Rev. Mat. Iberoam. **33** (2017), no. 4, 1285–1308
DOI 10.4171/RMI/972

Endpoint estimates for compact Calderón–Zygmund operators

Jan-Fredrik Olsen and Paco Villarroya

Abstract. We prove necessary and sufficient conditions for a Calderón–Zygmund operator to extend compactly at the endpoint from $L^1(\mathbb{R}^d)$ into $L^{1,\infty}(\mathbb{R}^d)$.

1. Introduction

The paper [10] introduced a new $T(1)$ theory to study compactness of singular integral operators. Its main result provided necessary and sufficient conditions for operators associated with classical Calderón–Zygmund kernels to be compact on $L^p(\mathbb{R})$ for all $1 < p < \infty$. This characterization was expressed in terms of three conditions: the decay of the derivative of the kernel along the direction of the diagonal, an appropriate 'weak compactness condition', and the membership of properly constructed $T(1)$ and $T^*(1)$ functions to the space $\mathrm{CMO}(\mathbb{R})$. Here, the latter space is defined as the closure in $\mathrm{BMO}(\mathbb{R})$ of the space of continuous functions vanishing at infinity. Later, in [6], the endpoint case of compactness from $L^\infty(\mathbb{R})$ into $\mathrm{CMO}(\mathbb{R})$ was obtained.

We note that, although the results in the two above-mentioned papers were proven in the context of functions defined on $\mathbb{R}$, the results and techniques developed also hold in the multi-dimensional setting. See, for instance, the preprint [11] which contains the proof of a global $T(b)$ theorem for compactness of singular integrals in $\mathbb{R}^d$.

A natural question is whether one can obtain the two remaining endpoint results, namely, compactness from $H^1(\mathbb{R}^d)$ into $L^1(\mathbb{R}^d)$ and from $L^1(\mathbb{R}^d)$ into $L^{1,\infty}(\mathbb{R}^d)$. A little bit of thought shows that the former case is an immediate consequence of [6] and Schauder's theorem, which states that an operator between two Banach spaces, $T\colon X \to Y$, is compact if and only if the same holds true for $T^*\colon Y^* \to X^*$ (see e.g. [8]). The point of this paper is to prove that the latter endpoint result also holds.

Mathematics Subject Classification (2010): 42B20, 42B25, 42C40.
Keywords: Compact singular integral operators, $T(1)$ theory.

Although both results are the natural extensions of the classical endpoint theorems for boundedness, the method used to prove compactness from $L^1(\mathbb{R}^d)$ to $L^{1,\infty}(\mathbb{R}^d)$ is very different from the standard one. It is true that the demonstration follows the same general scheme and shares identical initial steps as in the proof of boundedness. However, the standard reasoning comes to a halt when applied to the orthogonal projection operator, which is an element completely absent in the classical proof. This difficulty forces one to perform the operator analysis in a different way, more in accordance with the ideas carried out to show compactness at the non-endpoint case [10].

Since the current project is the continuation of [10], we often cite this paper for detailed references about the notation and the definitions we use, and also for proofs of those results that we merely state. And yet, we intend to present a paper as self-contained as possible.

2. Definitions

2.1. Notation

We say that $I = \prod_{i=1}^d [a_i, b_i)$ is a cube in $\mathbb{R}^d$ if the quantity $|b_i - a_i|$ remains constant for all indices $i \in \{1, \dots, d\}$. For every cube $I \subset \mathbb{R}^d$, we denote its centre by $c(I) = (2^{-1}(a_i + b_i))_{i=1}^d$, its side length by $\ell(I) = |b_i - a_i|$, and its volume by $|I| = \ell(I)^d$. For any $\lambda > 0$, we denote by λI the cube such that $c(\lambda I) = c(I)$ and $|\lambda I| = \lambda^d |I|$. Accordingly, we also write $\mathbb{B} = \mathbb{B}^d = (-1/2, 1/2)^d$ and $\mathbb{B}_\lambda = \lambda\mathbb{B} = (-\lambda/2, \lambda/2)^d$.

We denote by $|\cdot|_p$, with $0 < p \leq \infty$, the ℓ^p-norm in $\mathbb{R}^d$ and by $|\cdot|$ the modulus of a complex number. Hopefully, this notation will not cause any confusion with the one we use for the volume of a cube.

Given two cubes $I, J \subset \mathbb{R}^d$, we denote by $\langle I, J\rangle$ any cube with minimal side length containing $I \cup J$ and write its side length by $\mathrm{diam}(I \cup J)$. If there is more than one cube satisfying these conditions, we will simply select one and refer to it as $\langle I, J\rangle$ regardless of the choice.

We note that if $I = \prod_{i=1}^d I_i$, $J = \prod_{i=1}^d J_i$, with I_i, J_i intervals in $\mathbb{R}$, we have $\mathrm{diam}(I \cup J) = \max_i \mathrm{diam}(I_i \cup J_i)$, where $\mathrm{diam}(I_i \cup J_i)$ is the length of $\langle I_i, J_i\rangle$, the smallest interval containing I_i and J_i. Therefore, we have the following equivalences:

$$\mathrm{diam}(I \cup J) \approx \frac{\ell(I) + \ell(J)}{2} + |c(I) - c(J)|_\infty \approx \max(\ell(I), \ell(J)) + |c(I) - c(J)|_\infty.$$

We also define the relative distance between I and J by

$$\mathrm{rdist}(I, J) = \frac{\mathrm{diam}(I \cup J)}{\max(\ell(I), \ell(J))},$$

which is comparable to $\max(1, n)$, where n is the smallest number of times the larger cube needs to be shifted a distance equal to its side length so that it contains the smaller one. Note that, from the above, we have

$$\frac{1}{2}\Big(1+\frac{|c(I)-c(J)|_\infty}{\max(\ell(I),\ell(J))}\Big)\leq \mathrm{rdist}(I,J)\leq 1+\frac{|c(I)-c(J)|_\infty}{\max(\ell(I),\ell(J))}.$$

We also define the eccentricity of I and J to be

$$\mathrm{ecc}(I,J)=\frac{\min(|I|,|J|)}{\max(|I|,|J|)}.$$

Finally, we say that a cube I is dyadic if $I=2^j\prod_{i=1}^d[k_i,k_i+1)$ for some $j,k_1,\dots,k_d\in\mathbb{Z}$, and denote by $\mathcal{C}$ and $\mathcal{D}$ the families of all cubes and all dyadic cubes in $\mathbb{R}^d$, respectively.

Definition 2.1. For every $M\in\mathbb{N}$, we define $\mathcal{C}_M$ to be the family of all cubes in $\mathbb{R}^d$ such that $2^{-M}\leq\ell(I)\leq 2^M$ and $\mathrm{rdist}(I,\mathbb{B}_{2^M})\leq M$. We also define $\mathcal{D}_M$ to be the intersection of $\mathcal{C}_M$ with $\mathcal{D}$.

For every fixed M, we will call the cubes in $\mathcal{C}_M$ and $\mathcal{D}_M$ lagom[1] cubes and dyadic lagom cubes respectively.

Remark 2.2. Note that $I\in\mathcal{C}_M$ implies that $2^{-M}(2^M+|c(I)|_\infty)\leq M$, and so $|c(I)|_\infty\leq(M-1)2^M$. Therefore, in this case, $I\subset\mathbb{B}_{M2^M}$ with $2^{-M}\leq\ell(I)$.

On the other hand, $I\notin\mathcal{C}_M$ implies either $\ell(I)>2^M$ or $\ell(I)<2^{-M}$, or $2^{-M}\leq\ell(I)\leq 2^M$ with $|c(I)|_\infty>(M-1)2^M$.

2.2. Compact Calderón–Zygmund kernels and associated operators

We define the type of kernels that can be associated with compact operators.

Definition 2.3. Three bounded functions $L,S,D\colon[0,\infty)\to[0,\infty)$ constitute a set of admissible functions if the following limits hold:

$$\lim_{x\to\infty}L(x)=\lim_{x\to 0}S(x)=\lim_{x\to\infty}D(x)=0. \tag{2.1}$$

Remark 2.4. Since any fixed dilation of an admissible function is again admissible, we will often omit all universal constants appearing in the argument of these functions.

Definition 2.5. A function $K\colon(\mathbb{R}^d\times\mathbb{R}^d)\setminus\{(t,x)\in\mathbb{R}^d\times\mathbb{R}^d: t=x\}\to\mathbb{C}$ is called a compact Calderón–Zygmund kernel if it is bounded in its domain and there exist $0<\delta<1$, $C>0$, and admissible functions L,S,D such that

$$|K(t,x)-K(t',x')|\leq C\,\frac{(|t-t'|_\infty+|x-x'|_\infty)^\delta}{|t-x|_\infty^{d+\delta}}\,F(t,x), \tag{2.2}$$

whenever $2(|t-t'|_\infty+|x-x'|_\infty)<|t-x|_\infty$, where

$$F(t,x)=L(|t-x|_\infty)\,S(|t-x|_\infty)\,D(|t+x|_\infty).$$

[1] 'Lagom' is a Swedish word with the following meanings: adequate, moderate, in balance, just right.

We use the standard definition of multi-indices: $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$, $|\alpha| = \sum_{i=1}^d \alpha_i$ and $\partial^\alpha = \partial^{|\alpha|}/(\partial_{x_1}^{\alpha_1} \cdots \partial_{x_d}^{\alpha_d})$.

Definition 2.6. For every $N \in \mathbb{N}$, $N \geq 1$, we define $\mathcal{S}_N(\mathbb{R}^d)$ to be the set of all functions $f \in \mathcal{C}^N(\mathbb{R}^d)$ such that

$$\|f\|_{m,n} = \sup_{x \in \mathbb{R}^d} |x|^\beta |\partial^\alpha f(x)| < \infty$$

for all $\alpha, \beta \in \mathbb{N}^d$ with $|\alpha|, |\beta| \leq N$. Clearly, $\mathcal{S}_N(\mathbb{R}^d)$ equipped with the family of seminorms $\|\cdot\|_{\alpha,\beta}$ is a Fréchet space. Then, we can also define its dual space $\mathcal{S}_N(\mathbb{R}^d)'$ equipped with the dual topology which turns out to be a subspace of the space of multidimensional tempered distributions. We write $\mathcal{S}(\mathbb{R}^d)$ for the classical Schwartz space.

Definition 2.7. Let $T\colon \mathcal{S}_N(\mathbb{R}^d) \to \mathcal{S}_N(\mathbb{R}^d)'$ be a linear operator which is continuous with respect to the topology of $\mathcal{S}_N(\mathbb{R}^d)$ and the dual topology of $\mathcal{S}_N(\mathbb{R}^d)'$.

We say that T is associated with a compact Calderón–Zygmund kernel K if for all $f, g \in \mathcal{S}_N(\mathbb{R}^d)$ with disjoint compact supports, the action of Tf as a distribution satisfies the following integral representation:

$$\langle Tf, g \rangle = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(t) g(x)\, K(t,x)\, dt\, dx.$$

2.3. The weak compactness condition

Definition 2.8. For $0 < p \leq \infty$, we say that a function $\phi \in \mathcal{S}_N(\mathbb{R}^d)$ is an $L^p(\mathbb{R}^d)$-normalized bump function adapted to I with constant $C > 0$ and order $N \in \mathbb{N}$ if, for all multi-indices $0 \leq |\alpha| \leq N$, it holds that

$$|\partial^\alpha \phi(x)| \leq \frac{C}{|I|^{1/p} \ell(I)^{|\alpha|}} \Big(1 + \frac{|x - c(I)|_\infty}{\ell(I)}\Big)^{-N}.$$

Observe that, for $Np > d$, the bump functions in Definition 2.8 are normalized to be uniformly bounded in $L^p(\mathbb{R}^d)$. The order of the bump functions will always be denoted by N, even though its value might change from line to line. We will often use the Greek letters ϕ, φ for general bump functions while we reserve the use of ψ to denote bump functions with mean zero. If not otherwise stated, we will usually assume that the bump functions are $L^2(\mathbb{R}^d)$-normalized.

We now state the weak compactness condition.

Definition 2.9. A linear operator $T\colon \mathcal{S}_N(\mathbb{R}^d) \to \mathcal{S}_N(\mathbb{R}^d)'$ satisfies the weak compactness condition if there exist admissible functions L, S, D such that: for every $\epsilon > 0$, there exists $M \in \mathbb{N}$ so that for any cube $I \in \mathcal{D}$ and every pair ϕ_I, φ_I of L^2-normalized bump functions adapted to I with constant $C > 0$ and order N, we have

$$|\langle T\phi_I, \varphi_I \rangle| \lesssim C\Big(L\Big(\frac{\ell(I)}{2^M}\Big) \cdot S\big(2^M \ell(I)\big) \cdot D\Big(\frac{\mathrm{rdist}\big(I, \mathbb{B}_{2^M}\big)}{M}\Big) + \epsilon\Big),$$

where the implicit constant only depends on the operator T.

There are other alternative and less technical formulations of this concept. For example, we can say that T satisfies the weak compactness condition if and only if for all $I \in \mathcal{D}$ and for every pair ϕ_I, φ_I of L^2-normalized bump functions adapted to I, we have

$$\lim_{M\to\infty} \sup_{I\notin\mathcal{D}_M} |\langle T\phi_I, \varphi_I)\rangle| = 0,$$

where the lagom dyadic cubes $\mathcal{D}_M$ appear in Definition 2.1. However, we prefer the formulation used in Definition 2.9 because it is particularly well-suited for the calculations performed in [10] and thus, the ones carried out in the current paper.

We introduce the following notation to simplify otherwise cumbersome formulas, which appear both in the statement of Proposition 2.22 and in the proof of Theorem 4.1, below. Namely, we write

$$\begin{aligned} F(I;M) &= L_K\big(\ell(I)\big)\cdot S_K\big(\ell(I)\big)\cdot D_K\big(\mathrm{rdist}(I,\mathbb{B})\big) \\ &\quad + F_W\Big(\frac{\ell(I)}{2^M}\Big)\cdot S_W\big(2^M\ell(I)\big)\cdot D_W\Big(\frac{\mathrm{rdist}(I,\mathbb{B}_{2^M})}{M}\Big), \end{aligned}$$

where L_K, S_K, D_K are the functions appearing in the definition of a compact Calderón–Zygmund kernel, while L_W, S_W, D_W and the constant M are as in the definition of the weak compactness condition. We also set

$$\begin{aligned} F(I_1,\dots,I_n;M) &= \sum_{i=1}^n L_K\big(\ell(I_i)\big)\cdot\sum_{i=1}^n S_K\big(\ell(I_i)\big)\cdot\sum_{i=1}^n D_K\big(\mathrm{rdist}(I_i,\mathbb{B})\big) \\ &\quad + \sum_{i=1}^n L_W\Big(\frac{\ell(I_i)}{2^M}\Big)\cdot\sum_{i=1}^n S_W\big(2^M\ell(I_i)\big)\cdot\sum_{i=1}^n D_W\Big(\frac{\mathrm{rdist}(I_i,\mathbb{B})}{M}\Big). \end{aligned}$$

The following lemma is proven at the beginning of the proof of Theorem 2.21, below, as it is given in [10].

Lemma 2.10. *Given $\epsilon > 0$, then there exists exists M_0 so that for all $M > M_0$ we have $F(I_1,\dots,I_6;M_{T,\epsilon}) \lesssim \epsilon$ whenever all $I_i \in \mathcal{D}_M^c$.*

We end this subsection with two results that we will use to prove the reverse implication in our main result. Their proofs can be found in [4], Theorem 10.1, and [1], Theorem 3.1, respectively.

Theorem 2.11. *Let T be an operator with a standard Calderón–Zygmund kernel and bounded from $L^1(\mathbb{R}^d)$ into $L^{1,\infty}(\mathbb{R}^d)$. Then, T is bounded on $L^p(\mathbb{R}^d)$ for any $1 < p < \infty$ with $\|T\|_{L^p(\mathbb{R}^d)\to L^p(\mathbb{R}^d)} \lesssim \|T\|_{L^1(\mathbb{R}^d)\to L^{1,\infty}(\mathbb{R}^d)}$, and the implicit constant only depends on p and the dimension d.*

Theorem 2.12. *Let $A = (A_0, A_1)$ and $B = (B_0, B_1)$ be quasi-Banach couples and let $T\colon A \to B$ such that $T\colon A_0 \to B_0$ compactly. Then, for any $0 < \theta < 1$ and $0 < q \le \infty$, $T\colon (A_0, A_1)_{\theta,q} \to (B_0, B_1)_{\theta,q}$ is compact.*

2.4. Characterization of compactness and the lagom projection operator

The following characterization of compact operators in a Banach space with a Schauder basis (see for example [2]) was used in [10] to study compact Calderón–Zygmund operators.

Theorem 2.13. *Suppose that $\{e_n\}_{n\in\mathbb{N}}$ is a Schauder basis of a Banach space E. For each positive integer k, let P_k be the canonical projection*

$$P_k\Big(\sum_{n\in\mathbb{N}} \alpha_n\, e_n\Big) = \sum_{n\leq k} \alpha_n\, e_n.$$

Then, a bounded linear operator $T\colon E \to E$ is compact if and only if $P_k \circ T$ converges to T in operator norm.

Let E be one of the following Banach spaces: the Lebesgue space $L^p(\mathbb{R}^d)$, $1 < p < \infty$, the Hardy space $H^1(\mathbb{R}^d)$, or the space $\mathrm{CMO}(\mathbb{R}^d)$, defined in Subsection 2.5 below. In each of these cases, E is equipped with a wavelet basis which is also a Schauder basis (see [3] and Lemma 2.19). Moreover, in these cases, we can assume that the wavelets belong to $\mathcal{S}_N(\mathbb{R}^d)$ and, if needed, that they are compactly supported. However, we intentionally decide to use more general wavelets to explicitly show that our results hold in settings where, for example, compactly supported wavelets are not available.

Definition 2.14. Let E be one of the previously mentioned Banach spaces. Let $(\psi_I^i)_{I\in\mathcal{D},i=1,\ldots,2^d-1}$ be a normalized wavelet basis of E and $(\tilde{\psi}_I^i)_{\in\mathcal{D},i=1,\ldots,2^d-1}$ its dual wavelet basis. Then, for every $M \in \mathbb{N}$, we define the lagom projection operator

$$P_M f = \sum_{I\in\mathcal{D}_M} \sum_{i=1}^{2^d-1} \langle f, \tilde{\psi}_I^i\rangle\, \psi_I^i,$$

where $\langle f, \tilde{\psi}_I^i\rangle = \int_{\mathbb{R}^d} f(x)\, \overline{\tilde{\psi}_I^i(x)}dx$.

We also define $P_M^{\perp} f = f - P_M f$, and we remark that the equality

$$P_M^{\perp} f = \sum_{I\in\mathcal{D}_M^c} \sum_{i=1}^{2^d-1} \langle f, \tilde{\psi}_I^i\rangle\, \psi_I^i \tag{2.3}$$

is to be interpreted in the sense of Schauder bases, i.e.,

$$\lim_{M'\to\infty} \Big\| P_M^{\perp} f - \sum_{I\in\mathcal{D}_{M'}\setminus\mathcal{D}_M} \sum_{i=1}^{2^d-1} \langle f, \tilde{\psi}_I^i\rangle\, \psi_I^i \Big\|_E = 0.$$

In the language of the lagom projection, we can give yet another alternative formulation of weak compactness (Definition 2.9). Namely, an operator T is weakly compact if and only if, for all $I \in \mathcal{D}$ and for every pair ϕ_I, φ_I of L^2-normalized bump functions adapted to I, we have

$$\lim_{M\to\infty} |\langle (P_M^{\perp} \circ T)(\phi_I), \varphi_I)\rangle| = 0.$$

Strictly speaking, the characterization given in Theorem 2.13 is not sufficient for our purposes since in Section 4 we consider compact operators into $L^{1,\infty}(\mathbb{R}^d)$, which is a quasi-Banach space. This is addressed in Definition 2.15, where we define compact operators from $L^1(\mathbb{R}^d)$ into $L^{1,\infty}(\mathbb{R}^d)$ in the topological sense.

Definition 2.15. An operator $T\colon L^1(\mathbb{R}^d) \to L^{1,\infty}(\mathbb{R}^d)$ is compact if for every bounded set $A \subset L^1(\mathbb{R}^d)$, the set $T(A)$ is relatively compact in $L^{1,\infty}(\mathbb{R}^d)$.

Equivalently, T is compact if for each sequence $(f_n)_{n\in\mathbb{N}} \subset L^1(\mathbb{R}^d)$ with $\|f_n\|_{L^1(\mathbb{R}^d)} \lesssim 1$, there exists a subsequence $(f_{n_k})_{k\in\mathbb{N}}$ and $g \in L^{1,\infty}(\mathbb{R}^d)$ such that $\lambda m(\{x \in \mathbb{R}^d : |Tf_{n_k}(x) - g(x)| > \lambda\})$ tends to zero when k tends to infinity uniformly for all $\lambda > 0$.

Remark 2.16. Observe that finite rank operators are compact in this sense, and that the limit of finite rank operators is a compact operator.

We also note that, in light of Theorem 2.13, it would be natural to assume that the above definition is equivalent to asking that $P_M^{\perp} T$ converges to zero in the operator norm $\|\cdot\|_{L^1(\mathbb{R}^d)\to L^{1,\infty}(\mathbb{R}^d)}$. However, this is not the case as we see from the following example.

Example 2.17. Let $(\psi_I)_{I\in\mathcal{D}}$ be the Haar wavelet of $L^2(\mathbb{R})$ and P_M the associated lagom projection operator. Then, the operator defined by

$$Tf = \langle f, \psi_{[0,1]}\rangle\, \chi_{[0,1]}$$

is compact from $L^1(\mathbb{R})$ to $L^{1,\infty}(\mathbb{R})$ (since it is bounded and of finite rank), but $P_M^{\perp} T$ does not converge to zero in $L^{1,\infty}(\mathbb{R})$. Indeed, it follows from the computation below that $P_M^{\perp} T\psi_{[0,1]} = 2^{-M}\chi_{[0,2^M]}$, whence $\|P_M^{\perp} T\|_{L^1(\mathbb{R})\to L^{1,\infty}(\mathbb{R})} \geq 1$ for all $M \in \mathbb{N}$.

First, we observe that

$$P_M^{\perp}\, T\psi_{[0,1]} = P_M^{\perp}\, \chi_{[0,1]} = \chi_{[0,1]} - \sum_{I\in\mathcal{D}_M} \langle \chi_{[0,1]}, \psi_I\rangle\, \psi_I.$$

Now, $\langle \chi_{[0,1]}, \psi_I\rangle \neq 0$ if and only if $I = (0, 2^k)$ with $1 \leq k \leq M$ and, in that case, $\langle \chi_{[0,1]}, \psi_I\rangle = |I|^{-1/2}$. With this, we obtain

$$\begin{aligned} P_M^{\perp}\, T\psi_{[0,1]} &= \chi_{(0,1)} - \sum_{1\leq k\leq M} 2^{-k/2}\, 2^{-k/2}\, (\chi_{(0,2^{k-1})} - \chi_{(2^{k-1},2^k)}) \\ &= 2^{-M}\, \chi_{(0,1)} + \sum_{1\leq j\leq M} 2^{-M}\, \chi_{(2^{j-1},2^j)} = 2^{-M}\, \chi_{(0,2^M)}, \end{aligned}$$

as claimed.

2.5. The space $\mathrm{CMO}(\mathbb{R}^d)$ and the construction of $T(1)$

Definition 2.18. We define $\mathrm{CMO}(\mathbb{R}^d)$ as the closure in $\mathrm{BMO}(\mathbb{R}^d)$ of the space of continuous functions vanishing at infinity.

The next lemma gives two characterizations of $\mathrm{CMO}(\mathbb{R}^d)$: the first one in terms of the average deviation from the mean, and the second one in terms of a wavelet decomposition. See [7] and [5] for the proofs. We will only use the latter formulation.

Lemma 2.19. *The following statements are equivalent:*

(i) $f \in \mathrm{CMO}(\mathbb{R}^d)$,

(ii) $f \in \mathrm{BMO}(\mathbb{R}^d)$ *and*

$$\lim_{M\to\infty} \sup_{I\notin\mathcal{I}_M} \frac{1}{|I|}\int_I \Big| f(x) - \frac{1}{|I|}\int_I f(y)\,dy\Big|\,dx = 0,$$

(iii) $f \in \mathrm{BMO}(\mathbb{R}^d)$ *and*

$$\lim_{M\to\infty} \sup_{\Omega\subset\mathbb{R}^d} \Big(\frac{1}{|\Omega|}\sum_{I\notin\mathcal{D}_M, I\subset\Omega}\sum_{i=1}^{2^d-1} |\langle f,\psi_I^i\rangle|^2\Big)^{1/2} = 0,$$

where the supremum is taken over all measurable sets $\Omega\subset\mathbb{R}^d$.

Next, we state a technical lemma needed to give meaning to $T(1)$ and $T^*(1)$. To this end, we introduce some notation. For $a\in\mathbb{R}$ and $\lambda>0$, we define the translation operator as $T_a f(x) = f(x-a)$ and the dilation operator as $D_\lambda f(x) = f(x/\lambda)$. Let $\Phi\in\mathcal{S}(\mathbb{R}^d)$ be such that $\Phi(x)=1$ for $|x|_\infty\le 1$, $0<\Phi(x)<1$ for $1<|x|_\infty<2$ and $\Phi(x)=0$ for $|x|_\infty>2$.

Lemma 2.20. *Let T be a linear operator associated with a compact Calderón–Zygmund kernel K with parameter* $0<\delta<1$.

Let $I\subset\mathbb{R}^d$ be a cube and let $f\in\mathcal{S}_N(\mathbb{R}^d)$ have compact support in I and mean zero. Then, the limit

$$\mathcal{L}(f) = \lim_{k\to\infty} \langle T(\mathcal{T}_a\mathcal{D}_{2^k\ell(I)}\Phi), f\rangle$$

exists and is independent of the parameter $a\in\mathbb{R}$ and the cut-off function Φ.

The previous lemma allows one to define $T(1)$ as an element on the dual of the space of functions in $\mathcal{S}_N(\mathbb{R}^d)$ with compact support and mean zero. Namely, define $\langle T(1), f\rangle = \mathcal{L}(f)$ for all $f\in\mathcal{S}_N(\mathbb{R}^d)$.

2.6. Compactness on $L^p(\mathbb{R}^d)$

We now state the main result in [6], whose proof, although proven only for the one-dimensional case, also holds in the setting of several variables.

Theorem 2.21. *Let T be a linear operator associated with a standard Calderón–Zygmund kernel.*

Then, T extends to a compact operator on $L^p(\mathbb{R}^d)$, for any $1 < p < \infty$, if and only if T is associated with a compact Calderón–Zygmund kernel and it satisfies both the weak compactness condition and the cancellation conditions $T(1), T^*(1) \in \mathrm{CMO}(\mathbb{R}^d)$.

Under the same hypotheses, T is also compact as a map from $L^\infty(\mathbb{R}^d)$ into $\mathrm{CMO}(\mathbb{R}^d)$. Moreover, with the extra assumption $T(1) = T^*(1) = 0$, T is compact from $\mathrm{BMO}(\mathbb{R}^d)$ into $\mathrm{CMO}(\mathbb{R}^d)$.

We end this section stating the main auxiliary result in the proof of Theorem 2.21, which is also the starting point of the proof of the endpoint result in this paper. To this end, we provide the following definitions: given two cubes I, J, we denote $K_{\min} = J$, $K_{\max} = I$ if $\ell(J) \leq \ell(I)$, and $K_{\min} = I$, $K_{\max} = J$ otherwise. We denote by $\tilde{K}_{\max}$ the translate of $K_{\max}$ with the same centre as $K_{\min}$.

Proposition 2.22. *Let T be a linear operator associated with a compact Calderón–Zygmund kernel with parameter δ. We assume T satisfies the weak compactness condition and the special cancellation condition $T(1) = 0$ and $T^*(1) = 0$.*

Then, for any $\theta \in (0,1)$ small enough, there exist $0 < \delta' < \delta$, $N \geq 1$ and $C_{\delta'} > 0$ such that for every $\epsilon > 0$, all cubes I, J and all mean zero bump functions ψ_I, ψ_J, L^2-adapted to I and J respectively with constant $C > 0$ and order N, we have

$$|\langle T\psi_I, \psi_J\rangle| \leq C_{\delta'}\, C \,\frac{\mathrm{ecc}(I,J)^{1/2+\delta'/d}}{\mathrm{rdist}(I,J)^{d+\delta'}} \big(F(I_1,\dots,I_6; M_{T,\epsilon}) + \epsilon\big),$$

where $I_1 = I$, $I_2 = J$, $I_3 = \langle I, J\rangle$, $I_4 = \lambda_1 \tilde{K}_{\max}$, $I_5 = \lambda_2 \tilde{K}_{\max}$ and $I_6 = \lambda_2 K_{\min}$ with $\lambda_1 = \ell(K_{\max})^{-1}\mathrm{diam}(I \cup J)$, $\lambda_2 = \ell(K_{\min})^{-\theta}\mathrm{diam}(I \cup J)^\theta$.

3. Localization properties of bump functions

In this section, we prove two technical results. Lemma 3.1 concerns the localization of multi-variable bump functions while Lemma 3.2 estimates the interaction of bump functions with atoms. The proofs of both results in the one-dimensional case can be found in [9]. See also [10] for a more detailed proof of the latter result.

Lemma 3.1. *Let ϕ_I and ψ_J be bump functions L^2-adapted to I and J respectively with order $N \geq d$ and constant C. For $\ell(J) \leq \ell(I)$, then*

$$|\langle \phi_I, \psi_J\rangle| \lesssim C^2 \Big(\frac{|J|}{|I|}\Big)^{1/2}\Big(1 + \frac{|c(I) - c(J)|_\infty}{\ell(I)}\Big)^{-N}. \tag{3.1}$$

If, in addition, ϕ_I and ψ_J have order $N > d$ and ψ_J has vanishing mean, i.e., $\int \psi_J(x)dx = 0$, then

$$|\langle \phi_I, \psi_J\rangle| \lesssim C^2 \Big(\frac{|J|}{|I|}\Big)^{1/2+1/d}\Big(1 + \frac{|c(I) - c(J)|_\infty}{\ell(I)}\Big)^{-N+d}. \tag{3.2}$$

Proof. We start by proving inequality (3.1). Let c be the midpoint between $c(I)$ and $c(J)$, let L be the line going through $c(I)$ and $c(J)$ and let $H \subset \mathbb{R}^d$ be the hyperplane perpendicular to L passing through c. Let also H_I and H_J be the two half-spaces defined by the connected components of $\mathbb{R}^d \backslash H$ so that $c(I) \in H_I$ and $c(J) \in H_J$. We split

$$\langle \phi_I, \psi_J \rangle = \int_{H_I} \phi_I(x)\,\overline{\psi_J(x)}\,dx + \int_{H_J} \phi_I(x)\,\overline{\psi_J(x)}\,dx.$$

Applying Hölder's inequality and Definition 2.8, we get

$$\begin{aligned} |\langle \phi_I, \psi_J \rangle| &\le \|\phi_I\|_{L^1(\mathbb{R}^d)} \|\psi_J\|_{L^\infty(H_I)} + \|\phi_I\|_{L^\infty(H_J)} \|\psi_J\|_{L^1(\mathbb{R}^d)} \\ &\le C^2\Big(\frac{|I|}{|J|}\Big)^{1/2}\Big(1 + \frac{|c - c(J)|_\infty}{\ell(J)}\Big)^{-N} + C^2\Big(\frac{|J|}{|I|}\Big)^{1/2}\Big(1 + \frac{|c - c(I)|_\infty}{\ell(I)}\Big)^{-N}. \end{aligned} \tag{3.3}$$

Since $|c - c(I)|_\infty = |c - c(J)|_\infty = |c(I) - c(J)|_\infty/2$, $\ell(J) \le \ell(I)$ and $N \ge d$, we have that the first term is smaller than the second one, which is of the desired form.

To prove (3.2), we assume without loss of generality that $|c(I) - c(J)|_\infty = |c(I_1) - c(J_1)|$. Then, for all $x \in \mathbb{R}^d$ we write $x = (x_1, x')$ with $x' \in \mathbb{R}^{d-1}$. We define the operators

$$D_1^{-1}(\psi_J)(x) = \int_{-\infty}^{x_1} \psi_J(s, x')\,ds$$

and, for $t \in \mathbb{R}$,

$$D^{-1}(\psi_J)(t) = \int_{\mathbb{R}^{d-1}} \int_{-\infty}^{t} \psi_J(x_1, x')\,dx_1\,dx' = \int_{te_1 + H_1^-} \psi_J(x)\,dx,$$

where $H_1^- = \{x \in \mathbb{R}^d : x_1 \le 0\}$. Note that, due to the vanishing mean of ψ_J, we have

$$D^{-1}(\psi_J)(t) = -\int_{\mathbb{R}^{d-1}} \int_{t}^{\infty} \psi_J(x_1, x')\,ds\,dx' = -\int_{te_1 + H_1^+} \psi_J(x)\,dx, \tag{3.4}$$

where $H_1^+ = \{x \in \mathbb{R}^d : 0 \le x_1\}$. Then, it is readily checked that

$$\langle \phi_I, \psi_J \rangle = -\int_{\mathbb{R}^d} \partial_1 \phi_I(x) \cdot D_1^{-1}\overline{\psi_J}(x)\,dx.$$

Now, the function $D_1^{-1}\overline{\psi_J}$ can be expressed as the sum of four positive functions $D_1^{-1}\overline{\psi_J} = f_0 - f_2 + i(f_1 - f_3) = \sum_{k=0}^{3} i^k f_k$. Hence, applying the mean value theorem for integrals to each positive function f_k with respect to the variable $x' \in \mathbb{R}^{d-1}$, we obtain

$$\langle \phi_I, \psi_J \rangle = -\sum_{k=0}^{3} i^k \int_{\mathbb{R}} \partial_1 \phi_I\big(x_1, g_k(x_1)\big)\Big(\int_{\mathbb{R}^{d-1}} f_k(x)\,dx'\Big)\,dx_1,$$

where the functions g_k denote the dependence of all coordinates from x_1. Hence,

$$|\langle \phi_I, \psi_J \rangle| \le \int_{\mathbb{R}} \sup_k \big|\partial_1 \phi_I(x_1, g_k(x_1))\big|\,\Big(\sum_{k=0}^{3} \int_{\mathbb{R}^{d-1}} f_k(x)\,dx'\Big)\,dx_1.$$

Since

$$D^{-1}\overline{\psi_J}(t) = \int_{\mathbb{R}^{d-1}} D_1^{-1}\overline{\psi_J}(t,x')\,dx' = \sum_{k=0}^{3} i^k \int_{\mathbb{R}^{d-1}} f_k(t,x')\,dx' = \sum_{k=0}^{3} i^k F_k(t),$$

we have that

$$\sum_{k=0}^{3}\int_{\mathbb{R}^{d-1}} f_k(t,x')\,dx' \leq 2^{1/2}\Big(\Big(\sum_{k\text{ even}} F_k(t)\Big)^2 + \Big(\sum_{k\text{ odd}} F_k(t)\Big)^2\Big)^{1/2} = 2^{1/2}|D^{-1}\overline{\psi_J}(t)|.$$

Therefore, we can write

$$|\langle \phi_I, \psi_J\rangle_{L^2(\mathbb{R}^d)}| \lesssim \big\langle \sup_k |\partial_1\phi_I(t,g_k(t))|, |D^{-1}(\psi_J)(t)|\big\rangle_{L^2(\mathbb{R})}. \tag{3.5}$$

Now, on the one hand, we have by Definition 2.8,

$$\begin{aligned}|\partial_1\phi_I(t,g_k(t))| &\leq \frac{C}{|I|^{1/2}\ell(I)}\Big(1+\frac{|(t,g_k(t))-c(I)|_\infty}{\ell(I)}\Big)^{-N}\\ &\leq \frac{C}{|I|^{1/2}\ell(I)^{1/2}}\frac{1}{\ell(I_1)^{1/2}}\Big(1+\frac{|t-c(I_1)|}{\ell(I_1)}\Big)^{-N}.\end{aligned}$$

This is the decay estimate in Definition 2.8 of a function being adapted to the interval I_1 with constant $C|I|^{-1/2}\ell(I)^{-1/2}$.

On the other hand, to control the second factor in (3.5), we make the following computation: since $|\cdot|_1 \leq d|\cdot|_\infty$,

$$\begin{aligned}|D^{-1}(\overline{\psi_J})(t)| &\leq \frac{C}{|J|^{1/2}}\int_{te_1+H_1^+}\Big(1+\frac{|x-c(J)|_\infty}{\ell(J)}\Big)^{-N}\mathrm{d}x\\ &\leq \frac{C}{|J|^{1/2}}\int_{te_1-c(J)+H_1^+}\Big(1+\frac{|x|_1}{\ell(J)d}\Big)^{-N}\mathrm{d}x\\ &= \frac{C}{|J|^{1/2}}\int_{\mathbb{R}^{d-1}}\int_{t-c(J_1)}^{\infty}\Big(1+\frac{|x_1|+|x'|_1}{\ell(J)d}\Big)^{-N}\mathrm{d}x_1\mathrm{d}x'\\ &\lesssim \frac{C}{|J|^{1/2}}\ell(J)^d d^d\Big(1+\frac{|t-c(J_1)|}{\ell(J)d}\Big)^{-N+d} \leq Cd^N|J|^{1/2}\Big(d+\frac{|t-c(J_1)|}{\ell(J)}\Big)^{-N+d}\\ &\lesssim C|J|^{1/2}\ell(J)^{1/2}\frac{1}{\ell(J_1)^{1/2}}\Big(1+\frac{|t-c(J_1)|}{\ell(J_1)}\Big)^{-N+d}.\end{aligned}$$

Here, we tacitly assumed that $t-c(J_1)>0$. If the opposite is true, we use (3.4) in the first line of the argument to make the same calculation work. We note that this is the decay estimate in Definition 2.8 of a function being adapted to the interval I_1 with constant $C|J|^{1/2}\ell(J)^{1/2}$.

Now, we combine the above two estimates. Repeating the proof of (3.1) in the one-dimensional case for the expression in (3.5), starting with the splitting in (3.3), we obtain, for $N>d$, the bound

$$\begin{aligned}|\langle\phi_I,\psi_J\rangle| &\lesssim C^2\Big(\frac{|J|\ell(J)}{|I|\ell(I)}\Big)^{1/2}\Big(\frac{\ell(J)}{\ell(I)}\Big)^{1/2}\Big(1+\frac{|c(I_1)-c(J_1)|}{\ell(I)}\Big)^{-N+d}\\ &= C^2\Big(\frac{|J|}{|I|}\Big)^{1/2+1/d}\Big(1+\frac{|c(I)-c(J)|_\infty}{\ell(I)}\Big)^{-N+d}.\end{aligned}$$

□

Lemma 3.2. *Let I be a cube and f be an integrable function supported on I with mean zero. For each dyadic cube J, let ϕ_J be a bump function adapted to J with constant $C > 0$ and order N.*

Then, for all dyadic cubes J such that $\ell(J) \le \ell(I)$, we have

$$|\langle f, \phi_J\rangle|\,|J|^{1/2} \lesssim C\,\|f\|_{L^1(\mathbb{R}^d)}\Big(1 + \frac{|c(I) - c(J)|_\infty}{\ell(I)}\Big)^{-N}, \tag{3.6}$$

while for $\ell(I) \le \ell(J)$, we get

$$|\langle f, \phi_J\rangle|\,|J|^{1/2} \lesssim C\,\|f\|_{L^1(\mathbb{R}^d)}\,\frac{\ell(I)}{\ell(J)}\Big(1 + \frac{|c(I) - c(J)|_\infty}{\ell(J)}\Big)^{-N}. \tag{3.7}$$

Proof. The proofs of both inequalities follow the pattern from the previous lemma with the required modifications to take advantage of the compact support of f.

In order to prove (3.6), we divide the argument into two cases. When $|c(I) - c(J)|_\infty \le 2\ell(I)$, the inequality follows from Hölder:

$$|\langle f, \varphi_J\rangle| \le \|f\|_{L^1(\mathbb{R}^d)}\|\varphi_J\|_{L^\infty(\mathbb{R}^d)} \le C\,\|f\|_{L^1(\mathbb{R}^d)}\frac{1}{|J|^{1/2}}.$$

When $|c(I) - c(J)|_\infty > 2\ell(I)$, we denote by c the midpoint between $c(I)$ and $c(J)$, and by $H \subset \mathbb{R}^d$ the hyperplane passing through c and perpendicular to the line containing $c(I)$ and $c(J)$. Let also H_J be the half-space defined by the connected component of $\mathbb{R}^d\backslash H$ so that $c(J) \in H_J$. It can be readily checked that $I \cap H_J = \emptyset$, whence $\operatorname{supp} f \subset I \subset H_J^c$, and thus,

$$\begin{aligned}|\langle f, \phi_J\rangle| &\le \int_{H_J^c} |f(x)||\phi_J(x)|\,dx \le \|f\|_{L^1(\mathbb{R}^d)}\,\|\phi_J\|_{L^\infty(H_J^c)}\\ &\le \|f\|_{L^1(\mathbb{R}^d)}\,\frac{C}{|J|^{1/2}}\Big(1 + \frac{|c(I) - c(J)|_\infty}{\ell(J)}\Big)^{-N},\end{aligned}$$

which is smaller than the bound in (3.6) since $\ell(J) \le \ell(I)$.

To prove (3.7), we divide in two similar cases. We first assume that $|c(I) - c(J)|_\infty \ge 2\ell(J)$. Let $c = (c_1, c') \in \mathbb{R} \times \mathbb{R}^{d-1}$ be the midpoint between $c(I)$ and $c(J)$ and let H and H_J be as before. Since now $\ell(I) \le \ell(J)$, we have again that $c(J) \in H_J$ and $\operatorname{supp} f \subset H_J^c$.

As in the previous lemma, we assume without loss of generality that $|c(I) - c(J)|_\infty = |c(I_1) - c(J_1)|$. Then, we consider again the operator

$$D^{-1}(f)(t) = \int_{te_1 + H_1^-} f(x)\,dx = -\int_{te_1 + H_1^+} f(x)\,dx,$$

where $H_1^- = \{x \in \mathbb{R}^d : x_1 \le 0\}$ and $H_1^+ = \{x \in \mathbb{R}^d : 0 \le x_1\}$, due to the vanishing mean of f. Moreover, the support of $D^{-1}f$ is included in I_1, which is, in turn, included in $(-\infty, c_1)$.

From the computations developed in the proof of (3.2), we have that

$$\langle f, \phi_I\rangle = -\sum_{k=0}^{3} i^k \int_{\mathbb{R}} \Big(\int_{\mathbb{R}^{d-1}} f_k(x_1, x')\, dx' \Big) \partial_1 \overline{\phi_J}(x_1, g_k(x_1))\, dx_1,$$

where $D_1^{-1} f = \sum_{k=0}^{3} i^k f_k$, with f_k positive functions and the functions g_k denote the dependence of all coordinates from x_1. Now, as in the proof of Lemma 3.1, we have the inequalities

$$|\partial_1 \overline{\phi_J}(t, g_k(t))| \leq \frac{C}{|J|^{1/2}\, \ell(J)} \Big(1 + \frac{|t - c(J_1)|}{\ell(J)}\Big)^{-N},$$

and

$$\sum_{k=0}^{3} \int_{\mathbb{R}^{d-1}} f_k(t, x')\, dx' \leq 2^{1/2}\, |D^{-1} f(t)|.$$

Let $\varphi_J(t) = \sup_k |\partial_1 \overline{\phi_J}(t, g_k(t))|$. Then,

$$\begin{aligned} |\langle f, \phi_J\rangle| &\lesssim \int_{-\infty}^{c_1} |D^{-1} f(t)| \varphi_J(t)\, dt \leq \|D^{-1} f\|_{L^1(-\infty, c_1)} \|\varphi_J\|_{L^\infty(-\infty, c_1)} \\ &\leq \|D^{-1} f\|_{L^1(\mathbb{R})} \frac{C}{|J|^{1/2}\ell(J)} \Big(1 + \frac{|c(I_1) - c(J_1)|}{\ell(J)}\Big)^{-N}. \end{aligned}$$

Now, from the bound $\|D^{-1} f\|_{L^1(\mathbb{R})} \leq \ell(I_1) \|f\|_{L^1(\mathbb{R}^d)}$ and the assumption about the first coordinate, we obtain the bound stated in (3.7).

Finally, when $|c(I) - c(J)| \leq 2\ell(J)$, we use the easier estimate

$$\begin{aligned} |\langle f, \phi_J\rangle| &\lesssim \int_{-\infty}^{c_1} |D^{-1} f(t)|\, \varphi_J(t)\, dt \leq \|D^{-1} f\|_{L^1(\mathbb{R})} \|\varphi_J\|_{L^\infty(\mathbb{R})} \\ &\leq \ell(I)\, \|f\|_{L^1(\mathbb{R}^d)}\, C\, |J|^{-1/2}\, \ell(J)^{-1}. \end{aligned}$$

This ends the proof. □

4. Compactness from $L^1(\mathbb{R}^d)$ into $L^{1,\infty}(\mathbb{R}^d)$

In this section, we state and prove our main result.

Theorem 4.1. *Let T be a linear operator associated with a standard Calderón–Zygmund kernel. Then, T can be extended to a compact operator from $L^1(\mathbb{R}^d)$ into $L^{1,\infty}(\mathbb{R}^d)$ if and only if it is associated with a compact Calderón–Zygmund kernel satisfying the weak compactness condition and the cancellation conditions $T(1), T^*(1) \in \mathrm{CMO}(\mathbb{R}^d)$.*

Remark 4.2. If T is a linear operator with a standard Calderón–Zygmund kernel which can be extended compactly on $L^p(\mathbb{R}^d)$ for some $1 < p < \infty$, then we know by Theorem 2.21 that T satisfies the same three hypotheses for compactness of Theorem 4.1 and so, it can also be extended as a compact operator from $L^1(\mathbb{R}^d)$ into $L^{1,\infty}(\mathbb{R}^d)$.

We first justify the converse, which is essentially a consequence of Theorem 2.11 and compact real interpolation. We assume that T is compact from $L^1(\mathbb{R}^d)$ into $L^{1,\infty}(\mathbb{R}^d)$. Since T is bounded between the same spaces, by Theorem 2.11, we have that T is also bounded on, say, $L^4(\mathbb{R}^d)$. Then, by the interpolation Theorem 2.12 with $\theta = \frac{1-1/2}{1-1/4}$, we obtain that T is compact on $L^2(\mathbb{R}^d)$. Now, the reverse implication of Theorem 2.21 implies the required hypotheses: T has a compact Calderón–Zygmund kernel and it satisfies both the weak compactness condition and the cancellation conditions $T(1), T^*(1) \in \mathrm{CMO}(\mathbb{R}^d)$.

The remainder of the paper is devoted to show sufficiency. As in the study of boundedness, the proof is split into two cases: first a special case, when extra cancellation properties are assumed (Proposition 4.3); and second, the general case, which is dealt with by proving compactness of paraproducts (Proposition 4.4).

Proposition 4.3. *Let T be a linear operator associated with a compact Calderón–Zygmund kernel such that T satisfies the weak compactness condition and the cancellation conditions $T(1) = 0$ and $T^*(1) = 0$. Then, T is compact from $L^1(\mathbb{R}^d)$ into $L^{1,\infty}(\mathbb{R}^d)$.*

Proposition 4.4. *Given a function $b \in \mathrm{CMO}(\mathbb{R}^d)$, there exists a linear operator T_b associated with a compact Calderón–Zygmund kernel such that T_b and T_b^* are compact from $L^1(\mathbb{R}^d)$ into $L^{1,\infty}(\mathbb{R}^d)$ and satisfy $\langle T_b(1), g\rangle = \langle b, g\rangle$ and $\langle T_b(f), 1\rangle = 0$, for all $f, g \in \mathcal{S}(\mathbb{R}^d)$.*

We now remind the reader how to deduce Theorem 4.1 from these propositions. The argument follows the well-known scheme provided in the proof of the classical $T(1)$ theorem. Namely, when $b_1 = T(1)$, $b_2 = T^*(1)$ are functions in $\mathrm{CMO}(\mathbb{R}^d)$, we use Proposition 4.4 to construct the paraproduct operators T_{b_i}. As proved in [10], they have compact Calderón–Zygmund kernels, are compact operators on $L^2(\mathbb{R}^d)$ (and thus, they satisfy the weak compactness condition), and satisfy $T_{b_1}(1) = b_1$, $T_{b_2}(1) = b_2$ and $T_{b_1}^*(1) = T_{b_2}^*(1) = 0$. It now follows from Proposition 4.3 that the operator

$$T - T_{b_1} - T_{b_2}^*$$

is compact from $L^1(\mathbb{R}^d)$ into $L^{1,\infty}(\mathbb{R}^d)$. Finally, after proving that T_{b_1} and $T_{b_2}^*$ are compact from $L^1(\mathbb{R}^d)$ into $L^{1,\infty}(\mathbb{R}^d)$, we deduce that the same holds for the initial operator T.

4.1. Proof of Proposition 4.3

Let $(\psi_I^i)_{I\in\mathcal{D}, i=1,\ldots,2^d-1}$ be an orthogonal wavelet basis of $L^2(\mathbb{R}^d)$ such that every function ψ_I^i is adapted to a dyadic cube I with constant $C > 0$ and order N. We denote by P_M the lagom projection of Definition 2.14 associated with $(\psi_I^i)_{I\in\mathcal{D}, i=1,\ldots,2^d-1}$. Since the index $i \in \{1, \ldots, 2^d - 1\}$ and the dual wavelet play no significant role in the proof, in order to simplify notation, we will write the wavelet decomposition in $L^2(\mathbb{R}^d)$ simply as $f = \sum_{I\in\mathcal{D}} \langle f, \psi_I\rangle \psi_I$.

By the classical theory, we know that T extends to a bounded operator on $L^p(\mathbb{R}^d)$, and from $L^1(\mathbb{R}^d)$ into $L^{1,\infty}(\mathbb{R}^d)$. Therefore, for every $f \in \mathcal{S}(\mathbb{R}^d)$, Tf and also $P_M^\perp Tf$ are meaningful as functions in the intersection of $L^p(\mathbb{R}^d)$ and $L^{1,\infty}(\mathbb{R}^d)$.

Fix $1 < p < \infty$. By Theorem 2.21, we already know that T extends to a compact operator on $L^p(\mathbb{R}^d)$. Hence, for every $\epsilon > 0$, there exists an $M_0 \in \mathbb{N}$ such that, for all $M > M_0$ and $f \in \mathcal{S}(\mathbb{R}^d)$, we have

$$\|P_M^\perp T f\|_{L^p(\mathbb{R}^d)} \lesssim \epsilon \, \|f\|_{L^p(\mathbb{R}^d)}, \tag{4.1}$$

where the implicit constant depends only on T and p.

According to Remark 2.16, it suffices to prove that for any given $\epsilon > 0$ and its corresponding $M_1 \in \mathbb{N}$, we have for all $M > M_1$ with $M2^{-M\delta} + M^{-\delta} < \epsilon$,

$$m(\{x \in \mathbb{R}^d : |P_{2M}^\perp T f(x)| > \lambda\}) \lesssim \frac{\epsilon}{\lambda} \, \|f\|_{L^1(\mathbb{R}^d)} \tag{4.2}$$

for all $f \in \mathcal{S}(\mathbb{R}^d)$ and all $\lambda > 0$. The implicit constant is allowed to depend on $\delta > 0$, the parameter of the compact Calderón–Zygmund kernel, and the constant given by the wavelet basis, but is to be independent of ϵ, f and λ.

To prove (4.2), we perform a classical Calderón–Zygmund decomposition of f at level $\epsilon^{-1}\lambda > 0$. For this, we consider the collection $\mathcal{I}$ of maximal dyadic cubes I with respect to set inclusion such that

$$\frac{1}{|I|} \int_I |f(x)| \, dx > \frac{\lambda}{\epsilon}.$$

Let E be the disjoint union of all $I \in \mathcal{I}$, which satisfies $m(E) \leq \epsilon\lambda^{-1}\|f\|_{L^1(\mathbb{R}^d)}$. With this, we define the usual Calderón–Zygmund decomposition $f = \tilde{g} + \tilde{b}$, where

$$\tilde{g} = \sum_{I \in \mathcal{I}} m_I(f)\chi_I + f\chi_{E^c}, \quad \tilde{b} = \sum_{I \in \mathcal{I}} f_I = \sum_{I \in \mathcal{I}} \big(f - m_I(f)\big)\chi_I,$$

with $m_I(f) = |I|^{-1} \int_I f(x)\, dx$.

By standard arguments, it follows that $\|\tilde{g}\|_{L^\infty(\mathbb{R}^d)} \leq 2^d\lambda/\epsilon$, and moreover, that $\|\tilde{g}\|_{L^1(\mathbb{R}^d)} \leq \|f\|_{L^1(\mathbb{R}^d)}$. From this, the inequality (4.1), and the fact that $M > M_1 > M_0$, we get

$$\|P_{2M}^\perp \, T\tilde{g}\|_{L^p(\mathbb{R}^d)}^p \lesssim \epsilon^p \, \|\tilde{g}\|_{L^p(\mathbb{R}^d)}^p \lesssim \epsilon^p \int_{\mathbb{R}^d} |\tilde{g}(x)| \, \frac{\lambda^{p-1}}{\epsilon^{p-1}} \, dx \leq \epsilon \, \lambda^{p-1}\|f\|_{L^1(\mathbb{R}^d)}.$$

Whence,

$$m\big(\{x \in \mathbb{R}^d : |P_{2M}^\perp T\tilde{g}(x)| > \lambda/2\}\big) \lesssim \frac{1}{\lambda^p} \, \|P_{2M}^\perp \, T\tilde{g}\|_{L^p(\mathbb{R}^d)}^p \lesssim \frac{\epsilon}{\lambda} \, \|f\|_{L^1(\mathbb{R}^d)}.$$

Now we need to prove the same estimate for $\tilde{b}$. To do so, we define $\tilde{E}$ as the union of all cubes $10I$ with $I \in \mathcal{I}$. Writing $\mathbb{R}^d = \tilde{E} \cup \tilde{E}^C$, yields

$$m(\{x \in \mathbb{R}^d : |P_{2M}^\perp \, T\tilde{b}(x)| > \lambda/2\}) \lesssim m(\tilde{E}) + \frac{1}{\lambda}\|P_{2M}^\perp \, T\tilde{b}\|_{L^1(\tilde{E}^C)}.$$

Since $m(\tilde{E}) \lesssim \epsilon\lambda^{-1}\|f\|_{L^1(\mathbb{R}^d)}$, it remains to show that

$$\|P_{2M}^\perp \, T\tilde{b}\|_{L^1(\tilde{E}^C)} \lesssim \epsilon \, \|f\|_{L^1(\mathbb{R}^d)}.$$

To prove this, it suffices to show that for each $I \in \mathcal{I}$, we have

$$\|P_{2M}^{\perp} Tf_I\|_{L^1(\tilde{E}^C)} \lesssim \epsilon \, \|f_I\|_{L^1(\mathbb{R}^d)}. \tag{4.3}$$

Indeed, by sublinearity, this would imply

$$\|P_{2M}^{\perp} T\tilde{b}\|_{L^1(\tilde{E}^C)} \lesssim \epsilon \sum_{I\in\mathcal{I}} \|f_I\|_{L^1(\mathbb{R}^d)} \leq \epsilon \, \|f\|_{L^1(\mathbb{R}^d)},$$

whence,

$$m(\{x \in \mathbb{R}^d : |P_{2M}^{\perp} T\tilde{b}(x)| > \lambda/2\}) \lesssim \frac{\epsilon}{\lambda} \, \|f\|_{L^1(\mathbb{R}^d)},$$

which would complete the proof.

The remainder of the proof therefore deals with obtaining (4.3). To this end, since $f_I = (f - m_I(f))\chi_I$, we apply Fatou's lemma to obtain

$$\|P_{2M}^{\perp} Tf_I\|_{L^1(\tilde{E}^C)} \leq \liminf_{R\to\infty} \|P_{2M}^{\perp} Tf_I\|_{L^1(\tilde{E}^C\cap[-R,R]^d)}.$$

To estimate the last quantity, it suffices, by duality, to check that for all $g \in \mathcal{S}(\mathbb{R}^d)$ in the unit ball of $L^\infty(\mathbb{R}^d)$ with compact support in $K_R = \tilde{E}^C \cap [-R, R]^d$, we have

$$|\langle P_{2M}^{\perp} Tf_I, g\rangle| \lesssim \epsilon \, \|f_I\|_{L^1(\mathbb{R}^d)}. \tag{4.4}$$

We now justify this claim. Observe that since $f_I \in L^2(\mathbb{R}^d)$, it follows by the continuity of T and $P_{2M}^{\perp}$ on $L^2(\mathbb{R}^d)$ that $P_{2M}^{\perp} Tf_I \in L^2(\mathbb{R}^d)$. Hence, the function $h = \operatorname{sign}(P_{2M}^{\perp} Tf_I)\chi_{K_R}$ can be approximated in the norm of $L^2(K_R)$ by a function $g \in \mathcal{S}(\mathbb{R}^d)$ with compact support in K_R such that $\|g\|_{L^\infty(\mathbb{R}^d)} \leq \|h\|_{L^\infty(\mathbb{R}^d)} = 1$. With this, we have

$$\|P_{2M}^{\perp} Tf_I\|_{L^1(K_R)} \leq |\langle P_{2M}^{\perp} Tf_I, g\rangle| + |\langle P_{2M}^{\perp} Tf_I, h - g\rangle|,$$

where the last term can be bounded by a constant times

$$\|P_{2M}^{\perp} Tf_I\|_{L^2(\mathbb{R}^d)} \, \|h - g\|_{L^2(K_R)} \lesssim \|f_I\|_{L^2(\mathbb{R}^d)} \, \epsilon \, \frac{\|f_I\|_{L^1(\mathbb{R}^d)}}{1 + \|f_I\|_{L^2(\mathbb{R}^d)}}.$$

This ends the desired justification.

We work now to obtain (4.4). We start by justifying the equality

$$\langle P_{2M}^{\perp} Tf_I, g\rangle = \sum_{J\in\mathcal{D}} \sum_{K\in\mathcal{D}_{2M}^c} \langle f_I, \psi_J\rangle\langle g, \psi_K\rangle\langle T\psi_J, \psi_K\rangle, \tag{4.5}$$

for functions g as described above. Since $g \in \mathcal{S}(\mathbb{R}^d)$, we have that $P_{2M}^{\perp} g = g - \sum_{K\in\mathcal{D}_{2M}} \langle g, \psi_K\rangle\psi_K$ is a well defined bounded smooth function. Therefore, we can give sense to $\langle P_{2M}^{\perp} Tf_I, g\rangle = \langle Tf_I, P_{2M}^{\perp} g\rangle$.

Moreover, since $f_I \in L^2(\mathbb{R}^d)$, we can write $f_I = \sum_{J\in\mathcal{D}} \langle f_I, \psi_J\rangle\psi_J$ with convergence in $L^2(\mathbb{R}^d)$. Also $g \in L^2(\mathbb{R}^d)$ and so, according to Definition 2.14, we have

$P_{2M}^{\perp} g = \sum_{K\in\mathcal{D}_{2M}^c} \langle g, \psi_K\rangle \psi_K$ with convergence also in $L^2(\mathbb{R}^d)$. We now write for all $M', M'' > 2M$,

$$\begin{aligned}\Big|\langle Tf_I, P_{2M}^{\perp} g\rangle - \sum_{J\in\mathcal{D}_{M'}} \sum_{K\in\mathcal{D}_{M''}\setminus\mathcal{D}_{2M}} \langle f_I,\psi_J\rangle\langle g,\psi_K\rangle\langle T\psi_J,\psi_K\rangle\Big| \\ \leq \|Tf_I\|_{L^2(\mathbb{R}^d)} \Big\| P_{2M}^{\perp} g - \sum_{K\in\mathcal{D}_{M''}\setminus\mathcal{D}_{2M}} \langle g,\psi_K\rangle\psi_K\Big\|_{L^2(\mathbb{R}^d)} \\ + \|g\|_{L^2(\mathbb{R}^d)} \Big\| T\big(f_I - \sum_{J\in\mathcal{D}_{M'}} \langle f_I,\psi_J\rangle\psi_J\big)\Big\|_{L^2(\mathbb{R}^d)}\end{aligned}$$

By all the stated relationships and the continuity of T on $L^2(\mathbb{R}^d)$, both terms in previous inequality tend to zero when M', M'' tend to infinity. This justifies the equality (4.5).

Then, it follows from the triangle inequality that

$$\begin{aligned}|\langle P_{2M}^{\perp}\, Tf_I, g\rangle| &= |\langle Tf_I, P_{2M}^{\perp} g\rangle| \\ &\leq \sum_{J\in\mathcal{D}} \sum_{K\in\mathcal{D}_{2M}^c} |\langle f_I,\psi_J\rangle|\, |\langle g,\psi_K\rangle|\, |\langle T\psi_J,\psi_K\rangle|.\end{aligned}$$

Now, for any given $\epsilon > 0$ and $M_{T,\epsilon}\in\mathbb{N}$, we have by Proposition 2.22,

$$|\langle T\psi_J,\psi_K\rangle| \lesssim \frac{\mathrm{ecc}(J,K)^{1/2+\delta/d}}{\mathrm{rdist}(J,K)^{d+\delta}} \big(F(J_1,\dots,J_6; M_{T,\epsilon}) + \epsilon\big), \tag{4.6}$$

where we wrote the parameter δ' simply as δ, $J_1 = J$, $J_2 = K$, $J_3 = \langle J,K\rangle$, $J_4 = \lambda_1 \tilde{K}_{\max}$, $J_5 = \lambda_2 \tilde{K}_{\max}$ and $J_6 = \lambda_2 K_{\min}$, with parameters $\lambda_1,\lambda_2 \geq 1$ explicitly stated in the proposition.

To further simplify notation, we write the last factor as $F(J_i)+\epsilon$. Applying (4.6), we get

$$|\langle P_{2M}^{\perp}\, Tf_I, g\rangle| \lesssim \sum_{J\in\mathcal{D}} \sum_{K\in\mathcal{D}_{2M}^c} |\langle f_I,\psi_J\rangle|\, |\langle g,\psi_K\rangle|\, \frac{\mathrm{ecc}(J,K)^{1/2+\delta/d}}{\mathrm{rdist}(J,K)^{d+\delta}} \big(F(J_i)+\epsilon\big).$$

Now, we parametrise both sums according to the eccentricities and relative distances: first of J with respect to the fixed cube I and, later, of K with respect each cube J. To this end, for every $k\in\mathbb{Z}$ and $m\in\mathbb{N}$, $m\geq 1$, we define the family

$$I_{k,m} = \{J\in\mathcal{D} : \ell(I) = 2^k \ell(J), m \leq \mathrm{rdist}(I,J) < m+1\}.$$

We note that the cardinality of $I_{k,m}$ is $2^{\max(k,0)d}\, 2d\, (2m)^{d-1}$.

In the same way, for every $e\in\mathbb{Z}$ and $n\in\mathbb{N}$, $n\geq 1$, and every given cube $J\in I_{k,m}$, we define the family

$$J_{e,n} = \{K\in\mathcal{D} : \ell(J) = 2^e \ell(K), n \leq \mathrm{rdist}(J,K) < n+1\}$$

whose the cardinality is $2^{\max(e,0)d}\, 2d\,(2n)^{d-1}$. With all this, we have

$$\begin{aligned}|\langle P_{2M}^{\perp} Tf_I, g\rangle| &\lesssim \sum_{\substack{k\in\mathbb{Z}\\ m\in\mathbb{N}}}\sum_{\substack{e\in\mathbb{Z}\\ n\in\mathbb{N}}}\sum_{J\in I_{k,m}}\sum_{K\in J_{e,n}\cap\mathcal{D}_{2M}^c} |\langle f_I,\psi_J\rangle|\,|\langle g,\psi_K\rangle| \\ &\qquad\qquad \cdot 2^{-|e|d(1/2+\delta/d)}\, n^{-(d+\delta)}\,(F(J_i)+\epsilon) \\ &\lesssim \sum_{\substack{e\in\mathbb{Z}\\ n\in\mathbb{N}}} 2^{-|e|(d/2+\delta)}\, n^{-(d+\delta)}\, 2^{\max(e,0)d}\, n^{d-1} \\ &\qquad \cdot \sum_{\substack{k\in\mathbb{Z}\\ m\in\mathbb{N}}}\sum_{J\in I_{k,m}} |\langle f_I,\psi_J\rangle| \sup_{K\in J_{e,n}\cap\mathcal{D}_{2M}^c} |\langle g,\psi_K\rangle|\,(F(J_i)+\epsilon).\end{aligned}$$

A crude estimate yields

$$\begin{aligned}|\langle g,\psi_K\rangle| &\le \|g\|_{L^\infty(\mathbb{R}^d)}\int_{\tilde{E}^C} |K|^{-1/2}\Big(1+\frac{|x-c(K)|_\infty}{\ell(K)}\Big)^{-N} dx \\ &\lesssim |K|^{1/2}\Big(1+\frac{\operatorname{dist}(\tilde{E}^C,c(K))}{\ell(K)}\Big)^{-N} = |J|^{1/2}\,2^{-ed/2}\,w(\tilde{E}^C,K)^{-N}, \end{aligned} \tag{4.7}$$

where the expression $w(\tilde{E}^C,K)$ is defined by the last equality. Using this and $2^{-e/2}\,2^{\max(e,0)} = 2^{|e|/2}$, the above inequality becomes

$$\begin{aligned}|\langle P_{2M}^{\perp} Tf_I, g\rangle| &\lesssim \sum_{e\in\mathbb{Z}}\sum_{n\in\mathbb{N}} 2^{-|e|\delta}\, n^{-(1+\delta)} \\ &\quad \cdot\sum_{k\in\mathbb{Z}}\sum_{m\in\mathbb{N}}\sum_{J\in I_{k,m}} |\langle f_I,\psi_J\rangle|\,|J|^{1/2} \sup_{K\in J_{e,n}\cap\mathcal{D}_{2M}^c} w(\tilde{E}^C,K)^{-N}(F(J_i)+\epsilon).\end{aligned} \tag{4.8}$$

To keep the notation simple, we take the supremum over the empty set to be zero (recall that the support of g is contained in $\tilde{E}^C$). Also, observe that, even though it is hidden by our choice of notation, $F(J_i)$ depends on both J and K and thus, depends on k, m, e and n.

In order to estimate (4.8), we need to control the terms of the double inner sum. We split the argument into two cases, depending on the size of $F(J_i)$:

(I) $J_i \notin \mathcal{D}_M$ for all $i = 1,\dots,6$.

(II) $J_i \in \mathcal{D}_M$ for some $i = 1,\dots,6$.

The point here is that in case (I) we have by Lemma 2.10 that $F(J_i) < \epsilon$ and thus it suffices to merely bound the sum by some constant. On the other hand, in case (II) we only know that $F(J_i)$ is bounded and so, we need to use the size and location of the cubes J, K to deduce an estimate that depends on ϵ.

Proof of (I). As already noted, in this first case we have $F(J_i) < \epsilon$. We divide the study in two cases: $\ell(I) \le \ell(J)$ and $\ell(I) > \ell(J)$, which correspond to the cases $k \le 0$ and $k > 0$ respectively.

The first case follows directly from Lemma 3.2. Indeed, since $k \le 0$ and f_I is supported on I with zero mean, by (3.7) we have

$$(4.9)\quad |\langle f_I, \psi_J\rangle|\,|J|^{1/2} \lesssim \|f_I\|_{L^1(\mathbb{R}^d)} \frac{\ell(I)}{\ell(J)}\Big(1+\frac{|c(I)-c(J)|}{\ell(J)}\Big)^{-N} \lesssim \|f_I\|_{L^1(\mathbb{R}^d)} 2^k m^{-N}.$$

Moreover, the cardinality of $J \in I_{k,m}$ is comparable to m^{d-1} and so, in light of (4.9) and the inequality $F(J_i) < \epsilon$, the inequality (4.8) becomes

$$\begin{aligned}|\langle P_{2M}^{\perp} T f_I, g\rangle| &\lesssim \epsilon \sum_{e\in\mathbb{Z}}\sum_{n\in\mathbb{N}} 2^{-|e|\delta}\, n^{-(1+\delta)} \sum_{k\le 0}\sum_{m\in\mathbb{N}} 2^k\, \|f_I\|_{L^1(\mathbb{R}^d)}\, m^{d-1-N}\\ &\lesssim \epsilon\, \|f_I\|_{L^1(\mathbb{R}^d)},\end{aligned}$$

for $N > d$, with the implicit constant depending exponentially on d.

In the second case, however, we need to be more careful. Now we have $\ell(J) < \ell(I)$, or equivalently $k > 0$, and we further divide in two more cases: $\ell(I) < \ell(K)$ and $\ell(K) \le \ell(I)$. In the first case, we use (3.6) in Lemma 3.2, and so,

$$|\langle f_I, \psi_J\rangle|\,|J|^{1/2} \lesssim \|f_I\|_{L^1(\mathbb{R}^d)}\Big(1+\frac{|c(I)-c(J)|}{\ell(I)}\Big)^{-N} \lesssim \|f_I\|_{L^1(\mathbb{R}^d)}\, m^{-N}.$$

Moreover, we have that $\ell(I) < \ell(K) = 2^{-(e+k)}\ell(I)$ which implies $0 < k \le -e$. Note that $e \le 0$ in this situation since $\ell(J) \le \ell(I) \le \ell(K)$. Then, since $F(J_i) < \epsilon$, the bound for the corresponding terms of (4.8) becomes

$$\begin{aligned}&\epsilon \sum_{e\in\mathbb{Z}}\sum_{n\in\mathbb{N}} 2^{-|e|\delta}\, n^{-(1+\delta)} \sum_{k=1}^{|e|}\sum_{m\in\mathbb{N}} |\langle f_I, \psi_J\rangle|\,|J|^{1/2}\\ &\qquad \lesssim \epsilon \sum_{e\in\mathbb{Z}}\sum_{n\in\mathbb{N}} 2^{-|e|\delta}\, n^{-(1+\delta)}\, |e| \sum_{m\in\mathbb{N}} \|f_I\|_{L^1(\mathbb{R}^d)}\, m^{-N} \lesssim \epsilon\, \|f_I\|_{L^1(\mathbb{R}^d)}.\end{aligned}$$

We now assume that $\ell(K) \le \ell(I)$. By the definition of bump functions adapted to a cube, we have

$$|\langle f_I, \psi_J\rangle|\,|J|^{1/2} \le C \int_I |f_I(x)|\Big(1+\frac{|x-c(J)|_\infty}{\ell(J)}\Big)^{-N} dx.$$

Then, for every fixed eccentricity $0 < k$ and every fixed $x \in I$, we proceed to parametrise the cubes $J \in I_{k,m}$ associated with a fixed value of $\ell(J)^{-1}|x-c(J)|_\infty$. Since

$$m \le \operatorname{rdist}(I,J) \le 1 + \frac{|c(I)-c(J)|_\infty}{\ell(I)}$$

and

$$m+1 > \operatorname{rdist}(I,J) \ge \frac{1}{2}\Big(1+\frac{|c(I)-c(J)|_\infty}{\ell(I)}\Big),$$

we get $(m-1)\ell(I) \le |c(I)-c(J)|_\infty \le (2m+1)\ell(I)$. This way,

$$|x-c(J)|_\infty \ge |c(I)-c(J)|_\infty - |x-c(I)|_\infty \ge \Big(m-\frac{3}{2}\Big)\,\ell(I) \ge \frac{m-1}{2}\, 2^k\, \ell(J)$$

and

$$|x-c(J)|_\infty \le \Big(2m+\frac{3}{2}\Big)\,\ell(I) \le 2(m+1)\,2^k\,\ell(J).$$

Moreover, for every fixed integer $(m-1)2^{k-1} \le r \le 2^{k+1}(m+1)$, there are at most $d2^d(r+1)^{d-1}$ cubes $J\in I_{k,m}$ with $r \le \ell(J)^{-1}|x-c(J)|_\infty < r+1$. With all this, we have

$$\begin{aligned}\sum_{J\in I_{k,m}} |\langle f_I,\psi_J\rangle|\,|J|^{1/2} &\lesssim \int |f_I(x)| \sum_{r=(m-1)2^{k-1}}^{(m+1)2^{k+1}} (1+r)^{d-1-N} dx\\ &\lesssim \|f_I\|_{L^1(\mathbb{R}^d)}(1+(m-1)2^{k-1})^{d-N}. \qquad (4.10)\end{aligned}$$

For this reason, we again divide the argument into two cases: $J\cap 3I=\emptyset$ and $J\cap 3I\neq\emptyset$. In the former case, we have $m>1$, and so, with the estimate (4.10) and the inequality $F(J_i)<\epsilon$, the bound for the corresponding terms of (4.8) becomes

$$\begin{aligned}&\epsilon\sum_{e\in\mathbb{Z}}\sum_{n\in\mathbb{N}} 2^{-|e|\delta}\,n^{-(1+\delta)}\sum_{k\ge1}\sum_{m\ge2}\sum_{J\in I_{k,m}} |\langle f_I,\psi_J\rangle|\,|J|^{1/2}\\ &\quad\lesssim \epsilon\,\|f_I\|_{L^1(\mathbb{R}^d)}\sum_{e\in\mathbb{Z}}\sum_{n\in\mathbb{N}} 2^{-|e|\delta}\,n^{-(1+\delta)}\sum_{k\ge1}\sum_{m\ge2} m^{d-N}\,2^{k(d-N)} \lesssim \epsilon\,\|f_I\|_{L^1(\mathbb{R}^d)}\end{aligned}$$

as long as $N>d+1$.

In the case $J\cap 3I\neq\emptyset$, we have $m=1$ and $|c(I)-c(J)|_\infty \le 5\ell(I)/2$. To continue the analysis, we split into two further subcases: $K\cap 7I=\emptyset$ and $K\cap 7I\neq\emptyset$.

In the former case, we have that $|c(K)-c(I)|_\infty \ge 7\ell(I)/2$ and so,

$$\begin{aligned}n+1>\operatorname{rdist}(J,K) &\ge \frac{1}{2}\Big(1+\frac{|c(J)-c(K)|_\infty}{\max\big(\ell(J),\ell(K)\big)}\Big)\\ &\ge \frac{1}{2}\Big(1+\frac{|c(K)-c(I)|_\infty-|c(I)-c(J)|_\infty}{\max\big(\ell(J),\ell(K)\big)}\Big) \ge \frac{1}{2}\Big(1+\frac{\ell(I)}{\max\big(\ell(J),\ell(K)\big)}\Big).\end{aligned}$$

Since $\ell(J)=2^{-k}\ell(I)$ and $\ell(K)=2^{-e}\ell(J)=2^{-(e+k)}\ell(I)$, this yields

$$n+1>\frac{1}{2}\big(1+2^k\min(1,2^e)\big),$$

whence

$$1\le 2^k \le \frac{2n+1}{\min(1,2^e)} \le 3n(1+2^{-e}).$$

With the estimate (4.10) and the inequality $F(J_i)<\epsilon$, the bound for the corresponding terms of (4.8) becomes

$$\begin{aligned}&\epsilon\sum_{e\in\mathbb{Z}}\sum_{n\in\mathbb{N}} 2^{-|e|\delta}\,n^{-(1+\delta)}\sum_{k=1}^{\log(3n(1+2^{-e}))}\sum_{J\in I_{k,1}} |\langle f_I,\psi_J\rangle|\,|J|^{1/2}\\ &\quad\lesssim \epsilon\,\|f_I\|_{L^1(\mathbb{R}^d)}\sum_{e\in\mathbb{Z}}\sum_{n\in\mathbb{N}} 2^{-|e|\delta}\,n^{-(1+\delta)}\,\log\big(3n(1+2^{-e})\big) \lesssim \epsilon\,\|f_I\|_{L^1(\mathbb{R}^d)}\end{aligned}$$

as long as $N>d$.

On the other hand, when $K \cap 7I \neq \emptyset$ and $\ell(K) \leq \ell(I)$, we have $\operatorname{dist}(\tilde{E}^C, c(K)) \geq \ell(I)$. Moreover, $\ell(I) = 2^{k+e}\ell(K)$ with $k+e \geq 0$. Then,

$$w(\tilde{E}^C, K) = 1 + \frac{\operatorname{dist}(\tilde{E}^C, c(K))}{\ell(K)} \geq 1 + 2^{k+e}.$$

With this, the estimate (4.10), and the inequality $F(J_i) < \epsilon$, the bound for the corresponding terms of (4.8) becomes

$$\begin{aligned}\epsilon \sum_{e\in\mathbb{Z}} \sum_{n\in\mathbb{N}} 2^{-|e|\delta} n^{-(1+\delta)} \sum_{\substack{k\geq 1\\ k+e\geq 0}} \sum_{J\in I_{k,1}} |\langle f_I, \psi_J\rangle|\, |J|^{1/2}\, w(\tilde{E}^C, K)^{-N} \\ \lesssim \epsilon\, \|f_I\|_{L^1(\mathbb{R}^d)} \sum_{e\in\mathbb{Z}} \sum_{n\in\mathbb{N}} 2^{-|e|\delta} n^{-(1+\delta)} \sum_{\substack{k\geq 1\\ k+e\geq 0}} 2^{-(k+e)N} \lesssim \epsilon\, \|f_I\|_{L^1(\mathbb{R}^d)}.\end{aligned}$$

Combining all the obtained estimates, we get the desired bound for (4.8) under the assumption of case (I).

Proof of (II). As previously stated, in this case we use the size and location of the cubes J and K to deduce an estimate that depends on ϵ. This leads to the following sub-cases:

- (II$_1$) $J_1 = J \in \mathcal{D}_M$,
- (II$_2$) $J_2 = K \in \mathcal{D}_M$,
- (II$_3$) $J_3 = \langle J \cup K\rangle \in \mathcal{D}_M$,
- (II$_4$) $J \notin \mathcal{D}_M$ but $J_4 = \lambda_1 \tilde{K}_{\max} \in \mathcal{D}_M$,
- (II$_5$) $J \notin \mathcal{D}_M$ but $J_5 = \lambda_2 \tilde{K}_{\max} \in \mathcal{D}_M$,
- (II$_6$) $J \notin \mathcal{D}_M$ but $J_6 = \lambda_2 K_{\min} \in \mathcal{D}_M$.

We can use the fact that $K \in J_{e,n} \cap \mathcal{D}_{2M}^c$ to immediately rule out the case (II$_2$). We note that the property $K \notin \mathcal{D}_{2M}$ plays a crucial role in the remaining cases. We prove only the case (II$_1$) since, as explained in more detail in [10], all other cases can be dealt with by a similar reasoning.

(II$_1$). We recall that the cubes J and K in the sum (4.8) satisfy $\ell(J) = 2^e \ell(K)$ and $n \leq \operatorname{rdist}(J, K) < n+1$.

By assumption, we have $J \in \mathcal{D}_M$. That is, $2^{-M} \leq \ell(J) \leq 2^M$ and $\operatorname{rdist}(J, \mathbb{B}_{2^M}) \leq M$. Also, since F is bounded, we have $F(J_i) + \epsilon \lesssim 1$.

Since $K \in \mathcal{D}_{2M}^c$, we separate the study into three cases:

- (II$_{1.1}$) $\ell(K) > 2^{2M}$,
- (II$_{1.2}$) $\ell(K) < 2^{-2M}$
- (II$_{1.3}$) $2^{-2M} \leq \ell(K) \leq 2^{2M}$ with $\operatorname{rdist}(K, \mathbb{B}_{2^{2M}}) > 2M$.

Case (II$_{1.1}$). The inequalities $\ell(K) > 2^{2M}$ and $2^e \ell(K) = \ell(J) \leq 2^M$ imply $2^e \leq 2^M \ell(K)^{-1} \leq 2^{-M}$ and so, $e \leq -M$.

Using this and repeating the arguments from (I), the inequality (4.8) becomes

$$|\langle P_{2M}^{\perp} Tf_I, g\rangle| \lesssim \|f_I\|_{L^1(\mathbb{R}^d)} \sum_{e\leq -M}\sum_{n\in\mathbb{N}} 2^{-|e|\delta} n^{-(1+\delta)}(|e|+\log n)$$
$$\lesssim M 2^{-M\delta}\|f_I\|_{L^1(\mathbb{R}^d)} < \epsilon\,\|f_I\|_{L^1(\mathbb{R}^d)},$$

where the last inequality holds by the choice of M.

Case (II$_{1.2}$). The case $\ell(K) < 2^{-2M}$ is totally symmetrical with respect to the previous one, and amounts to changing $e \leq -M$ by $e \geq M$.

Case (II$_{1.3}$). When $2^{-2M} \leq \ell(K) \leq 2^{2M}$ and $\mathrm{rdist}(K, \mathbb{B}_{2^{2M}}) \geq 2M$, we have by Remark 2.2 that $|c(K)|_\infty \geq (2M-1)2^{2M}$. Moreover, since $J \in \mathcal{D}_M$, by the same remark we have $|c(J)|_\infty \leq (M-1)\,2^M$. Then,

$$|c(J)-c(K)|_\infty \geq |c(K)|_\infty - |c(J)|_\infty \geq M\,2^{2M}.$$

Furthermore, $\max(\ell(J),\ell(K)) \leq 2^{2M}$ and so,

$$n+1 > \mathrm{rdist}(J,K) \geq \frac{|c(J)-c(K)|_\infty}{\max(\ell(J),\ell(K))} \geq M.$$

Using this in combination with the arguments from (I), the inequality (4.8) now becomes

$$|\langle P_{2M}^{\perp} Tf_I, g\rangle| \lesssim \|f_I\|_{L^1(\mathbb{R}^d)} \sum_{e\in\mathbb{Z}}\sum_{n\geq M-1} 2^{-|e|\delta}\, n^{-(1+\delta)}\,(|e|+\log n)$$
$$\lesssim M^{-\delta}\,\|f_I\|_{L^1(\mathbb{R}^d)} < \epsilon\,\|f_I\|_{L^1(\mathbb{R}^d)},$$

again by the choice of M. □

4.2. Proof of Proposition 4.4

Let $(\psi_I^i)_{I\in\mathcal{D}, i=1,\dots,2^d-1}$ be an orthogonal wavelet basis of $L^2(\mathbb{R}^d)$ such that every function ψ_I^i is adapted to a dyadic cube I with constant $C>0$ and order N. As in the proof of Proposition 4.3, we suppress the dependence on the index i.

We denote by $\varphi \in \mathcal{S}(\mathbb{R}^d)$ a positive bump function adapted to $[-1/2,1/2]^d$ with order N and constant $C>0$ such that $\int_{\mathbb{R}^d}\varphi(x)dx = 1$. In particular, we have that $0 \leq \varphi(x) \leq C(1+|x|_\infty)^{-N}$ and $|\partial_i\varphi(x)| \leq C(1+|x|_\infty)^{-N}$ for all $i = 1,\dots,d$. Let $(\varphi_I)_{I\in\mathcal{D}}$ be the family of bump functions defined by $\varphi_I(x) = \frac{1}{|I|}\,\varphi\big(\frac{x-c(I)}{\ell(I)}\big)$.

Given a function $b \in \mathrm{CMO}(\mathbb{R}^d)$, we define the linear operator T_b by

$$\langle T_b f, g\rangle = \sum_{J\in\mathcal{D}} \langle b,\psi_J\rangle\langle f,\varphi_J\rangle\langle g,\psi_J\rangle,$$

for all $f, g \in \mathcal{S}(\mathbb{R}^d)$. It was shown in [10] that T_b and T_b^* are associated with a compact Calderón–Zygmund kernel, are compact on $L^p(\mathbb{R}^d)$ for every $1<p<\infty$, and they satisfy $\langle T_b(1), g\rangle = \langle b, g\rangle$ and $\langle T_b(f), 1\rangle = 0$.

Now, we prove that T_b, T_b^* are compact from $L^1(\mathbb{R}^d)$ into $L^{1,\infty}(\mathbb{R}^d)$. To prove compactness of the former operator, we show first the equality $P_M^\perp T_b = T_{P_M^\perp b}$. Let $f, g \in \mathcal{D}(\mathbb{R}^d)$. Since $P_M^\perp g = \sum_{J\in\mathcal{D}_M^c} \langle g, \psi_J\rangle\psi_J$,

$$\begin{aligned}\langle P_M^\perp T_b f, g\rangle = \langle T_b f_I, P_M^\perp g\rangle &= \sum_{J\in\mathcal{D}_M^c} \langle b, \psi_J\rangle\langle f, \varphi_J\rangle\langle g, \psi_J\rangle \\ &= \sum_{J\in\mathcal{D}} \langle P_M^\perp b, \psi_J\rangle\langle f, \varphi_J\rangle\langle g, \psi_J\rangle = \langle T_{P_M^\perp b}(f), g\rangle,\end{aligned}$$

where the second last equality holds because $b \in \mathrm{CMO}(\mathbb{R}^d)$ and so, we also have $P_M^\perp b = \sum_{J\in\mathcal{D}_M^c} \langle b, \psi_J\rangle\psi_J$.

Moreover, $b \in \mathrm{CMO}(\mathbb{R}^d)$ implies that for any given $\epsilon > 0$, there exists $M_0 \in \mathbb{N}$ such that $\|P_M^\perp b\|_{\mathrm{BMO}(\mathbb{R}^d)} < \epsilon$ for all $M > M_0$. Also, since $T_{P_M^\perp b}$ is a Calderón–Zygmund operator, we know by the classical theory that it is bounded from $L^1(\mathbb{R}^d)$ into $L^{1,\infty}(\mathbb{R}^d)$ with constant bounded by $\|P_M^\perp b\|_{\mathrm{BMO}(\mathbb{R}^d)}$. With all this we can write

$$\begin{aligned}m(\{x\in\mathbb{R}^d : |P_M^\perp T_b f(x)| > \lambda\}) &= m(\{x\in\mathbb{R}^d : |T_{P_M^\perp b} f(x)| > \lambda\}) \\ &\lesssim \frac{1}{\lambda}\|P_M^\perp b\|_{\mathrm{BMO}(\mathbb{R}^d)}\|f\|_{L^1(\mathbb{R}^d)} \lesssim \frac{\epsilon}{\lambda}\|f\|_{L^1(\mathbb{R}^d)},\end{aligned}$$

which is the result we seek.

Finally, we turn to the operator

$$T_b^* f(x) = \sum_{J\in\mathcal{D}} \langle b, \psi_J\rangle\,\langle f, \psi_J\rangle\,\varphi_J(x).$$

Our previous reasoning does not apply because, in general, $P_M^\perp T_b^*$ does not converge to zero. Namely, for $d = 1$, $b = \psi_{[0,1]}$ and $\varphi = \chi_{[0,1]}$, we have that $T_b^* f = \left\langle f, \psi_{[0,1]}\right\rangle \chi_{[0,1]}$, which is the operator we studied in example 2.17. As we saw, T_b^* is compact at the endpoint but $P_M T_b^*$ does not converge to T_b^* in $L^{1,\infty}(\mathbb{R})$.

However, by linearity we still have $T_b^* = T_{P_M b}^* + T_{P_M^\perp b}^*$. Now, $T_{P_M b}^*$ is of finite rank, and therefore compact. Moreover, a similar argument as before shows that $m(\{x\in\mathbb{R}^d : |T_{P_M^\perp b}^* f(x)| > \lambda\})$ can be made smaller than ϵ/λ by choosing M large. This proves compactness of T_b^*. □

Acknowledgements. We extend our most sincere gratitude to Ågot Holand Olsen for arranging an enjoyable and productive stay in Harstad, Norway, where the seeds of this project first sprouted. We also acknowledge Fernando Cobos, Michael Lacey and Xavier Tolsa for helpful comments which improved the final version of our work. The second author also would like to express his appreciation to 何琳 and George Janbing Lee for their invaluable help.

Last, we also thank the support received from the Center for Mathematical Sciences at Lund University, Sweden, and by 김자영 in Sunnyvale, USA, the places where this research was successfully conducted.

References

[1] Cobos, F. and Persson, L. E.: Real interpolation of compact operators between quasi-Banach spaces. *Math. Scan.* **82** (1998), no. 1, 138–160.

[2] Fabian, M., Habala, P., Hájek, P., Montesinos, V., and Zizler, V.: *Banach space theory. The basis for linear and nonlinear analysis.* CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, Springer, New York, 2011.

[3] Hernández, E. and Weiss, G.: *A first course on wavelets.* Studies in Advanced Mathematics, CRC Press Inc., Boca Raton, Florida, 1996.

[4] Nazarov, F., Treil, S. and Volberg, A.: Weak type estimates and Cotlar inequalities for Calderón–Zygmund operatorson nonhomogeneous spaces. *Internat. Math. Res. Notices* **1998** (1998), no. 9, 463–487.

[5] Lacey, M., Terwilleger, E. and Wick, B.: Remarks on product VMO. *Proc. Amer. Math. Soc.* **134** (2006), no. 2, 465–474.

[6] Perfekt, K. M., Pott, S. and Villarroya, P.: Endpoint compactness of singular integrals and perturbations of the Cauchy integral. *Kyoto J. Math.* **57** (2017), no. 2, 365–393.

[7] Rochberg, R.: Toeplitz and Hankel operators on the Paley–Wiener space. *Integral Equations Operator Theory* **10** (1987), no. 2, 187–235.

[8] Schauder, J.: Über lineare, vollstetige Funktionaloperationen. *Studia Math.* **2** (1930), 183–196.

[9] Thiele, C.: *Wave packet analysis.* CBMS Regional Conference Series in Mathematics 105, American Mathematical Society, Providence RI, 2006.

[10] Villarroya, P.: A characterization of compactness for singular integrals. *J. Math. Pures Appl. (9)* **104** (2015), no. 3, 485–532.

[11] Villarroya, P.: A global *Tb* theorem for boundedness and compactness. Preprint, arXiv: 1710.07869, 2017.

Received August 6, 2015.

Jan-Fredrik Olsen: Centre for Mathematical Sciences, University of Lund, Lund, 22 100, Sweden.
E-mail: jan-fredrik.olsen@math.lu.se

Paco Villarroya: Centre for Mathematical Sciences, University of Lund, Lund, 22 100, Sweden.
E-mail: paco.villarroya@maths.lth.se