SCIENCE CHINA
Mathematics

• ARTICLES •

June 2022 Vol. 65 No. 6: 1267–1292
https://doi.org/10.1007/s11425-020-1828-6

Variation operators for singular integrals and their commutators on weighted Morrey spaces and Sobolev spaces

Feng Liu* & Peng Cui

College of Mathematics and Systems Science, Shandong University of Science and Technology, Qingdao 266590*, China*

Email: fliu@sdust.edu.cn, cpengmath@126.com

Received September 26, 2020; accepted December 29, 2020; published online January 5, 2022

Abstract In this paper, we investigate the boundedness and compactness for variation operators of Calderón-Zygmund singular integrals and their commutators on weighted Morrey spaces and Sobolev spaces. To be precise, let $\rho > 2$ and K be a standard Calderón-Zygmund kernel. Denote by $\mathcal{V}_\rho(\mathcal{T}_K)$ and $\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)$ ($m \geqslant 1$) the ρ-variation operators of Calderón-Zygmund singular integrals and their m-th iterated commutators, respectively. By assuming that $\mathcal{V}_\rho(\mathcal{T}_K)$ satisfies an *a priori* estimate, i.e., the map $\mathcal{V}_\rho(\mathcal{T}_K): L^{p_0}(\mathbb{R}^n) \to L^{p_0}(\mathbb{R}^n)$ is bounded for some $p_0 \in (1, \infty)$, the bounds for $\mathcal{V}_\rho(\mathcal{T}_K)$ and $\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)$ on weighted Morrey spaces and Sobolev spaces are established. Meanwhile, the compactness properties of $\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)$ on weighted Lebesgue and Morrey spaces are also discussed. As applications, the corresponding results for the Hilbert transform, the Hermite Riesz transform, Riesz transforms and rough singular integrals as well as their commutators on the above function spaces are presented.

Keywords variation operator, Calderón-Zygmund singular integral, commutator, weighted Morrey space, Sobolev space

MSC(2020) 42B20, 42B99

Citation: Liu F, Cui P. Variation operators for singular integrals and their commutators on weighted Morrey spaces and Sobolev spaces. Sci China Math, 2022, 65: 1267–1292, https://doi.org/10.1007/s11425-020-1828-6

1 Introduction

The variational inequalities for various operators have been an active topic of current research. The first work was due to Lépingle [32] who established the variational inequality for general martingales (see also [42] for a simple proof). Motivated by the work [32], similar variation estimates were obtained by Bourgain [2] for the ergodic averages of a dynamic system. Since then, more and more scholars were devoted to studying variational inequalities for various operators. For examples, see [26, 27] for the ergodic averages, [35, 36] for the differential operators, [3, 18] for the Hilbert transform, [13, 18] for the Riesz transforms, [4, 5, 14, 28] for the singular integrals with rough kernels, [35, 36, 46, 47] for the Calderón-Zygmund singular integrals and their commutators, and [37,38] for the discrete singular integral operators.

* Corresponding author

Most of the function spaces considered in the above references are Lebesgue spaces and their weighted versions, and a natural question is whether variational inequalities hold for more general function spaces. In this paper, we establish the boundedness and compactness for variation operators of Calderón-Zygmund singular integrals and their commutators on weighted Morrey spaces and Sobolev spaces. This is the main motivation of this work.

1.1 Objectives of research

Let $\mathcal{T}=\{T_\epsilon\}_{\epsilon>0}$ be a family of bounded operators satisfying $\lim_{\epsilon\to0}T_\epsilon f(x)=Tf(x)$ almost everywhere for a certain class of functions f. For $\rho>2$, the ρ-variation operator of $\mathcal{T}$ is defined by

$$\mathcal{V}_\rho(\mathcal{T})(f)(x)=\sup_{\{\epsilon_i\}\searrow0}\bigg(\sum_{i=1}^\infty|T_{\epsilon_i}f(x)-T_{\epsilon_{i+1}}f(x)|^\rho\bigg)^{1/\rho},$$

where the supremum runs over all the sequences $\{\epsilon_i\}$ of positive numbers decreasing to zero.

Let $K(\cdot,\cdot)$ be a kernel defined on $\mathbb{R}^n\times\mathbb{R}^n\setminus\{(x,x):x\in\mathbb{R}^n\}$. We consider the following operator of Calderón-Zygmund type

$$T_K(f)(x)=\int_{\mathbb{R}^n}K(x,y)f(y)dy\quad\text{for all }x\notin\operatorname{supp}f. \tag{1.1}$$

Formally, the operator T_K can be rewritten as $T_K(f)(x)=\lim_{\epsilon\to0^+}T_{K,\epsilon}(f)(x)$, where $T_{K,\epsilon}$ is the truncated singular integral operator, i.e.,

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

The commutator of T_K with a suitable function b is defined as

$$T_{K,b}(f)(x):=[b,T_K](f)(x)=\int_{\mathbb{R}^n}(b(x)-b(y))K(x,y)f(y)dy=\lim_{\epsilon\to0^+}T_{K,b,\epsilon}(f)(x),$$

where

$$T_{K,b,\epsilon}(f)(x):=\int_{|x-y|>\epsilon}(b(x)-b(y))K(x,y)f(y)dy.$$

Let $T_{K,b}^1=T_{K,b}$. For $m\geqslant2$, the m-th iterated commutator $T_{K,b}^m$ is defined by

$$T_{K,b}^m(f)(x):=[b,T_{K,b}^{m-1}](f)(x)=\int_{\mathbb{R}^n}(b(x)-b(y))^mK(x,y)f(y)dy=\lim_{\epsilon\to0^+}T_{K,b,\epsilon}^m(f)(x), \tag{1.2}$$

where

$$T_{K,b,\epsilon}^m(f)(x):=\int_{|x-y|>\epsilon}(b(x)-b(y))^mK(x,y)f(y)dy.$$

Based on the above, the variation operators of Calderón-Zygmund singular integrals and their commutators can be defined as follows:

Definition 1.1 (Variation operators for singular integrals and their commutators). Let $\mathcal{T}_K=\{T_{K,\epsilon}\}_{\epsilon>0}$ and $\mathcal{T}_{K,b}^m=\{T_{K,b,\epsilon}^m\}_{\epsilon>0}$ with $m\geqslant1$. For $\rho>2$, the ρ-variation operator of $\mathcal{T}_K$ is defined by

$$\mathcal{V}_\rho(\mathcal{T}_K)(f)(x):=\sup_{\varepsilon_i\searrow0}\bigg(\sum_{i=1}^\infty\bigg|\int_{\varepsilon_{i+1}<|x-y|\leqslant\varepsilon_i}K(x,y)f(y)dy\bigg|^\rho\bigg)^{1/\rho}. \tag{1.3}$$

Similarly, the ρ-variation operator of $\mathcal{T}_{K,b}^m$ can be given as

$$\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f)(x):=\sup_{\varepsilon_i\searrow0}\bigg(\sum_{i=1}^\infty\bigg|\int_{\varepsilon_{i+1}<|x-y|\leqslant\varepsilon_i}(b(x)-b(y))^mK(x,y)f(y)dy\bigg|^\rho\bigg)^{1/\rho}, \tag{1.4}$$

where the above sup is taken over all the sequences $\{\varepsilon_i\}$ decreasing to zero. For convenience, we set $\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)=\mathcal{V}_\rho(\mathcal{T}_K)$ when $m=0$.

It should be pointed out that the operators defined in (1.1) and (1.2) have some classical models, which are listed as follows:

• When $n = 1$ and $K(x, y) = \frac{1}{x-y}$, T_K (resp. $T^m_{K,b}$) is (resp. the m-th order commutator of) the Hilbert transform. We define $\mathcal{T}_K = \mathcal{H}$ and $\mathcal{T}^m_{K,b} = \mathcal{H}^m_b$ for $m \geqslant 1$.

• When $n = 1$ and $K(x, y) = \mathfrak{R}^{\pm}(x, y)$, where $\mathfrak{R}^{\pm}(x, y)$ is an Hermite Riesz kernel whose expressions can be found in [43], T_K (resp. $T^m_{K,b}$) is (resp. the m-th order commutator of) the Hermite Riesz transform. We define $\mathcal{T}_K = \mathcal{R}^{\pm}$ and $\mathcal{T}^m_{K,b} = \mathcal{R}^m_{\pm,b}$ for $m \geqslant 1$.

• When $n \geqslant 2$ and $K(x, y) = R_j(x, y)$, where $R_j(x, y) := \Gamma(\frac{n+1}{2})\pi^{-\frac{n+1}{2}} \frac{x_j-y_j}{|x-y|^{n+1}}$ for $1 \leqslant j \leqslant n$, T_K (resp. $T^m_{K,b}$) is (resp. the m-th order commutator of) the Riesz transform. We define $\mathcal{T}_K = \mathcal{R}_j$ and $\mathcal{T}^m_{K,b} = \mathcal{R}^m_{j,b}$ for $m \geqslant 1$.

• When $n \geqslant 2$ and $K(x, y) = \frac{\Omega(x-y)}{|x-y|^n}$, where $\Omega \in L^1(\mathrm{S}^{n-1})$ is homogeneous of zero and satisfies $\int_{\mathrm{S}^{n-1}} \Omega(\theta)d\sigma(\theta) = 0$, T_K (resp. $T^m_{K,b}$) is just the usual (resp. the m-th order commutator of) singular integral operator with the rough kernel Ω. We define $\mathcal{T}_K = \mathcal{T}_\Omega$ and $\mathcal{T}^m_{K,b} = \mathcal{T}^m_{\Omega,b}$ for $m \geqslant 1$.

• When T_K is bounded on $L^2(\mathbb{R}^n)$ and the kernel K is a standard Calderón-Zygmund kernel, which satisfies the size condition

$$|K(x, y)| \leqslant \frac{A}{|x-y|^n} \quad \text{for } x \neq y \tag{1.5}$$

and the regularity conditions

$$|K(x, y) - K(z, y)| \leqslant \frac{A|x-z|^\delta}{|x-y|^{n+\delta}} \quad \text{for } |x-y| > 2|x-z|, \tag{1.6}$$

$$|K(y, x) - K(y, z)| \leqslant \frac{A|x-z|^\delta}{|x-y|^{n+\delta}} \quad \text{for } |x-y| > 2|x-z|, \tag{1.7}$$

where $\delta > 0$, then T_K (resp. $T^m_{K,b}$) is the (resp. m-th order commutator of) standard Calderón-Zygmund singular integral operator on $\mathbb{R}^n$.

Denote by $\mathrm{CZO}(\mathbb{R}^n)$ the set of all the Calderón-Zygmund singular integral operators on $\mathbb{R}^n$. It is well known that the Hilbert transform and the Hermite Riesz transform (see [43, Proposition 3.1]) belong to $\mathrm{CZO}(\mathbb{R})$ and the Riesz transform belongs to $\mathrm{CZO}(\mathbb{R}^n)$. Moreover, $T_\Omega \in \mathrm{CZO}(\mathbb{R}^n)$ when $\Omega \in \mathrm{Lip}_\alpha(\mathrm{S}^{n-1})$ for $\alpha > 0$.

Throughout this paper, we always assume that $\rho > 2$ since the ρ-variation in the case where $\rho \leqslant 2$ is often not bounded (see [1, 2]).

1.2 Variation operators for singular integrals

We attribute the developments of the variation operators for singular integrals to two stages.

Stage 1 ($n = 1$). The variation operators for singular integrals were originally studied by Campbell et al. [3] who proved that $\mathcal{V}_\rho(\mathcal{H})$ is of type (p, p) for $1 < p < \infty$ and of weak type $(1, 1)$. Subsequently, Gillespie and Torrea [18] extended the result of [3] to the weighted version and showed that $\mathcal{V}_\rho(\mathcal{H})$ is bounded on $L^p(w)$ for $1 < p < \infty$ and $w \in A_p(\mathbb{R})$. The same conclusion holds for $\mathcal{V}_\rho(\mathcal{R}^{\pm})$ (see [13, Theorem A]). Later on, Crescimbeni et al. [13] obtained the $L^1(w) \to L^{1,\infty}(w)$ bounds for $\mathcal{V}_\rho(\mathcal{H})$ and $\mathcal{V}_\rho(\mathcal{R}^{\pm})$ with $w \in A_1(\mathbb{R})$ (see [13, Theorem A]). Recently, Liu and Wu [33] established a weighted boundedness criterion: the operator $\mathcal{V}_\rho(\mathcal{T}_K)$ is bounded on $L^p(w)$ for $1 < p < \infty$ and $w \in A_p(\mathbb{R})$, provided that $n = 1$ and $\mathcal{V}_\rho(\mathcal{T}_K)$ is of type (p_0, p_0) for some $p_0 \in (1, \infty)$.

Stage 2 ($n \geqslant 1$). The higher-dimensional case began with Campbell et al. [4] in 2003 when they obtained the $L^p(\mathbb{R}^n)$ $(1 < p < \infty)$ bounds for $\mathcal{V}_\rho(\mathcal{T}_\Omega)$, provided that $\Omega \in L\log^+ L(\mathrm{S}^{n-1})$. This result was essentially improved by Ding et al. [14] to the case where $\Omega \in H^1(\mathrm{S}^{n-1})$ since $L\log^+ L(\mathrm{S}^{n-1}) \subsetneq H^1(\mathrm{S}^{n-1})$. The weighted result for $\mathcal{V}_\rho(\mathcal{T}_\Omega)$ was first considered by Ma et al. [36] who proved that $\mathcal{V}_\rho(\mathcal{T}_\Omega)$ is bounded on $L^p(w)$ for $1 < p < \infty$ and $w \in A_p(\mathbb{R}^n)$, provided that $\Omega \in \mathrm{Lip}_\alpha(\mathrm{S}^{n-1})$ for $\alpha > 0$. Later on, the above result was improved by Chen et al. [5] to the case where $\Omega \in L^q(\mathrm{S}^{n-1})$ for some $q > 1$. In [18], Gillespie and Torrea studied the variation operators for Riesz transforms and showed that $\mathcal{V}_\rho(\mathcal{R}_j)$ is bounded on $L^p(|x|^\alpha)$ for $1 < p < \infty$ and $-1 < \alpha < p - 1$. Recently, Zhang and Wu [47] extended the above result to the general A_p weight. Particularly, Ma et al. [36] established a criterion on weighted variation inequalities for Calderón-Zygmund singular integrals, which is the following theorem.

Theorem A (See [36]). *Let K be a kernel satisfying* (1.5)–(1.7) *and the a priori estimate*

$$\|\mathcal{V}_\rho(\mathcal{T}_K)(f)\|_{L^{p_0}(\mathbb{R}^n)} \lesssim_{n,p_0} \|f\|_{L^{p_0}(\mathbb{R}^n)} \tag{1.8}$$

hold for some $p_0 \in (1,\infty)$. Then $\mathcal{V}_\rho(\mathcal{T}_K)$ is bounded on $L^p(w)$ for all $1<p<\infty$ and $w \in A_p(\mathbb{R}^n)$.

1.3 Variation operators for commutators of singular integrals

It is well known that the commutator $[b,T](f) := bT(f) - T(bf)$ with the suitable operator T and the function b was initialized by Coifman et al. [11] who proved that the commutator $[b,\mathcal{R}_j]$ is bounded on $L^p(\mathbb{R}^n)$ for $1<p<\infty$ under the condition that $b \in \mathrm{BMO}(\mathbb{R}^n)$. This result was later improved by Uchiyama in his remarkable work [44], in which he showed that the commutator $[b,T_\Omega]$ with $\Omega \in \mathrm{Lip}_1(\mathrm{S}^{n-1})$ is bounded (resp. compactness) on $L^p(\mathbb{R}^n)$ for all $p \in (1,\infty)$ if and only if the symbol $b \in \mathrm{BMO}(\mathbb{R}^n)$ (resp. $b \in \mathrm{CMO}(\mathbb{R}^n)$). Here, $\mathrm{CMO}(\mathbb{R}^n)$ is the closure of $\mathcal{C}_c^\infty(\mathbb{R}^n)$ in the $\mathrm{BMO}(\mathbb{R}^n)$ topology, which coincides with the space of functions of vanishing mean oscillation. Since then, the boundedness and compactness of $[b,T]$ with sorts of operators T on various function spaces have been studied by many authors (see, e.g., [7,10,12,31]).

The variation operator for the commutators was first studied by Liu and Wu [33] who showed that $\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)$ is bounded on $L^p(w)$ for $1<p<\infty$ and $w \in A_p(\mathbb{R})$, provided that $m \geqslant 1$ and K satisfies (1.5)–(1.8) and $b \in \mathrm{BMO}(\mathbb{R})$. As applications, they obtained the $L^p(w)$ bounds for $\mathcal{V}_\rho(\mathcal{H}_b^m)$ and $\mathcal{V}_\rho(\mathcal{R}_{\pm,b}^m)$ for $1<p<\infty$ and $w \in A_p(\mathbb{R})$ if $b \in \mathrm{BMO}(\mathbb{R})$. Later on, the above results were extended by Zhang and Wu [46] to weighted Morrey spaces. In the general dimensional case where $n \geqslant 1$, Chen et al. [6] established the following result.

Theorem B (See [6]). *Let $m=1$, $\rho>2$, $b \in \mathrm{BMO}(\mathbb{R}^n)$ and K be a kernel satisfying* (1.5)–(1.8)*. Then for $1<p<\infty$ and $w \in A_p(\mathbb{R}^n)$,*

$$\|\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f)\|_{L^p(w)} \lesssim_{n,\rho,p} \|b\|_{\mathrm{BMO}(\mathbb{R}^n)}\|f\|_{L^p(w)}, \quad \forall\, f \in L^p(w).$$

It should be pointed out that Theorem B can be extended to arbitrary order commutators of Calderón-Zygmund singular integral operators by applying Theorem B, the method in the proof of [6, Theorem 1.1] and the induction arguments as in getting [15, Theorem 1]. Here, we only list the relevant result and leave the proof to the readers.

Theorem 1.2. *Let $m \geqslant 1$, $\rho>2$, $b \in \mathrm{BMO}(\mathbb{R}^n)$ and $\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)$ be defined as in* (1.4)*. Assume that the kernel K satisfies* (1.5)–(1.8)*. Then for $1<p<\infty$ and $w \in A_p(\mathbb{R}^n)$, we have*

$$\|\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f)\|_{L^p(w)} \lesssim_{n,\rho,p} \|b\|_{\mathrm{BMO}(\mathbb{R}^n)}^m\|f\|_{L^p(w)}, \quad \forall\, f \in L^p(w).$$

Recently, Guo et al. [21] studied the compactness for $\mathcal{V}_\rho(\mathcal{T}_{K,b})$ on $L^p(w)$, which can be formulated as follows.

Theorem C (See [21]). *Let $1<p<\infty$ and $w \in A_p(\mathbb{R}^n)$. Then $\mathcal{V}_\rho(\mathcal{T}_{K,b})$ is a compact operator on $L^p(w)$, provided that $b \in \mathrm{CMO}(\mathbb{R}^n)$ and K satisfies* (1.5)–(1.7) *and that $\mathcal{V}_\rho(\mathcal{T}_{K,b})$ is of type (p,p).*

1.4 Main motivations

The first question is based on Theorem 1.2 and Theorem C, which is the following question.

Question 1.3. Is the operator $\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)$ with $m \geqslant 2$ a compact operator on $L^p(w)$ for some $p \in (1,\infty)$ and $w \in A_p(\mathbb{R}^n)$ when $b \in \mathrm{CMO}(\mathbb{R}^n)$ and K satisfies (1.5)–(1.8)?

Our next question is related to weighted Morrey spaces. Let us recall the weighted Morrey spaces.

Definition 1.4 (Weighed Morrey spaces [30]). Let $1 \leqslant p < \infty$ and $0 \leqslant \beta < 1$. For a weight w defined on $\mathbb{R}^n$, the weighted Morrey space $M^{p,\beta}(w)$ is defined by

$$M^{p,\beta}(w) := \{f \in L^p_{\rm loc}(w) : \|f\|_{M^{p,\beta}(w)} < \infty\},$$

where

$$\|f\|_{M^{p,\beta}(w)} := \sup_{B \text{ balls in } \mathbb{R}^n} \left(\frac{1}{w(B)^\beta}\int_B |f(x)|^p w(x)dx\right)^{1/p}$$

with the supremum taken over all the balls in $\mathbb{R}^n$. Particularly, $M^{p,\beta}(w)$ is just the classical weighted Lebesgue space $L^p(w)$ when $\beta = 0$.

When $w \equiv 1$, $M^{p,\beta}(w)$ reduces to the classical Morrey space $M^{p,\beta}(\mathbb{R}^n)$, which was first introduced by Morrey [39] to study the local behavior of solutions to second order elliptic partial differential equations. The weighted Morrey spaces $M^{p,\beta}(w)$ were originally introduced by Komori and Shirai [30] who established the bounds for the Hardy-Littlewood maximal operator, the fractional integral operator and the Calderón-Zygmund singular integral operator on $M^{p,\beta}(w)$. Later on, more and more scholars have devoted to investigating the boundedness of various operators on $M^{p,\beta}(\mathbb{R}^n)$ (see, e.g., [16, 40, 41]).

Some questions arise naturally.

Question 1.5. Let $m \geqslant 1$, $p \in (1,\infty)$, $\beta \in (0,1)$, $w \in A_p(\mathbb{R}^n)$ and K satisfy (1.5)–(1.8). Are the operators $\mathcal{V}_\rho(\mathcal{T}_K)$ and $\mathcal{V}_\rho(\mathcal{T}^m_{K,b})$ bounded on $M^{p,\beta}(w)$? Is $\mathcal{V}_\rho(\mathcal{T}^m_{K,b})$ compact on $M^{p,\beta}(w)$?

The third question focuses on the regularity properties of variation operators of singular integrals and their commutators, which is inspired by the regularity theory of maximal operators. In 1997, Kinnunen [29] first studied the Sobolev regularity for maximal operators and showed that the centered Hardy-Littlewood maximal operator M is bounded on the Sobolev space $W^{1,p}(\mathbb{R}^n)$ for all $1 < p \leqslant \infty$, where $W^{1,p}(\mathbb{R}^n)$ is the first order Sobolev space, i.e.,

$$W^{1,p}(\mathbb{R}^n) := \{f : \mathbb{R}^n \to \mathbb{R} : \|f\|_{W^{1,p}(\mathbb{R}^n)} = \|f\|_{L^p(\mathbb{R}^n)} + \|\nabla f\|_{L^p(\mathbb{R}^n)} < \infty\},$$

where $\nabla f = (D_1 f, \ldots, D_n f)$ is the weak gradient of f. Later on, Kinnunen's work was extended and generalized to various variants. Particularly, Hajłasz and Onninen [23] established the following Sobolev boundedness criterion.

Theorem D (See [23]). *Let T be a sublinear operator and be bounded on $L^p(\mathbb{R}^n)$ for some $p \in (1,\infty)$. If T commutes with translations, i.e., $(Tf)_h(x) = T(f_h)(x)$ for all $x, h \in \mathbb{R}^n$, where $f_h(x) = f(x+h)$, then T is bounded on $W^{1,p}(\mathbb{R}^n)$.*

Note that the variation operators must not be linear, and it is natural to expect that it has similar regularity properties to maximal operators. Actually, it is not difficult to see that when the kernel K satisfies $K(x,y) = K(x-y)$, the operator $\mathcal{V}_\rho(\mathcal{T}_K)$ commutes with translations. Theorem D implies directly the $W^{1,p}$ bounds for $\mathcal{V}_\rho(\mathcal{T}_K)$ under the L^p bounds for $\mathcal{V}_\rho(\mathcal{T}_K)$ with some $p \in (1,\infty)$. Unfortunately, the operator $\mathcal{V}_\rho(\mathcal{T}^m_{K,b})$ does not commute with translations, even in the special case where $m = 1$ and $K(x,y) = K(x-y)$, which makes that Theorem D does not apply for $\mathcal{V}_\rho(\mathcal{T}^m_{K,b})$. Therefore, it is natural to ask the following question.

Question 1.6. Is the operator $\mathcal{V}_\rho(\mathcal{T}^m_{K,b})$ bounded on $W^{1,p}(\mathbb{R}^n)$?

1.5 Outline of this paper

The rest of this paper is organized as follows. In Section 2, we shall present some preliminary definitions and conclusions, including some properties for $A_p(\mathbb{R}^n)$ weights and $\mathrm{BMO}(\mathbb{R}^n)$ functions as well as a compactness criterion for operators on weighted Lebesgue and Morrey spaces, which are the main ingredients of concluding the bounds and compactness for variation operators of Calderón-Zygmund singular integrals and their commutators on weighted Lebesgue and Morrey spaces. In Section 3, we shall establish the boundedness for variation operators of Calderón-Zygmund singular integrals and their commutators on weighted Morrey spaces. The compactness for the above operators on weighted Lebesgue and Morrey spaces will be given in Section 4. Finally, the bounds for the above operators on Sobolev spaces will be proved in Section 5. The main ingredients of proving the bounds for the above operators on Sobolev spaces are to establish some suitable difference estimates for the above operators and some properties of Sobolev spaces.

2 Preliminaries

Throughout the paper, the letters C or c, sometimes with certain parameters, will stand for positive constants not necessarily the same at each occurrence, but are independent of the essential variables. If there exists a constant $c>0$ depending only on ϑ such that $A \leqslant cB$, we then write $A \lesssim_\vartheta B$ or $B \gtrsim_\vartheta A$, and if $A \lesssim_\vartheta$ and $B \lesssim_\vartheta A$, we then write $A \simeq_\vartheta B$. In what follows, we set $\mathbb{N} = \{0,1,\ldots\}$. The notation $Q(x,r)$ (resp. $B(x,r)$) represents the cube (resp. ball) in $\mathbb{R}^n$ centered at $x \in \mathbb{R}^n$ with the length of the side (resp. radius) $r>0$. For convenience, we define $|y|_\infty := \max_{1\leqslant j\leqslant n} |y_i|$ for every $y=(y_1,\ldots,y_n)$. For a cube Q and a function f defined on $\mathbb{R}^n$, we set $f_Q = \frac{1}{|Q|}\int_Q f(x)dx$.

We now give the definition of A_p weight class.

Definition 2.1 (A_p weight [19]). A weight is a nonnegative and locally integrable function on $\mathbb{R}^n$ that takes values in $(0,\infty)$ almost everywhere. For $1<p<\infty$, a weight w is said to be in the Muckenhoupt weight class $A_p(\mathbb{R}^n)$ if there exists a positive constant C such that

$$\sup_{Q \text{ cubes in } \mathbb{R}^n} \left(\frac{1}{|Q|}\int_Q w(x)dx\right)\left(\frac{1}{|Q|}\int_Q w(x)^{1-p'}dx\right)^{p-1} \leqslant C. \tag{2.1}$$

The smallest constant C in (2.1) is the corresponding A_p constant of w, which is denoted by $[w]_{A_p}$.

We now recall the definition of $\mathrm{BMO}(\mathbb{R}^n)$ space.

Definition 2.2 ($\mathrm{BMO}(\mathbb{R}^n)$ space [19]). The $\mathrm{BMO}(\mathbb{R}^n)$ space is given by

$$\mathrm{BMO}(\mathbb{R}^n) := \{f \in L^1_{\mathrm{loc}}(\mathbb{R}^n) : \|f\|_{\mathrm{BMO}(\mathbb{R}^n)} := \|M^\sharp f\|_{L^\infty(\mathbb{R}^n)} < \infty\},$$

where $M^\sharp f$ is the sharp maximal function, i.e., $M^\sharp f(x) = \sup_{Q\ni x} \frac{1}{|Q|}\int_Q |f(y)-f_Q|dy$ with the supremum taken over all the cubes Q in $\mathbb{R}^n$ that contain the given point x.

The following result presents some properties for $A_p(\mathbb{R}^n)$ weights and $\mathrm{BMO}(\mathbb{R}^n)$ functions, which are very useful in the proofs of our main results.

Lemma 2.3. *Let $1<p<\infty$ and $w \in A_p(\mathbb{R}^n)$. Then the following conclusions hold*:

(i) *There exists a constant $\theta \in (0,1)$ such that $w^{1+\theta} \in A_p(\mathbb{R}^n)$. Both θ and $[w^{1+\theta}]_{A_p}$ depend only on n, p and the A_p constant of w.*

(ii) *There exists a constant $t \in (1, \min\{p,2\})$ such that $w \in A_{p/t}(\mathbb{R}^n)$.*

(iii) *The measure $w(x)dx$ is doubling, i.e., for all $\lambda>1$ we have*

$$\sup_{\mathrm{Q\ cubes\ in\ } \mathbb{R}^n} \frac{w(\lambda Q)}{w(Q)} \leqslant [w]_{A_p}\lambda^{np}.$$

Moreover, there exists a constant $\gamma_w > 1$ such that

$$\inf_{\mathrm{Q\ cubes\ in\ } \mathbb{R}^n} \frac{w(2Q)}{w(Q)} \geqslant \gamma_w.$$

(iv) *There exists a constant $\epsilon>0$ depending only on n and $[w]_{A_p}$ such that for all cubes Q and all measurable subsets A of Q,*

$$\frac{w(A)}{w(Q)} \lesssim_{n,[w]_{A_p}} \left(\frac{|A|}{|Q|}\right)^\epsilon.$$

(v) *Let $b \in \mathrm{BMO}(\mathbb{R}^n)$. Then*

$$\sup_{Q \text{ cubes in } \mathbb{R}^n} \left(\frac{1}{w(Q)}\int_Q |b(x)-b_Q|^p w(x)dx\right)^{1/p} \simeq_{p,[w]_{A_p}} \|b\|_{\mathrm{BMO}(\mathbb{R}^n)}.$$

We remark that Lemma 2.3(i) follows from [9], Lemma 2.3(ii) follows from [30] and Lemma 2.3(iv) follows from [19]. Lemma 2.3(iii) follows by modifying the proofs of [34, Theorem 1.4.1 and Proposition 1.4.2(viii)], and the details are omitted. Lemma 2.3(v) is a well-known property for BMO functions (see [24, 25] for the recent developments).

For convenience, we always use the weighted Morrey spaces associated with cubes. Let $1 \leqslant p < \infty$ and $0 \leqslant \beta < 1$. For a weight w defined on $\mathbb{R}^n$, the weighted Morrey space associated with cubes is defined by $\widetilde{M}^{p,\beta}(w) := \{f \in L^p_{\rm loc}(w) : \|f\|_{\widetilde{M}^{p,\beta}(w)} < \infty\}$, where

$$\|f\|_{\widetilde{M}^{p,\beta}(w)} := \sup_{Q \text{ cubes in } \mathbb{R}^n} \left(\frac{1}{w(Q)^\beta}\int_Q |f(x)|^p w(x)dx\right)^{1/p}$$

with the supremum taken over all the cubes in $\mathbb{R}^n$.

Remark 2.4. If the weight w is doubling, then we have $\widetilde{M}^{p,\beta}(w) = M^{p,\beta}(w)$, i.e., $\|f\|_{\widetilde{M}^{p,\beta}(w)} \simeq \|f\|_{M^{p,\beta}(w)}$, which can be seen by the doubling property for w and the observation

$$Q(x_0, r) \subset B(x_0, \sqrt{n}/2r) \subset Q(x_0, \sqrt{n}r), \quad \forall\, x_0 \in \mathbb{R}^n, \quad r > 0.$$

To prove the compactness result, we need the following characterization that a subset in $M^{p,\beta}(w)$ is a strongly pre-compact set, which is a direct application of [22, Theorem 3.1].

Proposition 2.5. *Let $1 < p < \infty$, $0 \leqslant \beta < 1$ and $w \in A_p(\mathbb{R}^n)$. Then a subset $\mathfrak{F}$ of $M^{p,\beta}(w)$ is a strongly pre-compact set in $M^{p,\beta}(w)$ if $\mathfrak{F}$ satisfies the following conditions*:

(i) $\mathfrak{F}$ *is bounded, i.e.,* $\sup_{f\in\mathfrak{F}} \|f\|_{M^{p,\beta}(w)} < \infty$;

(ii) $\mathfrak{F}$ *uniformly vanishes at infinity, i.e.,*

$$\lim_{N\to+\infty} \|f\chi_{E_N}\|_{M^{p,\beta}(w)} = 0 \quad \textit{uniformly for all } f \in \mathfrak{F},$$

where $E_N = \{x \in \mathbb{R}^n;\ |x| > N\}$.

(iii) $\mathfrak{F}$ *is uniformly translation-continuous, i.e.,*

$$\lim_{r\to 0} \sup_{h\in B(0,r)} \|f(\cdot + h) - f(\cdot)\|_{M^{p,\beta}(w)} = 0 \quad \textit{uniformly for all } f \in \mathfrak{F}.$$

Remark 2.6. When $\beta = 0$, Proposition 2.5 is just the weighted Fréchet-Kolmogorov theorem following from [8]. When $0 < \beta < 1$, Proposition 2.5 represents a weight version of [7, Theorem 1.12].

We end this section by presenting a useful inequality

$$\left(\sum_{i=1}^{\infty}\left|\int_{\varepsilon_{i+1}<f(x,y)\leqslant\varepsilon_i} F(x,y)dy\right|^\rho\right)^{1/\rho} \leqslant \int_{\mathbb{R}^n} |F(x,y)|dy \tag{2.2}$$

for all $x \in \mathbb{R}^n$, and any arbitrary functions F and f defined on $\mathbb{R}^n \times \mathbb{R}^n$, where $\rho > 1$ and $\{\varepsilon_i\}$ is an increasing or decreasing sequence of positive numbers.

3 Boundedness on weighted Morrey spaces

3.1 A boundedness criterion on weighted Morrey spaces

The following is a criterion on the boundedness for variation operators of Calderón-Zygmund singular integrals and their commutators on weighted Morrey spaces.

Theorem 3.1. *Let $m \in \mathbb{N}$, $\rho > 2$, $b \in \mathrm{BMO}(\mathbb{R}^n)$ and $\mathcal{V}_\rho(\mathcal{T}^m_{K,b})$ be defined as in* (1.4)*. Assume that the kernel K satisfies* (1.5)–(1.8)*. Then for $1 < p < \infty$, $w \in A_p(\mathbb{R}^n)$ and $0 < \beta < 1$,*

$$\|\mathcal{V}_\rho(\mathcal{T}^m_{K,b})(f)\|_{M^{p,\beta}(w)} \lesssim_{n,m,\rho,p,[w]_{A_p}} \|b\|^m_{\mathrm{BMO}(\mathbb{R}^n)} \|f\|_{M^{p,\beta}(w)}, \quad \forall\, f \in M^{p,\beta}(w).$$

In fact, Theorem 3.1 can be deduced by Theorem 1.2 and the following proposition, which is of interest in its own right.

Proposition 3.2. *Let $m \in \mathbb{N}$, $\rho > 2$, $b \in \mathrm{BMO}(\mathbb{R}^n)$ and $\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)$ be defined as in* (1.4). *Assume that K satisfies* (1.5) *and*

$$\|\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f)\|_{L^p(w)} \leqslant B_p \|b\|_{\mathrm{BMO}(\mathbb{R}^n)}^m \|f\|_{L^p(w)} \tag{3.1}$$

for some $p \in (1,\infty)$ and $w \in A_p(\mathbb{R}^n)$. Here, $B_p := \frac{\|\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)\|_{L^p(w)\to L^p(w)}}{\|b\|_{\mathrm{BMO}(\mathbb{R}^n)}^m}$. Then for any $\beta \in (0,1)$,

$$\|\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f)\|_{M^{p,\beta}(w)} \lesssim_{n,p,\beta,m,A,[w]_{A_p},B_p} \|b\|_{\mathrm{BMO}(\mathbb{R}^n)}^m \|f\|_{M^{p,\beta}(w)}, \quad \forall f \in M^{p,\beta}(w). \tag{3.2}$$

Proof. Let $f \in \widetilde{M}^{p,\beta}(w)$ and $\beta \in (0,1)$. Fix a cube $Q = Q(x_0, r)$. To prove (3.2), by Remark 2.4, it suffices to show that

$$\left(\frac{1}{w(Q)^\beta}\int_Q |\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f)(x)|^p w(x)dx\right)^{1/p} \leqslant C\|b\|_{\mathrm{BMO}(\mathbb{R}^n)}^m \|f\|_{\widetilde{M}^{p,\beta}(w)}, \tag{3.3}$$

where $C > 0$ is independent of x_0, r and b.

Decompose f as $f = f\chi_{2Q} + f\chi_{(2Q)^c}$. It is clear that

$$\begin{aligned}\left(\frac{1}{w(Q)^\beta}\int_Q |\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f)(x)|^p w(x)dx\right)^{1/p} &\leqslant \left(\frac{1}{w(Q)^\beta}\int_Q |\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f\chi_{2Q})(x)|^p w(x)dx\right)^{1/p} \\ &\quad + \left(\frac{1}{w(Q)^\beta}\int_Q |\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f\chi_{(2Q)^c})(x)|^p w(x)dx\right)^{1/p} \\ &=: I_1 + I_2. \end{aligned} \tag{3.4}$$

By (3.1) and Lemma 2.3(iii), we get

$$\begin{aligned} I_1 &\lesssim_{p,B_p} \|b\|_{\mathrm{BMO}(\mathbb{R}^n)}^m \left(\frac{1}{w(Q)^\beta}\int_{2Q} |f(x)|^p w(x)dx\right)^{1/p} \\ &\lesssim_{p,B_p} \|b\|_{\mathrm{BMO}(\mathbb{R}^n)}^m \left(\left(\frac{w(2Q)}{w(Q)}\right)^\beta \frac{1}{w(2Q)^\beta}\int_{2Q} |f(x)|^p w(x)dx\right)^{1/p} \\ &\lesssim_{n,\beta,p,[w]_{A_p},B_p} \|b\|_{\mathrm{BMO}(\mathbb{R}^n)}^m \|f\|_{\widetilde{M}^{p,\beta}(w)}. \end{aligned} \tag{3.5}$$

We now estimate I_2. Fixing $x \in Q$, we get from (2.2) and (1.5) that

$$\begin{aligned} \mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f\chi_{(2Q)^c})(x) &\leqslant \int_{(2Q)^c} |b(x)-b(z)|^m |K(x,z)f(z)|dz \\ &\leqslant A\int_{(2Q)^c} \frac{|b(x)-b(z)|^m|f(z)|}{|x-z|^n}dz. \end{aligned} \tag{3.6}$$

Note that $|x-z| \geqslant |x-z|_\infty \geqslant |z-x_0|_\infty - |x-x_0|_\infty \geqslant \frac{1}{2}|z-x_0|_\infty$ for $z \in (2Q)^c$. By Lemma 2.3(i), there exists a constant $\theta \in (0,1)$ such that $w^{1+\theta} \in A_p(\mathbb{R}^n)$. Let $t = \frac{(1+\epsilon)p'}{(1+\epsilon)p'-\epsilon}$. Clearly, $t \in (1,p)$. By (3.6), we obtain

$$\begin{aligned} \mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f\chi_{(2Q)^c})(x) &\leqslant 2^n A\sum_{l=0}^\infty \int_{2^l r \leqslant |z-x_0|_\infty < 2^{l+1}r} \frac{|f(z)||b(x)-b(z)|^m}{|z-x_0|_\infty^n}dz \\ &\leqslant 2^n A\sum_{l=0}^\infty (2^l r)^{-n}\int_{2^{l+1}Q} |f(z)||b(x)-b(z)|^m dz \\ &\leqslant 2^{n+m} A\sum_{l=0}^\infty (2^l r)^{-n}\int_{2^{l+1}Q} |f(z)||b(z)-b_{2^{l+1}Q}|^m dz \end{aligned}$$

$$
\begin{aligned}
&\quad + 2^{n+m} A \sum_{l=0}^{\infty} (2^l r)^{-n} |b(x) - b_{2^{l+1}Q}|^m \int_{2^{l+1}Q} |f(z)| dz \\
&=: I_3 + I_4.
\end{aligned} \tag{3.7}
$$

For I_3, by Hölder's inequality,

$$
I_3 \leqslant 2^{n+m} A \sum_{l=0}^{\infty} (2^l r)^{-n} \bigg(\int_{2^{l+1}Q} |f(z)|^t dz \bigg)^{1/t} \bigg(\int_{2^{l+1}Q} |b(z) - b_{2^{l+1}Q}|^{mt'} dz \bigg)^{1/t'}. \tag{3.8}
$$

Since $b \in \mathrm{BMO}(\mathbb{R}^n)$, we have

$$
\bigg(\int_{2^{l+1}Q} |b(z) - b_{2^{l+1}Q}|^{mt'} dz \bigg)^{1/t'} \lesssim_{n,m,p} \|b\|_{\mathrm{BMO}(\mathbb{R}^n)}^m |2^{l+1}Q|^{1/t'}. \tag{3.9}
$$

Let $s = p/t$. Then $1/(s't) = 1/t - 1/p = 1/(p'(1+\epsilon))$. By Hölder's inequality, we get

$$
\begin{aligned}
\bigg(\int_{2^{l+1}Q} |f(z)|^t dz \bigg)^{1/t} &\leqslant \bigg(\int_{2^{l+1}Q} |f(z)|^p w(z) dz \bigg)^{1/p} \bigg(\int_{2^{l+1}Q} w(z)^{1-s'} dz \bigg)^{1/(s't)} \\
&\leqslant w(2^{l+1}Q)^{\beta/p} \|f\|_{\widetilde{M}^{p,\beta}(w)} \bigg(\int_{2^{l+1}Q} w(z)^{1-s'} dz \bigg)^{1/(p'(1+\epsilon))}.
\end{aligned} \tag{3.10}
$$

Since $w^{1+\epsilon} \in A_p(\mathbb{R}^n)$ and $(1+\epsilon)(1-p') = 1-s'$, we have

$$
\begin{aligned}
\int_{2^{l+1}Q} w(z)^{1-s'} dz &= \int_{2^{l+1}Q} w(z)^{(1+\epsilon)(1-p')} dz \\
&\leqslant [w^{1+\epsilon}]_{A_p}^{\frac{1}{p-1}} |2^{l+1}Q|^{p'} \bigg(\int_{2^{l+1}Q} w(x)^{1+\epsilon} dx \bigg)^{-1/(p-1)}.
\end{aligned} \tag{3.11}
$$

By Hölder's inequality, we get

$$
w(2^{l+1}Q) = \int_{2^{l+1}Q} w(x) dx \leqslant \bigg(\int_{2^{l+1}Q} w(x)^{1+\epsilon} dx \bigg)^{1/(1+\epsilon)} |2^{l+1}Q|^{\epsilon/(1+\epsilon)},
$$

which leads to $\int_{2^{l+1}Q} w(x)^{1+\epsilon} dx \geqslant w(2^{l+1}Q)^{1+\epsilon} |2^{l+1}Q|^{-\epsilon}$. This together with (3.11) leads to

$$
\begin{aligned}
\int_{2^{l+1}Q} w(z)^{1-s'} dz &\leqslant [w^{1+\epsilon}]_{A_p}^{\frac{1}{p-1}} |2^{l+1}Q|^{p'} (w(2^{l+1}Q)^{1+\epsilon} |2^{l+1}Q|^{-\epsilon})^{-\frac{1}{p-1}} \\
&\leqslant [w^{1+\epsilon}]_{A_p}^{\frac{1}{p-1}} w(2^{l+1}Q)^{-\frac{1+\epsilon}{p-1}} |2^{l+1}Q|^{p' + \frac{\epsilon}{p-1}}.
\end{aligned} \tag{3.12}
$$

Combining (3.12) with (3.10) implies

$$
\bigg(\int_{2^{l+1}Q} |f(z)|^t dz \bigg)^{1/t} \leqslant [w^{1+\epsilon}]_{A_p}^{\frac{1}{p(1+\epsilon)}} |2^{l+1}Q|^{\frac{1}{1+\epsilon} + \frac{\epsilon}{p(1+\epsilon)}} w(2^{l+1}Q)^{\frac{\beta-1}{p}} \|f\|_{\widetilde{M}^{p,\beta}(w)}. \tag{3.13}
$$

Then we get from (3.8), (3.9) and (3.13) that

$$
\begin{aligned}
I_3 &\lesssim_{n,p,m,A,[w]_{A_p}} \|b\|_{\mathrm{BMO}(\mathbb{R}^n)}^m \|f\|_{\widetilde{M}^{p,\beta}(w)} \\
&\quad \times \sum_{l=0}^{\infty} (2^l r)^{-n} |2^{l+1}Q|^{\frac{1}{t'}} |2^{l+1}Q|^{\frac{1}{1+\epsilon} + \frac{\epsilon}{p(1+\epsilon)}} w(2^{l+1}Q)^{\frac{\beta-1}{p}} \\
&\lesssim_{n,p,m,A,[w]_{A_p}} \|b\|_{\mathrm{BMO}(\mathbb{R}^n)}^m \|f\|_{\widetilde{M}^{p,\beta}(w)} \sum_{l=0}^{\infty} w(2^{l+1}Q)^{\frac{\beta-1}{p}},
\end{aligned} \tag{3.14}
$$

where in the last inequality of (3.14) we have used that $1/(1+\epsilon) + \epsilon/(p(1+\epsilon)) = 1/t$.

Fix $l \in \mathbb{N}$. By Hölder's inequality,

$$\int_{2^{l+1}Q}|f(z)|dz \leqslant \bigg(\int_{2^{l+1}Q}|f(z)|^p w(z)dz\bigg)^{1/p}\bigg(\int_{2^{l+1}Q}w(z)^{1-p'}dz\bigg)^{1/p'} \leqslant w(2^{l+1}Q)^{\beta/p}\|f\|_{\widetilde{M}^{p,\beta}(w)}\bigg(\int_{2^{l+1}Q}w(z)^{1-p'}dz\bigg)^{1/p'}.$$

Since $w \in A_p(\mathbb{R}^n)$, we have

$$\bigg(\int_{2^{l+1}Q}w(z)^{1-p'}dz\bigg)^{1/p'} \leqslant [w]_{A_p}^{1/p}|2^{l+1}Q|w(2^{l+1}Q)^{-1/p},$$

which gives

$$\int_{2^{l+1}Q}|f(z)|dz \leqslant [w]_{A_p}^{1/p}w(2^{l+1}Q)^{\frac{\beta-1}{p}}|2^{l+1}Q|\|f\|_{\widetilde{M}^{p,\beta}(w)}. \tag{3.15}$$

We get from (3.15) that

$$I_4 \lesssim_{n,p,m,A,[w]_{A_p}} \|f\|_{\widetilde{M}^{p,\beta}(w)}\sum_{l=1}^{\infty}|b(x)-b_{2^{l+1}Q}|^m w(2^{l+1}Q)^{\frac{\beta-1}{p}}. \tag{3.16}$$

It follows from (3.7), (3.14) and (3.16) that

$$\begin{aligned}
I_2 &\lesssim_{n,p,m,A,[w]_{A_p}} \|b\|_{\mathrm{BMO}(\mathbb{R}^n)}^m\|f\|_{\widetilde{M}^{p,\beta}(w)}\sum_{l=0}^{\infty}\bigg(\frac{w(2^{l+1}Q)}{w(Q)}\bigg)^{\frac{\beta-1}{p}} \\
&\quad + \|f\|_{\widetilde{M}^{p,\beta}(w)}\bigg(\frac{1}{w(Q)^{\beta}}\int_Q\bigg(\sum_{l=0}^{\infty}|b(x)-b_{2^{l+1}Q}|^m w(2^{l+1}Q)^{\frac{\beta-1}{p}}\bigg)^p w(x)dx\bigg)^{1/p} \\
&\lesssim_{n,p,m,A,[w]_{A_p}} \|b\|_{\mathrm{BMO}(\mathbb{R}^n)}^m\|f\|_{\widetilde{M}^{p,\beta}(w)}\sum_{l=0}^{\infty}\gamma_w^{-\frac{(1-\beta)(l+1)}{p}} \\
&\quad + \|f\|_{\widetilde{M}^{p,\beta}(w)}\bigg(\frac{1}{w(Q)}\int_Q\bigg(\sum_{l=0}^{\infty}|b(x)-b_{2^{l+1}Q}|^m\bigg(\frac{w(2^{l+1}Q)}{w(Q)}\bigg)^{\frac{\beta-1}{p}}\bigg)^p w(x)dx\bigg)^{1/p} \\
&\lesssim_{n,p,m,A,[w]_{A_p}} \|b\|_{\mathrm{BMO}(\mathbb{R}^n)}^m\|f\|_{\widetilde{M}^{p,\beta}(w)} \\
&\quad + \|f\|_{\widetilde{M}^{p,\beta}(w)}w(Q)^{-1/p}\bigg(\int_Q\bigg(\sum_{l=0}^{\infty}\gamma_w^{-\frac{(1-\beta)(l+1)}{p}}|b(x)-b_{2^{l+1}Q}|^m\bigg)^p w(x)dx\bigg)^{1/p}.
\end{aligned} \tag{3.17}$$

By Minkowski's inequality and Hölder's inequality we get

$$\bigg(\int_Q\bigg(\sum_{l=0}^{\infty}\gamma_w^{-\frac{(1-\beta)(l+1)}{p}}|b(x)-b_{2^{l+1}Q}|^m\bigg)^p w(x)dx\bigg)^{1/p} \leqslant \sum_{l=0}^{\infty}\gamma_w^{-\frac{(1-\beta)(l+1)}{p}}\bigg(\int_Q|b(x)-b_{2^{l+1}Q}|^{mp}w(x)dx\bigg)^{1/p}. \tag{3.18}$$

By Lemma 2.3(v) and the fact that $|b_Q-b_{2^{l+1}Q}| \lesssim_n (l+1)\|b\|_{\mathrm{BMO}(\mathbb{R}^n)}$, we get by Minkowski's inequality that

$$\begin{aligned}
&\bigg(\int_Q|b(x)-b_{2^{l+1}Q}|^{mp}w(x)dx\bigg)^{1/p} \\
&\quad\leqslant \bigg(w(Q)^{1/(mp)}|b_Q-b_{2^{l+1}Q}| + \bigg(\int_Q|b(x)-b_Q|^{mp}w(x)dx\bigg)^{1/(mp)}\bigg)^m \\
&\quad\lesssim_{n,m} (l+1)^m w(Q)^{1/p}\|b\|_{\mathrm{BMO}(\mathbb{R}^n)}^m,
\end{aligned}$$

which combining (3.17) and (3.18) implies

$$I_2 \lesssim_{n,p,m,A,[w]_{A_p}} \|b\|_{\mathrm{BMO}(\mathbb{R}^n)}^m \|f\|_{\widetilde{M}^{p,\beta}(w)} \sum_{l=0}^{\infty} \frac{(l+1)^m}{\gamma_w^{\frac{(1-\beta)(l+1)}{p}}} \lesssim_{n,p,m,A,[w]_{A_p}} \|b\|_{\mathrm{BMO}(\mathbb{R}^n)}^m \|f\|_{\widetilde{M}^{p,\beta}(w)}, \quad (3.19)$$

since $\gamma_w > 1$ and $0 < \beta < 1$. Combining (3.19) with (3.4) and (3.5) yields (3.3). This completes the proof of Proposition 3.2. □

Remark 3.3. (i) Theorem 3.1 for the case where $\beta = 0$ and $m = 0$ was proved by Ma et al. [36] (see Theorem A).

(ii) Theorem 3.1 implies [46, Theorem 1.6], which corresponds to the case where $n = 1$ and $m = 0$.

(iii) Theorem 3.1 is new, even in the unweighted case where $w \equiv 1$ or $m = 0$, $n \geqslant 2$.

3.2 Some applications

As applications of Theorem 3.1, we have the following corollaries.

Corollary 3.4. *Let $\rho > 2$, $1 < p < \infty$, $w \in A_p(\mathbb{R}^n)$ and $0 < \beta < 1$. Then*

$$\|\mathcal{V}_\rho(\mathcal{T})(f)\|_{M^{p,\beta}(w)} \leqslant C\|f\|_{M^{p,\beta}(w)}, \quad \forall\, f \in M^{p,\beta}(w)$$

holds, provided that one of the following conditions holds:

(a) $\mathcal{T} = \mathcal{R}_j$, $1 \leqslant j \leqslant n$;

(b) $\mathcal{T} = \mathcal{T}_\Omega$, $\Omega \in \mathrm{Lip}_\alpha(\mathrm{S}^{n-1})$ *for some* $\alpha > 0$.

Corollary 3.5. *Let $\rho > 2$ and $m \geqslant 1$. Then for $1 < p < \infty$, $w \in A_p(\mathbb{R}^n)$ and $0 \leqslant \beta < 1$, we have*

$$\|\mathcal{V}_\rho(\mathcal{T})(f)\|_{M^{p,\beta}(w)} \leqslant C\|b\|_{\mathrm{BMO}(\mathbb{R}^n)}^m \|f\|_{M^{p,\beta}(w)}, \quad \forall\, f \in M^{p,\beta}(w),$$

provided that one of the following conditions holds:

(a) $\mathcal{T} = \mathcal{R}_{j,b}^m$, $1 \leqslant j \leqslant n$;

(b) $\mathcal{T} = \mathcal{T}_{\Omega,b}^m$, $\Omega \in \mathrm{Lip}_\alpha(\mathrm{S}^{n-1})$ *for some* $\alpha > 0$.

Remark 3.6. (i) The corresponding result for $\mathcal{R}_j$ (resp. $\mathcal{T}_\Omega$) in the case where $\beta = 0$ in Corollary 3.4 was given in [47, Theorem 1.1] (resp. [36, Corollary 4]).

(ii) The bounds for $\mathcal{V}_\rho(\mathcal{H})$ and $\mathcal{V}_\rho(\mathcal{R}^\pm)$ on the weighted Morrey spaces were proved by Zhang and Wu [46, Theorem 1.1].

(iii) It was shown in [18] that $\mathcal{V}_\rho(\mathcal{R}_j)$ is bounded on $L^p(|x|^\alpha)$ for $1 < p < \infty$ and $-1 < \alpha < p-1$, which gives that $\mathcal{V}_\rho(\mathcal{R}_j)$ is bounded on $L^p(\mathbb{R}^n)$ for all $1 < p < \infty$. On the other hand, Campbell et al. [4] showed that $\mathcal{V}_\rho(\mathcal{T}_\Omega)$ is of type (p,p) for $1 < p < \infty$ when $\Omega \in \mathrm{Lip}_\alpha(\mathrm{S}^{n-1})$ for some $\alpha > 0$. These bounds together with Theorem 1.2 yield the conclusions in Corollary 3.4.

(iv) When $n \geqslant 2$, Corollary 3.4 is new, even in the unweighted case where $w \equiv 1$.

Remark 3.7. (i) Corollary 3.5 for $\mathcal{R}_{j,b}^m$ in the case where $\beta = 0$ was proved by Zhang and Wu [47, Corollary 1.11]. Corollary 3.5 for $\mathcal{T} = \mathcal{T}_{\Omega,b}^m$ in the case where $\beta = 0$ and $m = 1$ was shown by Chen et al. [6, Corollary 1.2].

(ii) By the known L^p $(1 < p < \infty)$ bounds for $\mathcal{V}_\rho(\mathcal{R}_j)$ with $1 \leqslant j \leqslant n$ and $\mathcal{V}_\rho(\mathcal{T}_\Omega)$ with $\Omega \in \mathrm{Lip}_\alpha(\mathrm{S}^{n-1})$ for some $\alpha > 0$ and Theorems 1.2 and 3.1, we can get the desired conclusions in Corollary 3.5.

(iii) Corollary 3.5 is new, even in the unweighted case where $w \equiv 1$ or $\beta = 0$.

4 Compactness on weighted Lebesgue and Morrey spaces

4.1 A compactness criterion on weighted Lebesgue and Morrey spaces

A compactness criterion on variation operators of commutators of Calderón-Zygmund singular integrals on weighted Lebesgue and Morrey spaces is as follows.

Theorem 4.1. *Let $m \geqslant 1$, $\rho > 2$, $1 < p < \infty$, $w \in A_p(\mathbb{R}^n)$ and $0 \leqslant \beta < 1$. If K satisfies (1.5)–(1.8) and $b \in \mathrm{CMO}(\mathbb{R}^n)$, then the operator $\mathcal{V}_\rho(\mathcal{T}^m_{K,b})$ defined as in (1.4) is a compact operator on $M^{p,\beta}(w)$.*

Proof. The proof of Theorem 4.1 will be divided into four steps:

Step 1 (Reduction via smooth truncated techniques)**.** We shall adopt the truncated techniques following from [31] to prove Theorem 4.1. Let $\varphi \in C^\infty([0,\infty))$ satisfy that $0 \leqslant \varphi \leqslant 1$, $\varphi(t) \equiv 1$ if $t \in [0,1]$ and $\varphi(t) \equiv 0$ if $t \in [2,\infty)$. For any $\eta > 0$, we define the function K_η by

$$K_\eta(x,y) = K(x,y)\bigg(1 - \varphi\bigg(\frac{2}{\eta}|x-y|\bigg)\bigg). \tag{4.1}$$

In what follows, let us fix $b \in \mathrm{CMO}(\mathbb{R}^n)$, $1 < p < \infty$, $w \in A_p(\mathbb{R}^n)$ and $0 \leqslant \beta < 1$. We shall prove that there exists a constant $C > 0$ independent of η such that

$$\|\mathcal{V}_\rho(\mathcal{T}^m_{K_\eta,b})(f) - \mathcal{V}_\rho(\mathcal{T}^m_{K,b})(f)\|_{M^{p,\beta}(w)} \leqslant C\eta\|f\|_{M^{p,\beta}(w)}, \quad \forall\, f \in M^{p,\beta}(w). \tag{4.2}$$

By (1.4), (1.5) and (2.2), we have

$$\begin{aligned}
&|\mathcal{V}_\rho(\mathcal{T}^m_{K_\eta,b})(f)(x) - \mathcal{V}_\rho(\mathcal{T}^m_{K,b})(f)(x)| \\
&\quad\leqslant \sup_{\epsilon_i \searrow 0}\bigg(\sum_{i=1}^\infty \bigg|\int_{\epsilon_{i+1}<|x-z|\leqslant\epsilon_i} (b(x)-b(z))^m(K_\eta(x,z)-K(x,z))f(z)dz\bigg|^\rho\bigg)^{1/\rho} \\
&\quad\leqslant \int_{\mathbb{R}^n} |(b(x)-b(z))^m(K_\eta(x,z)-K(x,z))f(z)|dz \\
&\quad= \int_{\mathbb{R}^n} |(b(x)-b(z))^m f(z)||K(x,z)|\varphi\bigg(\frac{2}{\eta}|x-z|\bigg)dz \\
&\quad\leqslant (2\|b\|_{L^\infty(\mathbb{R}^n)})^{m-1}\|\nabla b\|_{L^\infty(\mathbb{R}^n)}A\int_{|x-z|\leqslant\eta}\frac{|f(z)|}{|x-z|^{n-1}}dz \\
&\quad\leqslant (2\|b\|_{L^\infty(\mathbb{R}^n)})^{m-1}\|\nabla b\|_{L^\infty(\mathbb{R}^n)}A2^n\omega_n\eta Mf(x)
\end{aligned} \tag{4.3}$$

for almost every $x \in \mathbb{R}^n$, where $\omega_n = |B(0,1)|$. Combining (4.3) with the $M^{p,\beta}(w)$ boundedness for M implies (4.2).

By (4.2) and [45, p. 278, Theorem (iii)], the compactness for $\mathcal{V}_\rho(\mathcal{T}^m_{K,b})$ reduces to the compactness for $\mathcal{V}_\rho(\mathcal{T}^m_{K_\eta,b})$ when $\eta > 0$ is small enough, i.e., to prove the compactness for $\mathcal{V}_\rho(\mathcal{T}^m_{K,b})$, it suffices to prove that the set

$$\mathcal{F} := \{\mathcal{V}_\rho(\mathcal{T}^m_{K_\eta,b})(f) : \|f\|_{M^{p,\beta}(w)} \leqslant 1\}$$

is pre-compact when $\eta > 0$ is small enough. By Proposition 2.5, it is enough to verify that $\mathcal{F}$ satisfies Propositions 2.5(i)–2.5(iii) .

Step 2 (A verification for Proposition 2.5(i))**.** Let $\eta \in (0,1)$. By Theorems 1.2 and 3.1 and (4.2), we have

$$\begin{aligned}
\|\mathcal{V}_\rho(\mathcal{T}^m_{K_\eta,b})(f)\|_{M^{p,\beta}(w)} &\leqslant \|\mathcal{V}_\rho(\mathcal{T}^m_{K_\eta,b})(f) - \mathcal{V}_\rho(\mathcal{T}^m_{K,b})(f)\|_{M^{p,\beta}(w)} + \|\mathcal{V}_\rho(\mathcal{T}^m_{K,b})(f)\|_{M^{p,\beta}(w)} \\
&\leqslant C\|f\|_{M^{p,\beta}(w)} \leqslant C
\end{aligned}$$

when $\|f\|_{M^{p,\beta}(w)} \leqslant 1$. This yields that $\mathcal{F}$ satisfies Proposition 2.5(i).

Step 3 (A verification for Proposition 2.5(ii))**.** Assume that $b \in \mathcal{C}^\infty_0(\mathbb{R}^n)$ and is supported in a cube $Q = Q(0,r)$. Fix $f \in M^{p,\beta}(w)$ with $\|f\|_{M^{p,\beta}(w)} \leqslant 1$ and $E_N := \{x \in \mathbb{R}^n : |x| > N\}$ with $N \geqslant \max\{nr, 1\}$. Note that $|z| \leqslant n|z|_\infty \leqslant \frac{1}{2}nr \leqslant \frac{1}{2}N \leqslant \frac{1}{2}|x|$ when $x \in E_N$ and $z \in Q$. Then we have $|x-z| \geqslant |x|-|z| \geqslant \frac{1}{2}|x|$ when $x \in E_N$ and $z \in Q$. By (1.5), we have

$$|K_\eta(x,y)| \leqslant |K(x,y)| \leqslant \frac{A}{|x-y|^n} \quad \text{for } x \neq y. \tag{4.4}$$

Note that $b(x)=0$ when $x\in E_N$ since $N\geqslant nr$. By (4.4) and (2.2), we have

$$\mathcal{V}_\rho(\mathcal{T}^m_{K_\eta,b})(f)(x)\leqslant A\int_{\mathbb{R}^n}\frac{|(b(x)-b(z))^m f(z)|}{|x-z|^n}dz\leqslant 2^nA\|b\|^m_{L^\infty(\mathbb{R}^n)}|x|^{-n}\int_Q|f(z)|dz \tag{4.5}$$

for almost every $x\in E_N$. By the arguments similar to those used to derive (3.15),

$$\int_Q|f(z)|dz\leqslant[w]^{1/p}_{A_p}w(Q)^{\frac{\beta-1}{p}}|Q|\|f\|_{\widetilde{M}^{p,\beta}(w)}. \tag{4.6}$$

For a fixed cube $\tilde{Q}=\tilde{Q}(x_0,t)$, we get from (4.5) and (4.6) that

$$\begin{aligned}
&\frac{1}{w(\tilde{Q})^\beta}\int_{\tilde{Q}}|\mathcal{V}_\rho(\mathcal{T}^m_{K_\eta,b})(f)(x)\chi_{E_N}(x)|^pw(x)dx\\
&\quad\leqslant C_1w(Q)^{\beta-1}|Q|^p\frac{1}{w(\tilde{Q})^\beta}\int_{\tilde{Q}\cap E_N}|x|^{-np}w(x)dx\\
&\quad\leqslant C_1w(Q)^{\beta-1}|Q|^p\frac{1}{w(\tilde{Q})^\beta}\sum_{j=0}^\infty\int_{\tilde{Q}\cap(B(0,2^{j+1}N)\setminus B(0,2^jN))}|x|^{-np}w(x)dx\\
&\quad\leqslant C_1w(Q)^{\beta-1}|Q|^p\frac{1}{w(\tilde{Q})^\beta}\sum_{j=0}^\infty(2^jN)^{-np}w(\tilde{Q}\cap(B(0,2^{j+1}N)\setminus B(0,2^jN)))\\
&\quad\leqslant C_1w(Q)^{\beta-1}|Q|^p\sum_{j=0}^\infty(2^jN)^{-np}w(\tilde{Q}\cap(B(0,2^{j+1}N)\setminus B(0,2^jN)))^{1-\beta},
\end{aligned} \tag{4.7}$$

where $C_1=(2^nA\|b\|^m_{L^\infty(\mathbb{R}^n)}\|f\|_{\widetilde{M}^{p,\beta}(w)})^p[w]_{A_p}$. Since $w\in A_p(\mathbb{R}^n)$, there exists a constant $\epsilon\in(0,p)$ such that $w\in A_{p-\epsilon}(\mathbb{R}^n)$. Then by Lemma 2.3(iii), we have

$$\begin{aligned}
&w(\tilde{Q}\cap(B(0,2^{j+1}N)\setminus B(0,2^jN)))\\
&\quad\leqslant w(B(0,2^{j+1}N))\leqslant w(Q(0,2^{j+2}N))\leqslant[w]_{A_{p-\epsilon}}(2^{j+2}N)^{n(p-\epsilon)}w(Q(0,1)).
\end{aligned}$$

This together with (4.7) implies that

$$\begin{aligned}
&\frac{1}{w(\tilde{Q})^\beta}\int_{\tilde{Q}}|\mathcal{V}_\rho(\mathcal{T}^m_{K_\eta,b})(f)(x)\chi_{E_N}(x)|^pw(x)dx\\
&\quad\leqslant C_1[w]_{A_{p-\epsilon}}w(Q)^{\beta-1}|Q|^pw(Q(0,1))^{1-\beta}\sum_{j=0}^\infty(2^jN)^{-np}(2^{j+2}N)^{n(p-\epsilon)(1-\beta)}\\
&\quad\leqslant C_1[w]_{A_{p-\epsilon}}w(Q)^{\beta-1}|Q|^pw(Q(0,1))^{1-\beta}\sum_{j=0}^\infty(2^jN)^{-np\beta-\epsilon(1-\beta)n}\\
&\quad\leqslant C_1[w]_{A_{p-\epsilon}}w(Q)^{\beta-1}|Q|^pw(Q(0,1))^{1-\beta}N^{-np\beta-\epsilon(1-\beta)n},
\end{aligned}$$

which leads to

$$\begin{aligned}
&\|\mathcal{V}_\rho(\mathcal{T}^m_{K_\eta,b})(f)\chi_{E_N}\|_{M^{p,\beta}(w)}\\
&\quad\leqslant C_1^{1/p}[w]^{1/p}_{A_{p-\epsilon}}w(Q)^{\frac{\beta-1}{p}}|Q|w(Q(0,1))^{\frac{1-\beta}{p}}N^{-n\beta-\frac{\epsilon(1-\beta)n}{p}}.
\end{aligned}$$

This implies that $\mathcal{F}$ satisfies Proposition 2.5(ii).

Step 4 (A verification for Proposition 2.5(iii)**.** It suffices to show that

$$\lim_{|h|\to0}\|\mathcal{V}_\rho(\mathcal{T}^m_{K_\eta,b})(f)(\cdot+h)-\mathcal{V}_\rho(\mathcal{T}^m_{K_\eta,b})(f)(\cdot)\|_{M^{p,\beta}(w)}=0 \tag{4.8}$$

for a fixed $\eta\in(0,1)$.

At first, we claim that $K_\eta(x,y)$ satisfies

$$|K_\eta(x,y)-K_\eta(z,y)|\leqslant\frac{\tilde{A}|x-z|^\delta}{|x-y|^{n+\delta}}\quad\text{for }|x-y|>2|x-z|, \tag{4.9}$$

where $\tilde{A} = (2 + 2^{n+\delta+1}\|\nabla\varphi\|_{L^\infty([0,\infty))})A$.

When $|x-y| > 2|x-z|$, we consider the following different cases:

(i) ($|x-y| \geqslant \eta$ and $|z-y| \geqslant \eta$) In this case we have $K_\eta(x,y) = K(x,y)$ and $K_\eta(z,y) = K(z,y)$. This together with (1.6) yields (4.9).

(ii) ($|x-y| \geqslant \eta$ and $|z-y| < \eta$) In this case we have $K_\eta(x,y) = K(x,y)$ and $|z-y| > \frac{1}{2}|x-y|$ since $|x-y| > 2|x-z|$. These together with (1.5) and (1.6) imply that

$$\begin{aligned}|K_\eta(x,y) - K_\eta(z,y)| &\leqslant |K(x,y) - K(z,y)| + |K(z,y)|\left|\varphi\left(\frac{2}{\eta}|z-y|\right)\right| \\ &= |K(x,y) - K(z,y)| + |K(z,y)|\left|\varphi\left(\frac{2}{\eta}|z-y|\right) - \varphi\left(\frac{2}{\eta}|x-y|\right)\right| \\ &\leqslant \frac{A|x-z|^\delta}{|x-y|^{n+\delta}} + \frac{A}{|z-y|^n}\frac{2}{\eta}\|\nabla\varphi\|_{L^\infty([0,\infty))}|x-z| \\ &\leqslant (1 + 2^{n+\delta+1}\|\nabla\varphi\|_{L^\infty([0,\infty))})A\frac{|x-z|^\delta}{|x-y|^{n+\delta}},\end{aligned}$$

which proves (4.9).

(iii) ($|x-y| < \eta$ and $|z-y| \geqslant \eta$) The case is similar to the case (ii).

(iv) ($|x-y| < \eta$ and $|z-y| < \eta$) Without loss of generality we may assume that $|x-y| \geqslant |z-y|$. By (1.5) and (1.6) and the fact that $|y-z| > \frac{1}{2}|x-y|$, we get

$$\begin{aligned}|K_\eta(x,y) - K_\eta(z,y)| &\leqslant |K(x,y) - K(z,y)| + \left|K(x,y)\varphi\left(\frac{2}{\eta}|x-y|\right) - K(z,y)\varphi\left(\frac{2}{\eta}|z-y|\right)\right| \\ &\leqslant |K(x,y) - K(z,y)| + |K(x,y) - K(z,y)|\varphi\left(\frac{2}{\eta}|x-y|\right) \\ &\quad + |K(z,y)|\left|\varphi\left(\frac{2}{\eta}|z-y|\right) - \varphi\left(\frac{2}{\eta}|x-y|\right)\right| \\ &\leqslant \frac{2A|x-z|^\delta}{|x-y|^{n+\delta}} + \frac{A}{|z-y|^n}\frac{2}{\eta}\|\nabla\varphi\|_\infty|x-z| \\ &\leqslant (2 + 2^{n+\delta+1}\|\nabla\varphi\|_{L^\infty([0,\infty))})A\frac{|x-z|^\delta}{|x-y|^{n+\delta}}.\end{aligned}$$

This proves (4.9) in this case.

In what follows, we set $|h| < \frac{\eta}{8}$ and $\eta \in (0,1)$. By the definition of $\mathcal{V}_\rho(\mathcal{T}^m_{K_\eta,b})$, we have

$$\begin{aligned}&|\mathcal{V}_\rho(\mathcal{T}^m_{K_\eta,b})(f)(x+h) - \mathcal{V}_\rho(\mathcal{T}^m_{K_\eta,b})(f)(x)| \\ &\quad\leqslant \sup_{\varepsilon_i\searrow 0}\left(\sum_{i=1}^\infty\left|\int_{\varepsilon_{i+1}<|x+h-y|\leqslant\varepsilon_i}(b(x+h)-b(y))^m K_\eta(x+h,y)f(y)dy\right.\right. \\ &\qquad\left.\left. - \int_{\varepsilon_{i+1}<|x-y|\leqslant\varepsilon_i}(b(x)-b(y))^m K_\eta(x,y)f(y)dy\right|^\rho\right)^{1/\rho} \\ &\quad\leqslant \sup_{\varepsilon_i\searrow 0}\left(\sum_{i=1}^\infty\left|\int_{\varepsilon_{i+1}<|x-y|\leqslant\varepsilon_i}((b(x+h)-b(y))^m - (b(x)-b(y))^m)K_\eta(x,y)f(y)dy\right|^\rho\right)^{1/\rho} \\ &\qquad + \sup_{\varepsilon_i\searrow 0}\left(\sum_{i=1}^\infty\left|\int_{\varepsilon_{i+1}<|x-y|\leqslant\varepsilon_i}(b(x+h)-b(y))^m(K_\eta(x+h,y) - K_\eta(x,y))f(y)dy\right|^\rho\right)^{1/\rho} \\ &\qquad + \sup_{\varepsilon_i\searrow 0}\left(\sum_{i=1}^\infty\left|\int_{\mathbb{R}^n}(b(x+h)-b(y))^m K_\eta(x+h,y)f(y)\right.\right. \\ &\qquad\left.\left.\times(\chi_{\varepsilon_{i+1}<|x+h-y|\leqslant\varepsilon_i}(y) - \chi_{\varepsilon_{i+1}<|x-y|\leqslant\varepsilon_i}(y))dy\right|^\rho\right)^{1/\rho} \\ &\quad =: I_1 + I_2 + I_3.\end{aligned} \tag{4.10}$$

For I_1, it is easy to see that

$$(b(x+h)-b(y))^m-(b(x)-b(y))^m=\sum_{j=0}^{m-1}c_m^j(b(x+h)-b(x))^{m-j}(b(x)-b(y))^j$$

and

$$(b(x)-b(y))^j=\sum_{\mu=0}^{j}c_j^\mu b(x)^{j-\mu}(-1)^\mu b(y)^\mu,\quad \forall\, 0\leqslant j\leqslant m-1,$$

where $C_N^r=\frac{N!}{r!(N-r)!}$ for any r, $N\in\mathbb{N}$ with $r\leqslant N$. Therefore, we have

$$\begin{aligned}I_1\leqslant&\sum_{j=0}^{m-1}c_m^j\|\nabla b\|_{L^\infty(\mathbb{R}^n)}^{m-j}|h|^{m-j}\sum_{\mu=0}^{j}c_j^\mu\|b\|_{L^\infty(\mathbb{R}^n)}^{j-\mu}\\&\times\sup_{\varepsilon_i\searrow 0}\bigg(\sum_{i=1}^\infty\bigg|\int_{\varepsilon_{i+1}<|x-y|\leqslant\varepsilon_i}K_\eta(x,y)(b(y))^\mu f(y)dy\bigg|^\rho\bigg)^{1/\rho}.\end{aligned}\tag{4.11}$$

Given a decreasing sequence $\{\varepsilon_i\}_{i\geqslant 1}$ of positive numbers converging to 0, we set $N(\{\varepsilon_i\}):=\max\{i\geqslant 1:\varepsilon_i\geqslant\eta\}$. Note that $K_\eta(x,y)=0$ when $|x-y|\leqslant\frac{\eta}{2}$ and $K_\eta(x,y)=K(x,y)$ when $|x-y|\geqslant\eta$. By (4.4) and (2.2), we have that for $0\leqslant j\leqslant m-1$ and $0\leqslant\mu\leqslant j$,

$$\begin{aligned}&\bigg(\sum_{i=1}^\infty\bigg|\int_{\varepsilon_{i+1}<|x-y|\leqslant\varepsilon_i}K_\eta(x,y)b^\mu(y)f(y)dy\bigg|^\rho\bigg)^{1/\rho}\\&\leqslant\bigg(\sum_{i=1}^{N(\{\varepsilon_i\})-1}\bigg|\int_{\varepsilon_{i+1}<|x-y|\leqslant\varepsilon_i}K(x,y)b^\mu(y)f(y)dy\bigg|^\rho\bigg)^{1/\rho}\\&\quad+\bigg(\sum_{i=N(\{\varepsilon_i\})}^{\infty}\bigg|\int_{\varepsilon_{i+1}<|x-y|\leqslant\varepsilon_i}K_\eta(x,y)\chi_{\frac{\eta}{2}<|x-y|\leqslant\eta}(y)b^\mu(y)f(y)dy\bigg|^\rho\bigg)^{1/\rho}\\&\leqslant\mathcal{V}_\rho(\mathcal{T}_K)(b^\mu f)(x)+A\|b\|_{L^\infty(\mathbb{R}^n)}^\mu\int_{\frac{\eta}{2}<|x-y|\leqslant\eta}\frac{|f(y)|}{|x-y|^n}dy\\&\leqslant\mathcal{V}_\rho(\mathcal{T}_K)(b^\mu f)(x)+2^n\omega_n\|b\|_{L^\infty(\mathbb{R}^n)}^\mu AMf(x),\end{aligned}$$

which together with (4.11) implies

$$I_1\lesssim_{m,b}|h|\bigg(\sum_{\mu=0}^{m-1}\mathcal{V}_\rho(\mathcal{T}_K)(b^\mu f)(x)+2^n\omega_n\|b\|_{L^\infty(\mathbb{R}^n)}^\mu AMf(x)\bigg).\tag{4.12}$$

For I_2, due to $|h|<\frac{\eta}{8}$, we have $K_\eta(x+h,y)=K_\eta(x,y)=0$ when $|x-y|\leqslant\frac{\eta}{4}$. Moreover, $|x-y|>2|h|$ when $|x-y|>\frac{\eta}{4}$. In light of (2.2) and (4.9), we have

$$\begin{aligned}I_2&\leqslant\sup_{\varepsilon_i\searrow 0}\bigg(\sum_{i=1}^\infty\bigg|\int_{\epsilon_{i+1}<|x-y|\leqslant\epsilon_i}(b(x+h)-b(y))^m(K_\eta(x+h,y)-K_\eta(x,y))f(y)\chi_{|x-y|>\frac{\eta}{4}}(y)dy\bigg|^\rho\bigg)^{1/\rho}\\&\leqslant\int_{\mathbb{R}^n}|(b(x+h)-b(y))^m(K_\eta(x+h,y)-K_\eta(x,y))f(y)|\chi_{|x-y|>\frac{\eta}{4}}(y)dy\\&\leqslant 2^m\|b\|_{L^\infty(\mathbb{R}^n)}^m\tilde{A}|h|^\delta\int_{|x-y|>\frac{\eta}{4}}\frac{|f(y)|}{|x-y|^{n+\delta}}dy\\&\lesssim_{m,b,n,\delta,A,\varphi}\bigg(\frac{|h|}{\eta}\bigg)^\delta Mf(x)\end{aligned}\tag{4.13}$$

for almost every $x\in\mathbb{R}^n$.

It remains to estimate I_3. Note that $\chi_{\varepsilon_{i+1}<|x+h-y|\leqslant\varepsilon_i}(y)-\chi_{\varepsilon_{i+1}<|x-y|\leqslant\varepsilon_i}(y)\neq 0$ if and only if at least one of the following statements holds:

(a) $\varepsilon_{i+1} < |x+h-y| \leqslant \varepsilon_i$ and $|x-y| \leqslant \varepsilon_{i+1}$;
(b) $\varepsilon_{i+1} < |x+h-y| \leqslant \varepsilon_i$ and $|x-y| > \varepsilon_i$;
(c) $\varepsilon_{i+1} < |x-y| \leqslant \varepsilon_i$ and $|x+h-y| \leqslant \varepsilon_{i+1}$;
(d) $\varepsilon_{i+1} < |x-y| \leqslant \varepsilon_i$ and $|x+h-y| > \varepsilon_i$.

Then we get from (4.4) that

$$\begin{aligned}
&\left|\int_{\mathbb{R}^n}(b(x+h)-b(y))^m K_\eta(x+h,y)f(y)(\chi_{\varepsilon_{i+1}<|x+h-y|\leqslant\varepsilon_i}(y)-\chi_{\varepsilon_{i+1}<|x-y|\leqslant\varepsilon_i}(y))dy\right|\\
&\lesssim_{m,b,A}\int_{|x+h-y|\geqslant\frac{\eta}{2}}\frac{|f(y)|}{|x+h-y|^n}\chi_{\varepsilon_{i+1}<|x+h-y|\leqslant\varepsilon_i}(y)\chi_{|x-y|\leqslant\varepsilon_{i+1}}(y)dy\\
&\quad+\int_{|x+h-y|\geqslant\frac{\eta}{2}}\frac{|f(y)|}{|x+h-y|^n}\chi_{\varepsilon_{i+1}<|x+h-y|\leqslant\varepsilon_i}(y)\chi_{|x-y|>\varepsilon_i}(y)dy\\
&\quad+\int_{|x+h-y|\geqslant\frac{\eta}{2}}\frac{|f(y)|}{|x+h-y|^n}\chi_{\varepsilon_{i+1}<|x-y|\leqslant\varepsilon_i}(y)\chi_{|x+h-y|\leqslant\varepsilon_{i+1}}(y)dy\\
&\quad+\int_{|x+h-y|\geqslant\frac{\eta}{2}}\frac{|f(y)|}{|x+h-y|^n}\chi_{\varepsilon_{i+1}<|x-y|\leqslant\varepsilon_i}(y)\chi_{|x+h-y|>\varepsilon_i}(y)dy\\
&=: I_{3,1}+I_{3,2}+I_{3,3}+I_{3,4}.
\end{aligned}\tag{4.14}$$

By Lemma 2.3(ii), there exists $s\in(1,2)$ such that $w\in A_{p/s}(\mathbb{R}^n)$. We now estimate $I_{3,i}$, $i=1,2,3,4$, respectively.

For $I_{3,1}$, in the case (a) we have that $\varepsilon_{i+1}\geqslant|x-y|\geqslant|x+h-y|-|h|\geqslant\frac{\eta}{2}-\frac{\eta}{8}>\frac{\eta}{4}>2|h|$ and $|x+h-y|\geqslant|x-y|-|h|\geqslant\frac{1}{2}|x-y|$ when $|x+h-y|>\frac{\eta}{2}$. By Hölder's inequality,

$$\begin{aligned}
I_{3,1}&\lesssim_{m,b,A} C\int_{|x+h-y|\geqslant\frac{\eta}{2}}\frac{|f(y)|}{|x+h-y|^n}\chi_{\varepsilon_{i+1}<|x+h-y|\leqslant\varepsilon_i}(y)\chi_{\varepsilon_{i+1}<|x+h-y|\leqslant\varepsilon_{i+1}+|h|}(y)dy\\
&\lesssim_{m,b,A}\left(\int_{|x+h-y|\geqslant\frac{\eta}{2}}\frac{|f(y)|^s}{|x+h-y|^{ns}}\chi_{\varepsilon_{i+1}<|x+h-y|\leqslant\varepsilon_i}(y)dy\right)^{1/s}\\
&\quad\times\left(\int_{\mathbb{R}^n}\chi_{\varepsilon_{i+1}<|x+h-y|\leqslant\varepsilon_{i+1}+|h|}(y)dy\right)^{1/s'}\\
&\lesssim_{m,b,n,A,s}\left(\int_{|x+h-y|\geqslant\frac{\eta}{2}}\frac{|f(y)|^s}{|x+h-y|^{ns}}\chi_{\varepsilon_{i+1}<|x+h-y|\leqslant\varepsilon_i}(y)dy\right)^{1/s}\\
&\quad\times((\varepsilon_{i+1}+|h|)^n-\varepsilon_{i+1}^n)^{1/s'}\\
&\lesssim_{m,b,n,A,s}\left(\int_{|x+h-y|\geqslant\frac{\eta}{2}}\frac{|f(y)|^s}{|x+h-y|^{ns}}\chi_{\varepsilon_{i+1}<|x+h-y|\leqslant\varepsilon_i}(y)dy\right)^{1/s}\\
&\quad\times((\varepsilon_{i+1}+|h|)^{n-1}|h|)^{1/s'}\\
&\lesssim_{m,b,n,A,s}|h|^{1/s'}\left(\int_{|x-y|\geqslant\frac{\eta}{4}}\frac{|f(y)|^s}{|x-y|^{n+s-1}}\chi_{\varepsilon_{i+1}<|x+h-y|\leqslant\varepsilon_i}(y)dy\right)^{1/s},
\end{aligned}\tag{4.15}$$

where in the last inequality of (4.15) we have used the fact that

$$(\varepsilon_{i+1}+|h|)^{(n-1)/s'}\leqslant\left(\frac{5}{4}\right)^{(n-1)/s'}|x+h-y|^{(n-1)/s'}$$

and $|x+h-y|\geqslant\frac{1}{2}|x-y|$ when $|x+h-y|>\frac{\eta}{2}$.

For $I_{3,2}$, in the case (b) we have that $|x-y|\geqslant\varepsilon_i\geqslant|x+h-y|\geqslant\frac{\eta}{2}>4|h|$ and $|x+h-y|\geqslant|x-y|-|h|\geqslant\frac{|x-y|}{2}$ when $|x+h-y|>\frac{\eta}{2}$. By the arguments similar to those used to derive (4.15),

$$I_{3,2}\lesssim_{m,b,A}\int_{|x+h-y|\geqslant\frac{\eta}{2}}\frac{|f(y)|}{|x+h-y|^n}\chi_{\varepsilon_{i+1}<|x+h-y|\leqslant\varepsilon_i}(y)\chi_{\varepsilon_i<|x-y|\leqslant\varepsilon_i+|h|}(y)dy$$

$$\lesssim_{m,b,n,A,s} |h|^{1/s'}\bigg(\int_{|x-y|\geqslant\frac{\eta}{4}}\frac{|f(y)|^s}{|x-y|^{n+s-1}}\chi_{\varepsilon_{i+1}<|x+h-y|\leqslant\varepsilon_i}(y)dy\bigg)^{1/s}. \tag{4.16}$$

For $I_{3,3}$, in the case (c) we have that $|x-y|>\varepsilon_{i+1}\geqslant|x+h-y|\geqslant\frac{\eta}{2}>4|h|$ and $|x+h-y|\geqslant|x-y|-|h|\geqslant\frac{|x-y|}{2}$ when $|x+h-y|>\frac{\eta}{2}$. By the arguments similar to those used to derive (4.15),

$$\begin{aligned}I_{3,3}&\lesssim_{m,b,n,A,s}\int_{|x+h-y|\geqslant\frac{\eta}{2}}\frac{|f(y)|}{|x+h-y|^n}\chi_{\varepsilon_{i+1}<|x-y|\leqslant\varepsilon_i}(y)\chi_{\varepsilon_{i+1}<|x-y|\leqslant\varepsilon_{i+1}+|h|}(y)dy\\&\lesssim_{m,b,n,A,s} C|h|^{1/s'}\bigg(\int_{|x-y|\geqslant\frac{\eta}{4}}\frac{|f(y)|^s}{|x-y|^{n+s-1}}\chi_{\varepsilon_{i+1}<|x-y|\leqslant\varepsilon_i}(y)dy\bigg)^{1/s}.\end{aligned}\tag{4.17}$$

For $I_{3,4}$, in the case (d) we have that $\varepsilon_i\geqslant|x-y|\geqslant|x+h-y|-|h|\geqslant\frac{\eta}{2}-\frac{\eta}{8}\geqslant\frac{\eta}{4}>2|h|$ and $|x+h-y|\geqslant|x-y|-|h|\geqslant\frac{|x-y|}{2}$ when $|x+h-y|>\frac{\eta}{2}$. By the arguments similar to those used to derive (4.15),

$$\begin{aligned}I_{3,4}&\lesssim_{m,b,A}\int_{|x+h-y|\geqslant\frac{\eta}{2}}\frac{|f(y)|}{|x+h-y|^n}\chi_{\varepsilon_{i+1}<|x-y|\leqslant\varepsilon_i}(y)\chi_{\varepsilon_i<|x+h-y|\leqslant\varepsilon_i+|h|}(y)dy\\&\lesssim_{m,b,n,A,s} |h|^{1/s'}\bigg(\int_{|x-y|\geqslant\frac{\eta}{4}}\frac{|f(y)|^s}{|x-y|^{n+s-1}}\chi_{\varepsilon_{i+1}<|x-y|\leqslant\varepsilon_i}(y)dy\bigg)^{1/s}.\end{aligned}\tag{4.18}$$

Combining (4.18) with (2.2) and (4.14)–(4.17) implies

$$\begin{aligned}I_3&\lesssim_{m,b,n,A,s}|h|^{1/s'}\Bigg(\sup_{\varepsilon_i\searrow0}\bigg(\sum_{i=1}^{\infty}\bigg(\int_{|x-y|\geqslant\frac{\eta}{4}}\frac{|f(y)|^s}{|x-y|^{n+s-1}}\chi_{\varepsilon_{i+1}<|x+h-y|\leqslant\varepsilon_i}(y)dy\bigg)^{\rho/s}\bigg)^{1/\rho}\\&\quad+\sup_{\varepsilon_i\searrow0}\bigg(\sum_{i=1}^{\infty}\bigg(\int_{|x-y|\geqslant\frac{\eta}{4}}\frac{|f(y)|^s}{|x-y|^{n+s-1}}\chi_{\varepsilon_{i+1}<|x-y|\leqslant\varepsilon_i}(y)dy\bigg)^{\rho/s}\bigg)^{1/\rho}\Bigg)\\&\lesssim_{m,b,n,A,s}|h|^{1/s'}\bigg(\int_{|x-y|\geqslant\frac{\eta}{4}}\frac{|f(y)|^s}{|x-y|^{n+s-1}}dy\bigg)^{1/s}\\&\lesssim_{m,b,n,A,s}\bigg(\frac{|h|}{\eta}\bigg)^{1/s'}(M|f|^s)^{1/s}(x).\end{aligned}\tag{4.19}$$

It follows from (4.10), (4.12), (4.13) and (4.19) that

$$\begin{aligned}&|\mathcal{V}_\rho(\mathcal{T}^m_{K_\eta,b})(f)(x+h)-\mathcal{V}_\rho(\mathcal{T}^m_{K_\eta,b})(f)(x)|\\&\quad\lesssim_{m,b,n,A,s,\varphi}\bigg(|h|Mf(x)+|h|\sum_{\mu=0}^{m-1}\mathcal{V}_\rho(\mathcal{T}_K)(b^\mu f)(x)+\bigg(\frac{|h|}{\eta}\bigg)^{\delta}Mf(x)+\bigg(\frac{|h|}{\eta}\bigg)^{1/s'}(M|f|^s)^{1/s}(x)\bigg)\end{aligned}\tag{4.20}$$

for all $x\in\mathbb{R}^n$ when $|h|<\frac{\eta}{8}$. Note that $w\in A_{p/s}(\mathbb{R}^n)$. This together with the boundedness for M on weighted Morrey spaces implies $\|(M|f|^s)^{1/s}\|_{M^{p,\beta}(w)}\lesssim_{n,p,\beta}\|f\|_{M^{p,\beta}(w)}$. On the other hand, by our assumptions and Theorem A, we have that $\mathcal{V}_\rho(\mathcal{T}_K)$ is bounded on $L^p(w)$. This together with Theorem 3.1 yields that $\mathcal{V}_\rho(\mathcal{T}_K)$ is bounded on $M^{p,\beta}(w)$. Combining these with the boundedness of M on $M^{p,\beta}(w)$, Minkowski's inequality and (4.20) implies that

$$\begin{aligned}&\|\mathcal{V}_\rho(\mathcal{T}^m_{K_\eta,b})(f)(\cdot+h)-\mathcal{V}_\rho(\mathcal{T}^m_{K_\eta,b})(f)(\cdot)\|_{M^{p,\beta}(w)}\\&\quad\lesssim_{m,b,n,A,s,p,\beta}|h|\|Mf\|_{M^{p,\beta}(w)}+C|h|\sum_{\mu=0}^{m-1}\|\mathcal{V}_\rho(\mathcal{T}_K)(b^\mu f)\|_{M^{p,\beta}(w)}\\&\qquad+\bigg(\frac{|h|}{\eta}\bigg)^{\delta}\|Mf\|_{M^{p,\beta}(w)}+C\bigg(\frac{|h|}{\eta}\bigg)^{1/s'}\|(M|f|^s)^{1/s}\|_{M^{p,\beta}(w)}\\&\quad\lesssim_{m,b,n,A,s,p,\beta}\bigg(|h|+\bigg(\frac{|h|}{\eta}\bigg)^{\delta}+\bigg(\frac{|h|}{\eta}\bigg)^{1/s'}\bigg)\end{aligned}$$

when $|h|<\frac{\eta}{8}$, which leads to (4.8). This finishes the proof of Theorem 4.1. □

4.2 Applications

As applications of Theorem 4.1, we can get the following result.

Corollary 4.2. *Let $m \geqslant 1$, $\rho > 2$, $1 < p < \infty$, $w \in A_p(\mathbb{R}^n)$ and $0 \leqslant \beta < 1$. If $b \in \mathrm{CMO}(\mathbb{R}^n)$, then $\mathcal{V}_\rho(\mathcal{T}_{k,b}^m)$ is a compact operator on $M^{p,\beta}(w)$, provided that one of the following conditions holds*:

(i) $n = 1$ *and* $\mathcal{T}_{k,b}^m = \mathcal{H}_b^m$;

(ii) $n = 1$ *and* $\mathcal{T}_{k,b}^m = \mathcal{R}_{\pm,b}^m$;

(iii) $\mathcal{T}_{k,b}^m = \mathcal{R}_{j,b}^m$, $1 \leqslant j \leqslant n$;

(iv) $\mathcal{T}_{k,b}^m = \mathcal{T}_{\Omega,b}^m$, $\Omega \in \mathrm{Lip}_\alpha(\mathrm{S}^{n-1})$ *for some* $\alpha > 0$.

Remark 4.3. (i) It was shown in [13] that both $\mathcal{V}_\rho(\mathcal{H})$ and $\mathcal{V}_\rho(\mathcal{R}^\pm)$ are bounded on $L^p(w)$ for $1 < p < \infty$ and $w \in A_p(\mathbb{R})$. This together with Theorem 4.1 leads to the desired conclusions for $\mathcal{H}_b$ and $\mathcal{R}_{\pm,b}$ in Corollary 4.2.

(ii) By the known L^p $(1 < p < \infty)$ bounds for $\mathcal{V}_\rho(\mathcal{R}_j)$ with $1 \leqslant j \leqslant n$ and $\mathcal{V}_\rho(\mathcal{T}_\Omega)$ with $\Omega \in \mathrm{Lip}_\alpha(\mathrm{S}^{n-1})$ for some $\alpha > 0$ and Theorem 4.1, we can get the desired conclusions for $\mathcal{R}_{j,b}$ and $\mathcal{T}_{\Omega,b}$ in Corollary 4.2.

(iii) When $\alpha = 1$, $m = 1$ and $\beta = 0$, the corresponding result in Corollary 4.2 for the case (iv) was proved by Guo et al. [21, Corollary 1.5].

(iv) Theorem 4.1 and Corollary 4.2 are new, even in the unweighted case where $w \equiv 1$ or $\beta = 0$.

5 Boundedness on Sobolev spaces

This section is devoted to discussing the boundedness for variation operators of Calderón-Zygmund singular integrals and their commutators on Sobolev spaces. Before presenting our main results, let us introduce some notations. Let $e_l = (0, \dots, 0, 1, 0, \dots, 0)$ be the canonical l-th base vector in $\mathbb{R}^n$ for $l = 1, \dots, n$. For a fixed $f \in L^p(\mathbb{R}^n)$ with $p \geqslant 1$, all $h \in \mathbb{R}$ with $|h| > 0$ and $i = 1, \dots, n$, we define the function $f_{h,i}$ by setting

$$f_{h,i}(x) = \frac{f(x + he_i) - f(x)}{h}.$$

It is well known that for $p \geqslant 1$, $f_{h,i} \to D_i f$ in $L^p(\mathbb{R}^n)$ when $h \to 0$ if $f \in W^{1,p}(\mathbb{R}^n)$. For $x, y \in \mathbb{R}^n$, we denote by $\Delta_y f$ the 1st difference of f, i.e., $\Delta_y f(x) = f_y(x) - f(x)$, where $f_y(x) = f(x + y)$. Set

$$G(f; p) = \limsup_{|h| \to 0} \frac{\|\Delta_h f\|_{L^p(\mathbb{R}^n)}}{|h|}.$$

According to [17, Subsection 7.11], we have

$$u \in W^{1,q}(\mathbb{R}^n),\ 1 < q < \infty \Leftrightarrow u \in L^q(\mathbb{R}^n) \text{ and } G(u; q) < \infty. \tag{5.1}$$

5.1 Boundedness for variation operators of singular integrals

The following result represents a perfection of Theorem D.

Proposition 5.1. *Let T be a sublinear operator. Assume that T is bounded on $L^p(\mathbb{R}^n)$ for some $p \in (1, \infty)$ and commutes with translations. Then T is bounded on $W^{1,p}(\mathbb{R}^n)$. Moreover, if $f \in W^{1,p}(\mathbb{R}^n)$, then for any $i = 1, \dots, n$, we have*

$$|D_i(Tf)(x)| \leqslant |T(D_i f)(x)| \tag{5.2}$$

for almost every $x \in \mathbb{R}^n$. As an application of (5.2), *we have*

$$\|Tf\|_{W^{1,p}(\mathbb{R}^n)} \leqslant \|T\|_{L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)} \|f\|_{W^{1,p}(\mathbb{R}^n)}. \tag{5.3}$$

Proof. By Theorem D, we know that $Tf \in W^{1,p}(\mathbb{R}^n)$. Fix $i \in \{1, \dots, n\}$. Since $f \in W^{1,p}(\mathbb{R}^n)$ and $Tf \in W^{1,p}(\mathbb{R}^n)$, we have $f_{h,i} \to D_i f$ and $(Tf)_{h,i} \to D_i(Tf)$ in $L^p(\mathbb{R}^n)$ as $h \to 0$. By the sublinearity and the L^p boundedness for T and the fact that $f_{h,i} \to D_i f$ in $L^p(\mathbb{R}^n)$ as $k \to \infty$, we have that $T(f_{h,i}) \to T(D_i f)$ in $L^p(\mathbb{R}^n)$ as $h \to 0$. Therefore, there exist a sequence $\{h_k\}$ of real numbers with

$\lim_{k\to 0} h_k = 0$ and a measurable set E with $|\mathbb{R}^n \setminus E| = 0$ such that $(Tf)_{h_k,i}(x) \to D_i(Tf)(x)$ and $T(f_{h_k,i})(x) \to T(D_i f)(x)$ as $k \to \infty$ for $x \in E$.

Since T commutes with translations and is sublinear, for any $x \in E$ we have

$$\begin{aligned}|(Tf)_{h_k,i}(x)| &= \frac{|Tf(x+h_k e_i) - Tf(x)|}{|h_k|}\\ &= \frac{|Tf_{h_k e_i}(x) - Tf(x)|}{|h_k|} \leqslant \frac{|T(f_{h_k e_i} - f)(x)|}{|h_k|} = |T(f_{h_k,i})(x)|.\end{aligned}$$

This yields

$$|D_i(Tf)(x)| = \Big|\lim_{k\to\infty}(Tf)_{h_k,i}(x)\Big| \leqslant \lim_{k\to\infty}|T(f_{h_k,i})(x)| = |T(D_i f)(x)|$$

for all $x \in E$. This proves (5.2). By (5.2) and the arguments similar to those used to derive (2.4) in [29], we have $|\nabla(Tf)(x)| \leqslant |T(|\nabla f|)(x)|$ for almost every $x \in \mathbb{R}^n$. This together with our assumption implies (5.3). □

As an application of Proposition 5.1, we have the following theorem.

Theorem 5.2. *Let $\rho > 2$ and $\mathcal{V}_\rho(\mathcal{T}_K)$ be defined as in* (1.3). *Assume that $K(x,y) = K(x-y)$ and $\mathcal{V}_\rho(\mathcal{T}_K)$ is of type (p,p) for some $p \in (1,\infty)$. Then $\mathcal{V}_\rho(\mathcal{T}_K)$ is bounded on $W^{1,p}(\mathbb{R}^n)$. To be precise, if $f \in W^{1,p}(\mathbb{R}^n)$, then for any $i = 1,\ldots,n$, we have $|D_i(\mathcal{V}_\rho(\mathcal{T}_K)(f))(x)| \leqslant \mathcal{V}_\rho(\mathcal{T}_K)(D_i f)(x)$ for almost every $x \in \mathbb{R}^n$. Moreover, $\|\mathcal{V}_\rho(\mathcal{T}_K)(f)\|_{W^{1,p}(\mathbb{R}^n)} \leqslant \|\mathcal{V}_\rho(\mathcal{T}_K)\|_{L^p(\mathbb{R}^n)\to L^p(\mathbb{R}^n)}\|f\|_{W^{1,p}(\mathbb{R}^n)}$.*

As consequences of Theorem 5.2, we have the following corollaries.

Corollary 5.3. *Let $\rho > 2$ and $\mathcal{V}_\rho(\mathcal{T}_K)$ be defined as in* (1.3). *Assume that $K(x,y) = K(x-y)$ and K satisfies the conditions* (1.5)–(1.8). *Then for any $1 < p < \infty$, the operator $\mathcal{V}_\rho(\mathcal{T}_K)$ is bounded on $W^{1,p}(\mathbb{R}^n)$. To be precise, if $f \in W^{1,p}(\mathbb{R}^n)$, then for any $i = 1,\ldots,n$ and almost every $x \in \mathbb{R}^n$,*

$$|D_i(\mathcal{V}_\rho(\mathcal{T}_K)(f))(x)| \leqslant \mathcal{V}_\rho(\mathcal{T}_K)(D_i f)(x).$$

Moreover, $\|\mathcal{V}_\rho(\mathcal{T}_K)(f)\|_{W^{1,p}(\mathbb{R}^n)} \leqslant \|\mathcal{V}_\rho(\mathcal{T}_K)\|_{L^p(\mathbb{R}^n)\to L^p(\mathbb{R}^n)}\|f\|_{W^{1,p}(\mathbb{R}^n)}$.

Corollary 5.4. *Let $\rho > 2$ and $1 < p < \infty$. Assume that one of the following conditions holds*:

(i) *$n = 1$ and $\mathcal{T} = \mathcal{H}$*;

(ii) *$n = 1$ and $\mathcal{T} = \mathcal{R}^{\pm}$*;

(iii) *$\mathcal{T} = \mathcal{R}_j$, $1 \leqslant j \leqslant n$*;

(iv) *$\mathcal{T} = \mathcal{T}_\Omega$, $\Omega \in H^1(\mathrm{S}^{n-1})$ or $\Omega \in \bigcap_{\alpha>2}\mathcal{F}_\alpha(\mathrm{S}^{n-1})$, where $\mathcal{F}_\alpha(\mathrm{S}^{n-1})$ for $\alpha > 0$ denotes the set of all the integrable functions over S^{n-1} which satisfy the condition*

$$\sup_{\xi'\in \mathrm{S}^{n-1}}\int_{\mathrm{S}^{n-1}}|\Omega(y')|\bigg(\log^+\frac{1}{|\xi'\cdot y'|}\bigg)^\alpha d\sigma(y') < \infty.$$

Then the map $\mathcal{V}_\rho(\mathcal{T}): W^{1,p}(\mathbb{R}^n) \to W^{1,p}(\mathbb{R}^n)$ is bounded. To be precise, if $f \in W^{1,p}(\mathbb{R}^n)$, then for any $i = 1,\ldots,n$, $|D_i(\mathcal{V}_\rho(\mathcal{T})(f))(x)| \leqslant \mathcal{V}_\rho(\mathcal{T})(D_i f)(x)$ for almost every $x \in \mathbb{R}^n$. Moreover,

$$\|\mathcal{V}_\rho(\mathcal{T})(f)\|_{W^{1,p}(\mathbb{R}^n)} \leqslant \|\mathcal{V}_\rho(\mathcal{T})\|_{L^p(\mathbb{R}^n)\to L^p(\mathbb{R}^n)}\|f\|_{W^{1,p}(\mathbb{R}^n)}.$$

Remark 5.5. (i) Corollary 5.3 follows from Theorems A and 5.2. The corresponding results in Corollary 5.4 for the cases (i)–(iii) follow from Theorems A and 5.2.

(ii) The space $\mathcal{F}_\alpha(\mathrm{S}^{n-1})$ was introduced by Grafakos and Stefanov [20] in the study of the L^p boundedness of the singular integral operator with rough kernels. Clearly, $\bigcup_{q>1}L^q(\mathrm{S}^{n-1}) \subsetneq \mathcal{F}_\alpha(\mathrm{S}^{n-1})$ for any $\alpha > 0$. Moreover, the examples in [20] show that

$$\bigcap_{\alpha>1}\mathcal{F}_\alpha(\mathrm{S}^{n-1}) \nsubseteq H^1(\mathrm{S}^{n-1}) \nsubseteq \bigcup_{\alpha>1}\mathcal{F}_\alpha(\mathrm{S}^{n-1}).$$

(iii) It was shown in [14, Theorem 1.2 and Corollary 1.6] that $\mathcal{V}_\rho(\mathcal{T}_\Omega)$ is bounded on $L^p(\mathbb{R}^n)$ for all $p \in (1,\infty)$ under the condition that $\Omega \in H^1(\mathrm{S}^{n-1})$ or $\Omega \in \bigcap_{\alpha>2}\mathcal{F}_\alpha(\mathrm{S}^{n-1})$. This together with Theorem 5.2 yields the conclusions of Corollary 5.4 for the case (iv).

5.2 Boundedness for variation operators of commutators

As mentioned in Section 1, the operator $\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)$ does not commute with translations, even in the special case where $m=1$ and $K(x,y)=K(x-y)$, which makes that Proposition 5.1 does not apply for $\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)$. In order to establish the Sobolev regularity for variation operators of commutators of Calderón-Zygmund singular integrals, we shall make the most of the characterizations of Sobolev functions. At first, let us introduce the definition of Lipschitz spaces.

Definition 5.6 (Lipschitz space)**.** The *homogeneous* Lipschitz space $\dot{\Lambda}(\mathbb{R}^n)$ is given by

$$\dot{\Lambda}(\mathbb{R}^n):=\{f:\mathbb{R}^n\to\mathbb{C}\text{ continuous}:\|f\|_{\dot{\Lambda}(\mathbb{R}^n)}<\infty\},$$

where

$$\|f\|_{\dot{\Lambda}(\mathbb{R}^n)}:=\sup_{x\in\mathbb{R}^n}\sup_{h\in\mathbb{R}^n\setminus\{0\}}\frac{|f(x+h)-f(x)|}{|h|}<\infty.$$

The *inhomogeneous* Lipschitz space $\Lambda(\mathbb{R}^n)$ is defined by

$$\Lambda(\mathbb{R}^n):=\{f:\mathbb{R}^n\to\mathbb{C}\text{ continuous}:\|f\|_{\Lambda(\mathbb{R}^n)}:=\|f\|_{L^\infty(\mathbb{R}^n)}+\|f\|_{\dot{\Lambda}(\mathbb{R}^n)}<\infty\}.$$

Remark 5.7. Let $b\in\dot{\Lambda}(\mathbb{R}^n)$. Then the weak partial derivatives D_ib $(i=1,\ldots,n)$ exist almost everywhere. Moreover, for almost every $x\in\mathbb{R}^n$, we have

$$\lim_{h\to 0}b_{h,i}(x)=D_ib(x) \tag{5.4}$$

and

$$|D_ib(x)|\leqslant\|b\|_{\dot{\Lambda}(\mathbb{R}^n)}. \tag{5.5}$$

To see this, let us fix $i=1,\ldots,n$. Since b is Lipschitz continuous, b is differentiable almost everywhere by Rademacher's theorem. Therefore, the partial derivatives D_ib exist almost everywhere and (5.4) holds. For almost every $x\in\mathbb{R}^n$, we get from (5.4) that

$$|D_ib(x)|=\left|\lim_{h\to0}\frac{b(x+he_i)-b(x)}{h}\right|\leqslant\lim_{h\to0}\frac{|b(x+he_i)-b(x)|}{h}\leqslant\|b\|_{\dot{\Lambda}(\mathbb{R}^n)},$$

which gives (5.5).

The Sobolev regularity for the commutators is the following theorem.

Theorem 5.8. *Let $\rho>2$, $m\geqslant1$, $b\in\Lambda(\mathbb{R}^n)$ and $\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)$ be defined as in* (1.4)*. Assume that $K(x,y)=K(x-y)$ and $\mathcal{V}_\rho(\mathcal{T}_K)$ is of type (p,p) for some $p\in(1,\infty)$. Then $\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)$ is bounded on $W^{1,p}(\mathbb{R}^n)$. To be precise, if $f\in W^{1,p}(\mathbb{R}^n)$, then for any $i=1,\ldots,n$ and almost every $x\in\mathbb{R}^n$,*

$$\begin{aligned}D_i\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f)(x)&\leqslant\sum_{k=0}^m c_m^k\|b\|_{L^\infty(\mathbb{R}^n)}^{m-k}\mathcal{V}_\rho(\mathcal{T}_K)(b^kD_if)(x)\\&\quad+m\sum_{\mu=0}^{m-1}c_{m-1}^\mu\|b\|_{L^\infty(\mathbb{R}^n)}^{m-1-\mu}(\mathcal{V}_\rho(\mathcal{T}_K)(b^\mu D_ibf)(x)+|D_ib|(x)\mathcal{V}_\rho(\mathcal{T}_K)(b^\mu f)(x)).\end{aligned} \tag{5.6}$$

As an application of (5.6)*, we have*

$$\|\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f)\|_{W^{1,p}(\mathbb{R}^n)}\leqslant2^mmnA_p\|b\|_{\Lambda(\mathbb{R}^n)}^m\|f\|_{W^{1,p}(\mathbb{R}^n)}, \tag{5.7}$$

where $A_p:=\|\mathcal{V}_\rho(\mathcal{T}_K)\|_{L^p(\mathbb{R}^n)\to L^p(\mathbb{R}^n)}$.

Proof. We divide the proof of Theorem 5.8 into three steps:

Step 1. Proof of $\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f)\in W^{1,p}(\mathbb{R}^n)$. By the definition of $\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f)$, we conclude that for all $x\in\mathbb{R}^n$,

$$\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f)(x)\leqslant|b(x)|\mathcal{V}_\rho(\mathcal{T}_{K,b}^{m-1})(f)(x)+\mathcal{V}_\rho(\mathcal{T}_{K,b}^{m-1})(bf)(x),\quad\forall\,m\geqslant2. \tag{5.8}$$

When $m = 1$, one can easily check that

$$\mathcal{V}_\rho(\mathcal{T}_{K,b})(f)(x) \leqslant |b(x)|\mathcal{V}_\rho(\mathcal{T}_K)(f)(x) + \mathcal{V}_\rho(\mathcal{T}_K)(bf)(x) \tag{5.9}$$

for all $x \in \mathbb{R}^n$. Then (5.8) and (5.9) give that

$$\mathcal{V}_\rho(\mathcal{T}^m_{K,b})(f)(x) \leqslant \sum_{k=0}^{m} c^k_m |b^{m-k}(x)|\mathcal{V}_\rho(\mathcal{T}_K)(b^k f)(x) \tag{5.10}$$

for all $x \in \mathbb{R}^n$. Combining (5.10) with our assumptions and Minkowski's inequality implies

$$\begin{aligned}\|\mathcal{V}_\rho(\mathcal{T}^m_{K,b})(f)\|_{L^p(\mathbb{R}^n)} &\leqslant \sum_{k=0}^{m} c^k_m \|b\|^{m-k}_{L^\infty(\mathbb{R}^n)} \|\mathcal{V}_\rho(\mathcal{T}_K)(b^k f)\|_{L^p(\mathbb{R}^n)} \\ &\leqslant \sum_{k=0}^{m} c^k_m A_p \|b\|^{m-k}_{L^\infty(\mathbb{R}^n)} \|b^k f\|_{L^p(\mathbb{R}^n)} = 2^m A_p \|b\|^m_{L^\infty(\mathbb{R}^n)} \|f\|_{L^p(\mathbb{R}^n)}. \end{aligned} \tag{5.11}$$

Fix $x, h \in \mathbb{R}^n$. By a change of variables, we have

$$\begin{aligned}\mathcal{V}_\rho(\mathcal{T}^m_{K,b})(f)(x+h) &= \sup_{\epsilon_i \searrow 0} \left(\sum_{i=1}^{\infty} \left| \int_{\epsilon_{i+1} < |x+h-z| \leqslant \epsilon_i} (b(x+h) - b(z))^m K(x+h-z) f(z) dz \right|^\rho \right)^{1/\rho} \\ &= \sup_{\epsilon_i \searrow 0} \left(\sum_{i=1}^{\infty} \left| \int_{\epsilon_{i+1} < |x-z| \leqslant \epsilon_i} (b(x+h) - b(z+h))^m K(x-z) f(z+h) dz \right|^\rho \right)^{1/\rho} \\ &= \mathcal{V}_\rho(\mathcal{T}^m_{K,b_h})(f_h)(x). \end{aligned} \tag{5.12}$$

This yields that the operator $\mathcal{V}_\rho(\mathcal{T}^m_{K,b})$ does not commute with translations. By (5.12) and the sublinearity of $\mathcal{V}_\rho(\mathcal{T}^m_{K,b})$, one finds

$$\begin{aligned}|\Delta_h(\mathcal{V}_\rho(\mathcal{T}^m_{K,b})(f))(x)| &= |\mathcal{V}_\rho(\mathcal{T}^m_{K,b})(f)(x+h) - \mathcal{V}_\rho(\mathcal{T}^m_{K,b})(f)(x)| \\ &= |\mathcal{V}_\rho(\mathcal{T}^m_{K,b_h})(f_h)(x) - \mathcal{V}_\rho(\mathcal{T}^m_{K,b})(f)(x)| \\ &\leqslant \mathcal{V}_\rho(\mathcal{T}^m_{K,b_h})(\Delta_h f)(x) + |\mathcal{V}_\rho(\mathcal{T}^m_{K,b_h})(f)(x) - \mathcal{V}_\rho(\mathcal{T}^m_{K,b})(f)(x)|. \end{aligned} \tag{5.13}$$

By (5.10), we have

$$\mathcal{V}_\rho(\mathcal{T}^m_{K,b_h})(\Delta_h f)(x) \leqslant \sum_{k=0}^{m} c^k_m |b^{m-k}_h(x)| \mathcal{V}_\rho(\mathcal{T}_K)(b^k_h \Delta_h f)(x). \tag{5.14}$$

Note that

$$\begin{aligned}(b_h(x) - b_h(z))^m - (b(x) - b(z))^m &= (b(x) - b(z) + \Delta_h b(x) - \Delta_h b(z))^m - (b(x) - b(z))^m \\ &= \sum_{k=1}^{m} c^k_m (\Delta_h b(x) - \Delta_h b(z))^k (b(x) - b(z))^{m-k} \\ &= \sum_{k=1}^{m} c^k_m \sum_{\ell=0}^{k} c^\ell_k (\Delta_h b(x))^\ell (-\Delta_h b(z))^{k-\ell} (b(x) - b(z))^{m-k}. \end{aligned}$$

This together with (5.10) implies

$$\begin{aligned}&|\mathcal{V}_\rho(\mathcal{T}^m_{K,b_h})(f)(x) - \mathcal{V}_\rho(\mathcal{T}^m_{K,b})(f)(x)| \\ &\quad\leqslant \sup_{\epsilon_i \searrow 0} \left(\sum_{i=1}^{\infty} \left| \int_{\epsilon_{i+1} < |x-z| \leqslant \epsilon_i} ((b_h(x) - b_h(z))^m - (b(x) - b(z))^m) K(x,z) f(z) dz \right|^\rho \right)^{1/\rho} \\ &\quad\leqslant \sum_{k=1}^{m} c^k_m \sum_{\ell=0}^{k} c^\ell_k |\Delta_h b(x)|^\ell \mathcal{V}_\rho(\mathcal{T}^{m-k}_{K,b})((\Delta_h b)^{k-\ell} f) \end{aligned}$$

$$\leqslant \sum_{k=1}^{m} c_m^k \sum_{\ell=0}^{k} c_k^\ell |\Delta_h b(x)|^\ell \sum_{\mu=0}^{m-k} c_{m-k}^\mu |b^{m-k-\mu}(x)| \mathcal{V}_\rho(\mathcal{T}_K)(b^\mu (\Delta_h b)^{k-\ell} f)(x). \tag{5.15}$$

It follows from (5.13)–(5.15) that

$$\begin{aligned}|\Delta_h(\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f))(x)| &\leqslant \sum_{k=0}^{m} c_m^k |b_h^{m-k}(x)| \mathcal{V}_\rho(\mathcal{T}_K)(b_h^k \Delta_h f)(x)\\ &\quad + \sum_{k=1}^{m} c_m^k \sum_{\ell=0}^{k} c_k^\ell |\Delta_h b(x)|^\ell \sum_{\mu=0}^{m-k} c_{m-k}^\mu |b^{m-k-\mu}(x)| \mathcal{V}_\rho(\mathcal{T}_K)(b^\mu (\Delta_h b)^{k-\ell} f)(x). \end{aligned} \tag{5.16}$$

By our assumptions, (5.16) and Minkowski's inequality,

$$\begin{aligned}&\|\Delta_h(\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f))\|_{L^p(\mathbb{R}^n)}\\ &\quad\leqslant \sum_{k=0}^{m} c_m^k \|b\|_{L^\infty(\mathbb{R}^n)}^{m-k} \|\mathcal{V}_\rho(\mathcal{T}_K)(b_h^k \Delta_h f)\|_{L^p(\mathbb{R}^n)}\\ &\quad\quad + \sum_{k=1}^{m} c_m^k \sum_{\ell=0}^{k} c_k^\ell (\|b\|_{\dot{\Lambda}(\mathbb{R}^n)}|h|)^\ell \sum_{\mu=0}^{m-k} c_{m-k}^\mu \|b\|_{L^\infty(\mathbb{R}^n)}^{m-k-\mu} \|\mathcal{V}_\rho(\mathcal{T}_K)(b^\mu(\Delta_h b)^{k-\ell} f)\|_{L^p(\mathbb{R}^n)}\\ &\quad\leqslant \sum_{k=0}^{m} c_m^k A_p \|b\|_{L^\infty(\mathbb{R}^n)}^{m-k} \|b_h^k \Delta_h f\|_{L^p(\mathbb{R}^n)}\\ &\quad\quad + \sum_{k=1}^{m} A_p c_m^k \sum_{\ell=0}^{k} c_k^\ell (\|b\|_{\dot{\Lambda}(\mathbb{R}^n)}|h|)^\ell \sum_{\mu=0}^{m-k} c_{m-k}^\mu \|b\|_{L^\infty(\mathbb{R}^n)}^{m-k-\mu} \|b^\mu (\Delta_h b)^{k-\ell} f\|_{L^p(\mathbb{R}^n)}\\ &\quad\leqslant 2^m A_p \|b\|_{L^\infty(\mathbb{R}^n)}^m \|\Delta_h f\|_{L^p(\mathbb{R}^n)} + 2^m A_p \|f\|_{L^p(\mathbb{R}^n)} \bigg(\sum_{k=1}^{m} c_m^k \|b\|_{L^\infty(\mathbb{R}^n)}^{m-k} \|b\|_{\dot{\Lambda}(\mathbb{R}^n)}^k |h|^k\bigg).\end{aligned}$$

It follows that

$$\begin{aligned}G(\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f);p) &= \limsup_{|h|\to 0} \frac{\|\Delta_h(\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f))\|_{L^p(\mathbb{R}^n)}}{|h|}\\ &\leqslant 2^m A_p \|b\|_{L^\infty(\mathbb{R}^n)}^m G(f;p) + 2^m m A_p \|f\|_{L^p(\mathbb{R}^n)} \|b\|_{L^\infty(\mathbb{R}^n)}^{m-1} \|b\|_{\dot{\Lambda}(\mathbb{R}^n)} < \infty,\end{aligned} \tag{5.17}$$

where in the last inequality of (5.17) we have used the fact that $G(f;p) < \infty$ since $f \in W^{1,p}(\mathbb{R}^n)$. Combining (5.17) with (5.1) and (5.11) yields $\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f) \in W^{1,p}(\mathbb{R}^n)$.

Step 2. Proof of (5.6). Observe that

$$b_h^k(x) = (\Delta_h b(x) + b(x))^k = \sum_{\iota=0}^{k} c_k^\iota (\Delta_h b(x))^\iota b(x)^{k-\iota}, \quad b_h^{m-k}(x) = \sum_{\nu=0}^{m-k} c_{m-k}^\nu (\Delta_h b(x))^\nu b(x)^{m-k-\nu}.$$

These together with the sublinearity of $\mathcal{V}_\rho(\mathcal{T}_K)$ imply

$$\begin{aligned}&|b_h^{m-k}(x)| \mathcal{V}_\rho(\mathcal{T}_K)(b_h^k \Delta_h f)(x)\\ &\quad\leqslant \bigg(\sum_{\nu=0}^{m-k} c_{m-k}^\nu |\Delta_h b(x)|^\nu |b(x)|^{m-k-\nu}\bigg)\bigg(\sum_{\iota=0}^{k} c_k^\iota \mathcal{V}_\rho(\mathcal{T}_K)((\Delta_h b)^\iota b^{k-\iota} \Delta_h f)(x)\bigg).\end{aligned} \tag{5.18}$$

In light of (5.16) and (5.18) we would have that for all $x, h \in \mathbb{R}^n$,

$$\begin{aligned}&|\Delta_h(\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f))(x)|\\ &\quad\leqslant \sum_{k=0}^{m} c_m^k \bigg(\sum_{\nu=0}^{m-k} c_{m-k}^\nu |\Delta_h b(x)|^\nu |b(x)|^{m-k-\nu}\bigg)\bigg(\sum_{\iota=0}^{k} c_k^\iota \mathcal{V}_\rho(\mathcal{T}_K)((\Delta_h b)^\iota b^{k-\iota} \Delta_h f)(x)\bigg)\end{aligned}$$

$$+\sum_{k=1}^{m}c_m^k\sum_{\ell=0}^{k}c_k^\ell|\Delta_h b(x)|^\ell\sum_{\mu=0}^{m-k}c_{m-k}^\mu|b^{m-k-\mu}(x)|\mathcal{V}_\rho(\mathcal{T}_K)(b^\mu(\Delta_h b)^{k-\ell}f)(x). \tag{5.19}$$

Fix $i\in\{1,\ldots,n\}$. When $1\leqslant k\leqslant m$ and $1\leqslant\iota\leqslant k$, by the L^p boundedness and the sublinearity for $\mathcal{V}_\rho(\mathcal{T}_K)$ and the fact that $f_{h,i}\to D_if$ in $L^p(\mathbb{R}^n)$ as $k\to\infty$, we have

$$\begin{aligned}&\|\mathcal{V}_\rho(\mathcal{T}_K)((\Delta_{he_i}b)^\iota b^{k-\iota}f_{h,i})\|_{L^p(\mathbb{R}^n)}\\&\quad\leqslant A_p\|b\|_{\dot{\Lambda}(\mathbb{R}^n)}\|b\|_{L^\infty(\mathbb{R}^n)}^{k-\iota}|h|^\iota\|f_{h,i}\|_{L^p(\mathbb{R}^n)}\to 0\quad\text{as } h\to 0\end{aligned} \tag{5.20}$$

and

$$\begin{aligned}&\|\mathcal{V}_\rho(\mathcal{T}_K)(b^kf_{h,i})-\mathcal{V}_\rho(\mathcal{T}_K)(b^kD_if)\|_{L^p(\mathbb{R}^n)}\\&\quad\leqslant\|\mathcal{V}_\rho(\mathcal{T}_K)(b^k(f_{h,i}-D_if))\|_{L^p(\mathbb{R}^n)}\\&\quad\leqslant A_p\|b\|_{L^\infty(\mathbb{R}^n)}^k\|f_{h,i}-D_if\|_{L^p(\mathbb{R}^n)}\to 0\quad\text{as } h\to 0.\end{aligned} \tag{5.21}$$

When $2\leqslant k\leqslant m$ and $0\leqslant\mu\leqslant m-k$, it is easy to see that

$$\begin{aligned}&\|\mathcal{V}_\rho(\mathcal{T}_K)(b^\mu(\Delta_{he_i}b)^{k-1}b_{h,i}f)\|_{L^p(\mathbb{R}^n)}\\&\quad\leqslant A_p\|b^\mu(\Delta_{he_i}b)^{k-1}b_{h,i}f\|_{L^p(\mathbb{R}^n)}\\&\quad\leqslant A_p\|b\|_{L^\infty(\mathbb{R}^n)}^\mu\|b\|_{\dot{\Lambda}(\mathbb{R}^n)}^k|h|^{k-1}\|f\|_{L^p(\mathbb{R}^n)}\to 0\quad\text{as } h\to 0.\end{aligned} \tag{5.22}$$

By (5.4), we have

$$b^\mu(x)b_{h,i}(x)f(x)\to b^\mu(x)D_ib(x)f(x)$$

as $h\to\infty$ for almost every $x\in\mathbb{R}^n$. By (5.5), we have

$$\|b^\mu D_ibf\|_{L^p(\mathbb{R}^n)}\leqslant\|b\|_{L^\infty(\mathbb{R}^n)}^\mu\|b\|_{\dot{\Lambda}(\mathbb{R}^n)}\|f\|_{L^p(\mathbb{R}^n)},$$

which combining the dominated convergence theorem implies that $b^\mu b_{h,i}f\to b^\mu D_ibf$ in $L^p(\mathbb{R}^n)$ as $h\to\infty$. These together with the sublinearity and the L^p boundedness for $\mathcal{V}_\rho(\mathcal{T}_K)$ imply that

$$\|\mathcal{V}_\rho(\mathcal{T}_K)(b^\mu b_{h,i}f)-\mathcal{V}_\rho(\mathcal{T}_K)(b^\mu D_ibf)\|_{L^p(\mathbb{R}^n)}\leqslant A_p\|b^\mu b_{h,i}f-b^\mu D_ibf\|_{L^p(\mathbb{R}^n)}\to 0\quad\text{as } h\to 0 \tag{5.23}$$

for $0\leqslant\mu\leqslant m-1$. By the arguments similar to those used to derive (5.22),

$$\|\mathcal{V}_\rho(\mathcal{T}_K)(b^\mu(\Delta_{he_i}b)^{k-\ell}f)\|_{L^p(\mathbb{R}^n)}\to 0\quad\text{as } h\to 0 \tag{5.24}$$

for $1\leqslant k\leqslant m$ and $1\leqslant\ell\leqslant k-1$.

Since $f\in W^{1,p}(\mathbb{R}^n)$ and $\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f)\in W^{1,p}(\mathbb{R}^n)$, we have $(\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f))_{h,i}\to D_i(\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f))$ and $f_{h,i}\to D_if$ in $L^p(\mathbb{R}^n)$ when $h\to 0$. These together with (5.20)–(5.24) and Remark 5.7 imply that there exist a sequence $\{h_j\}$ of positive numbers satisfying $\lim_{j\to 0}h_j=0$ and a measurable set E with $|\mathbb{R}^n\setminus E|=0$ such that for all $x\in E$,

(i) $(\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f))_{h_j,i}(x)\to D_i(\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f))(x)$ as $j\to\infty$;

(ii) $|b(x)|\leqslant\|b\|_{L^\infty(\mathbb{R}^n)}$, $|D_ib|(x)\leqslant\|b\|_{\dot{\Lambda}(\mathbb{R}^n)}$ and $b_{h_j,i}(x)\to D_ib(x)$ as $j\to\infty$;

(iii) $\mathcal{V}_\rho(\mathcal{T}_K)((\Delta_{h_je_i}b)^\iota b^{k-\iota}f_{h_j,i})(x)\to 0$ as $j\to\infty$ for $1\leqslant k\leqslant m$ and $1\leqslant\iota\leqslant k$;

(iv) $\mathcal{V}_\rho(\mathcal{T}_K)(b^kf_{h_j,i})(x)\to\mathcal{V}_\rho(\mathcal{T}_K)(b^kD_if)(x)$ as $j\to\infty$ for $1\leqslant k\leqslant m$;

(v) $\mathcal{V}_\rho(\mathcal{T}_K)(b^\mu(\Delta_{h_je_i}b)^{k-1}b_{h_j,i}f)(x)\to 0$ as $j\to\infty$ for $2\leqslant k\leqslant m$ and $0\leqslant\mu\leqslant m-k$;

(vi) $\mathcal{V}_\rho(\mathcal{T}_K)(b^\mu b_{h_j,i}f)(x)\to\mathcal{V}_\rho(\mathcal{T}_K)(b^\mu D_ibf)(x)$ as $j\to\infty$ for $0\leqslant\mu\leqslant m-1$;

(vii) $\mathcal{V}_\rho(\mathcal{T}_K)(b^\mu(\Delta_{h_je_i}b)^{k-\ell}f)(x)\to 0$ as $j\to\infty$ for $1\leqslant k\leqslant m$ and $1\leqslant\ell\leqslant k-1$.

Hence, we get by (5.20) and the above (i)–(vii) that for any $x\in E$,

$$\begin{aligned}&|D_i(\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f))(x)|\\&\quad=\Big|\lim_{j\to\infty}(\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f))_{h_j,i}(x)\Big|\end{aligned}$$

$$
\begin{aligned}
&\leqslant \lim_{j\to\infty}\sum_{k=0}^{m} c_m^k\bigg(\sum_{\nu=0}^{m-k} c_{m-k}^{\nu}|\Delta_{h_je_i}b(x)|^{\nu}|b(x)|^{m-k-\nu}\bigg)\\
&\quad\times\bigg(\sum_{\iota=0}^{k} c_k^{\iota}\mathcal{V}_\rho(\mathcal{T}_K)((\Delta_{h_je_i}b)^{\iota}b^{k-\iota}f_{h_j,i})(x)\bigg)\\
&\quad+\lim_{j\to\infty}\sum_{k=1}^{m} c_m^k\sum_{\mu=0}^{m-k} c_{m-k}^{\mu}|b^{m-k-\mu}(x)|\mathcal{V}_\rho(\mathcal{T}_K)(b^{\mu}(\Delta_{h_je_i}b)^{k-1}b_{h_j,i}f)(x)\\
&\quad+\lim_{j\to\infty}\sum_{k=1}^{m} c_m^k\sum_{\ell=1}^{k} c_k^{\ell}|\Delta_{h_je_i}b(x)|^{\ell-1}|b_{h_j,i}|(x)\\
&\quad\times\sum_{\mu=0}^{m-k} c_{m-k}^{\mu}|b^{m-k-\mu}(x)|\mathcal{V}_\rho(\mathcal{T}_K)(b^{\mu}(\Delta_{h_je_i}b)^{k-\ell}f)(x)\\
&\leqslant \sum_{k=0}^{m} c_m^k\|b\|_{L^\infty(\mathbb{R}^n)}^{m-k}\mathcal{V}_\rho(\mathcal{T}_K)(b^kD_if)(x)\\
&\quad+m\sum_{\mu=0}^{m-1} c_{m-1}^{\mu}\|b\|_{L^\infty(\mathbb{R}^n)}^{m-1-\mu}(\mathcal{V}_\rho(\mathcal{T}_K)(b^{\mu}D_ibf)(x)+|D_ib|(x)\mathcal{V}_\rho(\mathcal{T}_K)(b^{\mu}f)(x)).
\end{aligned}
$$

This proves (5.6).

Step 3. Proof of (5.7). By (5.5), (5.6), our assumptions and Minkowski's inequality, we have

$$
\begin{aligned}
&\|D_i\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)(f)\|_{L^p(\mathbb{R}^n)}\\
&\quad\leqslant \sum_{k=0}^{m} c_m^k\|b\|_{L^\infty(\mathbb{R}^n)}^{m-k}\|\mathcal{V}_\rho(\mathcal{T}_K)(b^kD_if)\|_{L^p(\mathbb{R}^n)}\\
&\quad\quad+m\sum_{\mu=0}^{m-1} c_{m-1}^{\mu}\|b\|_{L^\infty(\mathbb{R}^n)}^{m-1-\mu}(\|\mathcal{V}_\rho(\mathcal{T}_K)(b^{\mu}D_ibf)\|_{L^p(\mathbb{R}^n)}+\|D_ib\mathcal{V}_\rho(\mathcal{T}_K)(b^{\mu}f)\|_{L^p(\mathbb{R}^n)})\\
&\quad\leqslant A_p2^m\|b\|_{L^\infty(\mathbb{R}^n)}^{m}\|D_if\|_{L^p(\mathbb{R}^n)}+2^m mA_p\|b\|_{L^\infty(\mathbb{R}^n)}^{m-1}\|b\|_{\dot{\Lambda}(\mathbb{R}^n)}\|f\|_{L^p(\mathbb{R}^n)}. \qquad (5.25)
\end{aligned}
$$

Combining (5.25) with (5.11) leads to (5.7). Then we finish the proof of Theorem 5.8. □

As several applications of Theorem 5.2, we have the following corollaries.

Corollary 5.9. *Let $\rho>2$, $m\geqslant 1$, $b\in\Lambda(\mathbb{R}^n)$ and $\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)$ be defined as in (1.4). Suppose that $K(x,y)=K(x-y)$ and K satisfies the conditions (1.5)–(1.8). Then for any $1<p<\infty$, the operator $\mathcal{V}_\rho(\mathcal{T}_{K,b}^m)$ is bounded on $W^{1,p}(\mathbb{R}^n)$. To be precise, if $f\in W^{1,p}(\mathbb{R}^n)$, then for any $i=1,\dots,n$ and almost every $x\in\mathbb{R}^n$,*

$$
\begin{aligned}
D_i\mathcal{V}_\rho(\mathcal{T}_{K,b}^mf)(x)\leqslant{}& \sum_{k=0}^{m} c_m^k\|b\|_{L^\infty(\mathbb{R}^n)}^{m-k}\mathcal{V}_\rho(\mathcal{T}_K)(b^kD_if)(x)\\
&+m\sum_{\mu=0}^{m-1} c_{m-1}^{\mu}\|b\|_{L^\infty(\mathbb{R}^n)}^{m-1-\mu}(\mathcal{V}_\rho(\mathcal{T}_K)(b^{\mu}D_ibf)(x)+|D_ib|(x)\mathcal{V}_\rho(\mathcal{T}_K)(b^{\mu}f)(x)).
\end{aligned}
$$

As an application of the above estimate, we have

$$
\|\mathcal{V}_\rho(\mathcal{T}_{K,b}^mf)\|_{W^{1,p}(\mathbb{R}^n)}\leqslant 2^m mn\|\mathcal{V}_\rho(\mathcal{T}_K)\|_{L^p(\mathbb{R}^n)\to L^p(\mathbb{R}^n)}\|b\|_{\Lambda(\mathbb{R}^n)}^m\|f\|_{W^{1,p}(\mathbb{R}^n)}.
$$

Corollary 5.10. *Let $m\geqslant 1$, $\rho>2$, $1<p<\infty$, $b\in\Lambda(\mathbb{R}^n)$ and one of the following conditions hold*:

(i) $n=1$ *and* $\mathcal{T}=\mathcal{H}_b^m$;

(ii) $n=1$ *and* $\mathcal{T}=\mathcal{R}_{\pm,b}^m$;

(iii) $\mathcal{T}=\mathcal{R}_{j,b}^m$, $1\leqslant j\leqslant n$;

(iv) $\mathcal{T} = \mathcal{T}^m_{\Omega,b}$, $\Omega \in H^1(\mathrm{S}^{n-1})$ *or* $\Omega \in \bigcap_{\alpha>2} \mathcal{F}_\alpha(\mathrm{S}^{n-1})$.

Then the map $\mathcal{V}_\rho(\mathcal{T}) : W^{1,p}(\mathbb{R}^n) \to W^{1,p}(\mathbb{R}^n)$ *is bounded. To be precise, if* $f \in W^{1,p}(\mathbb{R}^n)$*, we have*

$$\|\mathcal{V}_\rho(\mathcal{T})(f)\|_{W^{1,p}(\mathbb{R}^n)} \lesssim_{m,n,p,\rho,\Omega} \|b\|^m_{\Lambda(\mathbb{R}^n)} \|f\|_{W^{1,p}(\mathbb{R}^n)}.$$

Remark 5.11. (i) It should be pointed out that Corollary 5.9 follows from Theorems A and 5.8. The corresponding results in Corollary 5.10 for the cases (i)–(iii) follow from Theorems A and 5.8.

(ii) It was known that $\mathcal{V}_\rho(\mathcal{T}_\Omega)$ is bounded on $L^p(\mathbb{R}^n)$ for all $p \in (1,\infty)$ under the condition that $\Omega \in H^1(\mathrm{S}^{n-1})$ or $\Omega \in \bigcap_{\alpha>2} \mathcal{F}_\alpha(\mathrm{S}^{n-1})$. This together with Theorem 5.8 yields the conclusion of Corollary 5.10 for the case (iv).

Question 5.12. What happens when we consider the Calderón-Zygmund singular integrals of non-convolution type? To be more precise, do the corresponding results in Theorems 5.2 and 5.8 and Corollaries 5.3 and 5.9 hold when $K(x,y) \neq K(x-y)$?

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant No. 11701333). The authors express their sincerely thanks to the referees for their valuable remarks and suggestions, which made this paper more readable.

References

1 Akcoglu M, Jones R L, Schwartz P. Variation in probability, ergodic theory and analysis. Illinois J Math, 1998, 42: 154–177

2 Bourgain J. Pointwise ergodic theorems for arithmetic sets. Publ Math Inst Hautes Études Sci, 1989, 69: 5–45

3 Campbell J T, Jones R L, Reinhdd K, et al. Oscillation and variation for the Hilbert transform. Duke Math J, 2000, 105: 59–83

4 Campbell J T, Jones R L, Reinhdd K, et al. Oscillation and variation for singular integrals in higher dimensions. Trans Amer Math Soc, 2003, 355: 2115–2137

5 Chen Y, Ding Y, Hong G, et al. Weighted jump and variational inequalities for rough operators. J Funct Anal, 2018, 274: 2446–2475

6 Chen Y, Ding Y, Hong G, et al. Variational inequalities for the commutators of rough operators with BMO functions. Sci China Math, 2021, 64: 2437–2460

7 Chen Y, Ding Y, Wang X. Compactness of commutators for singular integrals on Morrey spaces. Canad J Math, 2012, 64: 257–281

8 Clop A, Cruz V. Weighted estimates for Beltrami equations. Ann Acad Sci Fenn Math, 2013, 38: 91–113

9 Coifman R, Fefferman C. Weighted norm inequalities for maximal functions and singular integrals. Studia Math, 1974, 51: 241–250

10 Coifman R, Meyer Y. Commutateurs d'intégrales singulières et opérateurs multilinéaires. Ann Inst Fourier Grenoble, 1978, 28: 177–202

11 Coifman R R, Rochberg R, Weiss G. Factorization theorems for Hardy spaces in several variables. Ann of Math (2), 1976, 103: 611–635

12 Cordes H O. On compactness of commutators of multiplications and convolutions, and boundedness of pseudodifferential operators. J Funct Anal, 1975, 18: 115–131

13 Crescimbeni R, Martin-reyes F J, Torre A L, et al. The ρ-variation of the Hermitian Riesz transform. Acta Math Sin Engl Ser, 2010, 26: 1827–1838

14 Ding Y, Hong G, Liu H. Jump and variational inequalities for rough operators. J Fourier Anal Appl, 2017, 23: 679–711

15 Ding Y, Lu S, Yabuta K. On commutators of Marcinkiewicz integrals with rough kernel. J Math Anal Appl, 2002, 275: 60–68

16 Duoandikoetexa J, Rosenthal M. Extension and boundedness of operators on Morrey spaces from extrapolation techniques and embeddings. J Geom Anal, 2018, 28: 3081–3108

17 Gilbarg D, Trudinger N S. Elliptic Partial Differential Equations of Second Order, 2nd ed. Berlin: Springer-Verlag, 1983

18 Gillespie T A, Torrea J L. Dimension free estimates for the oscillation of Riesz transforms. Israel J Math, 2004, 141: 125–144

19 Grafakos L. Modern Fourier Analysis, 2nd ed. Graduate Texts in Mathematics, vol. 249. New York: Springer, 2008

20 Grafakos L, Stefanov A. L^p bounds for singular integrals and maximal singular integrals with rough kernels. Indiana Univ Math J, 1998, 47: 455–469

21 Guo W, Wen Y, Wu H, et al. On the compactness of oscillation and variation of commutators. Banach J Math Anal, 2021, 15: 37
22 Guo W, Zhao G. On relatively compact sets in quasi-Banach function spaces. Proc Amer Math Soc, 2020, 148: 3359–3373
23 Hajłasz P, Onninen J. On boundedness of maximal functions in Sobolev spaces. Ann Acad Sci Fenn Math, 2004, 29: 167–176
24 Hart J, Torres R H. John-Nirenberg inequalities and weight invariant BMO spaces. J Geom Anal, 2019, 29: 1608–1648
25 Ho K-P. Characterizations of BMO spaces by A_p weights and p-convexity. Hiroshima Math J, 2011, 41: 153–165
26 Jones R L, Kaufman R, Rosenblatt J, et al. Oscillation in ergodic theory. Ergodic Theory Dynam Systems, 1998, 18: 889–935
27 Jones R L, Rosenblatt J, Wierdl M. Oscillation in ergodic theory: Higher dimensional results. Israel J Math, 2003, 135: 1–27
28 Jones R L, Seeger A, Wright J. Strong variational and jump inequalities in harmonic analysis. Trans Amer Math Soc, 2008, 360: 6711–6742
29 Kinnunen J. The Hardy-Littlewood maximal function of a Sobolev function. Israel J Math, 1997, 100: 117–124
30 Komori Y, Shirai S. Weighted Morrey spaces and a singular integral operator. Math Nachr, 2009, 282: 219–231
31 Krantz S G, Li S. Boundedness and compactness of integral operators on spaces of homogeneous type and applications, II. J Math Anal Appl, 2001, 258: 642–657
32 Lépingle D. La variation d'ordre p des semi-martingales. Z Wahrsch Verw Gebiete, 1976, 36: 295–316
33 Liu F, Wu H. A criterion on oscillation and variation for the commutators of singular integral operators. Forum Math, 2015, 27: 77–97
34 Lu S, Ding Y, Yan D. Singular Integral and Related Topics. Singapore: World Scientific Publishing, 2007
35 Ma T, Torrea J, Xu Q. Weighted variation inequalities for differential operators and singular integrals. J Funct Anal, 2015, 268: 376–416
36 Ma T, Torrea J L, Xu Q H. Weighted variation inequalities for differential operators and singular integrals in higher dimensions. Sci China Math, 2017, 60: 1419–1442
37 Mirek M, Stein E M, Trojan B. $\ell^p(\mathbb{Z}^d)$-estimates for discrete operators of Radon type: Variational estimates. Invent Math, 2017, 209: 665–748
38 Mirek M, Trojan B, Zorin-Kranich P. Variational estimates for averages and truncated singular integrals along the prime numbers. Trans Amer Math Soc, 2017, 69: 5403–5423
39 Morrey C B. On the solutions of quasi-linear elliptic partial differential equations. Trans Amer Math Soc, 1938, 43: 126–166
40 Nakamura S. Generalized weighted Morrey spaces and classical operators. Math Nachr, 2016, 289: 2235–2262
41 Nakamura S, Sawano Y, Tanaka H. The fractional operators on weighted Morrey spaces. J Geom Anal, 2018, 28: 1502–1524
42 Pisier G, Xu Q. The strong p-variation of martingales and orthogonal series. Probab Theory Related Fields, 1988, 77: 497–514
43 Stempak K, Torrea J L. Poisson integrals and Riesz transforms for Hermite function expansions with weights. J Funct Anal, 2003, 202: 443–472
44 Uchiyama A. On the compactness of operators of Hankel type. Tohoku Math J (2), 1978, 30: 163–171
45 Yosida K. Functional Analysis, 6th ed. Classics in Mathematics. Berlin: Springer-Verlag, 1995
46 Zhang J, Wu H. Oscillation and variation inequalities for singular integrals and commutators on weighted Morrey spaces. Front Math China, 2016, 11: 423–447
47 Zhang J, Wu H. Weighted oscillation and variation inequalities for singular integrals and commutators satisfying Hörmander type conditions. Acta Math Sin Engl Ser, 2017, 33: 1397–1420