j}}(x)} \Big(\frac{2^{-\gamma(i+s)}}{|x-z|^{n+\gamma}} + 2^{ni}1_{\mathcal{W}_{s,i}^x}(z)\Big)\big|k_{s,j}(z,y)\big|\,dz\Big]\,dx \\
& + \mathrm{c}_n \int_{\mathbb{R}^n} \Big[\int_{\mathbb{R}^n\setminus\mathsf{B}_{2\cdot 2^{-j}}(x)} \Big(\frac{2^{-\gamma(i+s)}}{|x-z|^{n+\gamma}} + 2^{ni}1_{\mathcal{W}_{s,i}^x}(z)\Big)\big|k_{s,j}(x,y)\big|\,dz\Big]\,dx \\
& = \mathrm{c}_n \int_{\mathbb{R}^n} \Big[\int_{\mathbb{R}^n\setminus\mathsf{B}_{2\cdot 2^{-j}}(z)} \Big(\frac{2^{-\gamma(i+s)}}{|x-z|^{n+\gamma}} + 2^{ni}1_{\mathcal{W}_{s,i}^z}(x)\Big)\,dx\Big]\big|k_{s,j}(z,y)\big|\,dz \\
& + \mathrm{c}_n \int_{\mathbb{R}^n} \Big[\int_{\mathbb{R}^n\setminus\mathsf{B}_{2\cdot 2^{-j}}(x)} \Big(\frac{2^{-\gamma(i+s)}}{|x-z|^{n+\gamma}} + 2^{ni}1_{\mathcal{W}_{s,i}^x}(z)\Big)\,dz\Big]\big|k_{s,j}(x,y)\big|\,dx, \\
& \le \mathrm{c}_n 2^{-3\gamma s}\,2^{-\gamma|i-j|}.
\end{aligned}$$

Here we have used Lemma 2.7, Fubini theorem and $1_{\mathcal{W}_{s,i}^x}(z) = 1_{\mathcal{W}_{s,i}^z}(x)$. The last inequality follows arguing as in Lemma 2.9. Let us now estimate the second $\mathcal{S}$ function. We may assume $i \neq j$ because otherwise $\mathcal{S}_{i,j,s}^{2,\beta} = 0$. Let us decompose

$$\big|k_{s,j}(z,y) - k_{s,j}(x,y)\big| \le \mathsf{A} + \mathsf{B},$$

where these terms are given by

$$\begin{aligned}
\mathsf{A} & = \big|k_{s,j}^1(z,y) - k_{s,j}^1(x,y)\big|, \\
\mathsf{B} & = \big|k_{s,j}^2(z,y) - k_{s,j}^2(x,y)\big|.
\end{aligned}$$

Moreover, we further decompose the A-term into

$$\begin{aligned}
\mathsf{A} & \le \int_{\widehat{Q}_y} 1_{\mathbb{R}^n\setminus\mathsf{B}_{4\cdot 2^{-j}}(z)}(w)\,\big|k(z,w)-k(x,w)\big|\,\big|\psi_{\widehat{Q}_y}(w)\big|\,dw \\
& + \int_{\widehat{Q}_y} \big|k(x,w)\big|\,\big|1_{\mathbb{R}^n\setminus\mathsf{B}_{4\cdot 2^{-j}}(z)}(w) - 1_{\mathbb{R}^n\setminus\mathsf{B}_{4\cdot 2^{-j}}(x)}(w)\big|\,\big|\psi_{\widehat{Q}_y}(w)\big|\,dw = \mathsf{A}_1 + \mathsf{A}_2.
\end{aligned}$$

This gives rise to

$$\begin{aligned}
& \mathcal{S}_{i,j,s}^{2,\beta}(y) \\
& \le \int_{\mathbb{R}^n} \Big(\int_{\mathsf{B}_{2\cdot 2^{-j}}(x)\setminus\mathsf{B}_{2\cdot 2^{-i}}(x)} \big|k_{s,i}(z,x)\big|\,\big(\mathsf{A}_1+\mathsf{A}_2+\mathsf{B}\big)\,dz\Big)\,dx \\
& \le \mathrm{c}_n \int_{\mathbb{R}^n} \Big(\int_{\mathsf{B}_{2\cdot 2^{-j}}(x)\setminus\mathsf{B}_{2\cdot 2^{-i}}(x)} \Big(\frac{2^{-\gamma(i+s)}}{|x-z|^{n+\gamma}} + 2^{ni}1_{\mathcal{W}_{s,i}^x}(z)\Big)\big(\mathsf{A}_1+\mathsf{A}_2+\mathsf{B}\big)\,dz\Big)\,dx \\
& = \mathcal{A}_1 + \mathcal{A}_2 + \mathcal{B}.
\end{aligned}$$

The $\mathcal{A}_1$-term. We have

$$\begin{aligned}
\mathsf{A}_1 & \le \int_{\widehat{Q}_y} 1_{\mathbb{R}^n\setminus\mathsf{B}_{4\cdot 2^{-j}}(z)}(w)\,\frac{|x-z|^\gamma}{|z-w|^{n+\gamma}}\,\big|\psi_{\widehat{Q}_y}(w)\big|\,dw \\
& \le \frac{\mathrm{c}_n 2^{nj}\,2^{\gamma j}}{|\widehat{Q}_y|} \int_{\widehat{Q}_y} \frac{|x-z|^\gamma}{\big(1+2^j|z-w|\big)^{n+\gamma}}\,dw \sim \mathrm{c}_n \frac{2^{nj}\,2^{\gamma j}\,|x-z|^\gamma}{\big(1+2^j|z-y|\big)^{n+\gamma}}.
\end{aligned}$$

Lipschitz smoothness is applicable since $z \in \mathsf{B}_{2\cdot 2^{-j}}(x) \setminus \mathsf{B}_{2\cdot 2^{-i}}(x)$. We then have

$$\begin{aligned}\mathcal{A}_1 &\le \mathrm{c}_n \int_{\mathbb{R}^n} \Big(\int_{\mathsf{B}_{2\cdot 2^{-j}}(x)\setminus \mathsf{B}_{2\cdot 2^{-i}}(x)} \frac{2^{-\gamma(i+s)}}{|x-z|^{n+\gamma}} \, \frac{2^{nj}\, 2^{\gamma j}\, |x-z|^\gamma}{\big(1+2^j|z-y|\big)^{n+\gamma}} \, dz \Big) \, dx \\ &+ \mathrm{c}_n \int_{\mathbb{R}^n} \Big(\int_{\mathsf{B}_{2\cdot 2^{-j}}(x)\setminus \mathsf{B}_{2\cdot 2^{-i}}(x)} 2^{ni} 1_{\mathcal{W}_{s,i}^x}(z) \, \frac{2^{nj}\, 2^{\gamma j}\, |x-z|^\gamma}{\big(1+2^j|z-y|\big)^{n+\gamma}} \, dz \Big) \, dx \\ &= \mathcal{A}_{11} + \mathcal{A}_{12}.\end{aligned}$$

The estimate of $\mathcal{A}_{11}$ is standard

$$\begin{aligned}\mathcal{A}_{11} &= \mathrm{c}_n \int_{\mathbb{R}^n} \frac{2^{-\gamma(i+s)}\, 2^{nj}\, 2^{\gamma j}}{\big(1+2^j|z-y|\big)^{n+\gamma}} \Big(\int_{\mathsf{B}_{2\cdot 2^{-j}}(z)\setminus \mathsf{B}_{2\cdot 2^{-i}}(z)} \frac{dx}{|x-z|^n} \Big) \, dz \\ &= \mathrm{c}_n |i-j| \int_{\mathbb{R}^n} \frac{2^{-\gamma(i+s)}\, 2^{nj}\, 2^{\gamma j}}{\big(1+2^j|z-y|\big)^{n+\gamma}} \, dz \ \sim \ \mathrm{c}_n |i-j| \, 2^{-\gamma s}\, 2^{-\gamma|i-j|}.\end{aligned}$$

The term $\mathcal{A}_{12}$ can be written as follows

$$\mathcal{A}_{12} = \mathrm{c}_n \int_{\mathbb{R}^n} \frac{2^{ni}\, 2^{nj}\, 2^{\gamma j}}{\big(1+2^j|z-y|\big)^{n+\gamma}} \Big(\int_{\mathsf{B}_{2\cdot 2^{-j}}(z)\setminus \mathsf{B}_{2\cdot 2^{-i}}(z)} |x-z|^\gamma 1_{\mathcal{W}_{s,i}^z}(x) \, dx \Big) \, dz.$$

Now, the presence of $1_{\mathcal{W}_{s,i}^z}(x)$ implies that

$$|x-z| \le 4 \cdot 2^{-i} + 2^{-(i+s-1)} \le 5 \cdot 2^{-i}.$$

Therefore we find

$$\mathcal{A}_{12} \ \le \ \mathrm{c}_n 2^{-\gamma|i-j|} \int_{\mathbb{R}^n} \frac{2^{ni}\, 2^{nj}\, \big|\mathcal{W}_{s,i}^z\big|}{\big(1+2^j|z-y|\big)^{n+\gamma}} \, dz \ \le \ \mathrm{c}_n \, 2^{-s}\, 2^{-\gamma|i-j|}.$$

This means that $\mathcal{A}_{11}$ dominates $\mathcal{A}_{12}$ and we conclude

$$\mathcal{A}_1 \le \mathrm{c}_n |i-j| \, 2^{-\gamma s}\, 2^{-\gamma|i-j|}. \tag{2.14}$$

The $\mathcal{A}_2$-term. Consider the symmetric difference

$$\mathcal{Z}_{x,z}^j = \mathsf{B}_{4\cdot 2^{-j}}(x) \triangle \mathsf{B}_{4\cdot 2^{-j}}(z) = \Big(\mathsf{B}_{4\cdot 2^{-j}}(x) \setminus \mathsf{B}_{4\cdot 2^{-j}}(z) \Big) \cup \Big(\mathsf{B}_{4\cdot 2^{-j}}(z) \setminus \mathsf{B}_{4\cdot 2^{-j}}(x) \Big).$$

Then we clearly have

$$\mathsf{A}_2 = \int_{\widehat{Q}_y \cap \mathcal{Z}_{x,z}^j} \big|k(x,w)\big| \big|\psi_{\widehat{Q}_y}(w)\big| \, dw \ \le \ \mathrm{c}_n 2^{nj} \frac{\big|\widehat{Q}_y \cap \mathcal{Z}_{x,z}^j\big|}{\big|\widehat{Q}_y\big|},$$

where the 2^{nj} comes from the size condition on the kernel and the inequality

$$|x-w| \ge \mathrm{dist}(x, \partial \mathcal{Z}_{x,z}^j) \ge 4 \cdot 2^{-j} - |x-z| \ge 2 \cdot 2^{-j},$$

which holds for any $w \in \mathcal{Z}_{x,z}^j$ and $z \in \mathsf{B}_{2\cdot 2^{-j}}(x) \setminus \mathsf{B}_{2\cdot 2^{-i}}(x)$. This allows us to write

$$\begin{aligned}\mathcal{A}_2 &\le \mathrm{c}_n 2^{nj} \int_{\mathbb{R}^n} \Big(\int_{\mathsf{B}_{2\cdot 2^{-j}}(x)\setminus \mathsf{B}_{2\cdot 2^{-i}}(x)} \frac{2^{-\gamma(i+s)}}{|x-z|^{n+\gamma}} \, \frac{\big|\widehat{Q}_y \cap \mathcal{Z}_{x,z}^j\big|}{\big|\widehat{Q}_y\big|} \, dz \Big) \, dx \\ &+ \mathrm{c}_n 2^{nj} \int_{\mathbb{R}^n} \Big(\int_{\mathsf{B}_{2\cdot 2^{-j}}(x)\setminus \mathsf{B}_{2\cdot 2^{-i}}(x)} 2^{ni} 1_{\mathcal{W}_{s,i}^x}(z) \frac{\big|\widehat{Q}_y \cap \mathcal{Z}_{x,z}^j\big|}{\big|\widehat{Q}_y\big|} \, dz \Big) \, dx \\ &= \mathcal{A}_{21} + \mathcal{A}_{22}.\end{aligned}$$

Before proceeding with the argument, we note

- If $|x-y| > 7 \cdot 2^{-j}$
$$|z-w| \ge |x-y| - |x-z| - |w-y| > 4 \cdot 2^{-j}$$
for all $(w,z) \in \widehat{Q}_y \times \big(\mathsf{B}_{2\cdot 2^{-j}}(x) \setminus \mathsf{B}_{2\cdot 2^{-i}}(x)\big)$. Similarly, we have
$$|x-w| \ge |x-y| - |w-y| > 6 \cdot 2^{-j}.$$
This implies that $w \notin \mathcal{Z}_{x,z}^j$ for any $w \in \widehat{Q}_y$, so that $\widehat{Q}_y \cap \mathcal{Z}_{x,z}^j = \emptyset$.
- If $|x-y| < 2^{-j}$
$$|z-w| \le |z-x| + |x-y| + |y-w| < 4 \cdot 2^{-j}$$
for all $(w,z) \in \widehat{Q}_y \times \big(\mathsf{B}_{2\cdot 2^{-j}}(x) \setminus \mathsf{B}_{2\cdot 2^{-i}}(x)\big)$. Similarly, we have
$$|x-w| \le |x-y| + |y-w| < 2 \cdot 2^{-j}.$$
This implies that $w \notin \mathcal{Z}_{x,z}^j$ for any $w \in \widehat{Q}_y$, so that $\widehat{Q}_y \cap \mathcal{Z}_{x,z}^j = \emptyset$.

In particular, we conclude that

$$\begin{aligned}
&\mathcal{A}_{21} + \mathcal{A}_{22} \\
&= \mathrm{c}_n 2^{nj} \int_{\mathsf{B}_{7\cdot 2^{-j}}(y)\setminus \mathsf{B}_{2^{-j}}(y)} \Big(\int_{\mathsf{B}_{2\cdot 2^{-j}}(x)\setminus \mathsf{B}_{2\cdot 2^{-i}}(x)} \frac{2^{-\gamma(i+s)}}{|x-z|^{n+\gamma}} \frac{\big|\widehat{Q}_y \cap \mathcal{Z}_{x,z}^j\big|}{\big|\widehat{Q}_y\big|} \, dz \Big) \, dx \\
&+ \mathrm{c}_n 2^{nj} \int_{\mathsf{B}_{7\cdot 2^{-j}}(y)\setminus \mathsf{B}_{2^{-j}}(y)} \Big(\int_{\mathsf{B}_{2\cdot 2^{-j}}(x)\setminus \mathsf{B}_{2\cdot 2^{-i}}(x)} 2^{ni} 1_{\mathcal{W}_{s,i}^x}(z) \frac{\big|\widehat{Q}_y \cap \mathcal{Z}_{x,z}^j\big|}{\big|\widehat{Q}_y\big|} \, dz \Big) \, dx.
\end{aligned}$$

Observe now that $\widehat{Q}_y$ behaves as a ball of radius $2^{-(j+s)}$ while $\mathcal{Z}_{x,z}^j$ behaves like an annulus of radius $4 \cdot 2^{-j}$ and width $|x-z|$. Therefore, the measure of the intersection can be estimated by

$$\big|\widehat{Q}_y \cap \mathcal{Z}_{x,z}^j\big| \ \le\ \mathrm{c}_n \min \Big\{ 2^{-(n-1)(j+s)} |x-z|, 2^{-n(j+s)} \Big\}.$$

This provides us with the estimate

$$\frac{\big|\widehat{Q}_y \cap \mathcal{Z}_{x,z}^j\big|}{\big|\widehat{Q}_y\big|} \ \le\ \mathrm{c}_n \min \Big\{ 2^{j+s} |x-z|, 1 \Big\}.$$

If $2^{-(j+s)} \le 2 \cdot 2^{-i}$

$$\mathcal{A}_{21} \le \mathrm{c}_n 2^{nj} 2^{-\gamma(i+s)} \int_{|x-y|<7\cdot 2^{-j}} \Big(\int_{|z-x|>2^{-(j+s)}} \frac{dz}{|x-z|^{n+\gamma}} \Big) \, dx \le \mathrm{c}_n 2^{-\gamma|i-j|}.$$

If $2^{-(j+s)} > 2 \cdot 2^{-i}$

$$\begin{aligned}
\mathcal{A}_{21} &\le \mathrm{c}_n 2^{nj} 2^{-\gamma(i+s)} \int_{|x-y|<7\cdot 2^{-j}} \Big(\int_{|z-x|>2^{-(j+s)}} \frac{dz}{|x-z|^{n+\gamma}} \Big) \, dx \\
&+ \mathrm{c}_n 2^{nj} 2^{-\gamma(i+s)} \int_{|x-y|<7\cdot 2^{-j}} \Big(\int_{\mathsf{B}_{2^{-(j+s)}}(x)\setminus \mathsf{B}_{2\cdot 2^{-i}}(x)} \frac{2^{j+s}|x-z|}{|x-z|^{n+\gamma}} \, dz \Big) \, dx \\
&\le \mathrm{c}_n 2^{-\gamma|i-j|} \ +\ \mathrm{c}_n \begin{cases} (i-j-s) 2^{-|i-j|} & \text{if } \gamma = 1, \\ \mathrm{c}_\gamma 2^{-\gamma|i-j|} & \text{if } 0 < \gamma < 1. \end{cases}
\end{aligned}$$

This gives $\mathcal{A}_{21} \le c_{n,\gamma}|i-j|2^{-\gamma|i-j|}$. On the other hand, we also have

$$\mathcal{A}_{22} \le c_n 2^{ni} 2^{nj} 2^{j+s} \int_{\mathsf{B}_{7\cdot 2^{-j}}(y)\setminus \mathsf{B}_{2^{-j}}(y)} \Big(\int_{\mathsf{B}_{2\cdot 2^{-j}}(x)\setminus \mathsf{B}_{2\cdot 2^{-i}}(x)} |x-z| 1_{\mathcal{W}_{s,i}^x}(z)\, dz \Big)\, dx.$$

Since we have $|x-z| < 5\cdot 2^{-i}$ for $z \in \mathcal{W}_{s,i}^x$ and $|\mathcal{W}_{s,i}^x| \le c_n 2^{-ni} 2^{-s}$, we get

$$\mathcal{A}_{22} \le c_n 2^{-|i-j|}.$$

Therefore, $\mathcal{A}_{21}$ dominates $\mathcal{A}_{22}$ and we conclude

$$\mathcal{A}_2 \le c_{n,\gamma} |i-j| 2^{-\gamma|i-j|}. \tag{2.15}$$

The $\mathcal{B}$-term. As usual, we decompose

$$\begin{aligned} \mathcal{B} &\le c_n \int_{\mathbb{R}^n} \Big(\int_{\mathsf{B}_{2\cdot 2^{-j}}(x)\setminus \mathsf{B}_{2\cdot 2^{-i}}(x)} \frac{2^{-\gamma(i+s)}}{|x-z|^{n+\gamma}} \mathsf{B}\, dz \Big)\, dx \\ &+ c_n \int_{\mathbb{R}^n} \Big(\int_{\mathsf{B}_{2\cdot 2^{-j}}(x)\setminus \mathsf{B}_{2\cdot 2^{-i}}(x)} 2^{ni} 1_{\mathcal{W}_{s,i}^x}(z) \mathsf{B}\, dz \Big)\, dx = \mathcal{B}_1 + \mathcal{B}_2, \end{aligned}$$

with $\mathsf{B} = \big| k_{s,j}^2(z,y) - k_{s,j}^2(x,y) \big|$. We have

$$\mathcal{B}_1 \le c_n \int_{\mathbb{R}^n} \Big(\int_{\mathsf{B}_{2\cdot 2^{-j}}(x)\setminus \mathsf{B}_{2\cdot 2^{-i}}(x)} \frac{2^{-\gamma s} 2^{ni}}{\big(1+2^i|x-z|\big)^{n+\gamma}} \mathsf{B}\, dz \Big)\, dx,$$

since for $z \in \mathsf{B}_{2\cdot 2^{-j}}(x) \setminus \mathsf{B}_{2\cdot 2^{-i}}(x)$ both integrands are comparable. Recalling that

$$k_{s,j}^2(x,y) = \Big\langle T_{4\cdot 2^{-j}} \psi_{\widehat{Q}_y}, \phi_{R_x} \Big\rangle,$$

we observe (as in our analysis of Φ_s) that $\mathsf{B} = \mathsf{E}_j(\mathsf{B})$ when regarded as a function of z. This means that $\mathsf{B} = 0$ for any $z \in R_x$. This, together with the fact that $\mathsf{B}_{2\cdot 2^{-j}}(x) \subset 5R_x$, implies

$$\mathcal{B}_1 \le c_n \int_{\mathbb{R}^n} \Big(\int_{5R_x \setminus R_x} \frac{2^{-\gamma s} 2^{ni}}{\big(1+2^i|x-z|\big)^{n+\gamma}} \big| k_{s,j}^2(z,y) - k_{s,j}^2(x,y) \big|\, dz \Big)\, dx.$$

Moreover, arguing as in Lemma 2.5

$$\begin{aligned} \mathcal{B}_1 &\le c_n \int_{\mathbb{R}^n} \Big(\int_{5R_x \setminus R_x} \mathsf{E}_j \Big[\frac{2^{-\gamma s} 2^{ni}}{\big(1+2^i|x-\cdot|\big)^{n+\gamma}} \Big](z) \big| k_{s,j}^2(z,y) - k_{s,j}^2(x,y) \big|\, dz \Big)\, dx \\ &\le c_n \int_{\mathbb{R}^n} \sum_{R \approx R_x} \Big(\int_R \frac{2^{-\gamma s} 2^{ni}}{\big(1+2^i|x-w|\big)^{n+\gamma}}\, dw \Big) \Big(\big| k_{s,j}^2(c_R,y) \big| + \big| k_{s,j}^2(x,y) \big| \Big)\, dx. \end{aligned}$$

Here we write $R \approx R_x$ to denote that R is a dyadic cube in $\mathcal{Q}_j$ contained in $5R_x \setminus R_x$. Then we apply the argument in Lemma 2.6

$$\mathcal{B}_1 \le c_n \int_{\mathbb{R}^n} \sum_{R \approx R_x} \frac{2^{-\gamma s}}{\big(1+2^i \delta_x\big)^{\gamma}} \Big(\big| k_{s,j}^2(c_R,y) \big| + \big| k_{s,j}^2(x,y) \big| \Big)\, dx.$$

Now, it is clear from (2.13) that we have

$$\big| k_{s,j}^2(x,y) \big| \le c_n \frac{2^{-\gamma s} 2^{nj}}{\big(1+2^j|x-y|\big)^{n+\gamma}}.$$

These estimates in conjunction with the argument in Lemma 2.6 give

$$\mathcal{B}_1 \le c_{n,\gamma} 2^{-2\gamma s} |i-j| 2^{-\gamma|i-j|}.$$

To estimate $\mathcal{B}_2$ we use that $\mathsf{B} = 0$ for any $z \in R_x$ and $\mathsf{B}_{2 \cdot 2^{-j}}(x) \subset 5R_x$

$$\mathcal{B}_2 \le \mathrm{c}_n 2^{ni} \int_{\mathbb{R}^n} \Big(\int_{(5R_x \setminus R_x) \cap \mathcal{W}_{s,i}^x} \big|k_{s,j}^2(z,y)\big| + \big|k_{s,j}^2(x,y)\big| \, dz \Big) \, dx.$$

Then we apply our estimate of $|k_{s,j}^2(\,\cdot\,,\cdot\,)|$ given above for (z,y) and (x,y)

$$\begin{aligned} \mathcal{B}_2 & \le \mathrm{c}_n 2^{ni} 2^{-\gamma s} 2^{nj} \int_{\mathbb{R}^n} \Big(\int_{(5R_x \setminus R_x) \cap \mathcal{W}_{s,i}^x} \frac{dz}{\big(1 + 2^j |z-y|\big)^{n+\gamma}} \Big) \, dx \\ & + \mathrm{c}_n 2^{ni} 2^{-\gamma s} 2^{nj} \int_{\mathbb{R}^n} \Big(\int_{(5R_x \setminus R_x) \cap \mathcal{W}_{s,i}^x} \frac{dz}{\big(1 + 2^j |x-y|\big)^{n+\gamma}} \Big) \, dx \\ & \sim \mathrm{c}_n 2^{ni} 2^{-\gamma s} 2^{nj} \int_{\mathbb{R}^n} \Big(\int_{(5R_x \setminus R_x) \cap \mathcal{W}_{s,i}^x} \frac{dz}{\big(1 + 2^j |x-y|\big)^{n+\gamma}} \Big) \, dx. \end{aligned}$$

Last equivalence follows from the presence of $\mathcal{W}_{s,i}^x$. Next, the set

$$(5R_x \setminus R_x) \cap \mathcal{W}_{s,i}^x$$

forces z to be outside R_x but at a distance of x controlled by $4 \cdot 2^{-i} + 2^{-(i+s-1)}$. Thus, the only $x \in \mathbb{R}^n$ for which the inner integral does not vanish are those x for which $\mathrm{dist}(x, \partial R_x) \le 4 \cdot 2^{-i} + 2^{-(i+s-1)}$. Notice that for $|i-j| \le 3$ this suppose no restriction but for $|i-j|$ large does. Given $R \in \mathcal{Q}_j$ we set

$$R_{s,i} = \Big\{ w \in R \, \Big| \, \mathrm{dist}(w, \partial R) \le 4 \cdot 2^{-i} + 2^{-(i+s-1)} \Big\}.$$

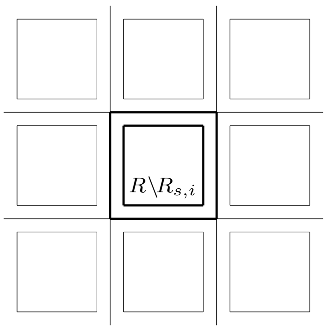


Figure IV
The factor $2^{-|i-j|}$ comes from $|R_{s,i}| \le \mathrm{c}_n 2^{-|i-j|} |R|$

Our considerations allows us to complete our estimate as follows

$$\begin{aligned} \mathcal{B}_2 & \le \mathrm{c}_n 2^{ni} 2^{-\gamma s} 2^{nj} \sum_{R \in \mathcal{Q}_j} \int_{R_{s,i}} \frac{\big|\mathcal{W}_{s,i}^x\big|}{\big(1 + 2^j |x-y|\big)^{n+\gamma}} \, dx \\ & = \mathrm{c}_n 2^{ni} 2^{-\gamma s} 2^{nj} \sum_{R \in \mathcal{Q}_j} \frac{|R_{s,i}|}{|R|} |R| \frac{1}{|R_{s,i}|} \int_{R_{s,i}} \frac{\big|\mathcal{W}_{s,i}^x\big|}{\big(1 + 2^j |x-y|\big)^{n+\gamma}} \, dx \\ & \le \mathrm{c}_n 2^{-(1+\gamma)s} 2^{nj} 2^{-|i-j|} \sum_{R \in \mathcal{Q}_j} |R| \frac{1}{|R_{s,i}|} \int_{R_{s,i}} \frac{1}{\big(1 + 2^j |x-y|\big)^{n+\gamma}} \, dx \end{aligned}$$

$$\sim \quad c_n 2^{-(1+\gamma)s} 2^{nj} 2^{-|i-j|} \int_{\mathbb{R}^n} \frac{dx}{\big(1+2^j|x-y|\big)^{n+\gamma}} \ \sim \ c_n 2^{-(1+\gamma)s} 2^{-|i-j|}.$$

Combining our estimates for $\mathcal{B}_1$ and $\mathcal{B}_2$ we get

$$\mathcal{B} \, \leq \, c_{n,\gamma} 2^{-2\gamma s} |i-j| 2^{-\gamma|i-j|} \tag{2.16}$$

Finally, the sum of (2.14), (2.15) and (2.16) produces

$$\mathcal{S}_{i,j,s}^{2,\beta}(y) \, \leq \, c_{n,\gamma} |i-j| 2^{-\gamma|i-j|}.$$

As we have proved that $\mathcal{S}_{i,j,s}^{2,\alpha}$ satisfies a better estimate, the proof is complete. □

Remark 2.11. The estimate given for $\mathcal{A}_1$ in the proof of Lemma 2.10 above is the only point in the whole argument for our pseudo-localization principle where the Lipschitz smoothness with respect to the x variable is used.

Conclusion. Now we have all the necessary estimates to complete the argument. Namely, a direct application of Lemmas 2.9 and 2.10 in conjunction with Schur lemma give us the following estimate

$$\big\|\Lambda_{s,i}^*\Lambda_{s,j}\big\|_{\mathcal{B}(L_2)} \, \leq \, c_{n,\gamma} \sqrt{2^{-2\gamma s}\big(1+|i-j|\big) 2^{-\gamma|i-j|}} \, \leq \, c_{n,\gamma} 2^{-\gamma s} \alpha_{i-j}^2$$

where $\alpha_k = (1+|k|)^{\frac{1}{4}} 2^{-\gamma|k|/4}$. Therefore, Cotlar lemma provides

$$\|\Psi_s\|_{\mathcal{B}(L_2)} \, = \, \Big\| \sum\nolimits_k (id - \mathsf{E}_k) T \Delta_{k+s} \Big\|_{\mathcal{B}(L_2)} \, \leq \, c_{n,\gamma} 2^{-\gamma s/2} \sum\nolimits_k \alpha_k \, = \, c_{n,\gamma} 2^{-\gamma s/2}.$$

3. Calderón-Zygmund decomposition

We now go back to the noncommutative setting and present a noncommutative form of Calderón-Zygmund decomposition. Let us recall from the Introduction that, for a given semifinite von Neumann algebra $\mathcal{M}$ equipped with a *n.s.f.* trace τ, we shall work on the weak-operator closure $\mathcal{A}$ of the algebra $\mathcal{A}_B$ of essentially bounded functions $f : \mathbb{R}^n \to \mathcal{M}$. Recall also the dyadic filtration $(\mathcal{A}_k)_{k\in\mathbb{Z}}$ in $\mathcal{A}$.

3.1. **Cuculescu revisited.** A difficulty inherent to the noncommutativity is the absence of maximal functions. It is however possible to obtain noncommutative maximal weak and strong inequalities. The strong inequalities follow by recalling that the L_p norm of a maximal function is an $L_p(\ell_\infty)$ norm. As observed by Pisier [45] and further studied by Junge [24], the theory of operator spaces is the right tool to define noncommutative $L_p(\ell_\infty)$ spaces; see [31] for a nice exposition. We shall be interested on weak maximal inequalities, which already appeared in Cuculescu's construction above and are simpler to describe. Indeed, given a sequence $(f_k)_{k\in\mathbb{Z}}$ of positive functions in L_1 and any $\lambda \in \mathbb{R}_+$, we are interested in describing the noncommutative form of the Lebesgue measure of

$$\Big\{ \sup_{k\in\mathbb{Z}} f_k > \lambda \Big\}.$$

If $f_k \in L_1(\mathcal{A})_+$ for $k \in \mathbb{Z}$, this is given by

$$\inf \Big\{ \varphi\big(\mathbf{1}_{\mathcal{A}} - q\big) \mid q \in \mathcal{A}_\pi, \, q f_k q \leq \lambda q \ \text{ for all } \ k \in \mathbb{Z} \Big\}.$$

Given a positive dyadic martingale $f = (f_1, f_2, \ldots)$ in $L_1(\mathcal{A})$ and looking one more time at Cuculescu's construction, it is apparent that the projection $q(\lambda)_k$ represents the following set

$$q(\lambda)_k \sim \Big\{ \sup_{1 \le j \le k} f_j \le \lambda \Big\}.$$

Therefore, we find

$$\mathbf{1}_{\mathcal{A}} - \bigwedge_{k \ge 1} q(\lambda)_k \sim \Big\{ \sup_{k \ge 1} f_k > \lambda \Big\}.$$

However, in this paper we shall be interested in the projection representing the set where $\sup_{k \in \mathbb{Z}} f_k > \lambda$ since we will work with the full dyadic filtration $(\mathcal{A}_k)_{k \in \mathbb{Z}}$, where $\mathcal{A}_k$ stands for $\mathsf{E}_k(\mathcal{A})$. We shall clarify below why it is not enough to work with the truncated filtration $(\mathcal{A}_k)_{k \ge 1}$. The construction of the right projection for $\sup_{k \in \mathbb{Z}} f_k > \lambda$ does not follow automatically from Cuculescu's construction, see Proposition 3.2 below. Moreover, given a general function $f \in L_1(\mathcal{A})_+$, we are not able at the time of this writing to construct the right projections $q_\lambda(f,k)$ which represent the sets

$$q_\lambda(f,k) \sim \Big\{ \sup_{j \in \mathbb{Z},\, j \le k} f_j \le \lambda \Big\}.$$

Indeed, the weak* limit procedure used in the proof of Proposition 3.2 below does not preserve the commutation relation i) of Cuculescu's construction and we are forced to work in the following dense class of $L_1(\mathcal{A})$

$$\mathcal{A}_{c,+} = \Big\{ f : \mathbb{R}^n \to \mathcal{M} \, \Big| \, f \in \mathcal{A}_+, \ \overrightarrow{\mathrm{supp}} f \ \text{ is compact} \Big\} \subset L_1(\mathcal{A}). \tag{3.1}$$

Here $\overrightarrow{\mathrm{supp}}$ means the support of f as a vector-valued function in $\mathbb{R}^n$. In other words, we have $\overrightarrow{\mathrm{supp}} f = \mathrm{supp} \|f\|_{\mathcal{M}}$. We employ this terminology to distinguish from $\mathrm{supp}\, f$ (the support of f as an operator in $\mathcal{A}$) defined in Section 1. Note that $\overrightarrow{\mathrm{supp}} f$ is a measurable subset of $\mathbb{R}^n$, while $\mathrm{supp}\, f$ is a projection in $\mathcal{A}$. In the rest of the paper we shall work with functions f in $\mathcal{A}_{c,+}$. This impose no restriction due to the density of $\mathrm{span}\, \mathcal{A}_{c,+}$ in $L_1(\mathcal{A})$. The following result is an adaptation of Cuculescu's construction which will be the one to be used in the sequel.

Lemma 3.1. *Let $f \in \mathcal{A}_{c,+}$ and $f_k = \mathsf{E}_k(f)$ for $k \in \mathbb{Z}$. The sequence $(f_k)_{k \in \mathbb{Z}}$ is a (positive) dyadic martingale in $L_1(\mathcal{A})$. Given any positive number λ, there exists a decreasing sequence $(q_\lambda(f,k))_{k \in \mathbb{Z}}$ of projections in $\mathcal{A}$ satisfying*

- i) *$q_\lambda(f,k)$ commutes with $q_\lambda(f,k-1)\, f_k\, q_\lambda(f,k-1)$ for each $k \in \mathbb{Z}$.*
- ii) *$q_\lambda(f,k)$ belongs to $\mathcal{A}_k$ for each $k \in \mathbb{Z}$ and $q_\lambda(f,k)\, f_k\, q_\lambda(f,k) \le \lambda q_\lambda(f,k)$.*
- iii) *The following estimate holds*

$$\varphi\Big(\mathbf{1}_{\mathcal{A}} - \bigwedge_{k \in \mathbb{Z}} q_\lambda(f,k) \Big) \le \frac{1}{\lambda} \|f\|_1.$$

Proof. Since $f \in \mathcal{A}_{c,+}$ we have for all $Q \in \mathcal{Q}_j$

$$f_Q = \frac{1}{|Q|} \int_Q f(x)\, dx \le 2^j \, \|f\|_{\mathcal{A}} \, \big|\overrightarrow{\mathrm{supp}} f \big| \, \mathbf{1}_{\mathcal{A}} \longrightarrow 0 \quad \text{as} \quad j \to -\infty.$$

In particular, given any $\lambda \in \mathbb{R}_+$, we have $f_j \le \lambda \mathbf{1}_{\mathcal{A}}$ for all $j < m_\lambda < 0$ and certain $m_\lambda \in \mathbb{Z} \setminus \mathbb{N}$ with $|m_\lambda|$ large enough. Then we define the desired projections by the following relations

$$q_\lambda(f,k) = \begin{cases} \mathbf{1}_{\mathcal{A}} & \text{if } k < m_\lambda, \\ \chi_{(0,\lambda]}(f_k) & \text{if } k = m_\lambda, \\ \chi_{(0,\lambda]}\big(q_\lambda(f,k-1)\, f_k\, q_\lambda(f,k-1)\big) & \text{if } k > m_\lambda. \end{cases}$$

To prove iii) we observe that our projections are exactly the ones obtained when applying Cuculescu's construction over the truncated filtration $(\mathcal{A}_k)_{k \ge m_\lambda}$. Thus we get

$$\varphi\Big(\mathbf{1}_{\mathcal{A}} - \bigwedge_{k\in\mathbb{Z}} q_\lambda(f,k)\Big) = \varphi\Big(\mathbf{1}_{\mathcal{A}} - \bigwedge_{k\ge m_\lambda} q(\lambda)_k\Big) \le \frac{1}{\lambda} \sup_{k\ge m_\lambda} \|f_k\|_1 = \frac{1}{\lambda}\|f\|_1.$$

The rest of the properties of the sequence $(q_\lambda(f,k))_{k\in\mathbb{Z}}$ are easily verifiable. □

3.2. **The maximal function.** We now recall the Hardy-Littlewood weak maximal inequality. In what follows it will be quite useful to have another expression for the $q_\lambda(f,k)$'s constructed in Lemma 3.1. It is not difficult to check that

$$q_\lambda(f,k) = \sum_{Q\in\mathcal{Q}_k} \xi_\lambda(f,Q) 1_Q$$

for $k \in \mathbb{Z}$, with $\xi_\lambda(f,Q)$ projections in $\mathcal{M}$ defined by

$$\xi_\lambda(f,Q) = \begin{cases} \mathbf{1}_{\mathcal{M}} & \text{if } k < m_\lambda, \\ \chi_{(0,\lambda]}(f_Q) & \text{if } k = m_\lambda, \\ \chi_{(0,\lambda]}\big(\xi_\lambda(f,\widehat{Q}) f_Q \xi_\lambda(f,\widehat{Q})\big) & \text{if } k > m_\lambda. \end{cases}$$

As for Cuculescu's construction, we have

- $\xi_\lambda(f,Q) \in \mathcal{M}_\pi$.
- $\xi_\lambda(f,Q) \le \xi_\lambda(f,\widehat{Q})$.
- $\xi_\lambda(f,Q)$ commutes with $\xi_\lambda(f,\widehat{Q}) f_Q\, \xi_\lambda(f,\widehat{Q})$.
- $\xi_\lambda(f,Q) f_Q\, \xi_\lambda(f,Q) \le \lambda \xi_\lambda(f,Q)$.

The noncommutative weak type $(1,1)$ inequality for the Hardy-Littlewood dyadic maximal function [36] follows as a consequence of this. We give a proof including some details (reported by Quanhua Xu to the author) not appearing in [36].

Proposition 3.2. *If $(f,\lambda) \in L_1(\mathcal{A}) \times \mathbb{R}_+$, there exists $q_\lambda(f) \in \mathcal{A}_\pi$ with*

$$\sup_{k\in\mathbb{Z}} \big\| q_\lambda(f) f_k q_\lambda(f) \big\|_{\mathcal{A}} \le 16\,\lambda \qquad \text{and} \qquad \varphi\big(\mathbf{1}_{\mathcal{A}} - q_\lambda(f)\big) \le \frac{8}{\lambda}\,\|f\|_1.$$

Proof. Let us fix an integer $m \in \mathbb{Z} \setminus \mathbb{N}$. Assume $f \in L_1(\mathcal{A})_+$ and consider the sequence $(q(\lambda)_{m,k})_{k\ge m}$ provided by Cuculescu's construction applied over the filtration $(\mathcal{A}_k)_{k\ge m}$. Define

$$q_m(\lambda) = \bigwedge_{k\ge m} q(\lambda)_{m,k} \quad \text{for each} \quad m \in \mathbb{Z}\setminus\mathbb{N}.$$

Let us look at the family $(q_m(\lambda))_{m\in\mathbb{Z}\setminus\mathbb{N}}$. By the weak* compactness of the unit ball $\mathsf{B}_{\mathcal{A}}$ and the positivity of our family, there must exists a cluster point $a \in \mathsf{B}_{\mathcal{A}_+}$. In particular, we may find a subsequence with $q_{m_j}(\lambda) \to a$ as $j \to \infty$ (note that

$m_j \to -\infty$ as $j \to \infty$) in the weak* topology. Then we set $q_\lambda(f) = \chi_{[1/2,1]}(a)$ and define positive operators $\delta(a)$ and $\beta(a)$ bounded by $2\mathbf{1}_\mathcal{A}$ and determined by

$$q_\lambda(f) = a\delta(a) = \delta(a)a,$$

$$\mathbf{1}_\mathcal{A} - q_\lambda(f) = \chi_{(1/2,1]}(\mathbf{1}_\mathcal{A} - a) = (\mathbf{1}_\mathcal{A} - a)\beta(a) = \beta(a)(\mathbf{1}_\mathcal{A} - a).$$

In order to prove the first inequality stated above, we note that

$$\big\|q_\lambda(f) f_k \, q_\lambda(f)\big\|_\mathcal{A} = \sup_{\|b\|_{L_1(\mathcal{A})} \le 1} \varphi\big(q_\lambda(f) f_k \, q_\lambda(f) b\big).$$

However, we have

$$\begin{aligned} \varphi\big(q_\lambda(f) f_k \, q_\lambda(f) b\big) & = & \varphi\big(a f_k a \, \delta(a) b \delta(a)\big) \\ & = & \lim_{j\to\infty} \varphi\big(q_{m_j}(\lambda) f_k \, q_{m_j}(\lambda) \delta(a) b \delta(a)\big) \\ & \le & \lim_{j\to\infty} \big\|q_{m_j}(\lambda) f_k \, q_{m_j}(\lambda)\big\|_\infty \|\delta(a) b \delta(a)\|_1. \end{aligned}$$

Therefore we conclude

$$\varphi\big(q_\lambda(f) f_k \, q_\lambda(f) b\big) \le \|b\|_1 \|\delta(a)\|_\infty^2 \lim_{j\to\infty} \big\|q(\lambda)_{m_j,k} f_k \, q(\lambda)_{m_j,k}\big\|_\infty \le 4\lambda.$$

This proves the first inequality, as for the second

$$\begin{aligned} \varphi\big(\mathbf{1}_\mathcal{A} - q_\lambda(f)\big) & = & \varphi\big((\mathbf{1}_\mathcal{A} - a)\beta(a)\big) \ \le \ 2\varphi(\mathbf{1}_\mathcal{A} - a) \\ & = & 2 \lim_{j\to\infty} \varphi\big(\mathbf{1}_\mathcal{A} - q_{m_j}(\lambda)\big) \ \le \ \frac{2}{\lambda} \, \|f\|_1. \end{aligned}$$

Finally, for a general $f \in L_1(\mathcal{A})$ we decompose

$$f = (f_1 - f_2) + i(f_3 - f_4)$$

with $f_j \in L_1(\mathcal{A})_+$ and define

$$q_\lambda(f) = \bigwedge_{1 \le j \le 4} q_\lambda(f_j).$$

Then, the estimate follows easily with constants $16\,\lambda$ and $8/\lambda$ respectively. □

3.3. **The good and bad parts.** If $f \in L_1$ positive and $\lambda \in \mathbb{R}_+$, define

$$M_d f(x) = \sup_{x \in Q \in \mathcal{Q}} \frac{1}{|Q|} \int_Q f(y) \, dy \quad \text{and} \quad \mathrm{E}_\lambda = \Big\{ x \in \mathbb{R}^n \, \big| \, M_d f(x) > \lambda \Big\}.$$

Writing $\mathrm{E}_\lambda = \bigcup_j Q_j$ as a disjoint union of maximal dyadic cubes with $f_Q \le \lambda < f_{Q_j}$ for all dyadic $Q \supset Q_j$, we may decompose $f = g + b$ where the good and bad parts are given by

$$g = f 1_{\mathrm{E}_\lambda^c} + \sum\nolimits_j f_{Q_j} 1_{Q_j} \quad \text{and} \quad b = \sum\nolimits_j (f - f_{Q_j}) 1_{Q_j}.$$

If $b_j = (f - f_{Q_j}) 1_{Q_j}$, we have

i) $\|g\|_1 \le \|f\|_1$ and $\|g\|_\infty \le 2^n \lambda$.
ii) $\mathrm{supp}\, b_j \subset Q_j$, $\int_{Q_j} b_j = 0$ and $\sum_j \|b_j\|_1 \le 2\|f\|_1$.

Before proceeding with the noncommutative Calderón-Zygmund decomposition, we simplify our notation for the projections $\xi_\lambda(f,Q)$ and $q_\lambda(f,k)$. Namely, (f,λ) will remain fixed in $\mathcal{A}_{c,+} \times \mathbb{R}_+$, see (3.1). These choices lead us to set

$$\big(\xi_Q, q_k, q\big) = \Big(\xi_\lambda(f,Q), q_\lambda(f,k), \bigwedge_{k\in\mathbb{Z}} q_\lambda(f,k)\Big).$$

Moreover, we shall write $(p_k)_{k\in\mathbb{Z}}$ for the projections

$$p_k = q_{k-1} - q_k = \sum_{Q\in\mathcal{Q}_k} (\xi_{\widehat{Q}} - \xi_Q) 1_Q = \sum_{Q\in\mathcal{Q}_k} \pi_Q 1_Q. \tag{3.2}$$

The terminology $\pi_Q = \xi_{\widehat{Q}} - \xi_Q$ will be frequently used below. Recall that the p_k's are pairwise disjoint and (according to our new terminology) we have $q_j = \mathbf{1}_\mathcal{A}$ for all $j < m_\lambda$. In particular, we find

$$\sum\nolimits_k p_k = \mathbf{1}_\mathcal{A} - q.$$

If we write p_∞ for q and $\widehat{\mathbb{Z}}$ stands for $\mathbb{Z}\cup\{\infty\}$, our noncommutative analogue for the Calderón-Zygmund decomposition can be stated as follows. If $f \in \mathcal{A}_{c,+}$ and $\lambda \in \mathbb{R}_+$, we consider the decomposition $f = g + b$ with

$$g = \sum_{i,j\in\widehat{\mathbb{Z}}} p_i f_{i\vee j} p_j \quad \text{and} \quad b = \sum_{i,j\in\widehat{\mathbb{Z}}} p_i (f - f_{i\vee j}) p_j, \tag{3.3}$$

where $i\vee j = \max(i,j)$. Note that $i \vee j = \infty$ whenever i or j is infinite. In particular, since $f = f_\infty$ by definition, the extended sum defining b is just an ordinary sum over $\mathbb{Z}\times\mathbb{Z}$. Note also that our expressions are natural generalizations of the classical good and bad parts stated in the classical decomposition. Indeed, recalling the orthogonality of the p_k's, all the off-diagonal terms vanish in the commutative setting and we find something like

$$g_d = qfq + \sum\nolimits_k p_k f_k p_k \quad \text{and} \quad b_d = \sum\nolimits_k p_k (f - f_k) p_k. \tag{3.4}$$

In this form, and recalling that for $\mathcal{M} = \mathbb{C}$ we have

$$q \sim \mathbb{R}^n \setminus \mathrm{E}_\lambda \quad \text{and} \quad p_k \sim \Big\{ Q_j \subset \mathrm{E}_\lambda \,\Big|\, Q_j \in \mathcal{Q}_k \Big\},$$

it is not difficult to see that we recover the classical decomposition.

Remark 3.3. In the following we shall use the square-diagram in Figure V below to think of our decomposition. Namely, we first observe that for any $f \in \mathcal{A}_{c,+}$ and for any $\lambda \in \mathbb{R}_+$ there will be an $m_\lambda \in \mathbb{Z}$ such that $f_j \le \lambda \mathbf{1}_\mathcal{A}$ for all $j < m_\lambda$, see the proof of Lemma 3.1 above. In particular, since f and λ will remain fixed, by a simple relabelling we may assume with no loss of generality that $m_\lambda = 1$. This will simplify very much the notation, since now we have $p_k = 0$ for all non-positive k. Therefore, the terms $p_i f_{i\vee j} p_j$ and $p_i(f - f_{i\vee j})p_j$ in our decomposition may be located in the (i,j)-th position of an $\infty\times\infty$ matrix where the 'last' row and column are devoted to the projection $q = p_\infty$.

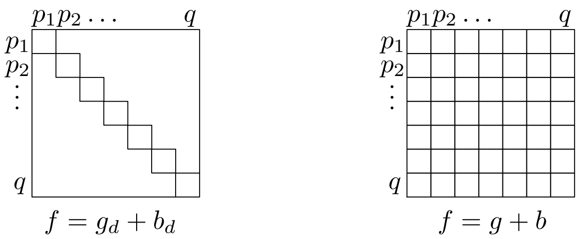


Figure V
Commutative and noncommutative decompositions

4. Weak type estimates for diagonal terms

In this section we start with the proof of Theorem A. Before that, a couple of remarks are in order. First, according to the classical theory it is clearly no restriction to assume that $q = 2$. In particular, since $L_2(\mathcal{A})$ is a Hilbert space valued L_2 space, boundedness in $L_2(\mathcal{A})$ will hold. Second, we may assume the function $f \in L_1(\mathcal{A})$ belongs to $\mathcal{A}_{c,+}$. Indeed, this follows by decomposing f as a linear combination $(f_1 - f_2) + i(f_3 - f_4)$ of positive functions $f_j \in L_1(\mathcal{A})_+$ and using the quasi-triangle inequality on $L_{1,\infty}(\mathcal{A})$ stated in Section 1. Then we approximate each $f_j \in L_1(\mathcal{A})_+$ by functions in $\mathcal{A}_{c,+}$. Third, since $f \geq 0$ by assumption, we may break it for any fixed $\lambda \in \mathbb{R}_+$ following our Calderón-Zygmund decomposition. In this section we prove our main result for the *diagonal* terms in (3.4). According to the quasi-triangle inequality for $L_{1,\infty}(\mathcal{A})$, this will reduce the problem to estimate the off-diagonal terms.

4.1. **Classical estimates.** The standard estimates i) and ii) satisfied by the good and bad parts of Calderón-Zygmund decomposition are satisfied by the diagonal terms (3.4). Indeed, since f is positive so is g_d and

$$\|g_d\|_1 = \varphi(qfq) + \sum_{k \geq 1} \varphi\big(p_k f_k p_k\big) = \varphi\big(fq + f(\mathbf{1}_\mathcal{A} - q)\big) \ = \ \|f\|_1.$$

On the other hand, by orthogonality we have

$$\|g_d\|_\infty \ = \ \max \Big\{ \|qfq\|_\infty, \sup_{k \geq 1} \big\|p_k f_k p_k\big\|_\infty \Big\}.$$

To estimate the first term, take $a \in L_1(\mathcal{A})$ of norm 1 with

$$\|qfq\|_\infty \leq \varphi(qfqa) + \delta.$$

Since $f_k \to f$ as $k \to \infty$ in the weak* topology, we deduce that

$$\begin{aligned} \|qfq\|_\infty & \leq \varphi(qfqa) + \delta \\ & = \lim_{k \to \infty} \varphi(qf_kqa) + \delta \\ & \leq \lim_{k \to \infty} \|qf_kq\|_\infty \|a\|_1 + \delta \ \leq \ \lambda + \delta, \end{aligned}$$

where the last inequality follows from $qf_kq = qq_kf_kq_kq \leq \lambda q$. Therefore, taking $\delta \to 0^+$ we deduce that $\|qfq\|_\infty \leq \lambda$. Let us now estimate the second terms. To that aim, we observe that

$$f_k = \sum_{Q \in \mathcal{Q}_k} \frac{1}{|Q|} \int_Q f(y)\, dy \, 1_Q = 2^n \sum_{Q \in \mathcal{Q}_k} \frac{1}{|\widehat{Q}|} \int_Q f(y)\, dy \, 1_Q \le 2^n f_{k-1}.$$

Therefore, we obtain

$$\big\| p_k f_k p_k \big\|_\infty \le 2^n \big\| q_{k-1} f_{k-1} q_{k-1} \big\|_\infty \le 2^n \lambda. \tag{4.1}$$

This completes the proof of our assertions for g_d. Let us now prove the assertions for b_d. If we take $b_{d,k}$ to be $p_k(f - f_k)p_k$, it is clear that $b_{d,k} = b_{d,k}^*$ and also that $\mathrm{supp}\, b_{d,k} \le p_k$. Moreover, recalling that

$$b_{d,k} = \sum_{Q \in \mathcal{Q}_k} (\xi_{\widehat{Q}} - \xi_Q)(f - f_Q)(\xi_{\widehat{Q}} - \xi_Q)\, 1_Q,$$

the following identity holds for any $Q_0 \in \mathcal{Q}_k$

$$\int_{Q_0} b_{d,k}(y)\, dy = (\xi_{\widehat{Q}_0} - \xi_{Q_0}) \Big(\int_{Q_0} f(y) - f_{Q_0}(y)\, dy \Big) (\xi_{\widehat{Q}_0} - \xi_{Q_0}) = 0.$$

Finally, we observe that

$$\sum_{k \ge 1} \|b_{d,k}\|_1 \le \sum_{k \ge 1} \varphi\big(p_k(f + f_k)p_k\big) = 2\varphi\big(f(\mathbf{1}_{\mathcal{A}} - q)\big) \le 2\|f\|_1.$$

This completes the proof of our assertions for the function b_d. As we shall see in the following section, the estimates for the off-diagonal terms require more involved arguments which do not appear in the classical (scalar-valued) theory.

Remark 4.1. It is important to note that the *doubling estimate* (4.1) is crucial for our further analysis and also that such inequality is the one which imposes to work with the full filtration $(\mathcal{A}_k)_{k \in \mathbb{Z}}$ instead with the truncated one $(\mathcal{A}_k)_{k \ge 1}$. Indeed, if we truncate at $k \ge 1$ (not at $k \ge m_\lambda$ as we have done), then condition (4.1) fails in general for $k = 1$. This is another difference with the approach in [44], where the doubling condition above was not needed.

4.2. **An $\mathbb{R}^n$-dilated projection.** As above, given a positive function $f \in L_1$, let E_λ be the set in $\mathbb{R}^n$ where the dyadic Hardy-Littlewood maximal function $M_d f$ is greater than λ. If we decompose $\mathrm{E}_\lambda = \bigcup_j Q_j$ as a disjoint union of maximal dyadic cubes, let us write $9\,\mathrm{E}_\lambda$ for the dilation

$$9\,\mathrm{E}_\lambda = \bigcup_j 9 Q_j.$$

As we pointed out in the Introduction, this is a key set to give a weak type estimate for the bad part in Calderón-Zygmund decomposition. On the other hand, we know from Cuculescu's construction that $\mathbf{1}_{\mathcal{A}} - q$ represents the noncommutative analog of E_λ, so that the noncommutative analog of $9\,\mathrm{E}_\lambda$ should look like '$9(\mathbf{1}_{\mathcal{A}} - q)$' in the sense that we dilate on $\mathbb{R}^n$ but not on $\mathcal{M}$. In the following result we construct the right noncommutative analog of $9\,\mathrm{E}_\lambda$.

Lemma 4.2. *There exists $\zeta \in \mathcal{A}_\pi$ such that*

i) $\lambda \varphi\big(\mathbf{1}_{\mathcal{A}} - \zeta\big) \le 9^n \|f\|_1$.

ii) *If* $Q_0 \in \mathcal{Q}$ *and* $x \in 9Q_0$, *then*

$$\zeta(x) \le \mathbf{1}_{\mathcal{M}} - \xi_{\widehat{Q}_0} + \xi_{Q_0}.$$

In particular, in this case we immediately find $\zeta(x) \le \xi_{Q_0}$.

Proof. Given $k \in \mathbb{Z}_+$, we define

$$\psi_k = \sum_{s=1}^{k} \sum_{Q \in \mathcal{Q}_s} (\xi_{\widehat{Q}} - \xi_Q) 1_{9Q} \quad \text{and} \quad \zeta_k = \mathbf{1}_{\mathcal{A}} - \operatorname{supp} \psi_k.$$

Since we have $\xi_Q \le \xi_{\widehat{Q}}$ for all dyadic cube Q, it turns out that $(\psi_k)_{k \ge 1}$ is an increasing sequence of positive operators. However, enlarging Q by its concentric father $9Q$ generates overlapping and the ψ_k's are not projections. This forces us to consider the associated support projections and define $\zeta_1, \zeta_2, \ldots$ as above. The sequence of projections $(\zeta_k)_{k \ge 1}$ is clearly decreasing and we may define

$$\zeta = \bigwedge_{k \ge 1} \zeta_k.$$

Now we are ready to prove the first estimate

$$\begin{aligned} \lambda \varphi\big(\mathbf{1}_{\mathcal{A}} - \zeta\big) &= \lambda \lim_{k \to \infty} \varphi\big(\mathbf{1}_{\mathcal{A}} - \zeta_k\big) \ \le \ \lambda \sum_{s=1}^{\infty} \sum_{Q \in \mathcal{Q}_s} \varphi\big((\xi_{\widehat{Q}} - \xi_Q) 1_{9Q}\big) \\ &= 9^n \lambda \sum_{s=1}^{\infty} \sum_{Q \in \mathcal{Q}_s} \varphi\big((\xi_{\widehat{Q}} - \xi_Q) 1_Q\big) \ = \ 9^n \lambda\, \varphi(\mathbf{1}_{\mathcal{A}} - q) \ \le \ 9^n \, \|f\|_1. \end{aligned}$$

Now fix $Q_0 \in \mathcal{Q}$, say $Q_0 \in \mathcal{Q}_{k_0}$ for some $k_0 \in \mathbb{Z}$. If $k_0 \le 0$, the assertion is trivial since we know from Remark 3.3 that $\xi_{Q_0} = \xi_{\widehat{Q}_0} = \mathbf{1}_{\mathcal{M}}$. Thus, we assume that $k_0 \ge 1$. Then we have

$$\begin{aligned} (\xi_{\widehat{Q}_0} - \xi_{Q_0}) 1_{9Q_0} \le \psi_{k_0} \quad &\Rightarrow \quad \zeta_{k_0} \le \mathbf{1}_{\mathcal{A}} - (\xi_{\widehat{Q}_0} - \xi_{Q_0}) 1_{9Q_0} \\ &\Rightarrow \quad \zeta(x) \le \zeta_{k_0}(x) \le \mathbf{1}_{\mathcal{M}} - \xi_{\widehat{Q}_0} + \xi_{Q_0} \end{aligned}$$

for any $x \in 9Q_0$. It remains to prove that in fact $\zeta(x) \le \xi_{Q_0}$. Let us write Q_j for the j-th dyadic antecessor of Q_0. In other words, Q_1 is the dyadic father of Q_0, Q_2 is the dyadic father of Q_1 and so on until $Q_{k_0-1} \in \mathcal{Q}_1$. Since the family $Q_0, Q_1, \ldots$ is increasing, the same happens for their concentric fathers and we find

$$x \in \bigcap_{j=0}^{k_0-1} 9Q_j.$$

In particular, applying the estimate proved so far

$$\zeta(x) \le \bigwedge_{j=0}^{k_0-1} \big(\mathbf{1}_{\mathcal{M}} - \xi_{\widehat{Q}_j} + \xi_{Q_j}\big) = \xi_{Q_0}.$$

The last identity easily follows from

$$\xi_{\widehat{Q}_{k_0-1}} = \mathbf{1}_{\mathcal{M}}.$$

Indeed, we have agreed in Remark 3.3 to assume $q_k = \mathbf{1}_{\mathcal{A}}$ for all $k \le 0$. □

4.3. **Chebychev's inequalities.** By Paragraph 4.1, we have

$$\|g_d\|_2^2 = \varphi\big(g_d^{\frac{1}{2}} g_d g_d^{\frac{1}{2}}\big) \le \|g_d\|_1 \|g_d\|_\infty \le 2^n \lambda \|f\|_1.$$

In particular, the estimate below follows from Chebychev's inequality

$$\lambda \varphi\Big\{|Tg_d| > \lambda\Big\} \le \frac{1}{\lambda} \|Tg_d\|_2^2 \lesssim \frac{1}{\lambda} \|g_d\|_2^2 \le 2^n \|f\|_1.$$

As it is to be expected, here we have used our assumption on the L_2-boundedness of T. Now we are interested on a similar estimate with b_d in place of g_d. Using the projection ζ introduced in Lemma 4.2, we may consider the following decomposition

$$Tb_d = (\mathbf{1}_\mathcal{A} - \zeta) T(b_d)(\mathbf{1}_\mathcal{A} - \zeta) + \zeta T(b_d)(\mathbf{1}_\mathcal{A} - \zeta) + (\mathbf{1}_\mathcal{A} - \zeta) T(b_d)\zeta + \zeta T(b_d)\zeta.$$

In particular, we find

$$\lambda \varphi\Big\{\big|Tb_d\big| > \lambda\Big\} \lesssim \lambda \varphi(\mathbf{1}_\mathcal{A} - \zeta) + \lambda \varphi\Big\{\big|\zeta T(b_d)\zeta\big| > \lambda\Big\}.$$

Indeed, according to our decomposition of Tb_d and the quasi-triangle inequality on $L_{1,\infty}(\mathcal{A})$, the estimate above reduces to observe that the first three terms in such decomposition are left or right supported by $\mathbf{1}_\mathcal{A} - \zeta$. Hence, since the quasi-norm in $L_{1,\infty}(\mathcal{A})$ is adjoint-invariant [14], we easily deduce it. On the other hand, according to the first estimate in Lemma 4.2, it suffices to study the last term above. Let us analyze the operator $\zeta T(b_d)\zeta$. In what follows we shall freely manipulate infinite sums with no worries of convergence. This is admissible because we may assume from the beginning (by a simple approximation argument) that $f \in \mathcal{A}_n$ for some finite $n \ge 1$. In particular, we could even think that all our sums are in fact finite sums. We may write

$$\zeta T(b_d)\zeta = \sum_{k \ge 1} \zeta T(b_{k,d})\zeta$$

with $b_{k,d} = p_k(f - f_k)p_k$ for all $k \ge 1$. Then, Chebychev's inequality gives

$$\lambda \varphi\Big\{\big|\zeta T(b_d)\zeta\big| > \lambda\Big\} \le \sum_{k=1}^\infty \big\|\zeta T(b_{d,k})\zeta\big\|_1.$$

According to Lemma 4.2 and using $\xi_Q \pi_Q = \pi_Q \xi_Q = 0$ (recall the definition of π_Q from (3.2) above), we have $\zeta(x) b_{d,k}(y) \zeta(x) = 0$ whenever x lies in the concentric father $9Q$ of the cube $Q \in \mathcal{Q}_k$ for which $y \in Q$. In other words, we know that x lives far away from the singularity of the kernel k and

$$\begin{aligned}
\Big[\zeta T(b_{d,k})\zeta\Big](x) &= \int_{\mathbb{R}^n} k(x,y)\big(\zeta(x) b_{d,k}(y) \zeta(x)\big)\, dy \\
&= \sum_{Q \in \mathcal{Q}_k} \Big(\int_Q k(x,y)\big(\zeta(x) b_{d,k}(y)\zeta(x)\big)\, dy\Big)\, 1_{(9Q)^c}(x) \\
&= \zeta(x)\Big(\sum_{Q \in \mathcal{Q}_k} \Big(\int_Q k(x,y) b_{d,k}(y)\, dy\Big)\, 1_{(9Q)^c}(x)\Big)\zeta(x).
\end{aligned}$$

Now we use the mean-zero condition of $b_{d,k}$ from Paragraph 4.1

$$\Big[\zeta T(b_{d,k})\zeta\Big](x) = \zeta(x)\Big(\sum_{Q \in \mathcal{Q}_k} \Big(\int_Q \big(k(x,y) - k(x,c_Q)\big) b_{d,k}(y)\, dy\Big)\, 1_{(9Q)^c}(x)\Big)\zeta(x),$$

where c_Q is the center of Q. Then we use the Lipschitz γ-smoothness to obtain

$$\begin{aligned}
\sum_{k=1}^\infty & \big\| \zeta T(b_{d,k}) \zeta \big\|_1 \\
& \le \sum_{k=1}^\infty \sum_{Q \in \mathcal{Q}_k} \int_Q \Big\| \big(k(\,\cdot\,,y) - k(\,\cdot\,,c_Q) \big) b_{d,k}(y) 1_{(9Q)^c}(\,\cdot\,) \Big\|_1 \, dy \\
& = \sum_{k=1}^\infty \sum_{Q \in \mathcal{Q}_k} \int_Q \tau \Big[\Big(\int_{(9Q)^c} \big| k(x,y) - k(x,c_Q) \big| \, dx \Big) |b_{d,k}(y)| \Big] \, dy \\
& \lesssim \sum_{k=1}^\infty \sum_{Q \in \mathcal{Q}_k} \int_Q \Big(\int_{(9Q)^c} \frac{|y - c_Q|^\gamma}{|x - c_Q|^{n+\gamma}} \, dx \Big) \tau |b_{d,k}(y)| \, dy \\
& \lesssim \sum_{k=1}^\infty \sum_{Q \in \mathcal{Q}_k} \int_Q \tau |b_{d,k}(y)| \, dy \ = \ \sum_{k=1}^\infty \int_{\mathbb{R}^n} \tau |b_{d,k}(y)| \, dy \ = \ \sum_{k=1}^\infty \|b_{d,k}\|_1 \ \le \ 2 \|f\|_1,
\end{aligned}$$

where the last inequality follows once more from Paragraph 4.1. This completes the argument for the diagonal part. Indeed, for any fixed $\lambda \in \mathbb{R}_+$ we have seen that the diagonal parts of g and b (which depend on the chosen λ) satisfy

$$\lambda \varphi \Big\{ |Tg_d| > \lambda \Big\} + \lambda \varphi \Big\{ |Tb_d| > \lambda \Big\} \le \mathrm{c}_n \, \|f\|_1. \tag{4.2}$$

5. Weak type estimates for off-diagonal terms

Given $\lambda \in \mathbb{R}_+$, we have broken f with our Calderón-Zygmund decomposition for such λ. In the last section, we have estimated the diagonal terms g_d and b_d. Let us now consider the off-diagonal terms g_{off} and b_{off} determined by $g = g_d + g_{\mathit{off}}$ and $b = b_d + b_{\mathit{off}}$. As we pointed out in the Introduction, it is paradoxical that the bad part behaves (when dealing with off-diagonal terms) better than the good one!

5.1. **An expression for g_{off}.** We have

$$\begin{aligned}
g_{\mathit{off}} & = \sum_{\substack{i \neq j \\ i,j \in \widehat{\mathbb{Z}}}} p_i f_{i \vee j} p_j \\
& = q f (\mathbf{1}_\mathcal{A} - q) + (\mathbf{1}_\mathcal{A} - q) f q + \sum_{s=1}^\infty \sum_{k=1}^\infty p_k f_{k+s} p_{k+s} + p_{k+s} f_{k+s} p_k.
\end{aligned}$$

Here we have restricted the sum $\sum_{k \in \mathbb{Z}}$ to $\sum_{k \ge 1}$ according to Remark 3.3. Applying property i) of Cuculescu's construction, we know that the projection q_j commutes with $q_{j-1} f_j q_{j-1}$ for all $j \ge 1$. Taking $i \wedge j = \min(i,j)$, this immediately gives that $p_i f_{i \wedge j} p_j = 0$ for $i \neq j$. Indeed, we have

$$\begin{array}{lclcll}
p_i f_{i \wedge j} p_j & = & p_i q_{i-1} f_i q_{i-1} p_j & = & 0 & \text{if} \quad i < j, \\
p_i f_{i \wedge j} p_j & = & p_i q_{j-1} f_j q_{j-1} p_j & = & 0 & \text{if} \quad i > j.
\end{array}$$

Using this property and inverting the order of summation, we deduce

$$\sum_{s=1}^\infty \sum_{k=1}^\infty p_k f_{k+s} p_{k+s} + p_{k+s} f_{k+s} p_k$$

$$= \sum_{s,k=1}^{\infty} p_k(f_{k+s} - f_k)p_{k+s} + p_{k+s}(f_{k+s} - f_k)p_k$$

$$= \sum_{s,k=1}^{\infty} \sum_{j=1}^{s} p_k df_{k+j} p_{k+s} + p_{k+s} df_{k+j} p_k = \sum_{j,k=1}^{\infty} \sum_{s=j}^{\infty} p_k df_{k+j} p_{k+s} + p_{k+s} df_{k+j} p_k.$$

Recall that we may use Fubini theorem since, as we observed in Paragraph 4.3, we may even assume that all our sums are finite sums. Now we can sum in s and apply the commutation property above to obtain

$$\begin{aligned}
&\sum_{s=1}^{\infty} \sum_{k=1}^{\infty} p_k f_{k+s} p_{k+s} + p_{k+s} f_{k+s} p_k \\
&\quad = \sum_{j,k=1}^{\infty} p_k df_{k+j}(q_{k+j-1} - q) + (q_{k+j-1} - q) df_{k+j} p_k \\
&\quad = \sum_{j,k=1}^{\infty} p_k df_{k+j} q_{k+j-1} + q_{k+j-1} df_{k+j} p_k - \sum_{k=1}^{\infty} p_k(f - f_k)q + q(f - f_k)p_k \\
&\quad = \sum_{j,k=1}^{\infty} p_k df_{k+j} q_{k+j-1} + q_{k+j-1} df_{k+j} p_k - \sum_{k=1}^{\infty} p_k f q + q f p_k \\
&\quad = \sum_{j,k=1}^{\infty} p_k df_{k+j} q_{k+j-1} + q_{k+j-1} df_{k+j} p_k - (\mathbf{1}_{\mathcal{A}} - q) f q + q f (\mathbf{1}_{\mathcal{A}} - q).
\end{aligned}$$

Indeed, we have used

$$p_k f_k q = p_k q_{k-1} f_k q_{k-1} q = 0 = q q_{k-1} f_k q_{k-1} p_k = q f_k p_k.$$

Combined the identities obtained so far, we get

$$g_{\mathit{off}} = \sum_{s=1}^{\infty} \sum_{k=1}^{\infty} p_k df_{k+s} q_{k+s-1} + q_{k+s-1} df_{k+s} p_k = \sum_{s=1}^{\infty} \sum_{k=1}^{\infty} g_{k,s}.$$

We shall use through out this expression for g_{off} in terms of the functions $g_{k,s}$.

5.2. **Noncommutative pseudo-localization.** Now we formulate and prove the noncommutative extension of our pseudo-localization principle. We need a weak notion of support from [44] which is quite useful when dealing with weak type inequalities. For a non-necessarily self-adjoint $f \in \mathcal{A}$, the *two-sided null projection* of f is the greatest projection q in $\mathcal{A}_\pi$ satisfying $qfq = 0$. Then we define the *weak support projection* of f as

$$\mathrm{supp}^* f = \mathbf{1}_{\mathcal{A}} - q.$$

It is clear that $\mathrm{supp}^* f = \mathrm{supp} f$ when $\mathcal{A}$ is abelian. Moreover, this notion is weaker than the usual support projection in the sense that we have $\mathrm{supp}^* f \le \mathrm{supp} f$ for any self-adjoint $f \in \mathcal{A}$ and $\mathrm{supp}^* f$ is a subprojection of both the left and right supports in the non-self-adjoint case.

Remark 5.1. Below we shall use the following characterization of the weak support projection. The projection $\mathrm{supp}^* f$ is the smallest projection p in $\mathcal{A}_\pi$ satisfying the identity

$$f = pf + fp - pfp.$$

Indeed, let q be the two-sided null projection of f and let $p = \mathbf{1}_\mathcal{A} - q$. Then we have $(\mathbf{1}_\mathcal{A} - p) f (\mathbf{1}_\mathcal{A} - p) = 0$ by definition. In other words, $f = pf + fp - pfp$ and p is the smallest projection with this property because q is the greatest projection satisfying the identity $qfq = 0$.

The following constitutes a noncommutative analog of the pseudo-localization principle that we have stated in the Introduction. The terminology has been chosen to fit with that of the noncommutative Calderón-Zygmund decomposition. This will make the exposition more transparent.

Theorem 5.2. *Let us fix a positive integer s. Given a function $f \in L_2(\mathcal{A})$ and any integer k, let us consider any projection q_k in $\mathcal{A}_\pi \cap \mathcal{A}_k$ satisfying that $\mathbf{1}_\mathcal{A} - q_k$ contains $\mathrm{supp}^* df_{k+s}$ as a subprojection. If we write*

$$q_k = \sum_{Q \in \mathcal{Q}_k} \xi_Q 1_Q$$

with $\xi_Q \in \mathcal{M}_\pi$, we may further consider the projection

$$\zeta_{f,s} = \bigwedge_{k \in \mathbb{Z}} \Big(\mathbf{1}_\mathcal{A} - \bigvee_{Q \in \mathcal{Q}_k} (\mathbf{1}_\mathcal{M} - \xi_Q) 1_{9Q} \Big).$$

Then we have the following localization estimate in $L_2(\mathcal{A})$

$$\Big(\int_{\mathbb{R}^n} \tau \Big(\big| \big[\zeta_{f,s} \, Tf \, \zeta_{f,s} \big](x) \big|^2 \Big) \, dx \Big)^{\frac12} \le \mathrm{c}_{n,\gamma} s \, 2^{-\gamma s/4} \Big(\int_{\mathbb{R}^n} \tau \big(|f(x)|^2 \big) \, dx \Big)^{\frac12},$$

for any L_2-normalized Calderón-Zygmund operator with Lipschitz parameter γ.

Proof. We shall reduce this result to its commutative counterpart. More precisely to the shifted form of the $T1$ theorem proved above. According to Remark 5.1 and the shift condition $\mathrm{supp}^* df_{k+s} \prec \mathbf{1}_\mathcal{A} - q_k$, we have

$$df_{k+s} = q_k^\perp df_{k+s} + df_{k+s} q_k^\perp - q_k^\perp df_{k+s} q_k^\perp$$

where we write $q_k^\perp = \mathbf{1}_\mathcal{A} - q_k$ for convenience. On the other hand, let

$$\zeta_k = \mathbf{1}_\mathcal{A} - \bigvee_{Q \in \mathcal{Q}_k} (\mathbf{1}_\mathcal{M} - \xi_Q) 1_{9Q},$$

so that $\zeta_{f,s} = \bigwedge_k \zeta_k$. Following Lemma 4.2, it is easily seen that $\mathbf{1}_\mathcal{A} - \zeta_k$ represents the $\mathbb{R}^n$-dilated projection associated to $\mathbf{1}_\mathcal{A} - q_k$ with a factor 9. Let $\mathcal{L}_a$ and $\mathcal{R}_a$ denote the left and right multiplication maps by the operator a. Let also $\mathcal{LR}_a$ stand for $\mathcal{L}_a + \mathcal{R}_a - \mathcal{L}_a \mathcal{R}_a$ Then our considerations so far and the fact that $\mathcal{L}_{\zeta_k}, \mathcal{R}_{\zeta_k}$ and $\mathcal{LR}_{q_k^\perp}$ commute with E_j for $j \ge k$ give

$$\begin{aligned} & \zeta_{f,s} \, Tf \, \zeta_{f,s} \\ & \quad = \ \mathcal{L}_{\zeta_{f,s}} \mathcal{R}_{\zeta_{f,s}} \Big(\sum\nolimits_k \mathsf{E}_k T \Delta_{k+s} \mathcal{LR}_{q_k^\perp} + \sum\nolimits_k (id - \mathsf{E}_k) \mathcal{L}_{\zeta_k} \mathcal{R}_{\zeta_k} T \mathcal{LR}_{q_k^\perp} \Delta_{k+s} \Big)(f). \end{aligned}$$

Now we claim that

$$\mathcal{L}_{\zeta_k} \mathcal{R}_{\zeta_k} T \mathcal{LR}_{q_k^\perp} = \mathcal{L}_{\zeta_k} \mathcal{R}_{\zeta_k} T_{4 \cdot 2^{-k}} \mathcal{LR}_{q_k^\perp}.$$

Indeed, this clearly reduces to see

$$\mathcal{L}_{\zeta_k} T \mathcal{L}_{q_k^\perp} = \mathcal{L}_{\zeta_k} T_{4 \cdot 2^{-k}} \mathcal{L}_{q_k^\perp} \qquad \text{and} \qquad \mathcal{R}_{\zeta_k} T \mathcal{R}_{q_k^\perp} = \mathcal{R}_{\zeta_k} T_{4 \cdot 2^{-k}} \mathcal{R}_{q_k^\perp}.$$

By symmetry, we just prove the first identity

$$\mathcal{L}_{\zeta_k} T \mathcal{L}_{q_k^\perp} f(x) \, = \sum_{Q \in \mathcal{Q}_k} \zeta_k(x)(\mathbf{1}_{\mathcal{M}} - \xi_Q) \int_Q k(x,y) f(y) \, dy.$$

Assume that $x \in 9Q$ for some $Q \in \mathcal{Q}_k$, then it easily follows from the definition of the projection ζ_k that $\zeta_k(x) \le \xi_Q$. Note that this is simpler than the argument in Lemma 4.2 because we do not need to prove here a property like i) there. In particular, we deduce from the expression above that for each $y \in Q$ we must have $x \in \mathbb{R}^n \setminus 9Q$. This implies that $|x-y| \ge 4 \cdot 2^{-k}$ as desired. Finally, since the operators $\mathcal{L}$ and $\mathcal{R}$ were created from properties of f and $\zeta_{f,s}$, we can eliminate them and obtain

$$\zeta_{f,s} \, Tf \, \zeta_{f,s} \, = \, \mathcal{L}_{\zeta_{f,s}} \mathcal{R}_{\zeta_{f,s}} \Big(\sum\nolimits_k \mathsf{E}_k T \Delta_{k+s} + \sum\nolimits_k (id - \mathsf{E}_k) T_{4 \cdot 2^{-k}} \Delta_{k+s} \Big)(f).$$

Assume that $T^*1 = 0$. According to our shifted form of the $T1$ theorem, we know that the operator inside the brackets has norm in $\mathcal{B}(L_2)$ controlled by $\mathrm{c}_{n,\gamma} s \, 2^{-\gamma s/4}$. In particular, the same happens when we tensor with the identity on $L_2(\mathcal{M})$, which is the case. This proves the assertion for convolution-type operators. When T^*1 is a non-zero element of BMO, we may follow verbatim the paraproduct argument given above since $\mathcal{L}\mathcal{R}_{q_k^\perp}$ commutes with E_k and $\zeta_{f,s} q_k^\perp = q_k^\perp \zeta_{f,s} = 0$. □

Remark 5.3. It is apparent that $\mathbf{1}_{\mathcal{A}} - q_k$ represents in the noncommutative setting the set Ω_k in the commutative formulation. Moreover, $\zeta_{f,s}$ and ζ_k represent $\mathbb{R}^n \setminus \Sigma_{f,s}$ and $\mathbb{R}_n \setminus 9\,\Omega_k$ respectively. The only significant difference is that in the commutative statement we take Ω_k to be the *smallest* $\mathcal{R}_k$-set containing $\mathrm{supp}\, df_{k+s}$. This is done to optimize the corresponding localization estimate. Indeed, the smaller are the Ω_k's the larger is $1_{\mathbb{R}^n \setminus \Sigma_{f,s}} Tf$. However, it is in general false that the smaller are the $\mathbf{1}_{\mathcal{A}} - q_k$'s the larger is $\zeta_{f,s} \, T \zeta_{f,s}$. That is why we consider *any* sequence of q_k's satisfying the shift condition.

5.3. **Estimation of Tg_{off}.** Our aim is to estimate

$$\lambda \varphi \Big\{ \big| T g_{\mathit{off}} \big| > \lambda \Big\}.$$

As usual, we decompose the term Tg_{off} in the following way

$$(\mathbf{1}_{\mathcal{A}} - \zeta) T(g_{\mathit{off}}) (\mathbf{1}_{\mathcal{A}} - \zeta) + \zeta T(g_{\mathit{off}})(\mathbf{1}_{\mathcal{A}} - \zeta) + (\mathbf{1}_{\mathcal{A}} - \zeta) T(g_{\mathit{off}}) \zeta + \zeta T(g_{\mathit{off}}) \zeta,$$

where ζ denotes the projection constructed in Lemma 4.2. According to this lemma and the argument in Paragraph 4.3, we are reduced to estimate the last term above. This will be done in several steps.

5.3.1. *Orthogonality.* It is not difficult to check that the terms $g_{k,s}$ in Paragraph 5.1 are pairwise orthogonal. It follows from the trace-invariance of conditional expectations and the mutual orthogonality of the p_k's. We first prove the following implication

$$\varphi(g_{k,s} g_{k',s'}^*) \neq 0 \;\Rightarrow\; k+s = k'+s'.$$

Indeed, if we assume w.l.o.g. that $k+s > k'+s'$, we get

$$\varphi(g_{k,s} g_{k',s'}^*) = \varphi\big(\mathsf{E}_{k+s-1}(g_{k,s} g_{k',s'}^*)\big) = \varphi\big(\mathsf{E}_{k+s-1}(g_{k,s}) g_{k',s'}^*\big) = 0.$$

Now, assume that $k \neq k'$ and $k+s = k'+s'$. By the orthogonality of the p_k's

$$\varphi(g_{k,s} g_{k',s'}^*) \quad = \quad \varphi\big(p_k df_{k+s} q_{k+s-1} p_{k'} df_{k+s} q_{k+s-1}\big)$$

$$\begin{aligned} & + \ \varphi\big(q_{k+s-1}df_{k+s}p_kq_{k+s-1}df_{k+s}p_{k'}\big) \\ & = \ \varphi\big(p_{k'}df_{k+s}q_{k+s-1}p_kdf_{k+s}q_{k+s-1}\big) \\ & + \ \varphi\big(q_{k+s-1}df_{k+s}p_kq_{k+s-1}df_{k+s}p_{k'}\big) \ = \ 0 \end{aligned}$$

since $p_kq_{k+s-1} = q_{k+s-1}p_k = 0$. This means that $\varphi(g_{k,s}g_{k',s'}^*) = 0$ unless $k = k'$ and $k+s = k'+s'$ or, equivalently, $(k,s) = (k',s')$. Therefore, the $g_{k,s}$'s are pairwise orthogonal and

$$\|g_{\mathit{off}}\|_2^2 = \sum_{s=1}^\infty \sum_{k=1}^\infty \|g_{k,s}\|_2^2.$$

5.3.2. *An $\ell_\infty(\ell_2)$ estimate.* Following the classical argument in Calderón-Zygmund decomposition or our estimate for the diagonal terms in Paragraph 4.1, it would suffice to prove that $\|g_{\mathit{off}}\|_2^2 \lesssim \lambda \|f\|_1$. According to the pairwise orthogonality of the $g_{k,s}$'s, that is to say

$$\sum_{s=1}^\infty \sum_{k=1}^\infty \|g_{k,s}\|_2^2 \lesssim \lambda \|f\|_1.$$

However, we just have the weaker inequality

$$\sup_{s \ge 1} \sum_{k=1}^\infty \|g_{k,s}\|_2^2 \lesssim \lambda \|f\|_1. \tag{5.1}$$

Let us prove this estimate before going on with the proof

$$\begin{aligned} \|g_{k,s}\|_2^2 \ & = \ 2\varphi\big(p_kdf_{k+s}q_{k+s-1}df_{k+s}p_k\big) \\ & = \ 2\varphi\big(p_kf_{k+s}q_{k+s-1}f_{k+s}p_k\big) \ - \ 2\varphi\big(p_kf_{k+s}q_{k+s-1}f_{k+s-1}p_k\big) \\ & - \ 2\varphi\big(p_kf_{k+s-1}q_{k+s-1}f_{k+s}p_k\big) \ + \ 2\varphi\big(p_kf_{k+s-1}q_{k+s-1}f_{k+s-1}p_k\big). \end{aligned}$$

By Cuculescu's construction ii) and $f_j \le 2^n f_{j-1}$ (see Paragraph 4.1), we find

$$\begin{aligned} \big\|f_{k+s}^{\frac12}\, q_{k+s-1} f_{k+s}^{\frac12}\big\|_\infty \ & = \ \big\|q_{k+s-1} f_{k+s}\, q_{k+s-1}\big\|_\infty \ && \lesssim \ \lambda, \\ \big\|f_{k+s-1}^{\frac12} q_{k+s-1} f_{k+s-1}^{\frac12}\big\|_\infty \ & = \ \big\|q_{k+s-1} f_{k+s-1} q_{k+s-1}\big\|_\infty \ && \le \ \lambda. \end{aligned}$$

The crossed terms require Hölder's inequality

$$\begin{aligned} \varphi\big(p_kf_{k+s}q_{k+s-1}f_{k+s-1}p_k\big) \ & \le \ \varphi\big(p_kf_{k+s}q_{k+s-1}f_{k+s}p_k\big)^{\frac12} \\ & \times \ \varphi\big(p_kf_{k+s-1}q_{k+s-1}f_{k+s-1}p_k\big)^{\frac12} \\ & \lesssim \ \lambda\, \varphi\big(p_kf_{k+s}p_k\big)^{\frac12} \varphi\big(p_kf_{k+s-1}p_k\big)^{\frac12} \ = \ \lambda \varphi\big(p_kfp_k\big), \end{aligned}$$

where the last identity uses the trace-invariance of the conditional expectations E_{k+s} and E_{k+s-1} respectively. The same estimate holds for the remaining crossed term. This proves that

$$\sup_{s\ge1} \sum_{k=1}^\infty \|g_{k,s}\|_2^2 \lesssim \lambda \sup_{s\ge1} \sum_{k=1}^\infty \varphi\big(p_kfp_k\big) \le \lambda \|f\|_1.$$

5.3.3. *The use of pseudo-localization.* Consider the function

$$g_{(s)} = \sum\nolimits_k g_{k,s}.$$

It is straightforward to see that $dg_{(s)}{}_{k+s} = g_{k,s}$. In particular, we have

$$\mathrm{supp}^* \, dg_{(s)}{}_{k+s} \le p_k = q_{k-1} - q_k \le \mathbf{1}_{\mathcal{A}} - q_k. \tag{5.2}$$

According to the terminology of Theorem 5.2, we consider the projection

$$\zeta_{g_{(s)},s} = \bigwedge_{k \ge 1} \Big(\mathbf{1}_{\mathcal{A}} - \bigvee_{Q \in \mathcal{Q}_k} (\mathbf{1}_{\mathcal{M}} - \xi_Q) 1_{9Q} \Big).$$

Notice that we are just taking $k \ge 1$ and not $k \in \mathbb{Z}$ as in Theorem 5.2. This is justified by the fact that the q_k's are now given by Cuculescu's construction applied to our $f \in \mathcal{A}_{c,+}$ and our assumption in Remark 3.3 implies that $\mathbf{1}_{\mathcal{M}} - \xi_Q = 0$ for all $Q \in \mathcal{Q}_k$ with $k < 1$. Now, if we compare this projection with the one provided by Lemma 4.2

$$\zeta = \bigwedge_{k \ge 1} \Big(\mathbf{1}_{\mathcal{A}} - \bigvee_{\substack{1 \le j \le k \\ Q \in \mathcal{Q}_j}} (\xi_{\widehat{Q}} - \xi_Q) 1_{9Q} \Big),$$

it becomes apparent that $\zeta \le \zeta_{g_{(s)},s}$. On the other hand, Chebychev's inequality gives

$$\lambda \varphi \Big\{ \big| \zeta T(g_{\mathit{off}}) \zeta \big| > \lambda \Big\} = \lambda \varphi \Big\{ \Big| \sum_{s=1}^{\infty} \zeta T(g_{(s)}) \zeta \Big| > \lambda \Big\} \le \frac{1}{\lambda} \Big[\sum_{s=1}^{\infty} \big\| \zeta T(g_{(s)}) \zeta \big\|_2 \Big]^2.$$

This automatically implies

$$\lambda \varphi \Big\{ \big| \zeta T(g_{\mathit{off}}) \zeta \big| > \lambda \Big\} \le \frac{1}{\lambda} \Big[\sum_{s=1}^{\infty} \big\| \zeta_{g_{(s)},s} T(g_{(s)}) \zeta_{g_{(s)},s} \big\|_2 \Big]^2.$$

Now, combining (5.1) and (5.2), we may use pseudo-localization and deduce

$$\begin{aligned} \lambda \varphi \Big\{ \big| \zeta T(g_{\mathit{off}}) \zeta \big| > \lambda \Big\} & \le \frac{\mathrm{c}_{n,\gamma}^2}{\lambda} \Big[\sum_{s=1}^{\infty} s \, 2^{-\gamma s/4} \|g_{(s)}\|_2 \Big]^2 \\ & = \frac{\mathrm{c}_{n,\gamma}^2}{\lambda} \Big[\sum_{s=1}^{\infty} s \, 2^{-\gamma s/4} \Big(\sum\nolimits_k \|g_{k,s}\|_2^2 \Big)^{\frac{1}{2}} \Big]^2 \ \le \ \mathrm{c}_{n,\gamma} \|f\|_1. \end{aligned}$$

This completes the argument for the off-diagonal terms of g.

5.4. **Estimation of Tb_{off}.** As above, it suffices to estimate

$$\zeta T(b_{\mathit{off}}) \zeta = \sum_{s=1}^{\infty} \sum_{k=1}^{\infty} \zeta T \Big(p_k (f - f_{k+s}) p_{k+s} + p_{k+s} (f - f_{k+s}) p_k \Big) \zeta.$$

In the sequel we use the following notation. For any dyadic cube $Q \in \mathcal{Q}_{k+s}$, we shall write Q_k to denote the s-th antecessor of Q. That is, Q_k is the only dyadic cube in $\mathcal{Q}_k$ containing Q. If we set

$$b_{k,s} = p_k (f - f_{k+s}) p_{k+s} + p_{k+s} (f - f_{k+s}) p_k,$$

the identity below follows from $\xi_{Q_k} \pi_{Q_k} = \pi_{Q_k} \xi_{Q_k} = 0$ and Lemma 4.2

$$\begin{aligned} & \zeta T(b_{k,s}) \zeta(x) \\ & \quad = \int_{\mathbb{R}^n} k(x,y) \big(\zeta(x) b_{k,s}(y) \zeta(x) \big) \, dy \end{aligned}$$

$$\begin{aligned}
&= \zeta(x)\Big(\sum_{Q\in\mathcal{Q}_{k+s}} \int_Q k(x,y)\pi_{Q_k}(f(y)-f_Q)\pi_Q\, dy\ 1_{(9Q_k)^c}(x)\Big)\zeta(x) \\
&+ \zeta(x)\Big(\sum_{Q\in\mathcal{Q}_{k+s}} \int_Q k(x,y)\pi_{Q}(f(y)-f_Q)\pi_{Q_k}\, dy\ 1_{(9Q_k)^c}(x)\Big)\zeta(x) \\
&= \zeta(x)\Big(\sum_{Q\in\mathcal{Q}_{k+s}} \int_Q k(x,y) b_{k,s}(y)\, dy\ 1_{(9Q_k)^c}(x)\Big)\zeta(x).
\end{aligned}$$

Before going on with the proof, let us explain a bit our next argument. Our terms $b_{s,k}$ are located in the $(s+1)$-th upper and lower diagonals and we want to compare their *size* with that of the main diagonal. To do so we write each $b_{k,s}$, located in the entries $(k,k+s)$ and $(k+s,k)$, as a linear combination of four diagonal boxes in a standard way. However, this procedure generates overlapping and we are forced to consider only those integers k congruent to a fixed $1 \le j \le s+1$ at a time. The figure below will serve as a model ($s=2$) for our forthcoming estimates.

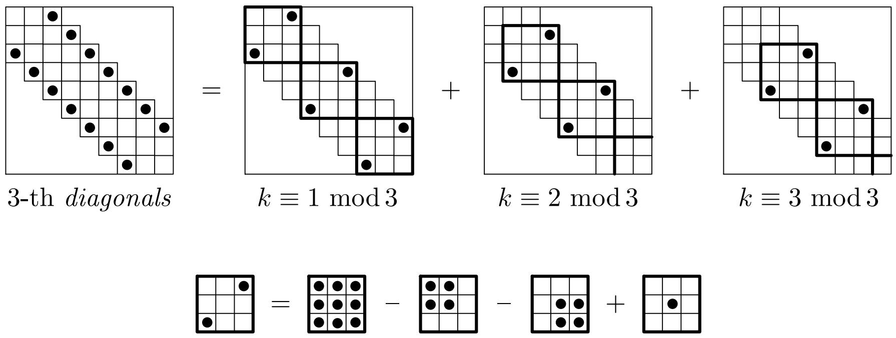


Figure VI
Decomposition into disjoint diagonal boxes for $s=2$

According to Chebychev's inequality we obtain

$$\begin{aligned}
\lambda\varphi\Big\{\Big|\sum_{s=1}^\infty\sum_{k=1}^\infty \zeta T(b_{k,s})\zeta\Big| > \lambda\Big\}
&\le \sum_{s=1}^\infty \Big\|\sum_{k=1}^\infty \zeta T(b_{k,s})\zeta\Big\|_1 \\
&\le \sum_{s=1}^\infty\sum_{k=1}^\infty \sum_{Q\in\mathcal{Q}_{k+s}} \Big\|\int_Q k(\,\cdot\,,y) b_{k,s}(y)\, dy\ 1_{(9Q_k)^c}(\,\cdot\,)\Big\|_1 \\
&\le \sum_{s=1}^\infty\sum_{j=0}^s \sum_{\substack{k\equiv j \\ \mathrm{mod}\ s+1}} \sum_{Q\in\mathcal{Q}_{k+s}} \Big\|\int_Q k(\,\cdot\,,y) b_{k,s}(y) dy\, 1_{(9Q_k)^c}(\,\cdot\,)\Big\|_1.
\end{aligned}$$

We now use the decomposition

$$b_{k,s} = \Big(\sum_{r=0}^s p_{k+r}\Big)(f-f_{k+s})\Big(\sum_{r=0}^s p_{k+r}\Big)$$

$$
\begin{aligned}
- & \quad \Big(\sum_{r=0}^{s-1} p_{k+r} \Big) (f - f_{k+s}) \Big(\sum_{r=0}^{s-1} p_{k+r} \Big) \\
- & \quad \Big(\sum_{r=1}^{s} p_{k+r} \Big) (f - f_{k+s}) \Big(\sum_{r=1}^{s} p_{k+r} \Big) \\
+ & \quad \Big(\sum_{r=1}^{s-1} p_{k+r} \Big) (f - f_{k+s}) \Big(\sum_{r=1}^{s-1} p_{k+r} \Big) \ = \ b_{k,s}^1 - b_{k,s}^2 - b_{k,s}^3 + b_{k,s}^4,
\end{aligned}
$$

of $b_{k,s}$ as a linear combination of four *diagonal* terms. Let us recall that the four projections $\sum_r p_{k+r}$ above (with $0 \preceq r \preceq s$ and $\preceq$ meaning either $<$ or $\leq$) belong to $\mathcal{A}_{k+s}$. In particular, since $\mathsf{E}_{k+s}(f - f_{k+s}) = 0$, the following identity holds for any $Q \in \mathcal{Q}_{k+s}$ and any $1 \leq i \leq 4$

$$
\int_Q b_{k,s}^i(y) \, dy = 0.
$$

Therefore, we find

$$
\begin{aligned}
\lambda \varphi & \Big\{ \Big| \sum_{s=1}^{\infty} \sum_{k=1}^{\infty} \zeta T(b_{k,s}) \zeta \Big| > \lambda \Big\} \\
\leq & \ \sum_{s=1}^{\infty} \sum_{i=1}^{4} \sum_{j=0}^{s} \sum_{\substack{k \equiv j \\ \mathrm{mod}\ s+1}} \sum_{Q \in \mathcal{Q}_{k+s}} \int_Q \Big\| \big(k(\,\cdot\,, y) - k(\,\cdot\,, c_Q) \big) \, b_{k,s}^i(y) \, 1_{(9Q_k)^c}(\,\cdot\,) \Big\|_1 \, dy.
\end{aligned}
$$

However, by Lipschitz γ-smoothness we have

$$
\begin{aligned}
\int_Q & \Big\| \big(k(\,\cdot\,, y) - k(\,\cdot\,, c_Q) \big) \, b_{k,s}^i(y) \, 1_{(9Q_k)^c}(\,\cdot\,) \Big\|_1 \, dy \\
& = \int_Q \tau \Big[\Big(\int_{(9Q_k)^c} \big| k(x,y) - k(x,c_Q) \big| \, dx \Big) |b_{k,s}^i(y)| \Big] \, dy \\
& \lesssim \int_Q \Big(\int_{(9Q_k)^c} \frac{|y - c_Q|^{\gamma}}{|x - c_Q|^{n+\gamma}} \, dx \Big) \, \tau |b_{k,s}^i(y)| \, dy \\
& \lesssim \ell(Q)^{\gamma} / \ell(Q_k)^{\gamma} \ \ \varphi \Big[\Big(\sum_{r=0}^{s} p_{k+r} \Big) (f + f_{k+s}) \Big(\sum_{r=0}^{s} p_{k+r} \Big) 1_Q \Big] \\
& \lesssim 2^{-\gamma s} \varphi \Big[\Big(\sum_{r=0}^{s} p_{k+r} \Big) f \Big(\sum_{r=0}^{s} p_{k+r} \Big) 1_Q \Big].
\end{aligned}
$$

Finally, summing over (s, i, j, k, Q) we get

$$
\begin{aligned}
\lambda \varphi & \Big\{ \Big| \sum_{s=1}^{\infty} \sum_{k=1}^{\infty} \zeta T(b_{k,s}) \zeta \Big| > \lambda \Big\} \\
& \leq \ \sum_{s=1}^{\infty} \sum_{i=1}^{4} \sum_{j=0}^{s} \sum_{\substack{k \equiv j \\ \mathrm{mod}\ s+1}} 2^{-\gamma s} \varphi \Big[\Big(\sum_{r=0}^{s} p_{k+r} \Big) f \Big(\sum_{r=0}^{s} p_{k+r} \Big) \Big] \\
& \leq \ \Big(\sum_{s=1}^{\infty} \sum_{i=1}^{4} \sum_{j=0}^{s} 2^{-\gamma s} \Big) \, \|f\|_1 \ = \ 4 \Big(\sum_{s=1}^{\infty} \frac{s+1}{2^{\gamma s}} \Big) \, \|f\|_1 = 4 \mathrm{c}_{\gamma} \, \|f\|_1.
\end{aligned}
$$

This completes the argument for the off-diagonal terms of b.

5.5. **Conclusion.** Combining the results obtained so far in Sections 4 and 5, we obtain the weak type inequality announced in Theorem A. The strong L_p estimates follow for $1 < p < 2$ from the real interpolation method, see e.g. [48] for more information on the real interpolation of noncommutative L_p spaces. In the case $2 < p < \infty$, our estimates follow from duality since our size/smoothness conditions on the kernel are symmetric in x and y.

Remark 5.4. Recent results in noncommutative harmonic analysis [25, 27, 28, 30] show the relevance of non-semifinite von Neumann algebras in the theory. The definition of the corresponding L_p spaces (so called Haagerup L_p spaces) is more involved, see [19, 56]. A well-known reduction argument due to Haagerup [18] allows us to extend our strong L_p estimates in Theorems A and B to functions $f : \mathbb{R}^n \to \mathcal{M}$ with $\mathcal{M}$ a type III von Neumann algebra $\mathcal{M}$. Indeed, if σ denotes the one-parameter unimodular group associated to $(\mathcal{A}, \varphi)$, we take the crossed product $\mathcal{R} = \mathcal{A} \rtimes_\sigma \mathrm{G}$ with the group $\mathrm{G} = \bigcup_{n \in \mathbb{N}} 2^{-n}\mathbb{Z}$. According to [18], $\mathcal{R}$ is the closure of a union of finite von Neumann algebras $\bigcup_{k \geq 1} \mathcal{A}_k$ directed by inclusion. We know that our result holds on $L_p(\mathcal{A}_k)$ for $1 < p < \infty$ and with constants independent of k. Therefore, the same will hold on $L_p(\mathcal{R})$. Then, using that $L_p(\mathcal{A})$ is a (complemented) subspace of $L_p(\mathcal{R})$, the assertion follows.

Remark 5.5. According to the classical theory, it seems that some hypotheses of Theorem A could be weakened. For instance, the size condition on the kernel is not needed for scalar-valued functions. Moreover, it is well-known that the classical theory only uses Lipschitz smoothness on the second variable to produce weak type $(1, 1)$ estimates. Going even further, it is unclear whether or not we can use weaker smoothness conditions, like Hörmander type conditions. Nevertheless, all these apparently extra assumptions become quite natural if we notice that all of them where used to produce our pseudo-localization principle, a key point in the whole argument. Under this point of view, we have just imposed the natural hypotheses which appear around the $T1$ theorem. This leads us to pose the following problem.

Problem. Can we weaken the hypotheses on the kernel as pointed above?

Remark 5.6. We believe that our methods should generalize if we replace $\mathbb{R}^n$ by any other space of homogeneous type. In other words, a metric space equipped with a non-negative Borel measure which is doubling with respect to the given metric. More general notions can be found in [8, 33, 34]. Of course, following recent results by Nazarov/Treil/Volberg and Tolsa, it is also possible to study extensions of non-doubling Calderón-Zygmund theory in our setting. It is not so clear that the methods of this paper can be easily adapted to this case.

6. Operator-valued kernels

We now consider Calderón-Zygmund operators associated to operator-valued kernels $k : \mathbb{R}^{2n} \setminus \Delta \to \mathcal{M}$ satisfying the canonical size/smoothness conditions. In other words, we replace the absolute value by the $\mathcal{M}$-norm, see the Introduction for details. We begin by constructing certain bad kernels which show that there is no hope to extend Theorem A in full generality to this context. Then we obtain positive results assuming some extra hypotheses.

6.1. **Negative results.** The origin of the counterexample we are constructing goes back to a lack (well-known to experts in the field) of noncommutative martingale transforms

$$\sum_k df_k \mapsto \sum_k \xi_{k-1} df_k.$$

Indeed, the boundedness of this operator on L_p might fail when the predictable sequence of ξ_k's is operator-valued. Here is a simple example. Let $\mathcal{A}$ be the algebra of $m \times m$ matrices equipped with the standard trace tr and consider the filtration $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_m$, where $\mathcal{A}_s$ denotes the subalgebra spanned by the matrix units e_{ij} with $1 \le i, j \le s$ and the matrix units e_{kk} with $k > s$.

- If $1 < p < 2$, we take $f = \sum_{k=2}^m e_{1k}$ and $\xi_k = e_{k1}$, so that
$$\Big\| \sum_k \xi_{k-1} df_k \Big\|_p = (m-1)^{1/p} \gg \sqrt{m-1} = \Big\| \sum_k df_k \Big\|_p.$$
- If $2 < p < \infty$, we take $f = \sum_{k=2}^m e_{k-1,k}$ and $\xi_k = e_{1k}$, so that
$$\Big\| \sum_k df_k \Big\|_p = (m-1)^{1/p} \ll \sqrt{m-1} = \Big\| \sum_k \xi_{k-1} df_k \Big\|_p.$$

Letting $m \to \infty$, we see that L_p boundedness might fail for any $p \neq 2$ even having L_2 boundedness. Our aim is to prove that the same phenomenon happens in the context of singular integrals with operator-valued kernels. The examples above show us the right way to proceed. Namely, we shall construct a similar operator using Littlewood-Paley type arguments. Note that a dyadic martingale approach is also possible here, but this would give rise to certain operators having non-smooth kernels and we want to show that smoothness does not help in this particular case.

Let $\mathcal{S}_\mathbb{R}$ be the Schwarz class in $\mathbb{R}$ and consider a non-negative function ψ in $\mathcal{S}_\mathbb{R}$ bounded above by 1, supported in $1 \le |\xi| \le 2$ and identically 1 in $5/4 \le |\xi| \le 7/4$. Define

$$\psi_k(\xi) = \psi(2^{-k}\xi).$$

Let Ψ denote the inverse Fourier transform of ψ, so that $\widehat{\Psi} = \psi$. If we construct the functions $\Psi_k(x) = 2^k \, \Psi(2^k x)$, we have $\widehat{\Psi}_k = \psi_k$ and we may define the following convolution-type operators

$$\begin{aligned} T_1 f(x) &= \sum_{k \ge 1} e_{k1} \, \Psi_k * f, \\ T_2 f(x) &= \sum_{k \ge 1} e_{1k} \, \Psi_k * f. \end{aligned}$$

In this case we are taking $\mathcal{M} = \mathcal{B}(\ell_2)$ and both T_1 and T_2 become contractive operators in $L_2(\mathcal{A})$. Indeed, let $\mathcal{F}_\mathcal{A} = \mathcal{F}_\mathbb{R} \otimes id_{L_2(\mathcal{M})}$ denote the Fourier transform on $L_2(\mathcal{A})$. According to Plancherel's theorem, $\mathcal{F}_\mathcal{A}$ is an isometry and the following inequality holds

$$\begin{aligned} \|T_1 f\|_2 &= \Big\| \sum_{k \ge 1} e_{k1} \widehat{\Psi_k * f} \Big\|_2 = \Big(\sum_{k \ge 1} \big\| e_{k1} \psi_k \widehat{f} \big\|_2^2 \Big)^{\frac12} \\ &\le \Big(\sum_{k \ge 1} \int_{2^k \le |\xi| \le 2^{k+1}} |\widehat{f}(\xi)|^2 \, d\xi \Big)^{\frac12} \le \|f\|_2. \end{aligned}$$

The same argument works for T_2. Now we show that the kernels of T_1 and T_2 also satisfy the expected size and smoothness conditions. These are convolution-type kernels given by

$$k_1(x,y) = \sum_{k\geq 1} e_{k1}\Psi_k(x-y) \quad \text{and} \quad k_2(x,y) = \sum_{k\geq 1} e_{1k}\Psi_k(x-y).$$

We clearly have

$$\begin{aligned}
\|k_1(x,y)\|_{\mathcal{M}} &= \Big\| \sum_{k\geq 1} e_{k1}\Psi_k(x-y) \Big\|_{\mathcal{M}} &= \Big(\sum_{k\geq 1} |\Psi_k(x-y)|^2 \Big)^{\frac12}, \\
\|k_2(x,y)\|_{\mathcal{M}} &= \Big\| \sum_{k\geq 1} e_{1k}\Psi_k(x-y) \Big\|_{\mathcal{M}} &= \Big(\sum_{k\geq 1} |\Psi_k(x-y)|^2 \Big)^{\frac12}.
\end{aligned}$$

Therefore, for the size condition it suffices to see that

$$\Big(\sum_{k\in\mathbb{Z}} |\Psi_k(x)|^2 \Big)^{\frac12} \lesssim \frac{1}{|x|}. \tag{6.1}$$

Similarly, using the mean value theorem in the usual way, the condition

$$\Big(\sum_{k\in\mathbb{Z}} |\Psi_k'(x)|^2 \Big)^{\frac12} \lesssim \frac{1}{|x|^2} \tag{6.2}$$

implies Lipschitz smoothness for any $0 < \gamma \leq 1$. The proof of (6.1) and (6.2) is standard. Namely, since Ψ and Ψ' belong to the Schwarz class $\mathcal{S}_{\mathbb{R}}$, there exist absolute constants c_1 and c_2 such that

$$|\Psi(x)| \leq \mathrm{c}_1 \min\Big\{1, \frac{1}{|x|^2}\Big\} \quad \text{and} \quad |\Psi'(x)| \leq \mathrm{c}_2 \min\Big\{1, \frac{1}{|x|^3}\Big\}.$$

If $2^{-j} \leq |x| < 2^{-j+1}$, we find the estimate

$$\Big(\sum_{k\in\mathbb{Z}} |\Psi_k(x)|^2 \Big)^{\frac12} \leq \Big(\mathrm{c}_1 \sum_{k\leq j} 2^{2k} + \mathrm{c}_1 |x|^{-4} \sum_{k>j} 2^{-2k} \Big)^{\frac12} \lesssim \Big(2^{2j} + \frac{1}{2^{2j}|x|^4} \Big)^{\frac12} \lesssim \frac{1}{|x|}.$$

Similarly, using that $\Psi_k'(x) = 2^{2k}\Psi'(2^k x)$, we have

$$\Big(\sum_{k\in\mathbb{Z}} |\Psi_k'(x)|^2 \Big)^{\frac12} \leq \Big(\mathrm{c}_2 \sum_{k\leq j} 2^{4k} + \mathrm{c}_2 |x|^{-6} \sum_{k>j} 2^{-2k} \Big)^{\frac12} \lesssim \Big(2^{4j} + \frac{1}{2^{2j}|x|^6} \Big)^{\frac12} \lesssim \frac{1}{|x|^2}.$$

Thus, T_1 and T_2 are bounded on $L_2(\mathcal{A})$ with operator-valued kernels satisfying the standard size and smoothness conditions. Now we shall see how the boundedness on $L_p(\mathcal{A})$ fails for $p \neq 2$. By definition, we know that

- ψ_k is supported by $2^k \leq |\xi| \leq 2^{k+1}$.
- ψ_k is identically 1 in $5 \cdot 2^k/4 \leq |\xi| \leq 7 \cdot 2^k/4$.

If $\mathcal{I}_0 = [5/4, 7/4]$ and $\mathcal{I}_k = \mathcal{I}_0 + \frac32(2^k - 1)$, it is easily seen that

$$\psi_k 1_{\mathcal{I}_k} = 1_{\mathcal{I}_k} \tag{6.3}$$

for all nonnegative integer k. Now we are ready to show the behavior of T_1 and T_2 on L_p. Indeed, let us fix an integer $m \geq 1$ and let g_k be the inverse Fourier transform of $1_{\mathcal{I}_k}$ for $1 \leq k \leq m$. Then we set

$$f_1 = \sum_{k=1}^m e_{1k}\, g_k \quad \text{and} \quad f_2 = \sum_{k=1}^m e_{kk}\, g_k.$$

By (6.3) we have $\Psi_j * g_k = \delta_{jk} g_k$ for $1 \le k \le m$. Moreover,

$$\widehat{g}_j(\xi) = \widehat{g}_k\Big(\xi + \frac{3}{2}(2^k - 2^j)\Big) \Rightarrow |g_j(x)| = |g_k(x)|.$$

These observations allow us to obtain the following identities

$$\begin{aligned} \frac{\|T_1 f_1\|_p}{\|f_1\|_p} &= \frac{\Big\| \sum_{k=1}^m e_{kk} g_k \Big\|_p}{\Big\| \sum_{k=1}^m e_{1k} g_k \Big\|_p} = \frac{\Big\| \Big(\sum_{k=1}^m |g_k|^p \Big)^{\frac1p} \Big\|_p}{\Big\| \Big(\sum_{k=1}^m |g_k|^2 \Big)^{\frac12} \Big\|_p} = m^{\frac1p - \frac12} \|g_1\|_p, \\ \frac{\|T_2 f_2\|_p}{\|f_2\|_p} &= \frac{\Big\| \sum_{k=1}^m e_{1k} g_k \Big\|_p}{\Big\| \sum_{k=1}^m e_{kk} g_k \Big\|_p} = \frac{\Big\| \Big(\sum_{k=1}^m |g_k|^2 \Big)^{\frac12} \Big\|_p}{\Big\| \Big(\sum_{k=1}^m |g_k|^p \Big)^{\frac1p} \Big\|_p} = m^{\frac12 - \frac1p} \|g_1\|_p. \end{aligned}$$

Therefore, letting $m \to \infty$ we see that T_1 and T_2 are not bounded on $L_p(\mathcal{A})$ for $1 < p < 2$ and $2 < p < \infty$ respectively. Since we have seen that both are bounded on $L_2(\mathcal{A})$ and are equipped with *good* kernels, we deduce that Theorem A does not hold for T_1 and T_2. This is a consequence of the matrix units we have included in the kernels of our operators.

Remark 6.1. We refer to [37] and [43] for a study of paraproducts associated to operator-valued kernels. There it is shown that certain classical estimates also fail when dealing with noncommuting operator-valued kernels. The results in [43] give new light to Carleson embedding theorem.

6.2. **The $L_\infty \to$ BMO boundedness.** In what follows we shall work under the hypotheses of Theorem B. In other words, with Calderón-Zygmund operators which are $\mathcal{M}$-bimodule maps bounded on $L_q(\mathcal{A})$ and are associated to operator-valued kernels satisfying the standard size/smoothness conditions, see the Introduction for further details. Let us define the noncommutative form of dyadic BMO associated to our von Neumann algebra $\mathcal{A}$. According to [36, 47], we may define the space $\mathrm{BMO}_\mathcal{A}$ as the closure of functions f in $L_{1,\mathrm{loc}}(\mathbb{R}^n; \mathcal{M})$ with

$$\|f\|_{\mathrm{BMO}_\mathcal{A}} = \max \Big\{ \|f\|_{\mathrm{BMO}_\mathcal{A}^r}, \|f\|_{\mathrm{BMO}_\mathcal{A}^c} \Big\} < \infty,$$

where the row and column BMO norms are given by

$$\begin{aligned} \|f\|_{\mathrm{BMO}_\mathcal{A}^r} &= \sup_{Q \in \mathcal{Q}} \Big\| \Big(\frac{1}{|Q|} \int_Q \big(f(x) - f_Q\big)\big(f(x) - f_Q\big)^* \, dx \Big)^{\frac12} \Big\|_\mathcal{M}, \\ \|f\|_{\mathrm{BMO}_\mathcal{A}^c} &= \sup_{Q \in \mathcal{Q}} \Big\| \Big(\frac{1}{|Q|} \int_Q \big(f(x) - f_Q\big)^* \big(f(x) - f_Q\big) \, dx \Big)^{\frac12} \Big\|_\mathcal{M}. \end{aligned}$$

In order to extend our pseudo-localization result to the framework of Theorem B, we shall need to work with the identity $\mathbf{1}_\mathcal{A}$ and show that $T^* \mathbf{1}_\mathcal{A}$ belongs to the noncommutative form of BMO. In fact, the (still unpublished) result below due to Tao Mei [38] gives much more.

Theorem 6.2. *If T is as above, then*

$$\|Tf\|_{\mathrm{BMO}_\mathcal{A}} \le \mathrm{c}_{n,\gamma} \|f\|_\mathcal{A}.$$

Mei's argument for Theorem 6.2 is short and nice for $q = 2$. The case $q \neq 2$ requires the noncommutative analog of John-Nirenberg theorem obtained by Junge and Musat in [26].

Remark 6.3. Let us fix an index $q < p < \infty$. By a recent result of Musat [41] adapted to our setting by Mei [36], we know that $L_q(\mathcal{A})$ and $\mathrm{BMO}_\mathcal{A}$ form an interpolation couple. Moreover, both the real and complex methods give the isomorphism

$$\big[\mathrm{BMO}_\mathcal{A}, L_q(\mathcal{A})\big]_{q/p} \simeq L_p(\mathcal{A})$$

with constant $\mathrm{c}_p \sim p$ for p large. The proof of the latter assertion was achieved in [26], refining the argument of [41]. In particular, the L_p estimates announced in Theorems A and B automatically follow from Theorem 6.2 combined with Musat's interpolation. Although this approach might look much simpler, the proof of the necessary interpolation results from [41] and of the noncommutative John-Nirenberg theorem (used in Mei's argument) are also quite technical.

Remark 6.4. It also follows from Theorem 6.2 that the problem posed in Remark 5.5 is only interesting for weak type inequalities. Indeed, if we are given a kernel with no size condition and only satisfying the Hörmander smoothness condition in the second variable, then we may obtain the strong L_p estimates provided by Theorems A and B for $1 < p \le 2$. We just need to apply Mei's argument for Theorem 6.2 (which works under these weaker assumptions) to the adjoint mapping and dualize backwards. A similar argument holds for Hörmander smooth kernels in the first variable and $2 \le p < \infty$.

6.3. **Proof of Theorem B.** Before proceeding with the argument, we set some preliminary results. According to Theorem 6.2 and the symmetry of the conditions on the kernel, we know that $T^*\mathbf{1}_\mathcal{A}$ belongs to $\mathrm{BMO}_\mathcal{A}$. In the following result, we shall write $\mathcal{H}_1$ for the Hardy space associated to the dyadic filtration on $\mathbb{R}^n$. That is, the predual of dyadic BMO, see [17].

Lemma 6.5. *If T is as above and $T^*\mathbf{1}_\mathcal{A} = 0$, then*

$$\int_{\mathbb{R}^n} Tf(x)\,dx = 0 \quad \textit{for any} \quad f \in \mathcal{H}_1.$$

Proof. Since $T^*\mathbf{1}_\mathcal{A} = 0$ vanishes as an element of $\mathrm{BMO}_\mathcal{A}$, we will have

$$\tau\Big(\int_{\mathbb{R}^n} T\phi(x)\,dx\Big) = \big\langle T\phi, \mathbf{1}_\mathcal{A}\big\rangle = \big\langle \phi, T^*\mathbf{1}_\mathcal{A}\big\rangle = 0 \tag{6.4}$$

for any $\phi \in \mathcal{H}_1(\mathcal{A})$, the Hardy space associated to the dyadic filtration $(\mathcal{A}_k)_{k\in\mathbb{Z}}$, see [47] for details and for the noncommutative analogue of Fefferman's duality theorem $\mathcal{H}_1(\mathcal{A})^* = \mathrm{BMO}_\mathcal{A}$. Given any projection $q \in \mathcal{M}_\pi$ of finite trace and $f \in \mathcal{H}_1$, it is clear that $\phi = fq \in \mathcal{H}_1(\mathcal{A})$. In particular, using $\mathcal{M}$-modularity again

$$\tau\Big(\int_{\mathbb{R}^n} T\phi(x)\,dx\Big) = \tau\Big(q\int_{\mathbb{R}^n} Tf(x)\,dx\Big) = 0$$

for any such projection q. Clearly, this immediately implies the assertion. □

Lemma 6.6. *Let T be as above for $q = 2$ and $L_2(\mathcal{A})$-normalized. Then, given $x_0 \in \mathbb{R}^n$ and $r_1, r_2 > 0$ with $r_2 > 2r_1$, the following estimate holds for any pair f, g*

of bounded scalar-valued functions respectively supported by $\mathsf{B}_{r_1}(x_0)$ *and* $\mathsf{B}_{r_2}(x_0)$

$$\Big\| \int_{\mathbb{R}^n} Tf(x)\, g(x)\, dx \Big\|_{\mathcal{M}} \le \mathrm{c}_n\, r_1^n \, \log(r_2/r_1)\, \|f\|_\infty \|g\|_\infty.$$

Proof. We proceed as in the proof of the localization estimate given in Paragraph 2.1. Let B denote the ball $\mathsf{B}_{3r_1/2}(x_0)$ and consider a smooth function ρ identically 1 on B and 0 outside $\mathsf{B}_{2r_1}(x_0)$. Taking $\eta = 1-\rho$, we may decompose

$$\Big\| \int_{\mathbb{R}^n} Tf(x)\, g(x)\, dx \Big\|_{\mathcal{M}} = \Big\| \int_{\mathbb{R}^n} Tf(x)\, \rho g(x)\, dx \Big\|_{\mathcal{M}} + \Big\| \int_{\mathbb{R}^n} Tf(x)\, \eta g(x)\, dx \Big\|_{\mathcal{M}}.$$

For the first term we adapt the commutative argument using the convexity of the function $a \mapsto |a|^2$. Indeed, if $\mathcal{M}$ embeds isometrically in $\mathcal{B}(\mathcal{H})$, it suffices to see that $a \mapsto \langle a^* a h, h\rangle_{\mathcal{H}}$ is a convex function for any $h \in \mathcal{H}$. However, this follows from the identity $\langle a^* a h, h\rangle_{\mathcal{H}} = \|ah\|_{\mathcal{H}}^2$. As an immediate consequence of this, we find the inequality

$$\Big| \frac{1}{|\mathsf{B}_{2r_1}(x_0)|} \int_{\mathsf{B}_{2r_1}(x_0)} Tf(x)\, \rho g(x)\, dx \Big|^2 \le \frac{1}{|\mathsf{B}_{2r_1}(x_0)|} \int_{\mathsf{B}_{2r_1}(x_0)} \big| Tf(x)\, \rho g(x) \big|^2 \, dx.$$

This combined with $\mathcal{M}$-modularity gives

$$\begin{aligned}
\Big\| \int_{\mathbb{R}^n} & Tf(x)\, \rho g(x)\, dx \Big\|_{\mathcal{M}} \\
&= |\mathsf{B}_{2r_1}(x_0)| \Big\| \frac{1}{|\mathsf{B}_{2r_1}(x_0)|} \int_{\mathsf{B}_{2r_1}(x_0)} Tf(x)\, \rho g(x)\, dx \Big\|_{\mathcal{M}} \\
&\le |\mathsf{B}_{2r_1}(x_0)| \Big\| \frac{1}{|\mathsf{B}_{2r_1}(x_0)|} \int_{\mathsf{B}_{2r_1}(x_0)} \big| Tf(x)\, \rho g(x) \big|^2 \, dx \Big\|_{\mathcal{M}}^{\frac12} \\
&= \mathrm{c}_n\, r_1^{n/2} \sup_{\|a\|_{L_2(\mathcal{M})} \le 1} \Big(\int_{\mathbb{R}^n} \tau \Big[1_{\mathsf{B}_{2r_1}(x_0)}(x) \big| Tf(x)\, \rho g(x) a \big|^2 \Big] \, dx \Big)^{\frac12} \\
&\le \mathrm{c}_n\, r_1^{n/2} \sup_{\|a\|_{L_2(\mathcal{M})} \le 1} \Big(\int_{\mathbb{R}^n} \tau \Big[\big| T(fa)(x) \big|^2 \Big] \, dx \Big)^{\frac12} \|g\|_\infty \\
&\le \mathrm{c}_n\, r_1^{n/2} \sup_{\|a\|_{L_2(\mathcal{M})} \le 1} \|f\|_2 \, \|a\|_2 \, \|g\|_\infty \ \le \ \mathrm{c}_n\, r_1^n \, \|f\|_\infty \, \|g\|_\infty,
\end{aligned}$$

since $\mathrm{supp} f \subset \mathsf{B}_{r_1}(x_0)$. On the other hand, the second term equals

$$\Big\| \int_{\mathbb{R}^n} Tf(x)\, \eta g(x)\, dx \Big\|_{\mathcal{M}} = \Big\| \int_{\mathsf{B}_{r_2}(x_0) \setminus \mathsf{B}} \Big(\int_{\mathsf{B}_{r_1}(x_0)} k(x,y)\, f(y)\, dy \Big)\, \eta g(x)\, dx \Big\|_{\mathcal{M}}.$$

This term is estimated exactly in the same way as in Paragraph 2.1. □

Sketch of the proof of Theorem B. As in the proof of Theorem A, we first observe that there is no restriction by assuming that $q = 2$. Indeed, according to Theorem 6.2 and Remark 6.3, it is easily seen that boundedness on $L_q(\mathcal{A})$ is equivalent to boundedness on $L_2(\mathcal{A})$. Moreover, we may assume that $f \in \mathcal{A}_{c,+}$ and decompose it for fixed $\lambda \in \mathbb{R}_+$ applying the noncommutative Calderón-Zygmund decomposition. This gives rise to $f = g + b$. The diagonal parts are estimated in the same way. Indeed, since we have $\|g_d\|_2^2 \le 2^n \lambda \|f\|_1$, the L_2-boundedness of T suffices for the good part. On the other hand, we use Lemma 4.2 for the bad part b_d in the usual way. This reduces the problem to estimate $\zeta T(b_d) \zeta$. By

$\mathcal{M}$-bimodularity, we can proceed verbatim with the argument given for this term in the proof of Theorem A. Moreover, exactly the same reasoning leads to control the off-diagonal part b_{off}. It remains to estimate the term associated to g_{off}. By Lemma 4.2 one more time, it suffices to study the quantity

$$\lambda \varphi \Big\{ \big| \zeta T(g_{\mathit{off}}) \zeta \big| > \lambda \Big\}.$$

As in the proof of Theorem A, we write $g_{\mathit{off}} = \sum_{k,s} g_{k,s}$ as a sum of martingale differences and use pseudo-localization. To justify our use of pseudo-localization we follow the argument in Theorem 5.2 using $\mathcal{M}$-bimodularity. This reduces the problem to study the validity of the *paraproduct argument* and of the *shifted form of the* $T1$ *theorem* for our new class of Calderón-Zygmund operators.

The paraproduct argument is simple. Indeed, since T is $\mathcal{M}$-bimodular, the same holds for T^* so that $T^* \mathbf{1}_\mathcal{A}$ becomes an element of $\mathrm{BMO}_{\mathcal{Z}_\mathcal{A}}$ where $\mathcal{Z}_\mathcal{A}$ denotes the center of $\mathcal{A}$. According to [38], the dyadic paraproduct Π_ξ associated to the term $\xi = T^* \mathbf{1}_\mathcal{A}$ defines a bounded map on $L_2(\mathcal{A})$. Moreover, since it is clear that Π_ξ is $\mathcal{M}$-bimodular, this allows us to consider the usual decomposition $T = T_0 + \Pi_\xi^*$. Now following the argument in Paragraph 2.3, with the characteristic functions $1_{\mathbb{R}^n \setminus \Sigma_{f,s}}$ and 1_{Ω_k} replaced by the corresponding projections provided by Theorem 5.2, we see that the estimate of the paraproduct also reduces here to the shifted $T1$ theorem. At this point we make crucial use of the fact that $\xi = T^* \mathbf{1}_\mathcal{A}$ is commuting, so that the same holds for $\Delta_j(\xi)$ for all $j \in \mathbb{Z}$.

Let us now sketch the main (slight) differences that appear when reproving the shifted $T1$ theorem for operator-valued kernels. Lemma 6.5 will play the role of the cancellation condition (2.1). On the other hand, we also have at our disposal the three auxiliary results (suitably modified) in Paragraph 2.1. Namely, Cotlar lemma as it was stated there will be used below with the only difference that we apply it over the Hilbert space $\mathcal{H} = L_2(\mathcal{A})$ instead of the classical L_2. Regarding Schur lemma, it is evident how to adapt it to the present setting. We just need to replace the Schur integrals by

$$\mathcal{S}_1(x) = \int_{\mathbb{R}^n} \|k(x,y)\|_\mathcal{M} \, dy \quad \text{and} \quad \mathcal{S}_2(y) = \int_{\mathbb{R}^n} \|k(x,y)\|_\mathcal{M} \, dx.$$

We leave the reader to complete the straightforward modifications in the original argument. Finally, Lemma 6.6 given above is the counterpart in our context of the localization estimate that we use several times in the proof of Theorem A. Once these tools are settled, the proof follows verbatim just replacing the absolute value $|\cdot|$ by the norm $\|\cdot\|_\mathcal{M}$ when corresponds. Maybe it is also worthy of mention that the two instances in the proof of the shifted $T1$ theorem where the Lebesgue differentiation theorem is mentioned, we should apply its noncommutative analog from [36]. This completes the proof. $\square$

Remark 6.7. The following question related to Theorem B was communicated to me by Tao Mei. It is clear that when the kernel takes values in the center $\mathcal{Z}_\mathcal{M}$ of $\mathcal{M}$, the corresponding Calderón-Zygmund operator is $\mathcal{M}$-bimodular and Theorem B applies. Assume now that the kernel k takes values in the commutant $\mathcal{M}'$. This gives rise to an $\mathcal{M}$-bimodular Calderón-Zygmund operator

$$T : L_2(\mathcal{A}) \to L_2\big(\mathbb{R}^n; L_2(\mathcal{B}(\mathcal{H}))\big)$$

where both $\mathcal{M}$ and $\mathcal{M}'$ embed in $\mathcal{B}(\mathcal{H})$. Assume further that such operator is bounded on $L_2(\mathcal{A})$. The question is whether our arguments in this paper can be suitably modified to produce the corresponding weak type inequality. We observe that some difficulties appear in Mei's argument for Theorem 6.2 and also in Lemmas 6.5 and 6.6. Note however that a positive answer to this question would produce, by real interpolation and duality, new strong L_p inequalities for a much wider class of operators.

Remark 6.8. After Theorem B, it is also natural to wonder about a vector-valued noncommutative Calderón-Zygmund theory. Let us be more precise, if the von Neumann algebra $\mathcal{M}$ is hyperfinite, Pisier's theory [45] allows us to consider the spaces $L_p(\mathcal{A}; \mathrm{X})$ with values in the operator space X. Here it is important to recall that we must impose on X an operator space structure since a Banach space structure is not rich enough. Then, we can consider vector-valued noncommutative singular integrals and study for which operator spaces we obtain weak type $(1,1)$ and/or strong type (p,p) inequalities. Of course, this is closely related to the geometry of the operator space in question and in particular to the notion of UMD_p operator spaces, also defined by Pisier. In this context a great variety of problems come into scene, like the independence of the UMD_p condition with respect to p (see [42] for some advances) or the operator space analog of Burkholder's geometric characterization of the UMD property in terms of ζ-convexity [5].

Remark 6.9. Another problem is the existence of $T1$ type theorems. This follows however from results by Hytönen [21] and Hytönen/Weis [23]. Namely, given a pair of Banach spaces (X, Y), they consider X-valued functions and $\mathcal{B}(\mathrm{X}, \mathrm{Y})$-valued kernels. In this general context, they need to impose $\mathcal{R}$-boundedness conditions on the kernel. However, in our setting $\mathrm{X} = L_2(\mathcal{M}) = \mathrm{Y}$ and $\mathcal{R}$-boundedness coincides with classical boundedness. Moreover, since our operators act by left or right multiplication, their norms in $\mathcal{B}(L_2(\mathcal{M}))$ coincide with the norm of the corresponding multiplier $k(x,y)$ in $\mathcal{M}$. Therefore, up to some extra conditions imposed in [21, 23], their results are applicable here.

Appendix A. On pseudo-localization

A.1. **Applicability.** We begin by analyzing how the pseudo-localization principle is applied to a given L_2-function. At first sight, it is only applicable to functions f in L_2 satisfying that $\mathsf{E}_m(f) = f_m = 0$ for some integer m. Indeed, according to the statement of the pseudo-localization principle we have

$$\operatorname{supp} f \subset \bigcup_{k \in \mathbb{Z}} \operatorname{supp} df_{k+s} \subset \bigcup_{k \in \mathbb{Z}} \Omega_k.$$

Given $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ with $\varepsilon_j = \pm 1$ for $1 \le j \le n$, let

$$\mathbb{R}^n_{(\varepsilon)} = \Big\{ x \in \mathbb{R}^n \ \big| \ \operatorname{sgn} x_j = \varepsilon_j \text{ for } 1 \le j \le n \Big\}$$

be the n-dimensional quadrant associated to ε and define $f_{(\varepsilon)}$ to be the restriction of f to such quadrant. If $f_m \neq 0$ for all $m \in \mathbb{Z}$, the same will happen to $f_{(\varepsilon)}$ for some index ε. Assume (with no loss of generality) that $\varepsilon = (1, 1, \ldots, 1)$ or, in other words, that f itself is supported by the first quadrant. Let Λ_f be the set of negative k's satisfying

$$\operatorname{supp} df_{k+s} \neq \emptyset.$$

Our hypothesis $f_m \neq 0$ for all $m \in \mathbb{Z}$ implies that Λ_f has infinitely many elements. According to the shift condition, we know that $\Omega_k \neq \emptyset$ for each $k \in \Lambda_f$ and therefore contains at least a cube in $\mathcal{Q}_k$, since Ω_k is an $\mathcal{R}_k$-set. In fact, for k small enough the $\mathcal{Q}_k$-cube in the first quadrant closest to the origin will be large enough to intersect the support of f. A moment of thought gives rise to the conclusion that Ω_k contains such cube for infinitely many negative k's and

$$\Sigma_{f,s} = \bigcup_{k \in \mathbb{Z}} 9\,\Omega_k = \mathbb{R}^n.$$

Therefore, our result does not provide any information in this case.

It is convenient to explain how to apply our result for an arbitrary function f in L_2 not satisfying the condition $f_m = 0$. By homogeneity, we may assume that $\|f\|_2 = 1$. On the other hand, if $\operatorname{supp} f$ is not compact we approximate f by a compactly supported function f_0 such that $\|f - f_0\|_2 \leq \mathrm{c}_{n,\gamma}\, s\, 2^{-\gamma s/4}$. This clearly reduces our problem to find the set $\Sigma_{f,s}$ around the support of f_0. Next we decompose $f_0 = \sum_{1 \leq j \leq 2^n} f_j$, with f_j being the restriction of f_0 to the j-th quadrant and work independently with each of these functions. In other words our localization problem reduces to study functions f in L_2 with compact support contained in the first n-dimensional quadrant. Let f be such a function and take Q to be the smallest dyadic cube containing the support of f. We have $Q \in \mathcal{Q}_m$ for some integer m. Then we find $f_m = \lambda 1_Q$ with $\lambda = \frac{1}{|Q|} \int_{\mathbb{R}^n} f(x)\, dx$ and thus we decompose

$$f = \big(f - \lambda 2^{-\gamma s/2} 1_{Q_s} \big) + \lambda 2^{-\gamma s/2} 1_{Q_s} = f^1 + f^2$$

where Q_s is a cube satisfying:

- Q_s contains Q.
- Q_s is contained in a dyadic antecessor of Q.
- The Lebesgue measure of Q_s is $|Q_s| = 2^{\gamma s/2} |Q|$.

It is clear that we have

$$\Big(\int_{\mathbb{R}^n} |f^2(x)|^2 dx \Big)^{\frac{1}{2}} = \lambda 2^{-\gamma s/2} 2^{\gamma s/4} \sqrt{|Q|} = 2^{-\gamma s/4} \|f_m\|_2 \leq 2^{-\gamma s/4}.$$

Therefore, f^2 is small enough for our aims. On the other hand, let $\widehat{Q}_s$ be the dyadic Q-antecessor of generation $m - j_0$ with j_0 being the smallest positive integer such that $j_0 \geq \gamma s/2n$. In other words, this cube is the smallest dyadic Q-antecessor containing Q_s. If we set $m_0 = m - j_0$, we clearly have $f^1_{m_0} = 0$. When $k \leq m_0 - s$ we have $df^1_{k+s} = 0$ and $\operatorname{supp} df^1_{k+s} = \emptyset$ so that there is no set to control. When $k + s > m_0$ we use

$$\operatorname{supp} df^1_{k+s} \subset \widehat{Q}_s.$$

Hence, we may choose Ω_k to be the smallest $\mathcal{R}_k$-set containing $\widehat{Q}_s$. In the worst case $k = m_0 - s + 1$ we are forced to take Ω_k as the $(s-1)$-th dyadic antecessor of $\widehat{Q}_s$. That is, the $(j_0 + s - 1)$-dyadic antecessor $\widehat{Q}(j_0 + s - 1)$ of Q. This gives rise to the set

$$\Sigma_{f,s} = \bigcup_{k \in \mathbb{Z}} 9\,\Omega_k = 9\widehat{Q}(j_0 + s - 1) \subset 9 \cdot 2^{j_0 + s}\, Q \sim 9 \cdot 2^{(1 + \frac{\gamma}{2n})s} \operatorname{supp} f$$

and completes the argument for arbitrary L_2 functions. To conclude, we should mention that the dependance on the n-dimensional quadrants, due to the geometry imposed by the standard dyadic filtration, is fictitious. Indeed, we can always translate the dyadic filtration, so that the role of the origin is played by another point which leaves the support of f in the *new* first quadrant.

Remark A.1. Given a function f in L_2 and a parameter $\delta \in \mathbb{R}_+$, we have analyzed so far how to find appropriate sets $\Sigma_{f,\delta}$ satisfying the localization estimate which motivated our pseudo-localization principle

$$\Big(\int_{\mathbb{R}^n \setminus \Sigma_{f,\delta}} |Tf(x)|^2 \, dx \Big)^{\frac12} \le \delta \Big(\int_{\mathbb{R}^n} |f(x)|^2 \, dx \Big)^{\frac12}.$$

Reciprocally, given a set Σ in $\mathbb{R}^n$ and $\delta \in \mathbb{R}_+$, it is quite simple to find functions $f_{\Sigma,\delta}$ satisfying such estimate on $\mathbb{R}^n \setminus \Sigma$. Indeed, let $s \ge 1$ be the smallest possible integer satisfying $\mathrm{c}_{n,\gamma} s 2^{-\gamma s/4} \le \delta$ and write $\Sigma = \bigcup_{k \in \mathbb{Z}} 9\Omega_k$ as a disjoint union of 9-dilations of maximal $\mathcal{R}_k$-sets. In this case, any function of the form

$$f_{\Sigma,\delta} = \sum_{k \in \mathbb{Z}} 1_{\Omega_k} dg_{k+s}$$

with $g \in L_2$ satisfies the hypotheses of our pseudo-localization principle with Σ as the final localization set. Indeed, we have $d(f_{\Sigma,\delta})_{k+s} = 1_{\Omega_k} dg_{k+s}$ because 1_{Ω_k} is $(k+s)$-predictable and we deduce that $f_{\Sigma,\delta}$ satisfies the shift condition.

A.2. **Decreasing rate of singular integrals in the L_2 metric.** As an immediate consequence of the pseudo-localization principle, we can give a *lower* estimate of how fast decreases a singular integral far away from a set Σ_f associated to f. To be more specific, the following result holds.

Corollary A.2. *Let f be in L_2 and define*

$$\Sigma_f = \bigcup_{k \in \mathbb{Z}} 9\Gamma_k \qquad \textit{with} \qquad \Gamma_k = \operatorname{supp} df_k \in \mathcal{R}_k.$$

Then, the following holds for any $\xi > 4$

$$\Big(\int_{\mathbb{R}^n \setminus \xi\Sigma_f} |Tf(x)|^2 \, dx \Big)^{\frac12} \ \le \ \mathrm{c}_{n,\gamma} \, \xi^{-\gamma/4} \log \xi \, \Big(\int_{\mathbb{R}^n} |f(x)|^2 dx \Big)^{\frac12}$$

and any L_2-normalized Calderón-Zygmund operator with Lipschitz parameter γ.

Proof. Let Ω_k be the smallest $\mathcal{R}_k$-set containing Γ_{k+s}. In the worst case, Γ_{k+s} can be written as a union $\bigcup_\alpha Q_\alpha$ of $\mathcal{Q}_{k+s}$-cubes. Taking $\widehat{Q}_\alpha(s)$ to be the s-th dyadic antecessor of Q_α, we observe that

$$\Omega_k \subset \bigcup_\alpha \widehat{Q}_\alpha(s) \subset 2^{s+1} \Gamma_{k+s}.$$

Then we construct

$$\Sigma_{f,s} = \bigcup_{k \in \mathbb{Z}} 9\,\Omega_k \subset 2^{s+1} \Sigma_f,$$

and the theorem above automatically gives

$$\Big(\int_{\mathbb{R}^n \setminus 2^{s+1}\Sigma_f} |Tf(x)|^2 \, dx \Big)^{\frac12} \ \le \ \mathrm{c}_{n,\gamma} \, s \, 2^{-\gamma s/4} \Big(\int_{\mathbb{R}^n} |f(x)|^2 dx \Big)^{\frac12}.$$

Since this holds for every positive integer s, the assertion follows. □

Remark A.3. All the considerations in Paragraph A.1 apply to this result.

Remark A.4. This result might be quite far from being optimal, see below.

A.3. **Atomic pseudo-localization in L_1.** Maybe the oldest localization result was already implicit in the Calderón-Zygmund decomposition. Indeed, let b denote the bad part of f associated to a fixed $\lambda > 0$ and let Σ_λ be the level set where the dyadic Hardy-Littlewood maximal function $M_d f$ is bigger than λ. Note that b is supported by Σ_λ. Then, we have

$$\int_{\mathbb{R}^n \setminus 2\Sigma_\lambda} |Tb(x)| \, dx \le \mathrm{c}_n \sum\nolimits_j \|b_j\|_1 \le \mathrm{c}_n \|f\|_1,$$

where the b_j's are the atoms in which we decompose b. In fact, this reduces to a well-known localization result for dyadic atoms in L_1. Namely, let a denote an atom supported by a dyadic cube Q_a. Then, the mean-zero of a gives the following estimate for any $\xi > 2$

$$\begin{aligned} \text{(A.1)} \qquad \int_{\mathbb{R}^n \setminus \xi Q_a} |Ta(x)| \, dx &= \int_{\mathbb{R}^n \setminus \xi Q_a} \Big| \int_{\mathbb{R}^n} \big[k(x,y) - k(x, c_{Q_a}) \big] \, a(y) \, dy \Big| \, dx \\ &\le \int_{\mathbb{R}^n \setminus \xi Q_a} \int_{\mathbb{R}^n} \frac{|y - c_{Q_a}|^\gamma}{|x-y|^{n+\gamma}} \, |a(y)| \, dy \, dx \le \mathrm{c}_n \, \xi^{-\gamma} \, \|a\|_1. \end{aligned}$$

Note that the only condition on T that we use is the γ-Lipschitz smoothness on the second variable, not even an a priori boundedness condition. Under these mild assumptions, we may generalize (A.1) in the language of our pseudo-localization principle. Namely, the following result (maybe known to experts) holds.

Theorem A.5. *Let us fix a positive integer s. Given a function f in L_1 and any integer k, we define Ω_k to be the smallest $\mathcal{R}_k$-set containing the support of df_{k+s} and consider the set*

$$\Sigma_{f,s} = \bigcup_{k \in \mathbb{Z}} 3\Omega_k.$$

Then, we have for any Calderón-Zygmund operator as above

$$\int_{\mathbb{R}^n \setminus \Sigma_{f,s}} |Tf(x)| \, dx \ \le \ \mathrm{c}_n \, 2^{-\gamma s} \int_{\mathbb{R}^n} |f(x)| dx.$$

Proof. We may clearly assume that $f_m = 0$ for some integer m. Namely, otherwise we can argue as in the previous paragraph to deduce that $\Sigma_{f,s} = \mathbb{R}^n$ and the assertion is vacuous. Define inductively

$$\begin{aligned} \mathsf{A}_1 &= \operatorname{supp} df_{m+1}, \\ \mathsf{A}_j &= \operatorname{supp} df_{m+j} \setminus \Big(\bigcup_{w<j} \mathsf{A}_w \Big). \end{aligned}$$

Use that $\operatorname{supp} f \subset \bigcup_j \mathsf{A}_j$ and pairwise disjointness of A_j's to obtain

$$\begin{aligned} \int_{\mathbb{R}^n \setminus \Sigma_{f,s}} |Tf(x)| \, dx &\le \sum\nolimits_j \sum_{\substack{Q \in \mathcal{Q}_{m+j} \\ Q \subset \mathsf{A}_j}} \int_{\mathbb{R}^n \setminus \Sigma_{f,s}} |T(f 1_Q)(x)| \, dx \\ &= \sum\nolimits_j \sum_{\substack{Q \in \mathcal{Q}_{m+j} \\ Q \subset \mathsf{A}_j}} \int_{\mathbb{R}^n \setminus \Sigma_{f,s}} \Big| T \Big(1_Q \sum_{k=m+j}^{\infty} df_k \Big)(x) \Big| \, dx. \end{aligned}$$

Let $\widehat{Q}_s$ be the s-th dyadic antecessor of Q. Since

$$Q \subset \mathsf{A}_j \subset \mathrm{supp}\, df_{m+j} \subset \Omega_{m+j-s}$$

and $Q \in \mathcal{Q}_{m+j}$, we deduce $2^s Q \subset 3\widehat{Q}_s \subset \Sigma_{f,s} \Rightarrow \mathbb{R}^n \setminus \Sigma_{f,s} \subset \mathbb{R}^n \setminus 2^s Q$ and

$$\int_{\mathbb{R}^n \setminus \Sigma_{f,s}} |Tf(x)|\, dx \ \le \ \sum\nolimits_j \sum_{\substack{Q \in \mathcal{Q}_{m+j} \\ Q \subset \mathsf{A}_j}} \int_{\mathbb{R}^n \setminus 2^s Q} \Big| T\Big(1_Q \sum_{k=m+j}^{\infty} df_k\Big)(x)\Big|\, dx.$$

On the other hand, by (A.1)

$$\begin{aligned} \int_{\mathbb{R}^n \setminus \Sigma_{f,s}} |Tf(x)|\, dx & \le \ \mathrm{c}_n 2^{-\gamma s} \sum\nolimits_j \sum_{\substack{Q \in \mathcal{Q}_{m+j} \\ Q \subset \mathsf{A}_j}} \Big\| 1_Q \sum_{k=m+j}^{\infty} df_k \Big\|_1 \\ & = \ \mathrm{c}_n 2^{-\gamma s} \sum\nolimits_j \Big\| 1_{\mathsf{A}_j} \sum_{k=m+j}^{\infty} df_k \Big\|_1 \ = \ \mathrm{c}_n 2^{-\gamma s} \sum\nolimits_j \|1_{\mathsf{A}_j} f\|_1. \end{aligned}$$

Using once more the pairwise disjointness of the A_j's we deduce the assertion. □

Remark A.6. Here we should notice that the condition $f_m = 0$ can not be removed as we did in the L_2 case and its applicability is limited to this atomic setting. On the other hand, if we try to use the argument of Theorem A.5 for $p = 2$, we will find a nice illustration of why the ideas around almost orthogonality that we have used in the paper come into play. In the L_1 framework, almost orthogonality is replaced by the triangle inequality.

A.4. **Other forms of pseudo-localization.** Once we have obtained results in L_1 and L_2, it is quite natural to wonder about L_p pseudo-localization for other values of p. If we only deal with atoms, it easily seen that (A.1) generalizes to any $p > 1$ in the following way

$$\text{(A.2)} \qquad \Big(\int_{\mathbb{R}^n \setminus \xi Q_a} |Ta(x)|^p\, dx \Big)^{\frac{1}{p}} \ \le \ \mathrm{c}_n\, \xi^{-(\gamma + n/p')} \Big(\int_{\mathbb{R}^n} |a(x)|^p\, dx \Big)^{\frac{1}{p}}.$$

This gives rise to two interesting problems:

i) In Theorem A.5 we showed that (A.1) generalizes to more general functions in L_1, those satisfying $f_m = 0$ for some integer m. On the other hand, as we have seen in Paragraph A.1, the condition $f_m = 0$ is not a serious restriction for $p = 2$, or any $p > 1$. Therefore, inequality (A.2) suggests that our pseudo-localization principle might hold with $s 2^{-\gamma s/4}$ replaced by the better constant $2^{-(\gamma + n/2)s}$. However, this result and its natural L_p generalization are out of the scope of this paper.

ii) Although the constant that we have obtained in our pseudo-localization principle on L_2 might be far from being optimal, it still makes a lot of sense to wonder whether or not the corresponding interpolated inequality holds for $1 < p < 2$. Below we give some guidelines which might lead to such a result. We have not checked details, since the necessary estimates might be quite technical, as those in the proof for $p = 2$. All our ideas below can be thought as problems for the interested reader.

The *interpolated inequality* that comes to mind is

$$\Big(\int_{\mathbb{R}^n \setminus \Sigma_{f,s}} |Tf(x)|^p \, dx \Big)^{\frac{1}{p}} \le \mathrm{c}_{n,\gamma} \frac{s \, 2^{-\gamma s/4}}{(s 2^{3\gamma s/4})^{\frac{2}{p}-1}} \Big(\int_{\mathbb{R}^n} |f(x)|^p \, dx \Big)^{\frac{1}{p}}.$$

However, by the presence of $\Sigma_{f,s}$, a direct interpolation argument does not apply and we need a more elaborated approach. Namely, following the proof of our result in L_2 verbatim, it suffices to find suitable upper bounds for Φ_s and Ψ_s in $\mathcal{B}(L_p)$. Here we might use Rubio de Francia's idea of extrapolation and content ourselves with a rough estimate (i.e. independent of s) for the norm of these operators from L_1 to $L_{1,\infty}$. Of course, by real interpolation this would give rise to the weaker inequality

$$\text{(A.3)} \qquad \Big(\int_{\mathbb{R}^n \setminus \Sigma_{f,s}} |Tf(x)|^p \, dx \Big)^{\frac{1}{p}} \le \mathrm{c}_{n,\gamma} \frac{(s \, 2^{-\gamma s/4})^{2-\frac{2}{p}}}{p-1} \Big(\int_{\mathbb{R}^n} |f(x)|^p \, dx \Big)^{\frac{1}{p}}.$$

However, this would be good enough for many applications. The Calderón-Zygmund method will be applicable to both Φ_s and Ψ_s if we know that their kernels satisfy a suitable smoothness estimate. The lack of regularity of E_k and Δ_{k+s} appears again as the main difficulty to overcome. In this case, it is natural to wonder if the Hörmander condition

$$\int_{|x|>2|y|} |k(x,y) - k(x,0)| \, dx \le \mathrm{c}_{n,\gamma},$$

holds for the kernels of Φ_s and Ψ_s. We believe this should be true. Anyway, a more in depth application of Rubio's extrapolation method (which we have not pursued so far) might be quite interesting here.

Remark A.7. According to the classical theory [54], it is maybe more natural to replace (in the shifted form of the $T1$ theorem) the dyadic martingale differences Δ_{k+s} by a Littlewood-Paley decomposition and the conditional expectations E_k by their partial sums. This result will be surely easier to prove since there is no lack of regularity as in the dyadic martingale setting. This alternative approach to the shifted $T1$ theorem might give rise to some sort of pseudo-localization result in terms of Littlewood-Paley decompositions. Although this is not helpful in the noncommutative setting (by our dependance on Cuculescu's construction), it makes the problem on the smoothness of the kernels of Φ_s and Ψ_s more accessible.

Remark A.8. If the argument sketched above for inequality (A.3) works, another natural question is whether results for $p > 2$ can be deduced by duality. On one hand, the operator Φ_s behaves well with respect to duality. In fact, the analysis of $\sum_k \Delta_{k+s} T \mathsf{E}_k$ just requires (in analogy with the $T1$ theorem) to assume first that we have $T1 = 0$. As pointed above, this kind of cancellation conditions are only necessary for Φ_s, since the presence of the terms $id - \mathsf{E}_k$ in Ψ_s produce suitable cancellations. However, this is exactly why the adjoint

$$\Psi_s^* = \sum\nolimits_k \Delta_{k+s} T^*_{4 \cdot 2^{-k}} (id - \mathsf{E}_k)$$

does not behave as expected. This leaves open the problem for $p > 2$.

Appendix B. On Calderón-Zygmund decomposition

B.1. **Weighted inequalities.** Given a positive function f in L_1 and $\lambda \in \mathbb{R}_+$, let us consider the Calderón-Zygmund decomposition $f = g + b$ associated to λ. As pointed out and well-known, the most significant inequalities satisfied by these functions are

$$\int_{\mathbb{R}^n} |g(x)|^2 \, dx \leq 2^n \lambda \|f\|_1 \quad \text{and} \quad \sum_j \int_{\mathbb{R}^n} |b_j(x)| \, dx \leq 2 \|f\|_1,$$

where the b_j's are the atoms in which b is decomposed. We already saw in Section 4 that these inequalities remain true for the diagonal terms of the noncommutative Calderón-Zygmund decomposition. However, we do not have at our disposal (see Paragraph B.2 below) such inequalities for the off-diagonal terms. As we have explained in the Introduction, our way to solve this lack has been to prove the off-diagonal estimates

- $\big\| \zeta T \big(\sum_k b_{k,s} \big) \zeta \big\|_1 \lesssim \alpha_s \|f\|_1,$
- $\big\| \zeta T \big(\sum_k g_{k,s} \big) \zeta \big\|_2^2 \lesssim \beta_s \lambda \|f\|_1,$

for some fast decreasing sequences α_s, β_s. The proof of these estimates has exploited the properties of the projection ζ in conjunction with our localization results. We have therefore hidden the actual inequalities satisfied by the off-diagonal terms which are independent of the behavior of $\zeta T(\cdot) \zeta$. Namely, we have

a) Considering the *atoms*

$$b_{k,s} = p_k (f - f_{k+s}) p_{k+s} + p_{k+s} (f - f_{k+s}) p_k$$

in $b_{\mathit{off}} = \sum_{k,s} b_{k,s}$, we have for any positive sequence $(\alpha_s)_{s \geq 1}$

$$\sum_s \sum_k \alpha_s \|b_{k,s}\|_1 \ \lesssim \ \Big(\sum_s s \, \alpha_s \Big) \|f\|_1.$$

b) Considering the *martingale differences*

$$g_{k,s} = p_k df_{k+s} q_{k+s-1} + q_{k+s-1} df_{k+s} p_k$$

in $g_{\mathit{off}} = \sum_{k,s} g_{k,s}$, we have for any positive sequence $(\beta_s)_{s \geq 1}$

$$\Big\| \sum_s \sum_k \beta_s \, g_{k,s} \Big\|_2^2 = \sum_s \sum_k \beta_s^2 \, \|g_{k,s}\|_2^2 \ \lesssim \ \Big(\sum_s \beta_s^2 \Big) \lambda \, \|f\|_1.$$

As the careful reader might have noticed, the proof of these estimates is implicit in our proof of Theorem A. It is still to be determined whether these estimates for the weights α_s and β_s are sharp. On the other hand, it is also possible to study weighted L_p estimates for the off-diagonal terms of the good part and $p > 1$. We have not pursued any of these lines.

B.2. **On the lack of a classical L_2 estimate.** The pseudo-localization approach of this paper has been motivated by the lack of the key estimate $\|g\|_2^2 \lesssim \lambda \|f\|_1$ in the noncommutative setting. Although we have not disproved such inequality so far, we end this paper by giving some evidences that it must fail. Recalling from Section 4 that the diagonal terms of g satisfy the estimate $\|g_d\|_2^2 \lesssim \lambda \|f\|_1$, it suffices disprove the inequality

$$\|g_{\mathit{off}}\|_2^2 \lesssim \lambda \|f\|_1.$$

By the original expression for g_{off}, we have

$$g_{\mathit{off}} = \sum_i \sum_{j<i} p_i f_i p_j + \sum_j \sum_{i<j} p_i f_j p_j = \sum_k p_k f_k (\mathbf{1}_{\mathcal{A}} - q_{k-1}) + (\mathbf{1}_{\mathcal{A}} - q_{k-1}) f_k p_k.$$

By orthogonality of the p_k's and the tracial property, it is easily seen that

$$\begin{aligned} \frac{1}{\lambda} \|g_{\mathit{off}}\|_2^2 &= \frac{2}{\lambda} \varphi\Big(\sum\nolimits_k (\mathbf{1}_{\mathcal{A}} - q_{k-1}) f_k p_k f_k (\mathbf{1}_{\mathcal{A}} - q_{k-1}) \Big) \\ &= 2 \sum\nolimits_k \varphi\Big(\frac{f_k p_k f_k}{\lambda} \Big) + 2 \sum\nolimits_k \varphi\Big(\frac{q_{k-1} f_k p_k f_k q_{k-1}}{\lambda} \Big) = \mathsf{A} + \mathsf{B}. \end{aligned}$$

By the tracial property

$$\mathsf{B} = \frac{2}{\lambda} \sum_k \varphi(p_k f_k q_{k-1} f_k p_k).$$

Moreover, we also have

$$\|f_k^{\frac{1}{2}} q_{k-1} f_k^{\frac{1}{2}}\|_\infty = \|q_{k-1} f_k q_{k-1}\|_\infty \le 2^n \|q_{k-1} f_{k-1} q_{k-1}\|_\infty \le 2^n \lambda.$$

Thus, we find the inequality

$$\mathsf{B} \le \mathrm{c}_n \sum_k \varphi(p_k f_k) \le \mathrm{c}_n \|f\|_1$$

and our problem reduces to disprove

$$\Big\| \sum\nolimits_k f_k p_k \Big\|_2^2 \lesssim \lambda \|f\|_1. \tag{B.1}$$

As in the argument given in Section 6 to find a bad-behaved noncommuting kernel, our motivation comes from a matrix construction. Namely, let $\mathcal{A}$ be the algebra of $2m \times 2m$ matrices equipped with the standard trace tr and consider the filtration $\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_{2m}$, where $\mathcal{A}_s$ denotes the subalgebra spanned by the matrix units e_{ij} with $1 \le i, j \le s$ and the matrix units e_{kk} with $k > s$. Let us set $\lambda = 1$ and define

$$f = \sum_{i,j=1}^{2m} e_{ij}.$$

It is easily checked that

$$\begin{aligned} q_1 &= \chi_{(0,1]}(f_1) = \mathbf{1}_{\mathcal{A}}, \\ q_2 &= \chi_{(0,1]}(q_1 f_2 q_1) = \sum\nolimits_{k>2} e_{kk}, \\ q_3 &= \chi_{(0,1]}(q_2 f_3 q_2) = \chi_{(0,1]}(q_2) = q_2, \\ q_4 &= \chi_{(0,1]}(q_3 f_4 q_3) = \sum\nolimits_{k>4} e_{kk}, \\ q_5 &= \chi_{(0,1]}(q_4 f_5 q_4) = \chi_{(0,1]}(q_4) = q_4, \\ q_6 &= \ldots \end{aligned}$$

Hence, $p_{2k-1} = 0$ and $p_{2k} = e_{2k-1,2k-1} + e_{2k,2k}$. This gives

$$\sum_{k=1}^{2m} f_k p_k = \sum_{k=1}^{m} f_{2k} p_{2k} = \sum_{k=1}^{m} \sum_{j=1}^{2k} e_{j,2k-1} + e_{j,2k}.$$

We have $\lambda = 1$ and it is clear that $\|f\|_1 = 2m$, while the L_2 norm is

$$\Big\| \sum_{k=1}^{2m} f_k p_k \Big\|_2^2 = \sum_{k=1}^{m} 4k = 2m(m+1).$$

Therefore, if we let $m \to \infty$ we see that (B.1) fails in this particular setting.

Problem. Adapt the construction given above to the usual von Neumann algebra $\mathcal{A}$ of operator-valued functions $f : \mathbb{R}^n \to \mathcal{M}$, equipped with the standard dyadic filtration. This would disprove (B.1) and thereby the inequality $\|g\|_2^2 \lesssim \lambda \|f\|_1$ in the noncommutative setting.

References

1. W.B. Arveson, *Analyticity in operator algebras.* Amer. J. Math. **89** (1967), 578-642.
2. J. Bourgain, *Some remarks on Banach spaces in which martingale difference sequences are unconditional.* Ark. Mat. **21** (1983), 163-168.
3. J. Bourgain, *Extension of a result of Benedek, Calderón and Panzone.* Ark. Mat. **22** (1984), 91-95.
4. J. Bourgain, *Vector valued singular integrals and the H^1-BMO duality.* Probability Theory and Harmonic Analysis. (Eds. Chao and Woyczynski) Decker (1986), 1-19.
5. D.L. Burkholder, *A geometrical characterization of Banach spaces in which martingale difference sequences are unconditional.* Ann. Probab. **9** (1981), 997-1011.
6. D.L. Burkholder, *A geometric condition that implies the existence of certain singular integrals of Banach space valued functions.* Conference on Harmonic Analysis in Honor of Antoni Zygmund. Wadsworth (1983), 270-286.
7. M. Christ, *Weak type $(1,1)$ bounds for rough operators.* Ann. of Math. **128** (1988), 19-42.
8. R. Coifman and G. Weiss, Analyse harmonique non-commutative sur certains espaces homogènes. Lecture Notes in Math. **242**, 1971.
9. I. Cuculescu, *Martingales on von Neumann algebras.* J. Multivariate Anal. **1** (1971), 17-27.
10. G. David and J.L. Journé, *A boundedness criterion for generalized Calderón-Zygmund operators.* Ann. of Math. **120** (1984), 371-397.
11. P.G. Dodds, T.K. Dodds and B. de Pagter, *Noncommutative Banach function spaces.* Math. Z. **201** (1989), 583-597.
12. J. Duoandikoetxea, Fourier Analysis. Translated and revised from the 1995 Spanish original by D. Cruz-Uribe. Grad. Stud. Math. **29**. American Mathematical Society, 2001.
13. J. Duoandikoetxea and J.L. Rubio de Francia, *Maximal and singular integral operators via Fourier transform estimates.* Invent. Math. **84** (1986), 541-561.
14. T. Fack and H. Kosaki, *Generalized s-numbers of τ-measurable operators.* Pacific J. Math. **123** (1986), 269-300.
15. T. Figiel, *Singular integral operators: a martingale approach.* Geometry of Banach Spaces. London Math. Soc. Lecture Notes Ser. **158**. (Ed. P.F.X. Müller and W. Schachermayer) Cambrideg Univ. Press (1989), 95-110.
16. J. Garnett and P. Jones, *BMO from dyadic BMO.* Pacific J. Math. **99** (1982), 351-371.
17. A.M. Garsia, Martingale inequalities: Seminar notes on recent progress. W. A. Benjamin, Inc., Reading, Mass.-London-Amsterdam, 1973, Mathematics Lecture Notes Series.
18. U. Haagerup, *Non-commutative integration theory.* Unpublished manuscript (1978). See also Haagerup's Lecture given at the Symposium in Pure Mathematics of the Amer. Math. Soc. Queens University, Kingston, Ontario, 1980.
19. U. Haagerup, *L_p spaces associated with an arbitrary von Neumann algebra.* Algèbres d'opérateurs et leurs applications en physique mathématique, CNRS (1979), 175-184.
20. A. Harcharras, *Fourier analysis, Schur multipliers on S^p and non-commutative $\Lambda(p)$-sets.* Studia Math. **137** (1999), 203-260.
21. T. Hytönen, *An operator-valued Tb theorem.* J. Funct. Anal. **234** (2006), 420-463.
22. T. Hytönen and L. Weis, *Singular convolution integrals with operator-valued kernels.* Math. Z. **255** (2007), 393-425.

23. T. Hytönen and L. Weis, *A $T1$ theorem for integral transformations with operator-valued kernel.* J. reine angew. Math. **599** (2006), 155-200.
24. M. Junge, *Doob's inequality for non-commutative martingales.* J. reine angew. Math. **549** (2002), 149-190.
25. M. Junge, *Embedding of the operator space* OH *and the logarithmic 'little Grothendieck inequality'.* Invent. Math. **161** (2005), 225-286.
26. M. Junge and M. Musat, *A non-commutative version of the John-Nirenberg theorem.* Trans. Amer. Math. Soc. **359** (2007), 115-142.
27. M. Junge and J. Parcet, Rosenthal's theorem for subspaces of noncommutatuve L_p. To appear in Duke Math. J.
28. M. Junge and J. Parcet, Operator space L_p embedding theory I and II. Preprint.
29. M. Junge, J. Parcet and Q. Xu, Rosenthal type inequalities for free chaos. Ann. Probab. **35** (2007), 1374-1437.
30. M. Junge and Q. Xu, *Noncommutative Burkholder/Rosenthal inequalities.* Ann. Probab. **31** (2003), 948-995.
31. M. Junge and Q. Xu, *Noncomutative maximal ergodic theorems.* J. Amer. Math. Soc. **20** (2007), 385-439.
32. R.V. Kadison and J.R. Ringrose, Fundamentals of the Theory of Operator Algebras I and II. Grad. Stud. Math. **15** and **16**. American Mathematical Society, 1997.
33. R.A. Macías and C. Segovia, *Lipschitz functions on spaces of homogeneous type.* Adv. in Math. **33** (1979), 257-270.
34. R.A. Macías and C. Segovia, *A decomposition into atoms of distributions on spaces of homogeneous type.* Adv. in Math. **33** (1979), 271-309.
35. T. Mei, BMO *is the intersection of two translates of dyadic* BMO. C.R. Math. Acad. Sci. Paris **336** (2003), 1003-1006.
36. T. Mei, *Operator valued Hardy spaces.* Mem. Amer. Math. Soc. **188** (2007).
37. T. Mei, *Notes on matrix valued paraproducts.* Indiana Univ. Math. J. **55** (2006), 747-760.
38. T. Mei. Personal communication.
39. Y. Meyer, Ondelettes et Opérateurs II: Opérateurs de Calderón-Zygmund. Hermann, 1990.
40. M. Mitrea, Clifford wavelets, singular integrals and Hardy spaces. Lecture Notes in Math. **1575**, 1994.
41. M. Musat, *Interpolation between non-commutative BMO and non-commutative Lp-spaces.* J. Functional Analysis **202** (2003), 195-225.
42. M. Musat, *On the operator space* UMD *property for noncommutative L_p-spaces.* Indiana Math. J. **55** (2006), 1857-1891.
43. F. Nazarov, G. Pisier, S. Treil and A. Volberg, *Sharp estimates in vector Carleson imbedding theorem and for vector paraproducts.* J. reine angew. Math. **542** (2002), 147-171.
44. J. Parcet and N. Randrianantoanina, *Gundy's decomposition for non-commutative martingales and applications.* Proc. London Math. Soc. **93** (2006), 227-252.
45. G. Pisier, Non-commutative vector valued L_p-spaces and completely p-summing maps. Astérisque (Soc. Math. France) **247**, 1998.
46. G. Pisier, Introduction to Operator Space Theory. Cambridge University Press, 2003.
47. G. Pisier and Q. Xu, *Non-commutative martingale inequalities.* Comm. Math. Phys. **189** (1997), 667-698.
48. G. Pisier and Q. Xu, *Non-commutative L_p-spaces.* Handbook of the Geometry of Banach Spaces II (Eds. W.B. Johnson and J. Lindenstrauss) North-Holland (2003), 1459-1517.
49. N. Randrianantoanina, *Hilbert transform associated with finite maximal subdiagonal algebras.* J. Austral. Math. Soc. Series A **65** (1998), 388-404.
50. N. Randrianantoanina, *Non-commutative martingale transforms*, J. Funct. Anal. **194** (2002), 181-212.
51. N. Randrianantoanina, *A weak type inequality for non-commutative martingales and applications.* Proc. London Math. Soc. **91** (2005), 509-544.
52. N. Randrianantoanina, *Conditioned square functions for non-commutative martingales.* Ann. Probab. **35** (2007), 1039-1070.
53. J.L. Rubio de Francia, F.J. Ruiz and J.L. Torrea, *Calderón-Zygmund theory for operator-valued kernels.* Adv. in Math. **62** (1986), 7-48.
54. E.M. Stein, Harmonic Analysis: Real Variable Methods, Orthogonality, and Oscillatory Integrals. Princeton Math. Ser. **43**. Princeton Univ. Press., NJ, 1993.

55. M. Takesaki, Theory of operator algebras I. Springer-Verlag, New York, 1979.
56. M. Terp, *L_p spaces associated with von Neumann algebras.* Math. Institute Copenhagen University, 1981.
57. A. Volberg, *Matrix A_p weights via S-functions.* J. Amer. Math. Soc. **10** (1997), 223-246.
58. Q. Xu, *Analytic functions with values in lattices and symmetric spaces of measurable operators.* Math. Proc. Cambridge Philos. Soc. **109** (1991), 541-563.

Javier Parcet
Departamento de Matemáticas
Instituto de Matemáticas y Física Fundamental
Consejo Superior de Investigaciones Científicas
Depto de Matemáticas, Univ. Autónoma de Madrid, 28049, Spain
javier.parcet@uam.es