

# Endpoint compactness of singular integrals and perturbations of the Cauchy integral

Karl-Mikael Perfekt, Sandra Pott, and Paco Villarroya

---

**Abstract** We prove sufficient and necessary conditions for the compactness of Calderón–Zygmund operators on the endpoint from  $L^\infty(\mathbb{R})$  into  $\text{CMO}(\mathbb{R})$ . We use this result to prove the compactness on  $L^p(\mathbb{R})$  with  $1 < p < \infty$  of a certain perturbation of the Cauchy integral on curves with normal derivatives satisfying a CMO-condition.

## 1. Introduction

In [13], we started a general theory to characterize the compactness of singular integral operators. More precisely, we showed that a Calderón–Zygmund operator  $T$  is compact on  $L^p(\mathbb{R})$  with  $1 < p < \infty$  if and only three conditions hold: the operator kernel satisfies the definition of a compact Calderón–Zygmund kernel, a strengthening of the smoothness condition of a standard Calderón–Zygmund kernel;  $T$  satisfies a new property of *weak compactness*, analogous to the classical *weak boundedness*; and the functions  $T(1)$  and  $T^*(1)$  belong to the space  $\text{CMO}(\mathbb{R})$ , the appropriate substitute of  $\text{BMO}(\mathbb{R})$ .

Now, the purpose of the current article is to continue this study in two different but related ways. First, we extend the results appearing in [13] to one of the endpoint cases, namely, from  $L^\infty(\mathbb{R})$  into  $\text{CMO}(\mathbb{R})$  (see Theorem 2.18). For this purpose, we follow a new approach, based on the study of boundedness of a modified martingale transform, Proposition 3.6, which substitutes the classical square function and, to the authors' knowledge, has not been studied before. Second, we use the latter result to provide an application of the general theory by showing how the methods devised in [13] allow one to prove the compactness on  $L^p(\mathbb{R})$  of a certain perturbation of the Cauchy integral operator defined over Lipschitz curves with CMO-smooth normal derivatives (see Proposition 4.2).

The article is structured as follows. In Section 2, we give the necessary definitions, state the main results of [13] that will be needed, and also state our main result in this article (see Theorem 2.19). In Section 3 we characterize the

---

*Kyoto Journal of Mathematics*, Vol. 57, No. 2 (2017), 365–393

DOI [10.1215/21562261-3821837](https://doi.org/10.1215/21562261-3821837), © 2017 by Kyoto University

Received April 2, 2015. Revised February 27, 2016. Accepted March 15, 2016.

*2010 Mathematics Subject Classification*: Primary 42B20, 42B25, 42C40; Secondary 47G10.

Pott's work partially supported by the Crafoord Foundation.

Villarroya's work partially supported by the Spanish Ministry of Economy and Competitiveness (project grants MTM2011-23164 and MTM2014-53009-P).

compactness of Calderón–Zygmund operators for the endpoint case  $L^\infty \rightarrow \text{CMO}$ . Finally, in Section 4 we provide an application of the theory by proving the compactness of the previously described perturbation of the Cauchy integral.

## 2. Definitions and statement of the main result

### 2.1. Definitions and notation

#### DEFINITION 2.1

We say that three bounded functions  $L, S, D : [0, \infty) \rightarrow [0, \infty)$  constitute a set of admissible functions if the following limits hold:

$$(1) \quad \lim_{x \rightarrow \infty} L(x) = \lim_{x \rightarrow 0} S(x) = \lim_{x \rightarrow \infty} D(x) = 0.$$

#### REMARK 2.2

Since any fixed dilation of an admissible function  $L_\lambda(x) = L(\lambda^{-1}x)$  is again admissible, we will often omit all universal constants appearing in the argument of these functions.

#### DEFINITION 2.3

Let  $\Delta$  be the diagonal of  $\mathbb{R}^2$ . Let  $L, S, D$  be admissible functions.

A function  $K : (\mathbb{R}^2 \setminus \Delta) \rightarrow \mathbb{C}$  is called a compact Calderón–Zygmund kernel if it is bounded on compact sets of  $\mathbb{R}^2 \setminus \Delta$  and, for some  $0 < \delta \leq 1$  and  $C > 0$ , we have

$$|K(t, x) - K(t', x')| \leq C \frac{(|t - t'| + |x - x'|)^\delta}{|t - x|^{1+\delta}} L(|t - x|) S(|t - x|) D(|t + x|),$$

whenever  $2(|t - t'| + |x - x'|) < |t - x|$ .

As shown in [13], it can be assumed without loss of generality that in Definition 2.3 the functions  $L$  and  $D$  are monotone nonincreasing while the function  $S$  is monotone nondecreasing.

We also remark that there is an equivalent definition of compact Calderón–Zygmund kernels which is more convenient to use in applications of the theory (see Section 4). In [13], we show that Definition 2.3 is equivalent to the existence of a bounded function  $B : \mathbb{R}^2 \rightarrow [0, \infty)$  such that

$$\lim_{|t-x| \rightarrow \infty} B(t, x) = \lim_{|t-x| \rightarrow 0} B(t, x) = \lim_{|t+x| \rightarrow \infty} B(t, x) = 0$$

and, for some  $0 < \delta \leq 1$  and  $C > 0$ ,

$$|K(t, x) - K(t', x')| \leq C \frac{(|t - t'| + |x - x'|)^\delta}{|t - x|^{1+\delta}} B(t, x),$$

whenever  $2(|t - t'| + |x - x'|) < |t - x|$ .

#### DEFINITION 2.4

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

such that

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

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

**DEFINITION 2.5**

Let  $T : \mathcal{S}_N(\mathbb{R}) \rightarrow \mathcal{S}'_N(\mathbb{R})$  be a linear operator which is continuous with respect to the topology of  $\mathcal{S}_N(\mathbb{R})$  for a fixed  $N \geq 1$ .

We say that  $T$  is associated with a compact Calderón–Zygmund kernel  $K$  if the action of  $T(f)$  as a distribution satisfies the following integral representation:

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

for all functions  $f, g \in \mathcal{S}_N(\mathbb{R})$  with disjoint compact supports.

**DEFINITION 2.6**

For  $0 < p \leq \infty$  and  $N \in \mathbb{N}$ , we say that a function  $\phi \in \mathcal{S}_N(\mathbb{R})$  is an  $L^p(\mathbb{R})$ -normalized bump function adapted to  $I$  with constant  $C > 0$  and order  $N$  if it satisfies

$$|\phi^{(n)}(x)| \leq C \frac{1}{|I|^{\frac{1}{p}+n}} \left(1 + \frac{|x - c(I)|}{|I|}\right)^{-N}, \quad 0 \leq n \leq N,$$

for every interval  $I \subset \mathbb{R}$ , where we denote its center by  $c(I)$  and its length by  $|I|$ .

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  $\phi, \varphi$  for general bump functions, while we reserve the use of  $\psi$  to denote bump functions with mean zero. If not otherwise stated, we will usually assume that bump functions are  $L^2(\mathbb{R})$ -normalized.

In forthcoming sections, we will use the following property of bump functions, whose proof can be found in [12].

**LEMMA 2.7**

Let  $I, J$  be intervals, and let  $\phi_I, \varphi_J$  be bump functions  $L^2$ -adapted to  $I$  and  $J$ , respectively, with order  $N$  and constant  $C > 0$ . Then,

$$|\langle \phi_I, \varphi_J \rangle| \leq C \left(\frac{\min(|I|, |J|)}{\max(|I|, |J|)}\right)^{1/2} \left(\frac{\text{diam}(I \cup J)}{\max(|I|, |J|)}\right)^{-N}.$$

Moreover, if  $|J| \leq |I|$  and  $\psi_J$  has mean zero, then

$$|\langle \phi_I, \psi_J \rangle| \leq C \left(\frac{|J|}{|I|}\right)^{3/2} \left(\frac{\text{diam}(I \cup J)}{|I|}\right)^{-(N-1)}.$$

## NOTATION 2.8

We now introduce some notation which will be frequently used throughout the article. We denote  $\mathbb{B} = [-1/2, 1/2]$  and  $\mathbb{B}_\lambda = \lambda\mathbb{B} = [-\lambda/2, \lambda/2]$ .

Given two intervals  $I, J \subset \mathbb{R}$ , we define  $\langle I, J \rangle$  as the smallest interval containing  $I \cup J$ , and we denote its measure by  $\text{diam}(I \cup J)$ . Notice that

$$\text{diam}(I \cup J) \approx |I|/2 + |c(I) - c(J)| + |J|/2.$$

We also define the relative distance between  $I$  and  $J$  by

$$\text{rdist}(I, J) = \frac{\text{diam}(I \cup J)}{\max(|I|, |J|)},$$

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

$$\text{rdist}(I, J) \approx 1 + \frac{|c(I) - c(J)|}{\max(|I|, |J|)}.$$

Finally, we define the eccentricity of  $I$  and  $J$  to be

$$\text{ec}(I, J) = \frac{\min(|I|, |J|)}{\max(|I|, |J|)}.$$

## DEFINITION 2.9

A linear operator  $T : \mathcal{S}_N(\mathbb{R}) \rightarrow \mathcal{S}'_N(\mathbb{R})$  with  $N \geq 1$  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 interval  $I$  and every pair  $\phi_I, \varphi_I$  of  $L^2$ -normalized bump functions adapted to  $I$  with constant  $C > 0$  and order  $N$ , we have

$$(2) \quad |\langle T(\phi_I), \varphi_I \rangle| \lesssim C(L(2^{-M}|I|)S(2^M|I|)D(M^{-1} \text{rdist}(I, \mathbb{B}_{2^M})) + \epsilon),$$

where the implicit constant only depends on the operator  $T$ .

## REMARK 2.10

We note that in the main results of the article, namely, Theorems 2.18 and 3.8, when we say that  $T$  satisfies the weak compactness condition, we mean that there is an integer  $N \geq 1$  sufficiently large depending on the operator or its kernel so that the operator can be defined as  $T : \mathcal{S}_N(\mathbb{R}) \rightarrow \mathcal{S}'_N(\mathbb{R})$ , it is continuous with respect to the topology in  $\mathcal{S}_N(\mathbb{R})$ , and it satisfies Definition 2.9 for that value of  $N$ .

In [13] we discuss other equivalent formulations of this definition.

From now on, we will denote

$$F_K(I) = L_K(|I|)S_K(|I|)D_K(\text{rdist}(I, \mathbb{B}))$$

and

$$F_W(I; M) = L_W(2^{-M}|I|)S_W(2^M|I|)D_W(M^{-1} \text{rdist}(I, \mathbb{B}_{2^M})),$$

where  $L_K, S_K$ , and  $D_K$  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. Note that the value  $M = M_{T,\epsilon}$  depends not only on  $T$  but also on  $\epsilon$ .

We will also denote  $F(I; M) = F_K(I) + F_W(I; M)$ ,

$$F_K(I_1, \dots, I_n) = \left( \sum_{i=1}^n L_K(|I_i|) \right) \left( \sum_{i=1}^n S_K(|I_i|) \right) \left( \sum_{i=1}^n D_K(\text{rdist}(I_i, \mathbb{B})) \right),$$

$$F_W(I_1, \dots, I_n; M) = \left( \sum_{i=1}^n L_W(2^{-M}|I_i|) \right) \left( \sum_{i=1}^n S_W(2^M|I_i|) \right) \\ \times \left( \sum_{i=1}^n D_W(M^{-1} \text{rdist}(I_i, \mathbb{B}_{2^M})) \right),$$

and  $F(I_1, \dots, I_n; M) = F_K(I_1, \dots, I_n) + F_W(I_1, \dots, I_n; M)$ .

### 2.2. Characterization of compactness: the lagom projection operator

To prove our results about compact singular integral operators, we will use the following characterization of compact operators in a Banach space with a Schauder basis (see [5]).

#### THEOREM 2.11

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

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

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

#### DEFINITION 2.12

For every  $M \in \mathbb{N}$ , let  $\mathcal{I}_M$  be the family of intervals such that  $2^{-M} \leq |I| \leq 2^M$  and  $\text{rdist}(I, \mathbb{B}_{2^M}) \leq M$ . Let  $\mathcal{D}$  be the family of dyadic intervals of the real line, and let  $\mathcal{D}_M$  be the intersection of  $\mathcal{I}_M$  with  $\mathcal{D}$ . We call the intervals in  $\mathcal{I}_M$  and  $\mathcal{D}_M$  lagom intervals and dyadic lagom intervals, respectively.

Note that  $I \in \mathcal{D}_M$  implies that  $2^{-M}(2^M + |c(I)|) \leq M$  and then  $|c(I)| \leq (M - 1)2^M$ . Therefore,  $I \subset \mathbb{B}_{M2^M}$  with  $2^{-M} \leq |I|$ .

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

Let  $E$  be one of the following Banach spaces: the Lebesgue space  $L^p(\mathbb{R})$ ,  $1 < p < \infty$ , the Hardy space  $H^1(\mathbb{R})$ , or the space  $\text{CMO}(\mathbb{R})$ , to be introduced later as the closure in  $\text{BMO}(\mathbb{R})$  of continuous functions vanishing at infinity. In each case,  $E$  is equipped with smooth wavelet bases which are also Schauder

bases (see [6] and Lemma 2.16). Moreover, in all cases, we have at our disposal smooth and compactly supported wavelet bases.

**DEFINITION 2.13**

Let  $E$  be one of the previously mentioned Banach spaces. Let  $(\psi_I)_{I \in \mathcal{D}}$  be a wavelet basis of  $E$ . Then, for every  $M \in \mathbb{N}$ , we define the lagom projection operator  $P_M$  by

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

where  $\langle f, \psi_I \rangle = \int_{\mathbb{R}} f(x) \overline{\psi_I(x)} dx$ .

We also define the orthogonal lagom projection operator as  $P_M^\perp(f) = f - P_M(f)$ .

**REMARK 2.14**

Without explicit mention, we will let the wavelet basis defining  $P_M$  vary from proof to proof to suit our technical needs.

We also note the use of the same notation for the action of  $T(f)$  as a distribution and the inner product. We hope that this will not cause confusion.

It is easy to see that both  $P_M$  and  $P_M^\perp$  are self-adjoint operators. We note the difference with the usual projection operator  $\mathcal{P}_Q$ , for every interval  $Q \subset \mathbb{R}$ , defined by

$$(3) \quad \mathcal{P}_Q(f) = \sum_{I \in \mathcal{D}, I \subset Q} \langle f, \psi_I \rangle \psi_I,$$

which we will also use in forthcoming sections.

Let  $S$  denote the square function operator associated with a wavelet basis  $(\psi_I)_{I \in \mathcal{D}}$ ,

$$S(f)(x) = \left( \sum_{I \in \mathcal{D}} \frac{|\langle f, \psi_I \rangle|^2}{|I|} \chi_I(x) \right)^{1/2}.$$

Since we trivially have the pointwise estimates  $S(P_M(f))(x) \leq S(f)(x)$  and  $S(P_M^\perp(f))(x) \leq S(f)(x)$ , by Littlewood–Paley theory, we deduce that the lagom projection operator and its orthogonal projection are both continuous on  $L^p(\mathbb{R})$  for all  $1 < p < \infty$ . Moreover, the estimate

$$\|P_M(f)\|_{\text{BMO}(\mathbb{R})} = \sup_{Q \subset \mathbb{R}} \left( \frac{1}{|Q|} \sum_{I \in \mathcal{D}_M, I \subset Q} |\langle f, \psi_I \rangle|^2 \right)^{1/2} \leq \|f\|_{\text{BMO}(\mathbb{R})}$$

shows that  $P_M$  is bounded on  $\text{BMO}(\mathbb{R})$  and, by duality, on  $H^1(\mathbb{R})$  with  $\|P_M\|_{\text{BMO}(\mathbb{R}) \rightarrow \text{BMO}(\mathbb{R})} \leq 1$  and  $\|P_M\|_{H^1(\mathbb{R}) \rightarrow H^1(\mathbb{R})} \leq 1$ . For similar reasons, we have  $\|P_M^\perp\|_{\text{BMO}(\mathbb{R}) \rightarrow \text{BMO}(\mathbb{R})} \leq 1$  and  $\|P_M^\perp\|_{H^1(\mathbb{R}) \rightarrow H^1(\mathbb{R})} \leq 1$ .

We remark that in  $E$  the equality

$$(4) \quad P_M^\perp(f) = \sum_{I \in \mathcal{D}_M^c} \langle f, \psi_I \rangle \psi_I$$

is to be interpreted in its Schauder basis sense,

$$\lim_{M' \rightarrow \infty} \left\| P_M^\perp(f) - \sum_{I \in \mathcal{D}_{M'} \setminus \mathcal{D}_M} \langle f, \psi_I \rangle \psi_I \right\|_E = 0.$$

Note that according to Theorem 2.11 an operator  $T : E \rightarrow E$  is compact if and only if

$$\lim_{M \rightarrow \infty} \|P_M^\perp \circ T\| = 0,$$

where  $\|\cdot\|$  is the operator norm.

### 2.3. The space $\text{CMO}(\mathbb{R})$

We now provide the definition and main properties of the space to which the function  $T(1)$  must belong if  $T$  is compact.

#### DEFINITION 2.15

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

We note that  $\text{CMO}(\mathbb{R})$  equipped with the norm  $\|\cdot\|_{\text{BMO}}$  is a Banach space. The next lemma gives two characterizations of  $\text{CMO}(\mathbb{R})$ : the first in terms of the average deviation from the mean, and the second in terms of a wavelet decomposition. See [11] for the first, and see [8] for the second characterization.

#### LEMMA 2.16

(i)  $f \in \text{CMO}(\mathbb{R})$  if and only if  $f \in \text{BMO}(\mathbb{R})$  and

$$(5) \quad \lim_{M \rightarrow \infty} \sup_{I \notin \mathcal{I}_M} \frac{1}{|I|} \int_I \left| f(x) - \frac{1}{|I|} \int_I f(y) dy \right| dx = 0.$$

(ii)  $f \in \text{CMO}(\mathbb{R})$  if and only if  $f \in \text{BMO}(\mathbb{R})$  and

$$(6) \quad \lim_{M \rightarrow \infty} \sup_{Q \subset \mathbb{R}} \left( \frac{1}{|Q|} \sum_{I \notin \mathcal{D}_M, I \subset Q} |\langle f, \psi_I \rangle|^2 \right)^{1/2} = 0,$$

where the supremum is calculated over all intervals  $Q \subset \mathbb{R}$ .

As a consequence of the previous lemma,  $(\psi_I)_{I \in \mathcal{D}}$  is a Schauder basis for  $\text{CMO}(\mathbb{R})$ . We will mainly be using the latter formulation.

#### REMARK 2.17

Considering the comment after Definition 2.12, we see that the preceding lemma is also true if we, in line with [11], replace  $\mathcal{I}_M$  by  $\mathcal{I}'_M$  consisting of those intervals  $I$  such that  $2^{-M} \leq |I| \leq 2^M$  and  $|c(I)| \leq M/2$  and replace  $\mathcal{D}_M$  by  $\mathcal{D}'_M = \mathcal{I}'_M \cap \mathcal{D}$ .

We note that the remarks about (4) work well for the spaces  $L^p(\mathbb{R})$ ,  $H^1(\mathbb{R})$ , and also  $\text{CMO}(\mathbb{R})$ , but not for  $\text{BMO}(\mathbb{R})$ . The latter space is not separable, and so, it does not contain an unconditional basis. However, the characterization of the

norm in  $BMO(\mathbb{R})$  by a wavelet basis implies that for every  $f \in BMO(\mathbb{R})$  we have the equality

$$f = \sum_{I \in \mathcal{D}} \langle f, \psi_I \rangle \psi_I$$

with convergence in the weak\* topology  $\sigma(BMO(\mathbb{R}), H^1(\mathbb{R}))$ . This, in turn, implies the equality

$$(7) \quad P_M^\perp(f) = \sum_{I \in \mathcal{D}_M^c} \langle f, \psi_I \rangle \psi_I$$

with convergence in the same topology, which is interpreted as

$$\lim_{M' \rightarrow \infty} \left| \langle P_M^\perp(f), g \rangle - \sum_{I \in \mathcal{D}_{M'} \setminus \mathcal{D}_M} \langle f, \psi_I \rangle \langle \psi_I, g \rangle \right| = 0$$

for all  $g \in H^1(\mathbb{R})$ . See [9] and [1] for proofs and more details.

## 2.4. Main results

We now give the statement of the main result in the article and also the results from [13] which we will need.

### THEOREM 2.18

*Let  $T$  be a linear operator associated with a standard Calderón–Zygmund kernel. Then,  $T$  extends to a compact operator from  $L^\infty(\mathbb{R})$  into  $CMO(\mathbb{R})$  if and only if  $T$  is associated with a compact Calderón–Zygmund kernel,  $T$  satisfies the weak compactness condition, and  $T(1), T^*(1) \in CMO(\mathbb{R})$ . Moreover, with the extra assumption  $T(1) = T^*(1) = 0$ ,  $T$  is compact from  $BMO(\mathbb{R})$  into  $CMO(\mathbb{R})$ .*

The analogous result appearing in [13] is the following theorem.

### THEOREM 2.19

*Let  $T$  be a linear operator associated with a standard Calderón–Zygmund kernel. Then,  $T$  extends to a compact operator on  $L^p(\mathbb{R})$  for  $1 < p < \infty$  if and only if  $T$  is associated with a compact Calderón–Zygmund kernel,  $T$  satisfies the weak compactness condition, and  $T(1), T^*(1) \in CMO(\mathbb{R})$ .*

We now state the key ingredient in the proof of Theorem 2.19 and also Theorem 2.18: the so-called bump lemma, which describes the action of the operator over functions adapted to two different intervals.

Given two intervals  $I$  and  $J$ , we will denote  $K_{\min} = J$  and  $K_{\max} = I$  if  $|J| \leq |I|$ , while  $K_{\min} = I$  and  $K_{\max} = J$  otherwise.

### PROPOSITION 2.20

*Let  $K$  be a compact Calderón–Zygmund kernel with parameter  $\delta$ . Let  $N$  be sufficiently large and depend on  $\delta$  and  $0 < \theta < 1$ , and let  $0 < \delta' < \delta$  depend on  $N$ . Let*

$T : \mathcal{S}_N \rightarrow \mathcal{S}'_N$  be a linear operator associated with  $K$  satisfying the weak compactness condition with parameter  $N$  and the special cancellation conditions  $T(1) = 0$  and  $T^*(1) = 0$ .

Then, there exists  $C_{\delta'} > 0$  such that, for every  $\epsilon > 0$ , all intervals  $I, J$ , and all mean zero bump functions  $\psi_I, \psi_J$ ,  $L^2$ -adapted to  $I$  and  $J$ , respectively, with order  $N$  and constant  $C > 0$ , we have

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

where  $I_1 = I, I_2 = J, I_3 = \langle I, J \rangle, I_4 = \lambda_1 \tilde{K}_{\max}, I_5 = \lambda_2 \tilde{K}_{\max}, I_6 = \lambda_2 K_{\min}, \lambda_1 = |K_{\max}|^{-1} \operatorname{diam}(I \cup J), \lambda_2 = (|K_{\min}|^{-1} \operatorname{diam}(I \cup J))^\theta$ , and  $\tilde{K}_{\max}$  is the translate of  $K_{\max}$  with the same center as  $K_{\min}$ .

### 3. Endpoint estimates

In this section, we extend the study of compactness for singular integral operators to the endpoint case. Namely, we characterize those Calderón–Zygmund operators that extend compactly as maps from  $L^\infty(\mathbb{R})$  to  $\operatorname{CMO}(\mathbb{R})$ .

#### 3.1. Necessity of the hypotheses

The necessity of the hypotheses of Theorem 2.18 is a direct consequence of the following result.

**PROPOSITION 3.1**

*Let  $T$  be a linear operator associated with a standard Calderón–Zygmund kernel. If  $T$  is compact from  $L^\infty(\mathbb{R})$  into  $\operatorname{CMO}(\mathbb{R})$ , then  $T$  is compact on  $L^p(\mathbb{R})$  for  $1 < p < \infty$ .*

*Proof*

Since  $T$  is bounded from  $L^\infty(\mathbb{R})$  into  $\operatorname{CMO}(\mathbb{R})$  and it is associated with a standard Calderón–Zygmund kernel, by [7, p. 49],  $T$  is bounded on  $L^p(\mathbb{R})$  for all  $1 < p < \infty$ . Therefore, by interpolation,  $T$  is compact on  $L^p(\mathbb{R})$ . □

Since, in particular,  $T$  is compact on  $L^2(\mathbb{R})$ , by the results in [13] we have that the hypotheses of Theorem 2.18 are satisfied, that is,  $T$  is associated with a compact Calderón–Zygmund kernel,  $T$  satisfies the weak compactness condition, and  $T(1), T^*(1) \in \operatorname{CMO}(\mathbb{R})$ .

#### 3.2. Wavelet basis

We devote the first part of this section to describing the way to choose a wavelet basis of  $L^p(\mathbb{R})$  and  $H^1(\mathbb{R})$  and how we use this basis to decompose the operators under study. To do this, we will use the results contained in the books [3] and [6].

For every function  $\psi$  and every dyadic interval  $I = 2^{-j}[k, k + 1]$ ,  $j, k \in \mathbb{Z}$ , we denote

$$\psi_I(x) = \mathcal{T}_{l(I)} \mathcal{D}_{|I|}^2 \psi(x) = 2^{j/2} \psi(2^j x - k),$$

where  $l(I) = \min\{x : x \in I\}$ .

**THEOREM 3.2**

Let  $\psi \in L^2(\mathbb{R})$  with  $\|\psi\|_{L^2(\mathbb{R})} = 1$ . Then,  $\{\psi_I\}_{I \in \mathcal{D}}$  is an orthonormal wavelet basis of  $L^2(\mathbb{R})$  if and only if

$$\sum_{k \in \mathbb{Z}} |\hat{\psi}(\xi + k)|^2 = 1, \quad \sum_{k \in \mathbb{Z}} \hat{\psi}(2^j(\xi + k)) \overline{\hat{\psi}(\xi + k)} = 0$$

for all  $\xi \in \mathbb{R}$  and all  $j \geq 1$ .

**DEFINITION 3.3**

For any function  $f : \mathbb{R} \rightarrow \mathbb{C}$ , we say that a bounded function  $W : [0, \infty) \rightarrow \mathbb{R}^+$  is a radial decreasing  $L^1$ -majorant of  $f$  if  $|f(x)| \leq W(|x|)$  and  $W$  satisfies the following three conditions:  $W \in L^1([0, \infty))$ ,  $W$  is decreasing, and  $W(0) < \infty$ .

**THEOREM 3.4**

Let  $\psi \in L^2(\mathbb{R})$  be differentiable and such that  $\{\psi_I\}_{I \in \mathcal{D}}$  is an orthonormal basis of  $L^2(\mathbb{R})$ . We further assume that  $\psi$  and its derivative  $\psi'$  have a common radial decreasing  $L^1$ -majorant  $W$  satisfying

$$\int_0^\infty xW(x) dx < \infty.$$

Then, the system  $(\psi_I)_{I \in \mathcal{D}}$  is an unconditional basis for  $L^p(\mathbb{R})$  with  $1 < p < \infty$  and for  $H^1(\mathbb{R})$ .

Now, for our particular purposes, we will take  $\psi$  satisfying the hypotheses of previous theorems with the additional conditions that  $\psi \in C^N(\mathbb{R})$  and it is adapted to  $[-1/2, 1/2]$  with constant  $C > 0$  and order  $N$ . Then, we remark that, for every interval  $I \in \mathcal{D}$ , every wavelet function  $\psi_I$  is a bump function adapted to  $I$  with the same constant  $C > 0$  and the same order  $N$ . Several examples of constructions of systems of wavelets with any required order of differentiability can also be found in [6].

In the described setting, the continuity of  $T$  with respect to the topology of  $\mathcal{S}_N(\mathbb{R})$  allows us to write

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

for every  $f, g \in \mathcal{S}(\mathbb{R})$ , where the sums run over the whole family of dyadic intervals in  $\mathbb{R}$  and convergence is understood in the topology of  $\mathcal{S}_N(\mathbb{R})$ . Furthermore, since

$$\langle P_M(T(f)), g \rangle = \langle T(f), P_M g \rangle = \sum_{I \in \mathcal{D}} \sum_{J \in \mathcal{D}_M} \langle f, \psi_I \rangle \langle g, \psi_J \rangle \langle T(\psi_I), \psi_J \rangle,$$

we have that

$$(8) \quad \langle P_M^\perp(T(f)), g \rangle = \sum_{I \in \mathcal{D}} \sum_{J \in \mathcal{D}_M^c} \langle f, \psi_I \rangle \langle g, \psi_J \rangle \langle T(\psi_I), \psi_J \rangle,$$

where the summation is performed as in (4).

**3.3. Boundedness of a martingale transform**

We now study a new martingale transform. Its definition and the proof of its boundedness on  $L^p(\mathbb{R})$  for  $1 < p < \infty$  appear in the preprint [10]. We include here the endpoint result.

**DEFINITION 3.5**

Let  $(\psi_I)_{I \in \mathcal{D}}$  be a wavelet basis of  $L^2(\mathbb{R})$ . Given  $k \in \mathbb{Z}$  and  $n \in \mathbb{N}$ ,  $n \geq 1$ , let  $T_{k,n}$  be the operator defined by

$$T_{k,n}(f)(x) = \sum_{I \in \mathcal{D}} \sum_{J \in I_{k,n}} \langle f, \psi_J \rangle \psi_I(x),$$

where, for each fixed dyadic interval  $I$ ,  $I_{k,n}$  is the family of all dyadic intervals  $J$  such that  $|I| = 2^k|J|$  and  $n \leq \text{rdist}(I, J) < n + 1$ .

We remind the reader that for every dyadic interval  $I$  and each  $n \in \mathbb{N}$  there are  $2^{\max(e,0)+1}$  dyadic intervals  $J$  such that  $|I| = 2^e|J|$  and  $n \leq \text{rdist}(I, J) < n + 1$ . This implies that the cardinality of  $I_{e,n}$  is comparable to  $2^{\max(e,0)}$ .

In the proposition below, we prove the boundedness of this modified martingale operator.

**PROPOSITION 3.6**

Let  $k \in \mathbb{Z}$  and  $n \in \mathbb{N}$ ,  $n \geq 1$ . Then,  $T_{k,n}$  is bounded on  $\text{BMO}(\mathbb{R})$ . Moreover,

$$\|T_{k,n}f\|_{\text{BMO}(\mathbb{R})} \lesssim 2^{\frac{|k|}{2}} (\log(n + 1) + \max(-k, 0) + 1)^{\frac{1}{2}} \|f\|_{\text{BMO}(\mathbb{R})}$$

with implicit constant independent of  $f$ ,  $k$ , and  $n$ .

**REMARK 3.7**

By duality and the fact that  $T_{k,n}^* = T_{-k,n}$ , we have that  $T_{k,n}$  is also bounded on  $H^1(\mathbb{R})$  with

$$\|T_{k,n}f\|_{H^1(\mathbb{R})} \lesssim 2^{\frac{|k|}{2}} (\log(n + 1) + \max(k, 0) + 1)^{\frac{1}{2}} \|f\|_{H^1(\mathbb{R})}.$$

*Proof of Proposition 3.6*

Since for any given  $f \in \mathcal{S}(\mathbb{R})$

$$\|T_{k,n}f\|_{\text{BMO}(\mathbb{R})} = \sup_{Q \subset \mathbb{R}} \left( |Q|^{-1} \sum_{I \subset Q} \left| \sum_{J \in I_{k,n}} \langle f, \psi_J \rangle \right|^2 \right)^{\frac{1}{2}},$$

where the supremum is calculated over all intervals  $Q \subset \mathbb{R}$ , we will show that

$$(9) \quad \sum_{I \subset Q} \left| \sum_{J \in I_{k,n}} \langle f, \psi_J \rangle \right|^2 \lesssim 2^{|k|} (\log(n+1) + \max(-k, 0) + 1) \|f\|_{\text{BMO}(\mathbb{R})}^2 |Q|.$$

To compute the double sum, we use an argument that distinguishes between large and small scales ( $k \geq 0$  and  $k \leq 0$ ), with a slightly different argument in each case.

We first assume  $k \geq 0$ . In this case, the cardinality of  $I_{k,n}$  is comparable to  $2^k$ , and so, every interval  $I \in \mathcal{D}$  is associated with  $2^k$  different intervals  $J \in I_{k,n}$ . Therefore, by Cauchy’s inequality, the contribution of those intervals collected in the sum in (9) can be bounded by

$$(10) \quad \sum_{I \subset Q} 2^k \sum_{J \in I_{k,n}} |\langle f, \psi_J \rangle|^2 = 2^k \sum_{j \in \mathbb{N}} \sum_{I \in Q^j} \sum_{J \in I_{k,n}} |\langle f, \psi_J \rangle|^2,$$

where  $Q^j = \{I \in \mathcal{D} : I \subset Q, 2^{-(j+1)}|Q| < |I| \leq 2^{-j}|Q|\}$ .

Now, we separate again into two different cases: when  $J \subset 3Q$  and when  $J \not\subset 3Q$ .

(1) In the first case, we start by showing that the intervals  $J$  in the inner sum of (10) only appear at most five times. This will be clear once we prove that given  $I \in \mathcal{D}$  and  $J \in I_{k,n}$  there exist at most four different intervals  $I' \in \mathcal{D}$ ,  $I' \neq I$ , such that  $J \in I'_{k,n}$ .

If  $J \in I_{k,n} \cap I'_{k,n}$ , then  $|I| = 2^k |J| = |I'|$ . Now, we denote  $I_n = (I + n|I|) \cup (I - n|I|)$ . Since  $k \geq 0$ , we have that  $J \subset I_n \cap I'_n \neq \emptyset$ . Then, if  $n > 1$ , this implies  $\text{rdist}(I, I') = 2n$ , and so,  $I' = I + 2n|I|$  or  $I' = I - 2n|I|$ . On the other hand, if  $n = 1$ , this implies  $\text{rdist}(I, I') \in \{1, 2\}$ , and so,  $I' = I + |I|$ ,  $I' = I - |I|$ ,  $I' = I + 2|I|$ , or  $I' = I - 2|I|$ . Therefore, the terms in the inner sum of (10) corresponding to this case can be bounded by a constant times

$$2^k \sum_{J \subset 3Q} |\langle f, \psi_J \rangle|^2 \lesssim 2^k \|f\|_{\text{BMO}(\mathbb{R})}^2 |Q|,$$

which is compatible with the stated bound.

(2) In the second case, for those intervals  $I, J$  such that  $I \subset Q$  and  $J \not\subset 3Q$ , we have  $\text{diam}(I \cup J) > |Q|$ . Then, for every  $I \in Q^j$  we get

$$n + 1 > \text{rdist}(I, J) = \frac{\text{diam}(I \cup J)}{|I|} > \frac{|Q|}{|I|} \geq 2^j,$$

where we have used that  $|J| \leq |I|$ .

We now show that, for every  $j$ , the union of the disjoint intervals  $J \in I_{k,n}$  when varying  $I \in Q^j$  has measure at most  $2|Q|$ . For fixed  $I$ , the union of the disjoint intervals  $J \in I_{k,n}$  measures  $2|I|$ . Moreover, the union of the disjoint intervals  $I \in Q^j$  measures at most  $|Q|$ . Therefore,

$$\left| \bigcup_{I \in Q^j} \bigcup_{J \in I_{k,n}} J \right| \leq \sum_{I \in Q^j} \sum_{J \in I_{k,n}} |J| \leq 2 \sum_{I \in Q^j} |I| \leq 2|Q|.$$

In this way, the relevant contribution of this case to the sum in (10) can be bounded by

$$\begin{aligned} & 2^k \sum_{j=0}^{\log(n+1)} \sum_{I \in Q^j, J \in I_{k,n}} |\langle f, \psi_J \rangle|^2 \\ & \lesssim 2^k \|f\|_{\text{BMO}(\mathbb{R})}^2 \sum_{j=0}^{\log(n+1)} \left| \bigcup_{I \in Q^j, J \in I_{k,n}} J \right| \\ & \lesssim 2^k \|f\|_{\text{BMO}(\mathbb{R})}^2 \sum_{j=0}^{\log(n+1)} |Q| \\ & = 2^k (1 + \log(n + 1)) \|f\|_{\text{BMO}(\mathbb{R})}^2 |Q|, \end{aligned}$$

which is the desired bound when  $k \geq 0$ .

For  $k \leq 0$  we reason as follows. The cardinality of  $I_{k,n}$  is now essentially one, and there are at most  $2^j$  intervals  $I \in Q^j$ . In addition, up to  $2^{-k}$  different intervals  $I$  of fixed size in the sum (9) are associated with the same interval  $J$  and so have the same coefficient  $\langle f, \psi_J \rangle$ . Then, if we denote  $Q_{k,n}^j = \{J \in I_{k,n} : I \in Q^j\}$ , we have that the terms in the sum (9) corresponding to this case can be bounded by

$$(11) \quad \sum_{j \in \mathbb{N}} \sum_{J \in Q_{k,n}^j} \min(2^j, 2^{-k}) |\langle f, \psi_J \rangle|^2,$$

where now the intervals  $J \in Q_{k,n}^j$  appearing in the sum are pairwise different. Moreover, since  $|I| = 2^k |J|$  and  $2^{-(j+1)} |Q| < |I| \leq 2^{-j} |Q|$  we get  $2^{j+k} \leq |Q|/|J| < 2^{j+k+1}$ .

We separate the study into the same two cases as before:  $J \subset 3Q$  and  $J \not\subset 3Q$ .

(1) When  $I \subset Q$  and  $J \subset 3Q$  we have

$$n \leq \text{rdist}(I, J) = \frac{\text{diam}(I \cup J)}{|J|} \leq \frac{3|Q|}{|J|} < 2^{j+k+4}.$$

Therefore,  $j > \log n - k - 4 \geq -k - 4$ , and so, the contribution of the intervals in this case to the sum (11) can be bounded by

$$\begin{aligned} & \sum_{-k-4 < j \leq -k} \sum_{J \in Q_{k,n}^j, J \subset 3Q} 2^j |\langle f, \psi_J \rangle|^2 + \sum_{-k \leq j} \sum_{J \in Q_{k,n}^j, J \subset 3Q} 2^{-k} |\langle f, \psi_J \rangle|^2 \\ & \lesssim 2^{-k} \sum_{J \subset 3Q} |\langle f, \psi_J \rangle|^2 \\ & \lesssim 2^{-k} \|f\|_{\text{BMO}(\mathbb{R})}^2 |Q|. \end{aligned}$$

(2) On the other hand, for those  $J$  such that  $J \not\subset 3Q$  we have that

$$n + 1 > \text{rdist}(I, J) = \frac{\text{diam}(I \cup J)}{|J|} > \frac{|Q|}{|J|} \geq 2^{j+k},$$

and so,  $j < \log(n + 1) - k$ . Then, the contribution to sum (11) can be estimated by

$$\begin{aligned}
 (12) \quad & \sum_{j=0}^{\log(n+1)+|k|} \min(2^j, 2^{-k}) \sum_{J \in Q_{k,n}^j} |\langle f, \psi_J \rangle|^2 \\
 & \leq \sum_{j=0}^{\log(n+1)+|k|} \min(2^j, 2^{-k}) \|f\|_{\text{BMO}(\mathbb{R})}^2 \left| \bigcup_{J \in Q_{k,n}^j} J \right|.
 \end{aligned}$$

We now calculate the measure of the union of those intervals  $J \in Q_{k,n}^j$ . If  $j \geq -k$ , then from the  $2^j$  different possible intervals  $I \in Q^j$ , up to  $2^{-k}$  of them are associated with the same interval  $J$ . Then, the union of those intervals  $J$  has measure  $\frac{2^j}{2^{-k}}|J| \leq |Q|$ . On the other hand, when  $j < -k$ , there is only a single interval  $J$  associated with all intervals  $I \in Q^j$ , which measures  $|J| \leq |Q|2^{-(k+j)}$ . Then, the union has measure at most  $\max(2^{-k-j}, 1)|Q|$ , and thus, we bound (12) by

$$\begin{aligned}
 & \|f\|_{\text{BMO}(\mathbb{R})}^2 \sum_{j=0}^{\log(n+1)+|k|} 2^j \min(1, 2^{-k-j}) \max(2^{-k-j}, 1) |Q| \\
 & = \|f\|_{\text{BMO}(\mathbb{R})}^2 \sum_{j=0}^{\log(n+1)+|k|} 2^{-k} |Q| \\
 & = 2^{-k} (\log(n + 1) + |k| + 1) \|f\|_{\text{BMO}(\mathbb{R})}^2 |Q|.
 \end{aligned}$$

This finishes the proof. □

### 3.4. Sufficiency of the hypotheses: proof of endpoint compactness

In this section, we prove the compactness of singular integral operators  $T$  as maps from  $L^\infty(\mathbb{R})$  to  $\text{CMO}(\mathbb{R})$ .

To prove this result, we follow the scheme of the original proof of the  $T(1)$  theorem. Namely, we first assume that the special cancellation property  $T(1) = T^*(1) = 0$  holds, and then we tackle the general case with the use of paraproducts. Actually, we prove that under the special cancellation conditions the operator  $T$  extends compactly from  $\text{BMO}(\mathbb{R})$  into  $\text{CMO}(\mathbb{R})$ .

#### 3.4.1. The special case: $T(1) = T^*(1) = 0$

We start by proving the main result under the special cancellation conditions.

#### THEOREM 3.8

*Let  $T$  be a linear operator associated with a compact Calderón-Zygmund kernel satisfying the weak compactness condition and the special cancellation conditions  $T(1) = 0$  and  $T^*(1) = 0$ . Then,  $T$  can be extended to a compact operator from  $\text{BMO}(\mathbb{R})$  into  $\text{CMO}(\mathbb{R})$ .*

*Proof*

Let  $(\psi_I)_{I \in \mathcal{D}}$  be a wavelet basis of  $L^2(\mathbb{R})$  and  $H^1(\mathbb{R})$  with  $L^2$ -normalized elements. Let  $P_M$  be the lagom projection operator defined by this basis. By the remarks at the end of Section 2.3, we have that all functions in  $BMO(\mathbb{R})$  and  $H^1(\mathbb{R})$  can be approximated by functions in  $\mathcal{S}_N(\mathbb{R})$  with convergence in the weak\* topology  $\sigma(BMO(\mathbb{R}), H^1(\mathbb{R}))$  and in the  $H^1(\mathbb{R})$ -norm, respectively.

Then, by Theorem 2.11 with  $E = CMO(\mathbb{R})$  equipped with the norm of  $BMO(\mathbb{R})$ , to show the compactness of  $T$ , we need to check that  $P_M^\perp(T_b)$  converges to zero in the operator norm  $\|\cdot\|_{BMO(\mathbb{R}) \rightarrow BMO(\mathbb{R})}$  when  $M$  tends to infinity. For this, it is enough to prove that  $\langle P_M^\perp(T(f)), g \rangle$  tends to zero uniformly for all  $f, g \in \mathcal{S}_N(\mathbb{R})$  in the unit ball of  $BMO(\mathbb{R})$  and  $H^1(\mathbb{R})$ , respectively.

For  $f, g \in \mathcal{S}_N(\mathbb{R})$ , we recall

$$(13) \quad \langle P_M^\perp(T(f)), g \rangle = \sum_{I \in \mathcal{D}} \sum_{J \in \mathcal{D}_M^c} \langle f, \psi_I \rangle \langle g, \psi_J \rangle \langle T(\psi_I), \psi_J \rangle.$$

Since this inequality is understood as a limit, we can assume that the sums run over finite but arbitrary in size families of dyadic intervals, and we will work to obtain bounds that are independent of the cardinality of the families.

We start by proving that for every  $\epsilon > 0$  there is  $M_0 \in \mathbb{N}$  such that, for any  $M > M_0$ , we have  $F(I_1, \dots, I_6; M_{T,\epsilon}) \lesssim \epsilon$  for  $I_i \in \mathcal{D}_M^c$ . This will follow as a consequence of the inequality  $F(I; M_{T,\epsilon}) \lesssim \epsilon$  for every  $I \in \mathcal{D}_M^c$ . We note that the implicit constants only depend on the admissible functions.

We first remind the reader that, by the definition of the weak compactness condition, for  $\epsilon > 0$  there is a constant  $M_{T,\epsilon} > 0$  such that, for any interval  $I$  and any  $\phi_I, \varphi_I$  bump functions adapted to  $I$  with constant  $C > 0$  and order  $N$ , we have

$$|\langle T(\phi_I), \varphi_I \rangle| \leq C(F_W(I; M_{T,\epsilon}) + \epsilon).$$

We now show that there is  $M'_0 \in \mathbb{N}$  depending on  $\epsilon$  such that, for any  $M > M'_0$ , we have  $F(I_i; M_{T,\epsilon}) = F(I_1, \dots, I_6; M_{T,\epsilon}) \lesssim \epsilon$  when all  $I_i \in \mathcal{D}_M^c$ .

By the limit properties of the admissible functions in Definition 2.1, we have that, for fixed  $M_{T,\epsilon} > 0$ , there is  $M'_0 \in \mathbb{N}$ , depending on  $\epsilon, M_{T,\epsilon}$ , with  $M'_0 > M_{T,\epsilon}$ , such that for any  $M > M'_0$  we get

$$L_K(2^M) + S_K(2^{-M}) + D_K(M) < \epsilon$$

and

$$L_W(2^{M-M_{T,\epsilon}}) + S_W(2^{-(M-M_{T,\epsilon})}) + D_W(M/M_{T,\epsilon}) < \epsilon.$$

Let  $I \in \mathcal{D}_M^c$ . The claim is proven by considering the following cases.

(1) If  $|I| > 2^M$ , then since  $L_K$  and  $L_W$  are nonincreasing, we have

$$\begin{aligned} F(I; M_{T,\epsilon}) &\lesssim L_K(|I|) + L_W(|I|/2^{M_{T,\epsilon}}) \\ &\leq L_K(2^M) + L_W(2^{M-M_{T,\epsilon}}) \lesssim \epsilon. \end{aligned}$$

(2) If  $|I| < 2^{-M}$ , then since  $S_K$  and  $S_W$  are nondecreasing, we have

$$\begin{aligned} F(I; M_{T,\epsilon}) &\lesssim S_K(|I|) + S_W(2^{M_{T,\epsilon}}|I|) \\ &\leq S_K(2^{-M}) + S_W(2^{-(M-M_{T,\epsilon})}) \lesssim \epsilon. \end{aligned}$$

(3) If  $2^{-M} \leq |I| \leq 2^M$  with  $\text{rdist}(I, \mathbb{B}_{2^M}) > M$ , then as we saw in the remark after Definition 2.12,  $|c(I)| > (M - 1)2^M$ . Therefore,

$$\begin{aligned} \text{rdist}(I, \mathbb{B}_{2^{M_{T,\epsilon}}}) &\geq 1 + \frac{|c(I)|}{\max(|I|, 2^{M_{T,\epsilon}})} \\ &\geq 1 + \frac{(M - 1)2^M}{\max(2^M, 2^{M_{T,\epsilon}})} \geq M. \end{aligned}$$

We can apply similar reasoning to show that we also have  $\text{rdist}(I, \mathbb{B}) > M$ . Then, since  $D_W$  is nonincreasing, we have

$$\begin{aligned} F(I; M_{T,\epsilon}) &\lesssim D_K(\text{rdist}(I, \mathbb{B})) + D_W(M_{T,\epsilon}^{-1} \text{rdist}(I, \mathbb{B}_{2^{M_{T,\epsilon}}})) \\ &\leq D_K(M) + D_W(M/M_{T,\epsilon}) \lesssim \epsilon. \end{aligned}$$

Therefore, there is finally  $M_0 \in \mathbb{N}$  depending on  $\epsilon$  such that, for any  $M > M_0$ , we have

- (1)  $F(I_i; M_{T,\epsilon}) = F(I_1, \dots, I_6; M_{T,\epsilon}) \lesssim \epsilon$  when all  $I_i \in \mathcal{D}_M^\epsilon$ ,
- (2)  $M^{-\frac{\delta}{2}} + M^{\frac{3}{2}}2^{-M\frac{\delta}{2}} + \sum_{e \geq M} 2^{-e\delta}e^{1/2} < \epsilon$ .

Now, for every  $\epsilon > 0$  and chosen  $M_0 \in \mathbb{N}$ , we are going to prove that for all  $M > M_0$  we have

$$|\langle P_{2M}^\perp(T(f)), g \rangle| \lesssim \epsilon,$$

with the implicit constant depending on  $\delta > 0$  and the constant given by the wavelet basis.

We first parameterize the terms in (13) according to eccentricity and relative distance to obtain

$$(14) \quad \langle P_{2M}^\perp(T(f)), g \rangle = \sum_{e \in \mathbb{Z}} \sum_{n \in \mathbb{N}} \sum_{J \in \mathcal{D}_{2^e}^\epsilon} \sum_{I \in J_{e,n}} \langle f, \psi_I \rangle \langle g, \psi_J \rangle \langle T(\psi_I), \psi_J \rangle,$$

where for fixed eccentricity  $e \in \mathbb{Z}$ , relative distance  $n \in \mathbb{N}$ , and every given interval  $J$

$$J_{e,n} = \{I : |I| = 2^e|J|, n \leq \text{rdist}(I, J) < n + 1\}.$$

By Proposition 2.20 we have

$$|\langle T(\psi_I), \psi_J \rangle| \lesssim 2^{-|e|(\frac{1}{2} + \delta)} n^{-(1 + \delta)} (F(I_i; M_{T,\epsilon}) + \epsilon),$$

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 parameters  $\lambda_1, \lambda_2 \geq 1$  explicitly stated in the mentioned proposition. To simplify notation, we will simply write  $F(I_i)$ . We also note that the implicit constant might depend on  $\delta$  and the wavelet basis, but it is universal otherwise.

Therefore,

$$(15) \quad \begin{aligned} & |\langle P_{2M}^\perp(T(f)), g \rangle| \\ & \lesssim \sum_{e \in \mathbb{Z}} \sum_{n \in \mathbb{N}} 2^{-|e|(\frac{1}{2} + \delta)} n^{-(1+\delta)} \sum_{J \in \mathcal{D}_{2M}^c} \sum_{I \in J_{e,n}} (F(I_i) + \epsilon) |\langle f, \psi_I \rangle| |\langle g, \psi_J \rangle|. \end{aligned}$$

Now, to estimate (15), we divide the study into six cases:

- (1)  $I_i \notin \mathcal{D}_M$  for all  $i = 1, \dots, 6$ ;
- (2)  $I \in \mathcal{D}_M$ ;
- (3)  $\langle I \cup J \rangle \in \mathcal{D}_M$ ;
- (4)  $I \notin \mathcal{D}_M$  but  $\lambda_1 \tilde{K}_{\max} \in \mathcal{I}_M$ ;
- (5)  $I \notin \mathcal{D}_M$  but  $\lambda_2 \tilde{K}_{\max} \in \mathcal{I}_M$ ;
- (6)  $I \notin \mathcal{D}_M$  but  $\lambda_2 K_{\min} \in \mathcal{I}_M$ .

(1) In the first case we have  $F(I_i) < \epsilon$ , and thus, we can bound the contribution of the corresponding terms to (14) by

$$(16) \quad \epsilon \sum_{e \in \mathbb{Z}} \sum_{n \in \mathbb{N}} 2^{-|e|(\frac{1}{2} + \delta)} n^{-(1+\delta)} \sum_{J \in \mathcal{D}_{2M}^c} \sum_{I \in J_{e,n}} |\langle f, \psi_I \rangle| |\langle g, \psi_J \rangle|.$$

Since, as said, we consider that  $I$  and  $J$  run over finite families of intervals, we can define  $\tilde{f} = \sum_I |\langle f, \psi_I \rangle| \psi_I$  and  $\tilde{g} = \sum_J |\langle g, \psi_J \rangle| \psi_J$ , so that  $\langle \tilde{f}, \psi_I \rangle = |\langle f, \psi_I \rangle|$  and similarly for  $\tilde{g}$ . Then, for any interval  $Q \subset \mathbb{R}$  we have

$$\|\mathcal{P}_Q(\tilde{f})\|_{L^2(\mathbb{R})}^2 \leq \sum_{I \subset Q} |\langle f, \psi_I \rangle|^2 = \|\mathcal{P}_Q(f)\|_{L^2(\mathbb{R})}^2,$$

where  $\mathcal{P}_Q(f) = \sum_{I \subset Q} \langle f, \psi_I \rangle \psi_I$  is the classical projection operator defined in (3). Therefore,  $\|\tilde{f}\|_{\text{BMO}(\mathbb{R})} \leq \|f\|_{\text{BMO}(\mathbb{R})}$ , and by a duality argument, we also have  $\|\tilde{g}\|_{H^1(\mathbb{R})} \leq \|g\|_{H^1(\mathbb{R})}$ .

With this, we get for the inner sums in (16)

$$\begin{aligned} & \sum_{J \in \mathcal{D}_{2M}^c} \sum_{I \in J_{e,n}} |\langle f, \psi_I \rangle| |\langle g, \psi_J \rangle| \\ & \leq \sum_{J \in \mathcal{D}} \sum_{I \in J_{e,n}} |\langle f, \psi_I \rangle| |\langle g, \psi_J \rangle| = \sum_{J \in \mathcal{D}} \sum_{I \in J_{e,n}} \langle \tilde{f}, \psi_I \rangle \langle \tilde{g}, \psi_J \rangle \\ & = \left\langle \tilde{g}, \sum_{J \in \mathcal{D}} \sum_{I \in J_{e,n}} \langle \tilde{f}, \psi_I \rangle \psi_J \right\rangle = \langle \tilde{g}, T_{e,n}(\tilde{f}) \rangle \leq \|\tilde{g}\|_{H^1(\mathbb{R})} \|T_{e,n}(\tilde{f})\|_{\text{BMO}(\mathbb{R})} \\ & \leq 2^{\frac{|e|}{2}} (\log(n+1) + |e| + 1)^{\frac{1}{2}} \|\tilde{f}\|_{\text{BMO}(\mathbb{R})} \|\tilde{g}\|_{H^1(\mathbb{R})}, \end{aligned}$$

where the last inequality is due to Proposition 3.6. Note also that  $\log(n+1) \leq 2 \log n \lesssim 2\delta^{-1} n^\delta$ .

In this way, (16) can be bounded by a constant times

$$\begin{aligned} & \epsilon \sum_{e \in \mathbb{Z}} \sum_{n \in \mathbb{N}} 2^{-|e|(\frac{1}{2} + \delta)} n^{-(1+\delta)} 2^{\frac{|e|}{2}} (n^\delta + |e| + 1)^{\frac{1}{2}} \|f\|_{\text{BMO}(\mathbb{R})} \|g\|_{H^1(\mathbb{R})} \\ & \lesssim \epsilon \sum_{e \in \mathbb{Z}} 2^{-|e|\delta} |e|^{\frac{1}{2}} \sum_{n \geq 1} n^{-(1+\delta-\frac{\delta}{2})} \|f\|_{\text{BMO}(\mathbb{R})} \|g\|_{H^1(\mathbb{R})} \lesssim \epsilon \|f\|_{\text{BMO}(\mathbb{R})} \|g\|_{H^1(\mathbb{R})}. \end{aligned}$$

In the remaining cases, we will not use the smallness of  $F$ . Instead, we will use the particular geometric disposition of the intervals  $I$  and  $J$ , which make either their eccentricity or their relative distance very extreme. We recall that the intervals  $I$  and  $J$  in the sum (15) satisfy  $|I| = 2^e |J|$  and  $n \leq \text{rdist}(I, J) < n + 1$ .

(2) We deal first with the case when  $I \in \mathcal{D}_M$ , that is, when  $2^{-M} \leq |I| \leq 2^M$  and  $\text{rdist}(I, \mathbb{B}_{2^M}) \leq M$ . Notice that, since  $F$  is bounded, we can estimate  $F(I_i) + \epsilon \lesssim 1$ .

Since  $J \in \mathcal{D}_{2^M}^c$ , we separate the study into three cases:  $|J| > 2^{2M}$ ,  $|J| < 2^{-2M}$ , and  $2^{-2M} \leq |J| \leq 2^{2M}$  with  $\text{rdist}(J, \mathbb{B}_{2^{2M}}) > 2M$ .

(2.1) In the case  $|J| > 2^{2M}$ , since  $2^e |J| = |I| \leq 2^M$ , we have  $2^e \leq 2^M |J|^{-1} \leq 2^{-M}$ , that is,  $e \leq -M$ . Therefore, the calculations developed in case (1) allow us to bound the corresponding terms in (15) by

$$\begin{aligned} & \sum_{e \leq -M} \sum_{n \geq 1} 2^{-|e|(\frac{1}{2} + \delta)} n^{-(1+\delta)} \sum_{J \in \mathcal{D}_{2^M}^c} \sum_{I \in J_{e,n}} |\langle f, \psi_I \rangle| |\langle g, \psi_J \rangle| \\ & \leq \sum_{e \leq -M} \sum_{n \geq 1} 2^{-|e|(\frac{1}{2} + \delta)} n^{-(1+\delta)} 2^{\frac{|e|}{2}} (n^\delta + |e| + 1)^{\frac{1}{2}} \|f\|_{\text{BMO}(\mathbb{R})} \|g\|_{H^1(\mathbb{R})} \\ & \lesssim \left( \sum_{e \leq -M} 2^{-|e|\delta} |e|^{\frac{1}{2}} \sum_{n \geq 1} n^{-(1+\frac{\delta}{2})} \right) \|f\|_{\text{BMO}(\mathbb{R})} \|g\|_{H^1(\mathbb{R})} \\ & \lesssim \epsilon \|f\|_{\text{BMO}(\mathbb{R})} \|g\|_{H^1(\mathbb{R})} \end{aligned}$$

by the choice of  $M$ . This finishes this case.

(2.2) The case  $|J| < 2^{-2M}$  is symmetric and amounts to changing  $e \leq -M$  to  $e \geq M$  in the previous case.

(2.3) In the case when  $2^{-2M} \leq |J| \leq 2^{2M}$  and  $\text{rdist}(J, \mathbb{B}_{2^{2M}}) \geq 2M$ , we have that  $|J| = 2^k$  with  $-2M \leq k \leq 2M$  and  $|c(J)| \geq (2M - 1)2^{2M}$ . Since  $I \in \mathcal{D}_M$ , we also have

$$\begin{aligned} M & \geq \text{rdist}(I, \mathbb{B}_{2^M}) = 2^{-M} \text{diam}(I \cup \mathbb{B}_{2^M}) \\ & \geq 2^{-M} (2^{M-1} + |I|/2 + |c(I)|) \geq 2^{-M} (2^{M-1} + |c(I)|), \end{aligned}$$

and then,  $|c(I)| \leq (M - 1/2)2^M$ . This implies

$$|c(I) - c(J)| \geq |c(J)| - |c(I)| \geq (2M - 1)2^{2M} - (M - 1/2)2^M \geq M2^{2M}.$$

In this way, since  $\max(|I|, |J|) \leq 2^{2M}$ , we get

$$n + 1 > \text{rdist}(I, J) = \frac{\text{diam}(I \cup J)}{\max(|I|, |J|)} \geq \frac{|c(I) - c(J)|}{\max(|I|, |J|)} \geq 2^{-2M} M 2^{2M} = M.$$

Therefore, as in the previous case, we bound the relevant terms in (15) by a constant times

$$\begin{aligned} & \sum_{e \in \mathbb{Z}} \sum_{n \geq M-1} 2^{-|e|(\frac{1}{2} + \delta)} n^{-(1+\delta)} \sum_{J \in \mathcal{D}_{2^M}^c} \sum_{I \in J_{e,n}} |\langle f, \psi_I \rangle| |\langle g, \psi_J \rangle| \\ & \lesssim \left( \sum_{e \in \mathbb{Z}} 2^{-|e|\delta} |e|^{\frac{1}{2}} \sum_{n \geq M-1} n^{-(1+\frac{\delta}{2})} \right) \|f\|_{\text{BMO}(\mathbb{R})} \|g\|_{H^1(\mathbb{R})} \end{aligned}$$

$$\begin{aligned} &\lesssim M^{-\frac{\delta}{2}} \|f\|_{\text{BMO}(\mathbb{R})} \|g\|_{H^1(\mathbb{R})} \\ &< \epsilon \|f\|_{\text{BMO}(\mathbb{R})} \|g\|_{H^1(\mathbb{R})}, \end{aligned}$$

again by the choice of  $M$ .

(3) Now, we deal with the case when  $\langle I, J \rangle \in \mathcal{D}_M$ , that is, when  $2^{-M} \leq |\langle I, J \rangle| \leq 2^M$  and  $\text{rdist}(\langle I, J \rangle, \mathbb{B}_{2^M}) \leq M$ . Both inequalities imply that  $2^{-M} \leq \text{diam}(I \cup J) \leq 2^M$  and  $|c(\langle I, J \rangle)| \leq M2^M$ .

Moreover, we have that  $c(\langle I, J \rangle) = 1/2(c(I) + c(J) + \alpha(|I| - |J|))$  with  $\alpha \in [-1, 1]$ . Then,

$$(17) \quad \begin{aligned} |c(I) + c(J)| &\leq 2|c(\langle I, J \rangle)| + ||I| - |J|| \\ &\leq 2M2^M + |\langle I, J \rangle| \leq (2M + 1)2^M. \end{aligned}$$

(3.1) When  $|J| > 2^{2M}$  we have that  $|\langle I, J \rangle| \geq |J| > 2^{2M}$  implies  $\langle I, J \rangle \notin \mathcal{D}_M$ , and so, we do not need to consider this case.

(3.2) When  $2^{-2M} \leq |J| \leq 2^{2M}$  with  $\text{rdist}(J, \mathbb{B}_{2^{2M}}) \geq 2M$ , we have that  $|c(J)| > (2M - 1)2^{2M} > M2^{2M}$ . If  $\text{sign } c(I) = -\text{sign } c(J)$ , we have

$$\begin{aligned} |\langle I, J \rangle| &= \text{diam}(I \cup J) \geq |c(I) - c(J)| \\ &= |c(I)| + |c(J)| > |c(J)| > M2^{2M}, \end{aligned}$$

which contradicts  $\langle I, J \rangle \in \mathcal{D}_M$ . Otherwise, if  $\text{sign } c(I) = \text{sign } c(J)$ , we have

$$|c(I) + c(J)| = |c(I)| + |c(J)| > M2^{2M},$$

which now contradicts (17). So, we do not need to consider this case either.

(3.3) The remaining case is when  $|J| < 2^{-2M}$ . If  $e \geq 0$  then,

$$n + 1 > |I|^{-1} \text{diam}(I \cup J) = 2^{-e}|J|^{-1} |\langle I, J \rangle| \geq 2^{-e}2^{2M}2^{-M} = 2^{M-e}.$$

Meanwhile, if  $e \leq 0$ , we have

$$n + 1 > \text{rdist}(I, J) = |J|^{-1} \text{diam}(I \cup J) \geq 2^{2M}2^{-M} = 2^M.$$

Therefore, we bound the relevant part of (15) by a constant times

$$\begin{aligned} &\sum_{e \geq 0} \sum_{n \geq \max(2^{M-e}-1, 1)} 2^{-|e|(\frac{1}{2}+\delta)} n^{-(1+\delta)} \sum_{J \in \mathcal{D}_{2^M}^e} \sum_{I \in J_{e,n}} |\langle f, \psi_I \rangle| |\langle g, \psi_J \rangle| \\ &+ \sum_{e \leq 0} \sum_{n \geq 2^{M-1}} 2^{-|e|(\frac{1}{2}+\delta)} n^{-(1+\delta)} \sum_{J \in \mathcal{D}_{2^M}^e} \sum_{I \in J_{e,n}} |\langle f, \psi_I \rangle| |\langle g, \psi_J \rangle| \\ &\leq \left( \sum_{0 \leq e \leq M-1} 2^{-|e|\delta} |e|^{\frac{1}{2}} \sum_{n \geq 2^{M-e}-1} n^{-(1+\frac{\delta}{2})} + \sum_{M \leq e} 2^{-|e|\delta} |e|^{\frac{1}{2}} \sum_{n \geq 1} n^{-(1+\frac{\delta}{2})} \right. \\ &\quad \left. + \sum_{e \leq 0} 2^{-|e|\delta} |e|^{\frac{1}{2}} \sum_{n \geq 2^{M-1}} n^{-(1+\frac{\delta}{2})} \right) \|f\|_{\text{BMO}(\mathbb{R})} \|g\|_{H^1(\mathbb{R})} \\ &\lesssim \left( \sum_{0 \leq e \leq M-1} 2^{-e\delta} |e|^{\frac{1}{2}} 2^{-(M-e)\frac{\delta}{2}} + \sum_{M \leq e} 2^{-e\delta} |e|^{\frac{1}{2}} + 2^{-M\frac{\delta}{2}} \right) \|f\|_{\text{BMO}(\mathbb{R})} \|g\|_{H^1(\mathbb{R})} \end{aligned}$$

$$\begin{aligned} &\lesssim \left( 2^{-M\frac{\delta}{2}} M^{\frac{3}{2}} + \sum_{M \leq e} 2^{-e\delta} |e|^{\frac{1}{2}} + 2^{-M\frac{\delta}{2}} \right) \|f\|_{\text{BMO}(\mathbb{R})} \|g\|_{H^1(\mathbb{R})} \\ &\lesssim \epsilon \|f\|_{\text{BMO}(\mathbb{R})} \|g\|_{H^1(\mathbb{R})} \end{aligned}$$

by the choice of  $M$ .

(6) We deal now with the case  $\lambda_2 K_{\min} \in \mathcal{I}_M$ , that is,  $2^{-M} \leq |\lambda_2 K_{\min}| \leq 2^M$  and  $\text{rdist}(\lambda_2 K_{\min}, \mathbb{B}_{2^M}) \leq M$ .

(6.1) When  $|J| > 2^{2M}$ , we have two cases. Whenever  $e > 0$ , then  $K_{\min} = J$ , and so,  $|\lambda_2 J| \geq |J| \geq 2^{2M}$ , which contradicts  $\lambda_2 J \in \mathcal{I}_M$ . On the other hand, when  $e \leq 0$ , we have  $K_{\min} = I$  and  $|I| \leq |\lambda_2 I| \leq 2^M$ . Then,  $2^e = |I|/|J| \leq 2^{-M}$ , and so,  $e \leq -M$ . Therefore, the arguments of case (2.1) show that the corresponding part of (15) can be bounded by  $\epsilon \|f\|_{\text{BMO}(\mathbb{R})} \|g\|_{H^1(\mathbb{R})}$ .

(6.2) When  $2^{-2M} \leq |J| \leq 2^{2M}$  with  $\text{rdist}(J, \mathbb{B}_{2^{2M}}) \geq 2M$ , we have  $|c(J)| > (2M - 1)2^{2M}$ . Now, we divide into the same two cases. When  $e \geq 0$ , we know  $K_{\min} = J$ , and so,  $2^{-M} \leq |\lambda_2 J| \leq 2^M$  with  $\text{rdist}(\lambda_2 J, \mathbb{B}_{2^M}) \leq M$ . This leads to the following contradiction:

$$M \geq \text{rdist}(\lambda_2 J, \mathbb{B}_{2^M}) > 2^{-M} |c(J)| \geq (2M - 1)2^M.$$

On the other hand, when  $e \leq 0$ , we have  $K_{\min} = I$ , and then  $|c(I)| = |c(\lambda_2 I)| \leq (M - 1)2^M$ . This implies  $|c(I) - c(J)| > M2^{2M}$  and

$$n + 1 > \text{rdist}(I, J) \geq \frac{|c(I) - c(J)|}{|J|} \geq M.$$

Then, the same arguments developed in case (2.3) provide the bound  $\epsilon \|f\|_{\text{BMO}(\mathbb{R})} \|g\|_{H^1(\mathbb{R})}$ .

(6.3) When  $|J| < 2^{-2M}$ , we proceed as follows. If  $e \geq 0$ , we have  $K_{\min} = J$ , and so,  $|\lambda_2 J| \geq 2^{-M}$ . This implies  $\lambda_2 \geq 2^{-M} |J|^{-1} > 2^M$  and

$$2^M < \lambda_2 = \left( \frac{\text{diam}(I \cup J)}{|J|} \right)^\theta = \left( \frac{|I|}{|J|} \right)^\theta \text{rdist}(I, J)^\theta < 2^{e\theta} (n + 1)^\theta.$$

Meanwhile, if  $e \leq 0$ , we have  $K_{\min} = I$ , and then  $|\lambda_2 I| \geq 2^{-M}$ . We also have  $|I| \leq |J| \leq 2^{-2M}$ . All this implies  $\lambda_2 \geq 2^{-M} |I|^{-1} > 2^M$  and

$$2^M < \lambda_2 = \left( \frac{\text{diam}(I \cup J)}{|I|} \right)^\theta = \left( \frac{|J|}{|I|} \right)^\theta \text{rdist}(I, J)^\theta < 2^{-e\theta} (n + 1)^\theta.$$

Then, since  $\theta < 1$ , we get  $n + 1 > 2^{-|e|} 2^{\frac{M}{\theta}} > 2^{-|e|} 2^M$ , and so, previous arguments show that the relevant part of (15) can be bounded by

$$\begin{aligned} &\left( \sum_{e \in \mathbb{Z}} 2^{-|e|\delta} |e|^{\frac{1}{2}} \sum_{n \geq 2^{-|e|} 2^{M-1}} n^{-(1+\frac{\delta}{2})} \right) \|f\|_{\text{BMO}(\mathbb{R})} \|g\|_{H^1(\mathbb{R})} \\ &\lesssim \left( \sum_{e \in \mathbb{Z}} 2^{-|e|\delta} |e|^{\frac{1}{2}} 2^{|e|\frac{\delta}{2}} 2^{-M\frac{\delta}{2}} \right) \|f\|_{\text{BMO}(\mathbb{R})} \|g\|_{H^1(\mathbb{R})} \\ &\lesssim 2^{-M\frac{\delta}{2}} \|f\|_{\text{BMO}(\mathbb{R})} \|g\|_{H^1(\mathbb{R})} \leq \epsilon \|f\|_{\text{BMO}(\mathbb{R})} \|g\|_{H^1(\mathbb{R})}. \end{aligned}$$

Finally, we note that similar types of calculations are enough to deal with the two remaining cases (4) and (5). This completely finishes the proof of Theorem 3.8.  $\square$

### 3.5. The general case

For the proof of compactness in the general case, that is, without the special cancellation conditions, we follow the same scheme as in the proof of the classical  $T(1)$  theorem. When  $b_1 = T(1)$  and  $b_2 = T^*(1)$  are arbitrary functions in  $\text{CMO}(\mathbb{R})$ , we construct compact paraproducts  $T_b$  associated with compact Calderón–Zygmund kernels such that  $T_{b_1}(1) = b_1$ ,  $T_{b_1}^*(1) = 0$ . Then, the operator

$$\tilde{T} = T - T_{b_1} - T_{b_2}^*$$

satisfies the hypotheses of Theorem 3.8, and so,  $\tilde{T}$  is compact from  $\text{BMO}(\mathbb{R})$  to  $\text{CMO}(\mathbb{R})$ . Finally, since the operators  $T_{b_1}$  and  $T_{b_2}^*$  are compact from  $L^\infty(\mathbb{R})$  to  $\text{CMO}(\mathbb{R})$  by construction, we deduce that the initial operator  $T$  is also compact from  $L^\infty(\mathbb{R})$  to  $\text{CMO}(\mathbb{R})$ . We remark that, as we will later see in full detail, the appropriate paraproducts are exactly the same ones as in the classical setting, with the only difference being that the parameter functions  $b_i$  belong to the space  $\text{CMO}(\mathbb{R})$  instead of  $\text{BMO}(\mathbb{R})$ .

As in Proposition 3.9, we use a wavelet basis  $(\psi_I)_{I \in \mathcal{D}}$  of  $L^2(\mathbb{R})$  and  $H^1(\mathbb{R})$  such that each  $\psi_I$  is an  $L^2$ -normalized bump function supported and adapted to  $I$  with constant  $C$  and order  $N$ . We now denote by  $\phi$  a positive bump function supported and adapted to  $[-1/2, 1/2]$  with order  $N$  and integral 1. Then, we have that  $0 \leq \phi(x) \leq C(1 + |x|)^{-N}$  and  $|\phi'(x)| \leq C(1 + |x|)^{-N}$ . Let  $(\phi_I)_{I \in \mathcal{D}}$  be the family of bump functions defined by  $\phi_I = \mathcal{T}_{c(I)} \mathcal{D}_{|I|}^1 \phi$ . Therefore, each  $\phi_I$  is an  $L^1$ -normalized bump function adapted to  $I$ , that is, it satisfies  $\phi_I(x) \leq C|I|^{-1}(1 + |I|^{-1}|x - c(I)|)^{-N}$  and  $|\phi'_I(x)| \leq C|I|^{-2}(1 + |I|^{-1}|x - c(I)|)^{-N}$ .

**PROPOSITION 3.9**

Given  $b \in \text{CMO}(\mathbb{R})$ , we define the operator

$$T_b(f) = \sum_{I \in \mathcal{D}} \langle b, \psi_I \rangle \langle f, \phi_I \rangle \psi_I,$$

where  $\psi_I$  and  $\phi_I$  are as described above. Then,  $T_b$  and  $T_b^*$  are associated with a compact Calderón–Zygmund kernel, and they are both compact from  $L^\infty(\mathbb{R})$  to  $\text{CMO}(\mathbb{R})$ . Furthermore,  $\langle T_b(1), g \rangle = \langle b, g \rangle$  and  $\langle T_b(f), 1 \rangle = 0$  for all  $f, g \in \mathcal{S}(\mathbb{R})$ .

*Proof*

In [13] we showed that  $T_b$  and  $T_b^*$  belong to the class of operators for which the theory applies, that is, the integral representation of Definition 2.5 holds with operator kernel satisfying Definition 2.3 of a compact Calderón–Zygmund kernel.

For the proof of compactness of  $T_b$ , it is sufficient to verify that  $\langle P_M^\perp(T_b)(f), g \rangle$  tends to zero for all  $f \in L^\infty(\mathbb{R})$  and  $g \in \mathcal{S}(\mathbb{R})$  uniformly in the unit ball of  $L^\infty(\mathbb{R})$  and  $H^1(\mathbb{R})$ , respectively. Since  $g \in H^1(\mathbb{R})$ , we have  $P_M^\perp(g) = \sum_{I \in \mathcal{D}_M^c} \langle g, \psi_I \rangle \psi_I$ .

We note that, by the classical  $T(1)$ -theory, we already know that the operator is bounded from  $L^\infty(\mathbb{R})$  to  $\text{CMO}(\mathbb{R})$ , and so, the expression  $T_b(f)$  is completely meaningful.

Moreover, since  $(\psi_I)_{I \in \mathcal{D}}$  can be chosen so that it is also a wavelet basis on  $\text{CMO}(\mathbb{R})$  (see the comment in Lemma 2.16), we have  $P_M^\perp(b) \in \text{BMO}(\mathbb{R})$  and  $P_M^\perp(b) = \sum_{I \in \mathcal{D}_M^{\text{fin}}} \langle b, \psi_I \rangle \psi_I$ . With this,

$$\begin{aligned} \langle P_M^\perp(T_b(f)), g \rangle &= \langle T_b(f), P_M^\perp(g) \rangle = \sum_{I \in \mathcal{D}} \langle b, \psi_I \rangle \langle f, \phi_I \rangle \langle P_M^\perp(g), \psi_I \rangle \\ &= \sum_{I \in \mathcal{D}_M^{\text{fin}}} \langle b, \psi_I \rangle \langle f, \phi_I \rangle \langle g, \psi_I \rangle = \sum_{I \in \mathcal{D}} \langle P_M^\perp(b), \psi_I \rangle \langle f, \phi_I \rangle \langle g, \psi_I \rangle, \end{aligned}$$

that is,

$$(18) \quad \langle P_M^\perp(T_b(f)), g \rangle = \langle T_{P_M^\perp(b)}(f), g \rangle.$$

Then, the boundedness of  $T_{P_M^\perp(b)}$  from  $L^\infty(\mathbb{R})$  to  $\text{BMO}(\mathbb{R})$  implies

$$|\langle P_M^\perp(T_b(f)), g \rangle| \lesssim \|P_M^\perp(b)\|_{\text{BMO}(\mathbb{R})} \|f\|_{L^\infty(\mathbb{R})} \|g\|_{H^1(\mathbb{R})}.$$

Since  $\lim_{M \rightarrow \infty} \|P_M^\perp(b)\|_{\text{BMO}(\mathbb{R})} = 0$ , the inequality above finally proves that  $T_b(f)$  is compact from  $L^\infty(\mathbb{R})$  into  $\text{CMO}(\mathbb{R})$ .

The proof that  $T_b^*$  is compact from  $L^\infty(\mathbb{R})$  to  $\text{CMO}(\mathbb{R})$  is slightly different since  $T_b^*$  does not satisfy the analogue to (18). Thus, we prove instead the dual compactness for  $T_b$ . By (18) and the boundedness of  $T_b$  from  $H^1(\mathbb{R})$  to  $L^1(\mathbb{R})$ , we have

$$|\langle P_M^\perp(T_b(f)), g \rangle| = |\langle T_{P_M^\perp(b)}(f), g \rangle| \lesssim \|P_M^\perp(b)\|_{\text{BMO}(\mathbb{R})} \|f\|_{H^1(\mathbb{R})} \|g\|_{L^\infty(\mathbb{R})}.$$

This proves that  $T_b$  is compact from  $H^1(\mathbb{R})$  to  $L^1(\mathbb{R})$ , and so, by duality  $T_b^*$  is compact from  $L^\infty(\mathbb{R})$  to  $\text{BMO}(\mathbb{R})$ . But this obviously implies that  $\lim_{M \rightarrow \infty} \|P_M^\perp(T_b^*(f))\|_{\text{BMO}(\mathbb{R})} = 0$  uniformly in the unit ball of  $H^1(\mathbb{R})$ , and thus, the range of  $T_b^*$  is actually in  $\text{CMO}(\mathbb{R})$ .  $\square$

#### 4. Compactness of a perturbation of the Cauchy transform

In this section we apply our main theorem to demonstrate the compactness of a certain perturbation of the Cauchy transform for Lipschitz paths in the complex plane satisfying a CMO-condition. The example illustrates with special clarity the scope and methodology of the new theory, since the computations involved are essentially variations of the well-known calculations pertaining to the study of the Cauchy transform in the classical  $T(1)$ -theory. We note that a  $T(b)$ -theorem for compactness in several dimensions is already under development, and it could be of further use in the compactness theory of Cauchy-type operators. We start by giving the following definition.

##### DEFINITION 4.1

We denote by  $L_{\text{CMO}}^\infty(\mathbb{R})$  the closed subspace  $L^\infty(\mathbb{R}) \cap \text{CMO}(\mathbb{R})$  of  $L^\infty(\mathbb{R})$ .

Let  $A : \mathbb{R} \rightarrow \mathbb{R}$  be an absolutely continuous function such that  $A' \in L^\infty_{\text{CMO}}(\mathbb{R})$ , and let  $\Gamma \subset \mathbb{C}$  be the curve given by the parameterization  $z(t) = t + iA(t)$ ,  $t \in \mathbb{R}$ . Given points  $z(x), z(t) \in \Gamma$ , we denote by  $\sigma_{z(x),z(t)}, \tau_{z(x),z(t)} \in \Gamma$  the points

$$\sigma_{z(x),z(t)} = z\left(x - \frac{1}{4}(x - t)\right), \quad \tau_{z(x),z(t)} = z\left(x - \frac{3}{4}(x - t)\right)$$

lying in between  $z(x)$  and  $z(t)$  with respect to the parameterization of  $\Gamma$ .

The application we present concerns a perturbation of the Cauchy transform associated with  $\Gamma$ . Namely, define  $T_\Gamma : L^p(\Gamma) \rightarrow L^p(\Gamma)$  by

$$T_\Gamma f(z) = 2 \int_\Gamma \frac{f(w)}{z - w + 2\overline{(\sigma_{z,w} - \tau_{z,w})}} ds(w), \quad f \in L^p(\Gamma), z \in \Gamma,$$

where  $ds$  denotes the arc length measure on  $\Gamma$ . Note that if  $z = z(x)$  and  $w = z(t)$ , then

$$\operatorname{Re}(z - w + 2\overline{(\sigma_{z,w} - \tau_{z,w})}) = 2\operatorname{Re}(z - w)$$

and

$$\begin{aligned} & \operatorname{Im}(z - w + 2\overline{(\sigma_{z,w} - \tau_{z,w})}) \\ (19) \quad & = A(x) - A(t) - 2A\left(x - \frac{1}{4}(x - t)\right) + 2A\left(x - \frac{3}{4}(x - t)\right). \end{aligned}$$

In analogy to the Hilbert transform, we also introduce the operator  $H_\Gamma : L^p(\Gamma) \rightarrow L^p(\Gamma)$ ,

$$H_\Gamma f(z) = \int_\Gamma \frac{f(w)}{\operatorname{Re}(z - w)} ds(w), \quad f \in L^p(\Gamma), z \in \Gamma.$$

One might surmise that there is a sufficient amount of cancellation in (19) to cause  $T_\Gamma - H_\Gamma$  to be compact on  $L^p(\Gamma)$ . We will apply the results of this article to prove exactly this when  $\|A'\|_\infty$  is sufficiently small.

**PROPOSITION 4.2**

*Suppose that  $A : \mathbb{R} \rightarrow \mathbb{R}$  is absolutely continuous, and suppose that  $A' \in L^\infty_{\text{CMO}}(\mathbb{R})$ . Then, there exists an  $\eta > 0$  such that  $T_\Gamma - H_\Gamma$  is compact on  $L^p(\Gamma)$ ,  $1 < p < \infty$ , whenever  $\|A'\|_\infty < \eta$ .*

*Proof*

Moving over to the real line, we formally have that

$$(20) \quad (T_\Gamma f - H_\Gamma f)(z(t)) = \sum_{n=1}^\infty \left(\frac{-i}{2}\right)^n T_n(f \cdot \sqrt{1 + |A'|^2})(t),$$

where  $T_n : L^p(\mathbb{R}) \rightarrow L^p(\mathbb{R})$  is the operator associated to the kernel

$$K_n(x, t) = \frac{(A(x) - A(t) - 2A(x - \frac{1}{4}(x - t)) + 2A(x - \frac{3}{4}(x - t)))^n}{(x - t)^{n+1}}.$$

The expression for  $K_1$  is reminiscent of a double difference of  $A$ . Operators associated to such kernels have received attention by Coifman and Meyer [4].

Note that  $K_n$  is antisymmetric for each  $n$ . In what follows, we will prove that each  $K_n$  is a compact Calderón–Zygmund kernel and that  $T_n$  satisfies the weak compactness condition with appropriate bounds. Moreover, through an inductive procedure, we shall compute  $T_n(1)$  and check its membership in CMO. In this way, we will deduce that  $T_n$  is compact and obtain the existence of a constant  $C > 0$  such that  $\|T_n\|_{L^p \rightarrow L^p} \leq C^n \|A'\|_\infty^n$ . Then, by setting  $\eta = 2/C$ , we will have finally proven Proposition 4.2.

Note that we may write

$$K_n(x, t) = \frac{1}{x - t} \left( \frac{\int_t^x A'(z) dz}{x - t} - \frac{\int_{x-\frac{3}{4}(x-t)}^{x-\frac{1}{4}(x-t)} A'(z) dz}{\frac{1}{2}(x-t)} \right)^n.$$

The inner expression can be interpreted as the difference of two averages of  $A'$ . From the estimate

$$|f_I - f_J| \leq \frac{1}{|J|} \int_J |f - f_I| dt \leq \frac{2}{|I|} \int_I |f - f_I| dt,$$

where  $J \subset I$  are two intervals such that  $|I| = 2|J|$  and  $f_I$  denotes the average of  $f$  on  $I$ , we deduce

$$(21) \quad |K_n(x, t)| \lesssim 2^n \|A'\|_{\text{BMO}}^n \frac{1}{|x - t|}.$$

Demonstrating the smoothness condition of Definition 2.3 is more involved. Let  $x, t, t' \in \mathbb{R}$  with  $0 < 2|t - t'| \leq |x - t|$ . We denote  $G_n(x, t) = (x - t)^{n+1} K_n(x, t)$  for notational convenience and note that

$$K_n(x, t) - K_n(x, t') = G_n(x, t') \left( \frac{(x - t')^{n+1} - (x - t)^{n+1}}{(x - t)^{n+1}(x - t')^{n+1}} \right) + \frac{G_n(x, t) - G_n(x, t')}{(x - t)^{n+1}}.$$

Regarding the first term of this decomposition, there exists, by the mean value theorem, a  $\lambda$  between  $t$  and  $t'$ , and therefore satisfying  $|t - \lambda| < |t - t'|$ , such that

$$G_n(x, t') \left( \frac{(x - t')^{n+1} - (x - t)^{n+1}}{(x - t)^{n+1}(x - t')^{n+1}} \right) = (n + 1) \frac{G_n(x, t')}{(x - t')^n} \frac{(x - \lambda)^n}{(x - t)^n} \frac{t - t'}{(x - t)(x - t')}.$$

For  $M > 0$ , let

$$F_{1,n}(M) = \sup_{\substack{I_{x,t} \in \mathcal{I}_M^c \\ 2|t-t'| \leq |x-t|}} \left| \frac{G_n(x, t')}{(x - t')^n} \right|,$$

where  $I_{x,t}$  is the interval with endpoints  $x$  and  $t$ , and  $\mathcal{I}_M$  is the set of intervals  $I$  with center  $c(I)$  such that  $2^{-M} < |I| < 2^M$  and  $|c(I)| < \frac{M}{2}$ . Clearly,  $F_{1,n}$  is decreasing, and from the assumption that  $A' \in \text{CMO}$  in conjunction with the estimate (21), it follows that  $\|F_{1,n}\|_\infty \lesssim 2^n \|A'\|_{\text{BMO}}^n$  and  $\lim_{M \rightarrow \infty} F_{1,n}(M) = 0$ . This gives us control of the first term,

$$\begin{aligned} & \left| G_n(x, t') \left( \frac{(x - t')^{n+1} - (x - t)^{n+1}}{(x - t)^{n+1}(x - t')^{n+1}} \right) \right| \\ & \lesssim (n + 1) \left( \frac{3}{2} \right)^n F_{1,n}(\max(|\log_2(|x - t||, |x + t|)|)) \frac{|t - t'|}{|x - t|^2}. \end{aligned}$$

To deal with the second term we will consider the cases  $n = 1$  and  $n \geq 2$  separately. Suppose first that  $n \geq 2$ . By applying the mean value theorem, there exists a  $\lambda$  with

$$\begin{aligned} & \left| A(t) + 2A\left(x - \frac{1}{4}(x - t)\right) - 2A\left(x - \frac{3}{4}(x - t)\right) - \lambda \right| \\ & < |G_1(x, t) - G_1(x, t')| \end{aligned}$$

and such that

$$\frac{G_n(x, t) - G_n(x, t')}{(x - t)^{n+1}} = n \frac{G_1(x, t) - G_1(x, t')}{t - t'} \frac{(A(x) - \lambda)^{n-1}}{(x - t)^{n-1}} \frac{t - t'}{(x - t)^2}.$$

At this point the condition  $A' \in L^\infty$  comes into play, since it is necessary for estimating the first factor:

$$\left| \frac{G_1(x, t) - G_1(x, t')}{t - t'} \right| \lesssim \|A'\|_\infty.$$

On the other hand, introducing

$$F_{2,n}(M) = \sup_{\substack{I_{x,t} \in \mathcal{I}_M^c \\ 2|t-t'| \leq |x-t|}} \left| \frac{A(x) - \lambda}{x - t} \right|^{n-1},$$

we have that  $A' \in \text{CMO}$  again implies that  $\|F_{2,n}\|_\infty \lesssim 4^n \|A'\|_{\text{BMO}}^n$  and  $\lim_{M \rightarrow \infty} F_{2,n}(M) = 0$ . Therefore,

$$\left| \frac{G_n(x, t) - G_n(x, t')}{(x - t)^{n+1}} \right| \lesssim n \|A'\|_\infty F_{2,n}(\max(|\log_2(|x - t|)|, |x + t|)) \frac{|t - t'|}{|x - t|^2}.$$

When  $n = 1$ , the previous argument fails. Instead, we pick a  $\delta$ ,  $0 < \delta < 1$ , and write

$$\left| \frac{G_1(x, t) - G_1(x, t')}{(x - t)^2} \right| = \left| \frac{G_1(x, t) - G_1(x, t')}{t - t'} \right| \frac{|t - t'|^{1-\delta}}{|x - t|^{1-\delta}} \frac{|t - t'|^\delta}{|x - t|^{1+\delta}}.$$

Define

$$F_{2,1}(M) = \sup_{\substack{I_{x,t} \in \mathcal{I}_M^c \\ 2|t-t'| \leq |x-t|}} \left| \frac{G_1(x, t) - G_1(x, t')}{t - t'} \right| \frac{|t - t'|^{1-\delta}}{|x - t|^{1-\delta}}.$$

It is clear that  $\|F_{2,1}\|_\infty \lesssim \|A'\|_\infty$ . We prove now that  $\lim_{M \rightarrow \infty} F_{2,1}(M) = 0$ . Suppose that  $\overline{\lim}_{M \rightarrow \infty} F_{2,1}(M) = \ell > 0$ . Then there exists a sequence  $(M_k)$  with  $M_k \rightarrow \infty$  and corresponding sequences  $(x_k)$ ,  $(t_k)$ , and  $(t'_k)$  such that  $I_{x_k, t_k} \in \mathcal{I}_{M_k}^c$ ,  $2|t_k - t'_k| \leq |x_k - t_k|$ , and

$$(22) \quad \left| \frac{G_1(x_k, t_k) - G_1(x_k, t'_k)}{t_k - t'_k} \right| \frac{|t_k - t'_k|^{1-\delta}}{|x_k - t_k|^{1-\delta}} > \ell/2.$$

There could not exist a constant  $C > 0$  such that

$$\frac{|t_k - t'_k|}{|x_k - t_k|} \geq C$$

for all  $k$ , for then

$$\begin{aligned} & \left| \frac{G_1(x_k, t_k) - G_1(x_k, t'_k)}{t_k - t'_k} \right| \frac{|t_k - t'_k|^{1-\delta}}{|x_k - t_k|^{1-\delta}} \\ & \leq \frac{1^{2-\delta}}{C} \left| \frac{G_1(x_k, t_k) - G_1(x_k, t'_k)}{x_k - t_k} \right| \rightarrow 0, \quad k \rightarrow \infty, \end{aligned}$$

by the fact that  $A' \in \text{CMO}$ . Hence, it must be that

$$\liminf_{k \rightarrow \infty} \frac{|t_k - t'_k|}{|x_k - t_k|} = 0.$$

This also contradicts (22), however, since the first factor of the left-hand side is bounded, seeing as  $A' \in L^\infty$ . By this contradiction we conclude that  $\lim_{M \rightarrow \infty} F_{1,2}(M) = 0$ .

Appealing to these estimates and the antisymmetry of  $K_n$ , we may easily construct a set of admissible functions  $L_n, S_n$ , and  $D_n$  so that the conditions of Definition 2.3 are fulfilled with  $\delta = 1$  for  $n \geq 2$  and every  $\delta < 1$  for  $n = 1$ . Hence,  $K_n$  is a compact Calderón–Zygmund kernel.

We turn now to the verification of the weak compactness condition. For every compact interval  $I$  with center  $c(I)$ , we introduce the kernel

$$K_n^I(x, t) = |I|K_n(|I|x + c(I), |I|t + c(I))$$

and note that

$$(23) \quad |K_n^I(x, t)| \leq \frac{1}{|x - t|} F_{3,n}(\max(|\log_2(|I||x - t)|, |I|(x + t) + 2c(I))),$$

where

$$F_{3,n}(M) = \sup_{I_{x,t} \in \mathcal{T}_M^c} \left| \frac{G_n(x, t)}{(x - t)^n} \right|.$$

As before it is clear that  $F_{3,n}$  is a decreasing function with  $\|F_{3,n}\|_\infty \lesssim 2^n \|A'\|_{\text{BMO}}^n$  and  $\lim_{M \rightarrow \infty} F_{3,n}(M) = 0$ . For  $\phi \in \mathcal{S}(\mathbb{R})$ , we write  $\phi_I(x) = |I|^{-1/2} \phi(\frac{x - c(I)}{|I|})$ . Given  $\phi, \varphi \in \mathcal{S}(\mathbb{R})$ , we have by the antisymmetry of  $K_n$  that

$$\begin{aligned} 2\langle T_n \varphi_I, \phi_I \rangle &= \int_{\mathbb{R}^2} K_n(x, t) (\phi_I(x) \varphi_I(t) - \phi_I(t) \varphi_I(x)) \, dx \, dt \\ &= \int_{\mathbb{R}^2} K_n^I(x, t) (\phi(x) \varphi(t) - \phi(t) \varphi(x)) \, dx \, dt. \end{aligned}$$

Since  $|K_n^I(x, t)| \lesssim 2^n |x - t|^{-1}$  uniformly in  $I$ , there exists for each pair  $(\varphi, \phi)$  a constant  $C_{\varphi, \phi}$ , depending only on a finite number of Schwarz class seminorms, with the following property: for every  $\varepsilon > 0$  there is an  $M \geq 1$ , independent of  $\varphi$  and  $\phi$ , such that

$$\begin{aligned} & \left| \int_{\{|x-t|>M\} \cup \{|x-t|<\frac{1}{M}\} \cup \{|x+t|>M\}} K_n^I(x, t) (\phi(x) \varphi(t) - \phi(t) \varphi(x)) \, dx \, dt \right| \\ & \leq 2^n C_{\varphi, \phi} \varepsilon. \end{aligned}$$

To see this, simply note that

$$\frac{\phi(x)\varphi(t) - \phi(t)\varphi(x)}{x - t} \in \mathcal{S}(\mathbb{R}^2)$$

is a Schwarz function of two variables. Furthermore, in view of (23) we have

$$(24) \quad \left| \int_{\{\frac{1}{M} < |x-t| < M\} \cap \{|x+t| < M\}} K_n^I(x, t) (\phi(x)\varphi(t) - \phi(t)\varphi(x)) dx dt \right| \leq C_n C'_{\varphi, \phi} F_{3,n} \left( \max \left( \log_2 \left( \frac{|I|}{M} \right), -\log_2 (|I|M), \frac{\text{rdist}(I, \mathbb{D}_1)}{M} \right) \right),$$

for some constants  $C_n$  and  $C'_{\varphi, \phi}$ , depending only on  $n$  and a finite number of seminorms of  $\varphi$  and  $\phi$ , respectively. Note, in particular, that  $|x + t| < M$  and  $|c(I)| > M|I|$  imply that

$$||I|(x + t) + 2c(I)| > |c(I)| \gtrsim \frac{\text{rdist}(I, \mathbb{D}_1) - 1}{M}.$$

Together with the trivial facts that  $|I||x - t| > |I|/M$  and  $|I||x - t| < |I|M$  when  $(x, t)$  lies in the domain of integration of (24) and that  $\text{rdist}(I, \mathbb{D}_1)/M \leq 2$  when  $|c(I)| \leq M|I|$ , we obtain the desired estimate in (24) for an appropriate  $C_n$ . We conclude that  $T_n$  satisfies the weak compactness condition. For future reference we also record the implied bound on the weak boundedness constant of  $T_n$  present in the above considerations. Namely,

$$|\langle T_n \varphi_I, \phi_I \rangle| \lesssim 2^n \|A'\|_{\text{BMO}}^n C'_{\varphi, \phi}.$$

Finally, we shall show that  $T_n(1)$  belongs to CMO by evaluating it inductively in a principal value sense. The justifications for these computations are analogous to those that appear in considerations of Cauchy-type operators in connection with the classical  $T(1)$ -theory (see, e.g., Christ [2]).

For  $x$  with  $|x| < r < R$  and  $\varepsilon > 0$ , integrate by parts to obtain that

$$\begin{aligned} \int_{\substack{|t| < R \\ |t-x| > \varepsilon}} K_n(x, t) dt &= \frac{1}{n} \left[ \frac{G_n(x, t)}{(x-t)^n} \right]_{t=-R}^{t=R} - \frac{1}{n} \left[ \frac{G_n(x, t)}{(x-t)^n} \right]_{t=x-\varepsilon}^{t=x+\varepsilon} \\ &+ \int_{\substack{|t| < R \\ |t-x| > \varepsilon}} \frac{G_{n-1}(x, t)}{(x-t)^n} \\ &\times \left( A'(t) + \frac{1}{2} A' \left( x - \frac{1}{4}(x-t) \right) - \frac{3}{2} A' \left( x - \frac{3}{4}(x-t) \right) \right) dt, \end{aligned}$$

with the understanding that  $G_0 \equiv 1$ . Splitting the latter integral into three parts according to its summands and making the linear changes of variables  $x - t = 4(x - z)$  and  $x - t = \frac{4}{3}(x - w)$  in the last two terms we find that

$$(25) \quad \int_{\substack{|t| < R \\ |t-x| > \varepsilon}} K_n(x, t) dt = \frac{1}{n} \left[ \frac{G_n(x, t)}{(x-t)^n} \right]_{t=-R}^{t=R} - \frac{1}{n} \left[ \frac{G_n(x, t)}{(x-t)^n} \right]_{t=x-\varepsilon}^{t=x+\varepsilon} + \int_{\substack{|t| < R \\ |t-x| > \varepsilon}} \frac{G_{n-1}(x, t)}{(x-t)^n} A'(t) dt$$

$$\begin{aligned}
 &+ \frac{2}{4^n} \int_{\substack{|x-4(x-z)| < R \\ |z-x| > \varepsilon/4}} \frac{G_{n-1}(x, x-4(x-z))}{(x-z)^n} A'(z) dz \\
 &- 2 \left(\frac{3}{4}\right)^n \int_{\substack{|x-\frac{4}{3}(x-w)| < R \\ |w-x| > 3\varepsilon/4}} \frac{G_{n-1}(x, x-\frac{4}{3}(x-w))}{(x-w)^n} A'(w) dw.
 \end{aligned}$$

Since  $A' \in \text{CMO}$  it is clear that the first two terms tend to zero, uniformly for  $|x| < r$ , as  $\varepsilon \rightarrow 0$  and  $R \rightarrow \infty$ .

Suppose now that  $n = 1$ . Seeing as  $r$  is arbitrary, we then find in the limit that

$$T_1(1) = H(A') + \frac{1}{2}H(A') - \frac{3}{2}H(A') = 0,$$

where  $H$  denotes the usual Hilbert transform. Note that the Hilbert transform is bounded as a map  $H : \text{CMO} \rightarrow \text{CMO}$ .

At this point we have verified the compactness of  $T_1$  on  $L^p$ ,  $1 < p < \infty$ , and as a map  $T_1 : L^\infty_{\text{CMO}} \rightarrow \text{CMO}$ . We now proceed with the inductive step to prove the same for  $T_n$ ,  $n \geq 2$ . In this case, passing to the limit in (25) gives

$$(26) \quad T_n(1) = T_{n-1}(A') + \frac{2}{4^n} \tilde{T}_{n-1}(A') - 2 \left(\frac{3}{4}\right)^n \hat{T}_{n-1}(A'),$$

where  $\tilde{T}_{n-1}$  and  $\hat{T}_{n-1}$  are the operators associated to the kernels

$$\begin{aligned}
 \tilde{K}_{n-1}(x, z) &= \frac{G_{n-1}(x, x-4(x-z))}{(x-z)^n}, \\
 \hat{K}_{n-1}(x, w) &= \frac{G_{n-1}(x, x-\frac{4}{3}(x-w))}{(x-w)^n}.
 \end{aligned}$$

These kernels are very similar in character to  $K_{n-1}$ , and all computations performed up to this point can be repeated with minor modifications for them. In particular,  $\tilde{K}_{n-1}$  and  $\hat{K}_{n-1}$  are compact Calderón–Zygmund kernels,  $\tilde{T}_{n-1}$  and  $\hat{T}_{n-1}$  satisfy the weak compactness condition,  $\tilde{T}_1(1) = \hat{T}_1(1) = 0$ , and for  $n \geq 2$ , both  $\tilde{T}_n(1)$  and  $\hat{T}_n(1)$  are linear combinations of  $T_{n-1}(A')$ ,  $\tilde{T}_{n-1}(A')$ , and  $\hat{T}_{n-1}(A')$  with coefficients exponential in  $n$ .

Using these results and the fact that  $A' \in L^\infty \cap \text{CMO}$ , we obtain by induction that  $T_n : L^p \rightarrow L^p$ ,  $1 < p < \infty$ , and  $T_n : L^\infty_{\text{CMO}} \rightarrow \text{CMO}$  are compact maps for  $n \geq 1$ . Furthermore, by inspecting the constants in the above calculations and appealing to classical  $T(1)$ -theory (see [2]), we obtain bounds on the corresponding operator norms; there exists a constant  $C > 0$  such that

$$\|T_n\|_{L^p \rightarrow L^p} \leq C^n \|A'\|_\infty^n, \quad \|T_n\|_{L^\infty_{\text{CMO}} \rightarrow \text{CMO}} \leq C^n \|A'\|_\infty^n.$$

We conclude that  $T_\Gamma - H_\Gamma : L^p(\Gamma) \rightarrow L^p(\Gamma)$  is compact when  $\|A'\|_\infty < 2/C$ , hence finishing the proof of Proposition 4.2. □

## References

- [1] P. Auscher and P. Tchamitchian, *Bases d'ondelettes sur les courbes corde-arc, noyau de Cauchy et espaces de Hardy associés*, Rev. Mat. Iberoamericana **5** (1989), 139–170. [MR 1077208](#). [DOI 10.4171/RMI/89](#).
- [2] M. Christ, *Lectures on Singular Integral Operators*, CBMS Reg. Conf. Ser. Math. **77**, Amer. Math. Soc., Providence, 1990. [MR 1104656](#).
- [3] C. K. Chui, *An Introduction to Wavelets*, Academic Press, Boston, 1992. [MR 1150048](#).
- [4] R. R. Coifman and Y. Meyer, *On the Cauchy integral and related integral operators*, Zap. Nauchn. Sem. LOMI **81** (1978), 220–221.
- [5] M. Fabian, P. Habala, P. Hájek, V. Montesinos, and V. Zizler, *Banach Space Theory: The Basis for Linear and Nonlinear Analysis*, CMS Books Math./Ouvrages Math. SMC, Springer, New York, 2011. [MR 2766381](#). [DOI 10.1007/978-1-4419-7515-7](#).
- [6] E. Hernández and G. Weiss, *A First Course on Wavelets*, CRC Press, Boca Raton, FL, 1996. [MR 1408902](#). [DOI 10.1201/9781420049985](#).
- [7] J. L. Journé, *Calderón–Zygmund Operators, Pseudo-Differential Operators and the Cauchy Integral of Calderón*, Lecture Notes in Math. **994**, Springer, Berlin, 1983. [MR 0706075](#). [DOI 10.1007/BFb0061458](#).
- [8] M. T. Lacey, E. Terwilleger, and B. D. Wick, *Remarks on product VMO*, Proc. Amer. Math. Soc. **134** (2006), 465–474. [MR 2176015](#). [DOI 10.1090/S0002-9939-05-07974-8](#).
- [9] P. G. Lemarié and Y. Meyer, *Ondelettes et bases hilbertiennes*, Rev. Mat. Iberoamericana **2** (1986), 1–18. [MR 0864650](#). [DOI 10.4171/RMI/22](#).
- [10] S. Pott and P. Villarroya, *A  $T(1)$  theorem on product spaces*, preprint, [arXiv:1105.2516v6](#) [math.CA].
- [11] R. Rochberg, *Toeplitz and Hankel operators on the Paley–Wiener space*, Integral Equations Operator Theory **10** (1987), 187–235. [MR 0878246](#). [DOI 10.1007/BF01199078](#).
- [12] C. Thiele, *Wave Packet Analysis*, CBMS Reg. Conf. Ser. Math. **105**, Amer. Math. Soc., Providence, 2006. [MR 2199086](#). [DOI 10.1090/cbms/105](#).
- [13] P. Villarroya, *A characterization of compactness for singular integrals*, J. Math. Pures Appl. (9) **104** (2015), 485–532. [MR 3383175](#). [DOI 10.1016/j.matpur.2015.03.006](#).

*Perfekt*: Department of Mathematical Sciences, Norwegian University of Science and Technology (NTNU), NO-7491 Trondheim, Norway; [karl-mikael.perfekt@math.ntnu.no](mailto:karl-mikael.perfekt@math.ntnu.no)

*Pott*: Centre for Mathematical Sciences, University of Lund, 22100 Lund, Sweden; [sandra@maths.lth.se](mailto:sandra@maths.lth.se)

*Villarroya*: Centre for Mathematical Sciences, University of Lund, 22100 Lund, Sweden; [paco.villarroya@maths.lth.se](mailto:paco.villarroya@maths.lth.se)