Appl. Math. J. Chinese Univ.
2019, 34(1): 55-75

Compactness for the commutators of multilinear singular integral operators with non-smooth kernels

BU Rui[1] CHEN Jie-cheng[*,2]

Abstract. In this paper, the behavior for commutators of a class of bilinear singular integral operators associated with non-smooth kernels on the product of weighted Lebesgue spaces is considered. By some new maximal functions to control the commutators of bilinear singular integral operators and CMO($\mathbf{R}^n$) functions, compactness for the commutators is proved.

§1 Introduction

In recent decades, the study of multilinear analysis becomes an active topic in harmonic analysis. The first important work, among several pioneer papers, is the famous work by Coifman and Meyer in [8,9], where they established a bilinear multiplier theorem on the Lebesgue spaces. Note that a multilinear multiplier actually is a convolution operator. Naturally one will study the non-convolution operator

$$T(f_1,\dots,f_m)(x)=\int_{(\mathbf{R}^n)^m}K(x,y_1,\dots,y_m)f_1(y_1)\cdots f_m(y_m)\mathrm{d}y_1\cdots\mathrm{d}y_m, \qquad (1)$$

where $K(x,y_1,\dots,y_m)$ is a locally integral function defined away from the diagonal $x=y_1=\cdots=y_m$ in $(\mathbf{R}^n)^{m+1}$, $x\notin\cap_{j=1}^m\mathrm{supp}\, f_j$ and $f_1,\dots,f_m$ are bounded functions with compact supports. Precisely,

$$T:\mathcal{S}(\mathbf{R}^n)\times\cdots\times\mathcal{S}(\mathbf{R}^n)\mapsto\mathcal{S}'(\mathbf{R}^n)$$

is an m-linear operator associated with the kernel $K(x,y_1,\dots,y_m)$. If there exist positive constants A and $\gamma\in(0,1]$ such that K satisfies the size condition

$$|K(x,y_1,\dots,y_m)|\le\frac{A}{(|x-y_1|+\cdots+|x-y_m|)^{mn}} \qquad (2)$$

Received: 2016-10-13. Revised: 2018-12-10.
MR Subject Classification: 42B20, 42B25.
Keywords: singular integral operator, maximal function, weighted norm inequality, commutator, compact operator.
Digital Object Identifier(DOI): https://doi.org/10.1007/s11766-019-3501-z.
Supported by the Natural Science Foundation of Shandong Province(Nos. ZR2018PA004 and ZR2016AB07), the National Natural Science Foundation of China(Nos. 11571306 and 11671363).
*Corresponding author.

for all $(x, y_1, \ldots, y_m) \in (\mathbf{R}^n)^{m+1}$ with $x \neq y_j$ for some $j \in \{1, 2, \ldots, m\}$; and the smoothness condition

$$|K(x, y_1, \ldots, y_m) - K(x', y_1, \ldots, y_m)| \leq \frac{A|x-x'|^\gamma}{(\sum_{i=1}^m |x-y_i|)^{mn+\gamma}},$$

whenever $|x-x'| \leq \frac{1}{2}\max_{1\leq j\leq m}|x-y_j|$ and also for each j,

$$|K(x, y_1, \ldots, y_j, \ldots, y_m) - K(x, y_1, \ldots, y_j', \ldots, y_m)| \leq \frac{A|y_j-y_j'|^\gamma}{(\sum_{i=1}^m |x-y_i|)^{mn+\gamma}}, \tag{3}$$

whenever $|y_j - y_j'| \leq \frac{1}{2}\max_{1\leq j\leq m}|x-y_j|$, then we say that K is a Calderón-Zygmund kernel and denote it by $K \in m - CZK(A, \gamma)$. Also, T is called the multilinear Calderón-Zygmund operator associated with the kernel K. In [16], Grafakos and Torres established the multilinear $T1$ theorem, so that they obtained the strong type boundedness on product of $L^p(\mathbf{R}^n)$ spaces and endpoint weak type estimates of operators T associated with kernels $K \in m - CZK(A, \gamma)$. Furthermore, the $A_p(\mathbf{R}^n)$ weights (see Definition 1.2) on the operator T and on the corresponding maximal operator were considered in [15]. After then, the study of multilinear Calderón-Zygmund operator is fruitful. The readers can refer to [14-16,21,23,25-28,32] and the references therein.

However, there are some multilinear singular integral operators, including the Calderón commutator, whose kernels do not satisfy (3). Here, the Calderón commutator is defined by

$$\mathcal{C}_{m+1}(f, a_1, \ldots, a_m)(x) = \int_{\mathbf{R}} \frac{\prod_{j=1}^m (A_j(x) - A_j(y))}{(x-y)^{m+1}} f(y)\mathrm{d}y,$$

where $A_j' = a_j$ for all $j \in \{1, 2, \cdots, m\}$. This operator first appeared in the study of Cauchy integral along Lipschitz curves and, in fact, led to the first proof of the L^2-boundedness of the latter. Moreover, the operator $\mathcal{C}_{m+1}$ can be viewed as an $(m+1)$-linear operator. Define

$$e(x) = \begin{cases} 1 & : \quad x > 0, \\ 0 & : \quad x < 0. \end{cases}$$

Since $A_j' = a_j$, the $(m+1)$-linear operator $\mathcal{C}_{m+1}(f, a_1, \ldots, a_m)$ can be written as

$$\mathcal{C}_{m+1}(f, a_1, \ldots, a_m)(x) = \int_{(\mathbf{R})^{m+1}} K(x, y_1, \ldots, y_{m+1}) a_1(y_1) \cdots a_m(y_m) f(y_{m+1}) \mathrm{d}y_1 \cdots \mathrm{d}y_{m+1},$$

where the kernel K is

$$K(x, y_1, \ldots, y_{m+1}) = \frac{(-1)^{e(y_{m+1}-x)m}}{(x-y_{m+1})^{m+1}} \prod_{j=1}^m \chi_{(\min(x, y_{m+1}), \max(x, y_{m+1}))}(y_j). \tag{4}$$

In [11], the authors pointed out that the kernel K in (4) does not satisfy (3). In order to discuss these operators whose kernels don't satisfy (3), Duong et al. [10,11] introduced a class of multilinear singular integral operators whose kernels satisfy the following assumptions (H1),(H2) and (H3).

Let $\{A_t\}_{t>0}$ be a class of integral operators, which play the role of an approximation to the identity as in [12]. We always assume that the operators A_t are associated with kernels $a_t(x, y)$ in the sense that

$$A_t f(x) = \int_{\mathbf{R}^n} a_t(x, y) f(y)\mathrm{d}y$$

for all $f \in L^p(\mathbf{R}^n)$, $1 \leq p \leq \infty$, and the kernels $a_t(x,y)$ satisfy

$$|a_t(x,y)| \leq h_t(x,y) = t^{-n/s} h(\frac{|x-y|}{t^{1/s}}), \tag{5}$$

where s is a fixed positive constant and h is a bounded, positive, decreasing function satisfying

$$\lim_{r\to\infty} r^{n+\iota} h(r^s) = 0 \tag{6}$$

for some $\iota > 0$.

Assumption (H1). Assume that for each $j \in \{1,2,\cdots,m\}$, there exist operators $\{A_t^{(j)}\}_{t>0}$ with kernels $a_t^{(j)}(x,y)$ that satisfy conditions (5) and (6) with constants s and ι, and there exist kernels $K_t^{(j)}(x,y_1,\cdots,y_m)$ such that

$$< T(f_1,\cdots,A_t^{(j)} f_j,\cdots,f_m), g >= \int_{\mathbf{R}^n}\int_{(\mathbf{R}^n)^m} K_t^{(j)}(x,y_1,\cdots,y_m) f_1(y_1)\cdots f_m(y_m) g(x) \mathrm{d}\vec{y}\mathrm{d}x$$

for all $f_1,\cdots,f_m, g \in \mathcal{S}(\mathbf{R}^n)$ with $\cap_{j=1}^m \mathrm{supp}\, f_j \cap \mathrm{supp}\, g = \emptyset$, and there exist a function $\Phi \in \mathcal{C}(\mathbf{R})$ with $\mathrm{supp}\Phi \subset [-1,1]$ and a constant $\epsilon > 0$ such that for all $x, y_1,\cdots,y_m \in \mathbf{R}^n$ and $t>0$, we have

$$|K(x,y_1,\cdots,y_m) - K_t^{(j)}(x,y_1,\cdots,y_m)| \leq \frac{C}{(|x-y_1|+\cdots+|x-y_m|)^{mn}} \sum_{k=1,k\neq j}^{m} \Phi(\frac{|y_j-y_k|}{t^{1/s}}) + \frac{Ct^{\epsilon/s}}{(|x-y_1|+\cdots+|x-y_m|)^{mn+\epsilon}}$$

for some $C>0$, whenever $2t^{1/s} \leq |x-y_j|$.

Kernels K satisfying condition (2) and assumption (H1) with parameters m, C, s, ι, ϵ are called generalized Calderón-Zygmund kernels, and their collection is denoted by $m{-}GCZK_0(C, s,\iota,\epsilon)$. It was proved in [11] that the assumption (H1) is weaker than condition (3). But it is sufficient for the weak type endpoint estimate.

Recall that the j-th transpose T^{*j} of T is defined via

$$< T^{*j}(f_1,\cdots,f_m), h >=< T(f_1,\cdots,f_{j-1},h,f_{j+1},\cdots,f_m), f_j >$$

for all $f_1,\cdots,f_m, h \in \mathcal{S}(\mathbf{R}^n)$. It is easy to check that the kernel K^{*j} of T^{*j} is related to the kernel K of T via the identity

$$K^{*j}(x,y_1,\cdots,y_{j-1},y_j,y_{j+1},\cdots,y_m) = K(y_j,y_1,\cdots,y_{j-1},x,y_{j+1},\cdots,y_m).$$

Notice that if a multilinear operator T maps a product of Banach spaces $X_1\times\cdots\times X_m$ to another Banach space X, then T^{*j} maps the product of Banach spaces $X_1\times\cdots\times X_{j-1}\times X^*\times X_{j+1}\times\cdots\times X_m$ to Banach space X_j^*. Moreover, the norms of T and T^{*j} are equal. For notational convenience, we may occasionally denote T by T^{*0} and K by K^{*0}.

Assumption (H2). Assume that for every $i \in \{1,\cdots,m\}$, there exist operators $\{A_t^{(i)}\}_{t>0}$ with kernels $a_t^{(i)}(x,y)$ that satisfy conditions (5) and (6) with constants s and ι, and there also exist kernels $K_t^{*j,(i)}(x,y_1,\cdots,y_m)$ so that

$$< T^{*j}(f_1,\cdots,A_t^{(i)} f_i,\cdots,f_m), g >= \int_{\mathbf{R}^n}\int_{(\mathbf{R}^n)^m} K_t^{*j,(i)}(x,y_1,\cdots,y_m) \prod_{k=1}^m f_k(y_k) g(x) \mathrm{d}\vec{y}\mathrm{d}x$$

for each $j \in \{0,1,\cdots,m\}$, and all $f_1,\cdots,f_m \in \mathcal{S}(\mathbf{R}^n)$ with $\cap_{k=1}^m \mathrm{supp}\, f_k \cap \mathrm{supp}\, g = \emptyset$. Also assume that there exist a function $\Phi \in \mathcal{C}(\mathbf{R})$ with $\mathrm{supp}\Phi \subset [-1,1]$ and a constant $\epsilon > 0$ so that for all $x, y_1, \cdots, y_m \in \mathbf{R}^n$, $t > 0$, each $j \in \{0, 1, \cdots, m\}$, and every $i \in \{1, \cdots, m\}$, we have

$$|K^{*j}(x, y_1, \cdots, y_m) - K_t^{*j,(i)}(x, y_1, \cdots, y_m)| \le \frac{C}{(|x - y_1| + \cdots + |x - y_m|)^{mn}} \sum_{k=1, k\neq i}^{m} \Phi(\frac{|y_i - y_k|}{t^{1/s}}) + \frac{Ct^{\epsilon/s}}{(|x - y_1| + \cdots + |x - y_m|)^{mn+\epsilon}},$$

whenever $2t^{1/s} \le |x - y_i|$.

Kernels K satisfying condition (2) and assumption (H2) with parameters m, C, s, ι, ϵ are also called generalized Calderón-Zygmund kernels, and their collection is denoted by $m - GCZK(C, s, \iota, \epsilon)$. We say that T is of class $m - GCZO(C, s, \iota, \epsilon)$ if the kernel of it is in $m - GCZK(C, s, \iota, \epsilon)$. In [11], the authors obtained the strong type boundedness on product of $L^p(\mathbf{R}^n)$ spaces and endpoint weak type estimates of operators $T \in m - GCZO(C, s, \iota, \epsilon)$. Furthermore, they pointed out that the m-th order Calderón commutator $\mathcal{C}_{m+1}$ is an $(m+1)$-linear operator associated with a kernel K in $(m+1) - GCZK(C, 1, 1, 1)$. In this way they first proved that for $p_1, \ldots, p_{m+1} \in [1, \infty]$ and $p \in (0, \infty)$ with $\frac{1}{p} = \sum_{j=1}^{m+1} \frac{1}{p_j}$,

$$\|\mathcal{C}_{m+1}(f, a_1, \ldots, a_m)\|_{L^{p,\infty}(\mathbf{R})} \le C\|f\|_{L^{p_{m+1}}(\mathbf{R})} \prod_{j=1}^{m} \|a_j\|_{L^{p_j}(\mathbf{R})},$$

and if $\min_{1\le j\le m+1} p_j > 1$, then

$$\|\mathcal{C}_{m+1}(f, a_1, \ldots, a_m)\|_{L^{p}(\mathbf{R})} \le C\|f\|_{L^{p_{m+1}}(\mathbf{R})} \prod_{j=1}^{m} \|a_j\|_{L^{p_j}(\mathbf{R})}.$$

Assumption (H3). Assume that there exist operators $\{B_t\}_{t>0}$ with kernels $b_t(x, y)$ that satisfy conditions (5) and (6) with constants s and ι, and there also exist kernels $K_t^{(0)}(x, y_1, \cdots, y_m)$ such that

$$K_t^{(0)}(x, y_1, \cdots, y_m) = \int_{\mathbf{R}^n} K(z, y_1, \cdots, y_m) b_t(x, z) \mathrm{d}z$$

for all $x, y_1, \cdots, y_m \in \mathbf{R}^n$. Also assume that there exist a function $\Phi \in \mathcal{C}(\mathbf{R})$ with supp$\Phi \subset [-1, 1]$ and a constant $\epsilon > 0$ so that for all $x, y_1, \cdots, y_m \in \mathbf{R}^n$ and $t > 0$, we have

$$|K(x, y_1, \cdots, y_m) - K_t^{(0)}(x, y_1, \cdots, y_m)| \le \frac{C}{(|x - y_1| + \cdots + |x - y_m|)^{mn}} \sum_{k=1}^{m} \Phi(\frac{|x - y_k|}{t^{1/s}}) + \frac{Ct^{\epsilon/s}}{(|x - y_1| + \cdots + |x - y_m|)^{mn+\epsilon}}$$

for some $C > 0$, whenever $2t^{1/s} \le \max_{1\le j\le m} |x - y_j|$. Moreover, assume that for all $x, y_1, \cdots, y_m \in \mathbf{R}^n$,

$$|K_t^{(0)}(x, y_1, \cdots, y_m)| \le \frac{C}{(|x - y_1| + \cdots + |x - y_m|)^{mn}},$$

whenever $2t^{1/s} \le \min_{1\le j\le m} |x - y_j|$, and for all $x, x', y_1, \cdots, y_m \in \mathbf{R}^n$,

$$|K_t^{(0)}(x, y_1, \cdots, y_m) - K_t^{(0)}(x', y_1, \cdots, y_m)| \le \frac{Ct^{\epsilon/s}}{(|x - y_1| + \cdots + |x - y_m|)^{mn+\epsilon}},$$

whenever $2|x - x'| \le t^{1/s}$ and $2t^{1/s} \le \min_{1\le j\le m} |x - y_j|$.

In [10], under assumption (H2) and (H3), some weighted estimates of operator T are obtained, and the authors also proved that the m-th order Calderón commutator $\mathcal{C}_{m+1}$ satisfies assumption (H3). Furthermore, the weighted estimates, including the multiple weights, of the maximal Calderón commutator were considered in [10] and [14]. Moreover, there are a large amount of work related to singular integral operators with non-smooth kernels. The readers may refer [12], [18] and [19], among many interesting works.

In this article, we are interested in the compactness for the commutators of multilinear singular integral operators with non-smooth kernels and $\mathrm{CMO}(\mathbf{R}^n)$ functions, where $\mathrm{CMO}(\mathbf{R}^n)$ denotes the closure of $C_c^\infty(\mathbf{R}^n)$ in the $\mathrm{BMO}(\mathbf{R}^n)$ topology. For the sake of convenience, we will write out the case of compactness in a bilinear setting. In particular, We will study the compactness for the commutator of T, where we assume that T is a bilinear singular integral operator associated with kernel K in the sense of (1) and satisfying (2), and

(i) T is bounded from

$$L^1(\mathbf{R}^n)\times L^1(\mathbf{R}^n)\to L^{1/2,\infty}(\mathbf{R}^n); \tag{7}$$

(ii) for $x,x',y_1,y_2\in\mathbf{R}^n$ with $8|x-x'|<\min_{1\le j\le 2}|x-y_j|$,

$$|K(x,y_1,y_2)-K(x',y_1,y_2)|\le\frac{D\tau^\gamma}{(|x-y_1|+|x-y_2|)^{2n+\gamma}}, \tag{8}$$

where D is a constant and τ is a number such that $2|x-x'|<\tau$ and $4\tau<\min_{1\le j\le 2}|x-y_j|$. It was pointed out in [20] that the condition (8) is weaker than, and indeed a consequence of, assumption (H3). For $b\in\mathrm{BMO}(\mathbf{R}^n)$, we consider commutators

$$T_b^1(f_1,f_2)=[b,T]_1(f_1,f_2)=bT(f_1,f_2)-T(bf_1,f_2),$$
$$T_b^2(f_1,f_2)=[b,T]_2(f_1,f_2)=bT(f_1,f_2)-T(f_1,bf_2).$$

For $\vec{b}=(b_1,b_2)\in\mathrm{BMO}(\mathbf{R}^n)\times\mathrm{BMO}(\mathbf{R}^n)$, we consider the iterated commutator

$$T_{\vec{b}}(f_1,f_2)=[b_2,[b_1,T]_1]_2(f_1,f_2),$$

and, in the sense of (1),

$$[b,T]_1(f_1,f_2)(x)=\int_{\mathbf{R}^n}\int_{\mathbf{R}^n}K(x,y_1,y_2)(b(x)-b(y_1))f_1(y_1)f_2(y_2)\mathrm{d}y_1\mathrm{d}y_2,$$
$$[b,T]_2(f_1,f_2)(x)=\int_{\mathbf{R}^n}\int_{\mathbf{R}^n}K(x,y_1,y_2)(b(x)-b(y_2))f_1(y_1)f_2(y_2)\mathrm{d}y_1\mathrm{d}y_2,$$
$$T_{\vec{b}}(f_1,f_2)(x)=\int_{\mathbf{R}^n}\int_{\mathbf{R}^n}K(x,y_1,y_2)(b_1(x)-b_1(y_1))(b_2(x)-b_2(y_2))f_1(y_1)f_2(y_2)\mathrm{d}y_1\mathrm{d}y_2.$$

Our aim is to obtain the compactness for the above commutators. Before stating our results, we briefly describe the background and our motivation. In [3], Calderón first proposed the concept of compactness in the multilinear setting and Bényi and Torres put forward an equivalent one in [2]. Bényi and Torres extended the result of compactness for linear singular integrals by Uchiyama [30] to the bilinear setting and obtained that $[b,T]_1$, $[b,T]_2$, $[b_2,[b_1,T]_1]_2$ are compact bilinear operators from $L^{p_1}(\mathbf{R}^n)\times L^{p_2}(\mathbf{R}^n)$ to $L^p(\mathbf{R}^n)$ when $b,b_1,b_2\in\mathrm{CMO}(\mathbf{R}^n)$, $1<p_1,p_2<\infty$ and $1/p_1+1/p_2=1/p\le 1$. Recently, Clop and Cruz [7] considered the compactness for the linear commutator on weighted spaces. For the bilinear case, Bényi et al. [1] extended the result of [2] to the weighted case, and they obtained that all $[b,T]_1$, $[b,T]_2$, $[b_2,[b_1,T]_1]_2$ are compact operators from $L^{p_1}(w_1)\times L^{p_2}(w_2)$ to $L^p(\nu_{\vec{w}})$ when $1<p_1,p_2<\infty$,

$1/p_1 + 1/p_2 = 1/p < 1$, $\vec{w} \in A_p(\mathbf{R}^n) \times A_p(\mathbf{R}^n)$ and $b, b_1, b_2 \in \mathrm{CMO}(\mathbf{R}^n)$. We note that in [1], T is a Calderón-Zygmund operator with smooth kernel. Hence, in this article, we will consider the same compactness for these commutators by assuming T is an operator associated with non-smooth kernel. Although we will adopt the concept of compactness proposed in [2](The reader can refer to [2] and [31] for more properties of compact and precompact) and some basic ideas used in [2,4,5,6,20,22,24,28], our proof meet some special difficulties so that some new ideas and estimates must be bought in. Particularly, some specific maximal functions will be involved.

We denote the closed ball of radius r centered at the origin in the normed space X as $B_{r,X} = \{x \in X : \|x\| \le r\}$.

Definition 1.1. A bilinear operator $T : X \times Y \mapsto Z$ is called compact if $T(B_{1,X} \times B_{1,Y})$ is precompact in Z.

Definition 1.2. A weight w belongs to the class $A_p(\mathbf{R}^n)$, $1 < p < \infty$, if

$$\sup_Q \left(\frac{1}{|Q|}\int_Q w(y)\mathrm{d}y\right)\left(\frac{1}{|Q|}\int_Q w(y)^{1-p'}\mathrm{d}y\right)^{p-1} < \infty.$$

A weight w belongs to the class $A_1(\mathbf{R}^n)$ if there is a constant C such that

$$\frac{1}{|Q|}\int_Q w(y)\mathrm{d}y \le C \inf_{x\in Q} w(x).$$

Definition 1.3. Let $\vec{p} = (p_1, p_2)$ and $1/p = 1/p_1 + 1/p_2$ with $1 \le p_1, p_2 < \infty$. Given $\vec{w} = (w_1, w_2)$, set $\nu_{\vec{w}} = \prod_{j=1}^2 w_j^{p/p_j}$. We say that $\vec{w}$ satisfies the $A_{\vec{p}}(\mathbf{R}^{2n})$ condition if

$$\sup_Q \left(\frac{1}{|Q|}\int_Q \nu_{\vec{w}}\right)^{1/p} \prod_{j=1}^2 \left(\frac{1}{|Q|}\int_Q w_j^{1-p_j'}\right)^{1/p_j'} < \infty.$$

Here, $\left(\frac{1}{|Q|}\int_Q w_j^{1-p_j'}\right)^{1/p_j'}$ is understood as $(\inf_Q w_j)^{-1}$,when $p_j = 1$.

The following two theorems are our main results:

Theorem 1.1. *Let T be a bilinear operator satisfying condition* (7) *and its kernel K satisfies* (2), (8). *Assume $b \in \mathrm{CMO}(\mathbf{R}^n)$, $p_1, p_2 \in (1, \infty)$, $p \in (1, \infty)$ such that $1/p = 1/p_1 + 1/p_2$ and $\vec{w} = (w_1, w_2) \in A_{\vec{p}}(\mathbf{R}^{2n})$ such that $\nu_{\vec{w}} \in A_p(\mathbf{R}^n)$. Then $[b, T]_1$, $[b, T]_2$ are compact from $L^{p_1}(w_1) \times L^{p_2}(w_2)$ to $L^p(\nu_{\vec{w}})$.*

In order to prove Theorem 1.1, we need the following result which has independent interest.

Theorem 1.2. *Let T be a bilinear operator satisfying condition* (7) *and its kernel K satisfies* (2), (8). *Assume $b \in \mathrm{BMO}(\mathbf{R}^n)$, $p_1, p_2 \in (1, \infty)$, $p \in (0, \infty)$ such that $1/p = 1/p_1 + 1/p_2$, $\vec{w} = (w_1, w_2) \in A_{\vec{p}}(\mathbf{R}^{2n})$. Then*

$$\|[b,T]_1(f_1,f_2)\|_{L^p(\nu_{\vec{w}})}, \|[b,T]_2(f_1,f_2)\|_{L^p(\nu_{\vec{w}})} \le C\|b\|_{\mathrm{BMO}(\mathbf{R}^n)}\|f_1\|_{L^{p_1}(w_1)}\|f_2\|_{L^{p_2}(w_2)}.$$

Remark 1.1. Theorem 1.1, Theorem 1.2 are also true for the iterated commutator $[b_2, [b_1, T]_1]_2$, and their proofs are similar to the proofs of Theorem 1.1 and Theorem 1.2. We leave the details to the interested readers.

We make some conventions. In this paper, we always denote a positive constant by C which is independent of the main parameters and its value may differ from line to line. We use the symbol $A \lesssim B$ to denote that there exists a positive constant C such that $A \le CB$. For a measurable set E, χ_E denotes its characteristic function. For a fixed p with $p \in [1, \infty)$, p' denotes the dual index of p. We also denote $\vec{f} = (f_1, \cdots, f_m)$ with scalar functions f_j $(j = 1, 2, ..., m)$. Given $\alpha > 0$ and a cube Q, $\ell(Q)$ denotes the side length of Q, and αQ denotes the cube which is the same center as Q and $\ell(\alpha Q) = \alpha \ell(Q)$. f_Q denotes the average of f over Q. Let M be the standard Hardy-Littlewood maximal operator. For $0 < \delta < \infty$, M_δ is the maximal operator defined by

$$M_\delta f(x) = M(|f|^\delta)^{1/\delta}(x) = \left(\sup_{Q \ni x} \frac{1}{|Q|} \int_Q |f(y)|^\delta \mathrm{d}y\right)^{1/\delta},$$

$M^{\#}$ is the sharp maximal operator defined by Fefferman and Stein [13],

$$M^{\#} f(x) = \sup_{Q \ni x} \inf_c \frac{1}{|Q|} \int_Q |f(y) - c| \mathrm{d}y \approx \sup_{Q \ni x} \frac{1}{|Q|} \int_Q |f(y) - f_Q| \mathrm{d}y,$$

and

$$M^{\#}_\delta f(x) = M^{\#}(|f|^\delta)^{1/\delta}(x).$$

It is known that, when $0 < p, \delta < \infty$, $w \in A_\infty(\mathbf{R}^n)$, there exists a constant $C > 0$ such that

$$\int_{\mathbf{R}^n} (M_\delta f(x))^p w(x) \mathrm{d}x \le C \int_{\mathbf{R}^n} (M^{\#}_\delta f(x))^p w(x) \mathrm{d}x \tag{9}$$

for any function f for which the left-hand side is finite.

§2 A multilinear maximal operator

We need some basic facts about Orlicz spaces, for more information about these spaces the readers may consult [29]. For

$$\Phi(t) = t(1 + \log^+ t)$$

and a cube Q in $\mathbf{R}^n$, we define

$$\|f\|_{L(\log L), Q} = \inf\{\lambda > 0 : \frac{1}{|Q|} \int_Q \Phi\left(\frac{|f(x)|}{\lambda}\right) \mathrm{d}x \le 1\}.$$

It is obvious that $\|f\|_{L(\log L), Q} > 1$ if and only if $\frac{1}{|Q|} \int_Q \Phi(|f(x)|) \mathrm{d}x > 1$. The generalized Hölder inequality in Orlicz space together with the John-Nirenberg inequality imply that

$$\frac{1}{|Q|} \int_Q |b(y) - b_Q| f(y) \mathrm{d}y \le C \|b\|_{\mathrm{BMO}(\mathbf{R}^n)} \|f\|_{L(\log L), Q}.$$

Define the maximal operator $\mathcal{M}_{L(\log L)}$ by

$$\mathcal{M}_{L(\log L)}(\vec{f})(x) = \sup_{Q \ni x} \prod_{j=1}^{2} \|f_j\|_{L(\log L), Q},$$

where the supremum is taken over all the cubes containing x. The following boundedness for $\mathcal{M}_{L(\log L)}(\vec{f})$ was proved in [22].

Lemma 2.1. *If $1 < p_1, p_2 < \infty$, $\frac{1}{p} = \sum_{j=1}^{2} \frac{1}{p_j}$, and $\vec{w} = (w_1, w_2) \in A_{\vec{p}}(\mathbf{R}^{2n})$, then $\mathcal{M}_{L(\log L)}(\vec{f})$ is bounded from $L^{p_1}(w_1) \times L^{p_2}(w_2)$ to $L^p(\nu_{\vec{w}})$.*

Lemma 2.1 is helpful in the proof of Theorem 1.2. Besides this maximal operator, we need several other maximal operators in the following.

In [22], a maximal function $\mathcal{M}(\vec{f})$ was introduced, and its definition is

$$\mathcal{M}(\vec{f})(x) = \sup_{Q \ni x} \prod_{j=1}^{2} \left(\frac{1}{|Q|} \int_Q |f_j(y_j)| \mathrm{d}y_j \right),$$

where the supremum is taken over all cubes Q containing x. The boundedness of $\mathcal{M}(\vec{f})$ on weighted spaces was considered in [22, Theorem 3.3].

Furthermore, Grafakos, Liu, and Yang [14] introduced some new multilinear maximal operators:

$$\mathcal{M}_{2,1}(\vec{f})(x) = \sup_{Q \ni x} \sum_{k=0}^{\infty} 2^{-kn} \left(\frac{1}{|Q|} \int_Q |f_1(y_1)| \mathrm{d}y_1 \right) \left(\frac{1}{|2^k Q|} \int_{2^k Q} |f_2(y_2)| \mathrm{d}y_2 \right),$$

$$\mathcal{M}_{2,2}(\vec{f})(x) = \sup_{Q \ni x} \sum_{k=0}^{\infty} 2^{-kn} \left(\frac{1}{|Q|} \int_Q |f_2(y_2)| \mathrm{d}y_2 \right) \left(\frac{1}{|2^k Q|} \int_{2^k Q} |f_1(y_1)| \mathrm{d}y_1 \right),$$

where $\vec{f} = (f_1, f_2)$ and each f_j ($j \in \{1,2\}$) is a locally integrable function. The following boundedness of $\mathcal{M}_{2,1}$ and $\mathcal{M}_{2,2}$ were proved in [14].

Lemma 2.2. *Let* $1 < p_1, p_2 < \infty$, $\frac{1}{p} = \sum_{j=1}^{2} \frac{1}{p_j}$, *and* $\vec{w} = (w_1, w_2) \in A_{\vec{p}}(\mathbf{R}^{2n})$. *Then* $\mathcal{M}_{2,1}$ *and* $\mathcal{M}_{2,2}$ *are bounded from* $L^{p_1}(w_1) \times L^{p_2}(w_2)$ *to* $L^p(\nu_{\vec{w}})$.

In addition, Hu [17] introduced another kind of bilinear maximal operators $\mathcal{M}_\beta^1$ and $\mathcal{M}_\beta^2$ which was defined by

$$\mathcal{M}_\beta^1(\vec{f})(x) = \sup_{Q \ni x} \frac{1}{|Q|} \int_Q |f_1(y_1)| \mathrm{d}y_1 \sum_{k=1}^{\infty} 2^{-kn} 2^{k\beta} \frac{1}{|2^k Q|} \int_{2^k Q} |f_2(y_2)| \mathrm{d}y_2,$$

$$\mathcal{M}_\beta^2(\vec{f})(x) = \sup_{Q \ni x} \frac{1}{|Q|} \int_Q |f_2(y_2)| \mathrm{d}y_2 \sum_{k=1}^{\infty} 2^{-kn} 2^{k\beta} \frac{1}{|2^k Q|} \int_{2^k Q} |f_1(y_1)| \mathrm{d}y_1,$$

where $\beta \in \mathbf{R}$ and the supremum is taken over all cubes Q containing x. As it is well known, a weight $w \in A_\infty(\mathbf{R}^n)$ implies that there exists a $\theta \in (0,1)$ such that for all cubes Q and any set $E \subset Q$,

$$\frac{w(E)}{w(Q)} \le C \left(\frac{|E|}{|Q|} \right)^\theta. \tag{10}$$

For a fixed $\theta \in (0,1)$, set

$$R_\theta = \{ w \in A_\infty(\mathbf{R}^n) : w \text{ satisfies } (10) \}.$$

In [17], the following boundedness of $\mathcal{M}_\beta^1$ and $\mathcal{M}_\beta^2$ were proved.

Lemma 2.3. *Let* $1 < p_1, p_2 < \infty$, $\frac{1}{p} = \sum_{j=1}^{2} \frac{1}{p_j}$, $\vec{w} = (w_1, w_2) \in A_{\vec{p}}(\mathbf{R}^{2n})$ *and* $\nu_{\vec{w}} \in R_\theta$ *for some* θ *such that* $\beta < n\theta \min\{1/p_1, 1/p_2\}$. *Then* $\mathcal{M}_\beta^1$ *and* $\mathcal{M}_\beta^2$ *are bounded from* $L^{p_1}(w_1) \times L^{p_2}(w_2)$ *to* $L^p(\nu_{\vec{w}})$.

§3 Proof of Theorem 1.2

The proof of Theorem 1.2 will depend on some pointwise estimates using sharp maximal functions. The pointwise estimates are the following:

Lemma 3.1. *Let T be a bilinear operator satisfying condition* (7) *and its kernel K satisfies* (2), (8). *If $0<\delta<\frac{1}{2}$, then for all $\vec{f}$ in any product of $L^{p_j}(\mathbf{R}^n)$ spaces with $1<p_j<\infty$,*

$$M_\delta^{\#}(T(\vec{f}))(x)\le C\mathcal{M}(\vec{f})(x)+C\sum_{i=1}^{2}\mathcal{M}_{2,i}(\vec{f})(x).$$

The proof of this Lemma uses some ideas of [22, Theorem 3.2] and the following Lemma 3.2. Its proof is not difficult, so we omit.

Lemma 3.2. *Let T be a bilinear operator satisfying condition* (7) *and its kernel K satisfies* (2), (8). *If T_b^1, T_b^2 be commutators with $b\in\mathrm{BMO}(\mathbf{R}^n)$. For $0<\delta<\epsilon$ with $0<\delta<1/2$, let $r>1$ and $0<\beta<n$. Then, there exists a constant $C>0$, depending on δ and ϵ, such that*

$$\sum_{i=1}^{2}M_\delta^{\#}(T_b^i(\vec{f}))(x)\le C\|b\|_{\mathrm{BMO}(\mathbf{R}^n)}\Big(\mathcal{M}_{L(\log L)}(\vec{f})(x)+M_\epsilon(T(\vec{f}))(x)+\sum_{i=1}^{2}\{\mathcal{M}_\beta^i(f_1^r,f_2^r)(x)\}^{1/r}\Big)$$

for all $\vec{f}=(f_1,f_2)$ of bounded functions with compact support.

Proof. We only write out the estimate of $M_\delta^{\#}(T_b^1(\vec{f}))(x)$, since the other can be obtained similarly. In our proof we will use some ideas of [28]. For a fixed $x\in\mathbf{R}^n$, a cube Q centered at x and constants c,λ, because $0<\delta<1/2$,

$$\begin{aligned}
&\inf_c\left(\frac{1}{|Q|}\int_Q\big||T_b^1(\vec{f})(z)|^\delta-|c|^\delta\big|\mathrm{d}z\right)^{1/\delta}\\
&\le\inf_c\left(\frac{1}{|Q|}\int_Q|T_b^1(\vec{f})(z)-c|^\delta\mathrm{d}z\right)^{1/\delta}\\
&\le\inf_c\left(\frac{C}{|Q|}\int_Q|(b(z)-\lambda)T(f_1,f_2)(z)-T((b-\lambda)f_1,f_2)(z)-c|^\delta\mathrm{d}z\right)^{1/\delta}\\
&\le\left(\frac{C}{|Q|}\int_Q|(b(z)-\lambda)T(f_1,f_2)(z)|^\delta\mathrm{d}z\right)^{1/\delta}+\inf_c\left(\frac{C}{|Q|}\int_Q|T((b-\lambda)f_1,f_2)(z)-c|^\delta\mathrm{d}z\right)^{1/\delta}\\
&=I_1+I_2.
\end{aligned}$$

Let $Q^*=8^nQ$, $\lambda=b_{Q^*}$. The estimate of the first part I_1 is the same as [28, Theorem 3.1]. Therefore, we omit the proof, and from [28, Theorem 3.1], we obtain that

$$I_1\le C\|b\|_{\mathrm{BMO}(\mathbf{R}^n)}M_\epsilon(T(f_1,f_2))(x).$$

Now, we turn our attention to I_2. We decompose f_1,f_2 as

$$f_1=f_1^1+f_1^2=f_1(x)\chi_{Q^*}+f_1(x)\chi_{\mathbf{R}^n\setminus Q^*},$$
$$f_2=f_2^1+f_2^2=f_2(x)\chi_{Q^*}+f_2(x)\chi_{\mathbf{R}^n\setminus Q^*}.$$

Let $c = c_1 + c_2 + c_3$ and

$$c_1 = T((b-\lambda)f_1^1, f_2^2)(x),$$
$$c_2 = T((b-\lambda)f_1^2, f_2^1)(x),$$
$$c_3 = T((b-\lambda)f_1^2, f_2^2)(x).$$

Therefore,

$$\begin{aligned} I_2 \le& \left(\frac{C}{|Q|}\int_Q |T((b-\lambda)f_1^1, f_2^1)(z)|^\delta \mathrm{d}z\right)^{1/\delta} \\ &+ \left(\frac{C}{|Q|}\int_Q |T((b-\lambda)f_1^1, f_2^2)(z) - c_1|^\delta \mathrm{d}z\right)^{1/\delta} \\ &+ \left(\frac{C}{|Q|}\int_Q |T((b-\lambda)f_1^2, f_2^1)(z) - c_2|^\delta \mathrm{d}z\right)^{1/\delta} \\ &+ \left(\frac{C}{|Q|}\int_Q |T((b-\lambda)f_1^2, f_2^2)(z) - c_3|^\delta \mathrm{d}z\right)^{1/\delta} \\ =& I_2^1 + I_2^2 + I_2^3 + I_2^4. \end{aligned}$$

We choose $1 < q < 1/(2\delta)$. By Hölder's inequality and the fact that T satisfies condition (7), we have

$$\begin{aligned} I_2^1 \le& \left(\frac{C}{|Q|}\int_Q |T((b-\lambda)f_1^1, f_2^1)(z)|^{q\delta} \mathrm{d}z\right)^{1/q\delta} \\ \le& C\|T((b-\lambda)f_1^1, f_2^1)\|_{L^{1/2,\infty}(Q,\frac{dz}{|Q|})} \\ \le& C\left(\frac{1}{|Q|}\int_Q |(b(z)-\lambda)f_1^1(z)|\mathrm{d}z\right)\left(\frac{1}{|Q|}\int_Q |f_2^1(z)|\mathrm{d}z\right) \\ \le& C\|b\|_{\mathrm{BMO}(\mathbf{R}^n)}\|f_1\|_{L(\log L),Q}\|f_2\|_{L(\log L),Q} \\ \le& C\|b\|_{\mathrm{BMO}(\mathbf{R}^n)}\mathcal{M}_{L(\log L)}(f_1, f_2)(x). \end{aligned}$$

For I_2^2, by generalized Jensen's inequality, we get

$$\begin{aligned} &|T((b-\lambda)f_1^1, f_2^2)(z) - T((b-\lambda)f_1^1, f_2^2)(x)| \\ \le& \int_{\mathbf{R}^{2n}} |K(z, y_1, y_2)||(b-\lambda)f_1^1(y_1)||f_2^2(y_2)|\mathrm{d}y_1\mathrm{d}y_2 \\ &+ \int_{\mathbf{R}^{2n}} |K(x, y_1, y_2)||(b-\lambda)f_1^1(y_1)||f_2^2(y_2)|\mathrm{d}y_1\mathrm{d}y_2 \\ \lesssim& \int_{\mathbf{R}^{2n}} \frac{1}{(|z-y_1|+|z-y_2|)^{2n}}|(b-\lambda)f_1^1(y_1)||f_2^2(y_2)|\mathrm{d}y_1\mathrm{d}y_2 \\ &+ \int_{\mathbf{R}^{2n}} \frac{1}{(|x-y_1|+|x-y_2|)^{2n}}|(b-\lambda)f_1^1(y_1)||f_2^2(y_2)|\mathrm{d}y_1\mathrm{d}y_2 \end{aligned}$$

$$
\begin{aligned}
&\lesssim \int_{Q^*}|(b-\lambda)f_1^1(y_1)|\mathrm{d}y_1\int_{\mathbf{R}^n\setminus Q^*}\frac{1}{|z-y_2|^{2n}}|f_2^2(y_2)|\mathrm{d}y_2\\
&\quad+\int_{Q^*}|(b-\lambda)f_1^1(y_1)|\mathrm{d}y_1\int_{\mathbf{R}^n\setminus Q^*}\frac{1}{|x-y_2|^{2n}}|f_2^2(y_2)|\mathrm{d}y_2\\
&\lesssim \int_{Q^*}|(b-\lambda)f_1^1(y_1)|\mathrm{d}y_1\sum_{k=1}^{\infty}\int_{2^kQ^*\setminus 2^{k-1}Q^*}\frac{1}{|z-y_2|^{2n}}|f_2^2(y_2)|\mathrm{d}y_2\\
&\quad+\int_{Q^*}|(b-\lambda)f_1^1(y_1)|\mathrm{d}y_1\sum_{k=1}^{\infty}\int_{2^kQ^*\setminus 2^{k-1}Q^*}\frac{1}{|x-y_2|^{2n}}|f_2^2(y_2)|\mathrm{d}y_2\\
&\lesssim \|b\|_{\mathrm{BMO}(\mathbf{R}^n)}\|f_1\|_{L(\log L),Q^*}\sum_{k=1}^{\infty}2^{-kn}\left(\frac{1}{|2^kQ^*|}\int_{2^kQ^*}|f_2^2(y_2)|\mathrm{d}y_2\right)\\
&\lesssim \|b\|_{\mathrm{BMO}(\mathbf{R}^n)}\|f_1\|_{L(\log L),Q^*}\sum_{k=1}^{\infty}2^{-kn}\|f_2\|_{L(\log L),2^kQ^*}\\
&\lesssim \|b\|_{\mathrm{BMO}(\mathbf{R}^n)}\left(\frac{1}{|Q^*|}\int_{Q^*}|f_1(y_1)|^r\mathrm{d}y_1\right)^{\frac{1}{r}}\sum_{k=1}^{\infty}2^{-kn}\left(\frac{1}{|2^kQ^*|}\int_{2^kQ^*}|f_2(y_2)|^r\mathrm{d}y_2\right)^{\frac{1}{r}}\\
&\lesssim \|b\|_{\mathrm{BMO}(\mathbf{R}^n)}\{\mathcal{M}_\beta^1(f_1^r,f_2^r)(x)\}^{\frac{1}{r}},
\end{aligned}
$$

where $r>1$. Based on the above estimates, we deduce that

$$
I_2^2\lesssim \|b\|_{\mathrm{BMO}(\mathbf{R}^n)}\{\mathcal{M}_\beta^1(f_1^r,f_2^r)(x)\}^{\frac{1}{r}}.
$$

For I_2^3, we obtain that

$$
\begin{aligned}
&|T((b-\lambda)f_1^2,f_2^1)(z)-T((b-\lambda)f_1^2,f_2^1)(x)|\\
&\lesssim \int_{\mathbf{R}^{2n}}\frac{1}{(|z-y_1|+|z-y_2|)^{2n}}|(b-\lambda)f_1^2(y_1)||f_2^1(y_2)|\mathrm{d}y_1\mathrm{d}y_2\\
&\quad+\int_{\mathbf{R}^{2n}}\frac{1}{(|x-y_1|+|x-y_2|)^{2n}}|(b-\lambda)f_1^2(y_1)||f_2^1(y_2)|\mathrm{d}y_1\mathrm{d}y_2\\
&\lesssim \int_{Q^*}|f_2^1(y_2)|\mathrm{d}y_2\sum_{k=1}^{\infty}\int_{2^kQ^*\setminus 2^{k-1}Q^*}\frac{|(b-\lambda)f_1^2(y_1)|}{|z-y_1|^{2n}}\mathrm{d}y_1\\
&\quad+\int_{Q^*}|f_2^1(y_2)|\mathrm{d}y_2\sum_{k=1}^{\infty}\int_{2^kQ^*\setminus 2^{k-1}Q^*}\frac{|(b-\lambda)f_1^2(y_1)|}{|x-y_1|^{2n}}\mathrm{d}y_1\\
&\lesssim \frac{1}{|Q^*|}\int_{Q^*}|f_2^1(y_2)|\mathrm{d}y_2\sum_{k=1}^{\infty}2^{-kn}\frac{1}{|2^kQ^*|}\int_{2^kQ^*}|(b-\lambda)f_1^2(y_1)|\mathrm{d}y_1\\
&\lesssim \frac{1}{|Q^*|}\int_{Q^*}|f_2^1(y_2)|\mathrm{d}y_2\sum_{k=1}^{\infty}2^{-kn}\frac{1}{|2^kQ^*|}\int_{2^kQ^*}|(b-b_{2^kQ^*})f_1^2(y_1)|\mathrm{d}y_1\\
&\quad+\frac{1}{|Q^*|}\int_{Q^*}|f_2^1(y_2)|\mathrm{d}y_2\sum_{k=1}^{\infty}2^{-kn}\frac{1}{|2^kQ^*|}\int_{2^kQ^*}|(b_{2^kQ^*}-b_{Q^*})f_1^2(y_1)|\mathrm{d}y_1\\
&\lesssim \|b\|_{\mathrm{BMO}(\mathbf{R}^n)}\|f_2\|_{L(\log L),Q^*}\sum_{k=1}^{\infty}2^{-kn}\|f_1\|_{L(\log L),2^kQ^*}
\end{aligned}
$$

$$
\begin{aligned}
&+\|b\|_{\mathrm{BMO}(\mathbf{R}^n)}\|f_2\|_{L(\log L),Q^*}\sum_{k=1}^{\infty}2^{-kn}k\|f_1\|_{L(\log L),2^kQ^*}\\
&\lesssim \|b\|_{\mathrm{BMO}(\mathbf{R}^n)}\left(\frac{1}{|Q^*|}\int_{Q^*}|f_2(y_2)|^r\mathrm{d}y_2\right)^{\frac{1}{r}}\sum_{k=1}^{\infty}2^{-kn}\left(\frac{1}{|2^kQ^*|}\int_{2^kQ^*}|f_1(y_1)|^r\mathrm{d}y_1\right)^{\frac{1}{r}}\\
&\quad+\|b\|_{\mathrm{BMO}(\mathbf{R}^n)}\left(\frac{1}{|Q^*|}\int_{Q^*}|f_2(y_2)|^r\mathrm{d}y_2\right)^{\frac{1}{r}}\sum_{k=1}^{\infty}2^{-kn}k\left(\frac{1}{|2^kQ^*|}\int_{2^kQ^*}|f_1(y_1)|^r\mathrm{d}y_1\right)^{\frac{1}{r}}\\
&\lesssim \|b\|_{\mathrm{BMO}(\mathbf{R}^n)}\{\mathcal{M}_\beta^2(f_1^r,f_2^r)(x)\}^{\frac{1}{r}},
\end{aligned}
$$

where $r>1$. This implies that

$$I_2^3\lesssim \|b\|_{\mathrm{BMO}(\mathbf{R}^n)}\{\mathcal{M}_\beta^2(f_1^r,f_2^r)(x)\}^{\frac{1}{r}}.$$

Finally, we use condition (8) to estimate I_2^4. Note that for any $x,z\in Q$ and $y_1,y_2\in\mathbf{R}^n\setminus Q^*$,

$$|x-z|\le n\ell(Q)\le \frac{1}{8}\min\{|z-y_1|,|z-y_2|\}.$$

So, it is easy to verify that,

$$
\begin{aligned}
&|T((b-\lambda)f_1^2,f_2^2)(z)-T((b-\lambda)f_1^2,f_2^2)(x)|\\
&\le \int_{\mathbf{R}^{2n}}|K(z,y_1,y_2)-K(x,y_1,y_2)||(b-\lambda)f_1^2(y_1)||f_2^2(y_2)|\mathrm{d}y_1\mathrm{d}y_2\\
&\lesssim \int_{\mathbf{R}^n\setminus Q^*}\int_{\mathbf{R}^n\setminus Q^*}\frac{\ell(Q)^\gamma}{(|z-y_1|+|z-y_2|)^{2n+\gamma}}|(b-\lambda)f_1^2(y_1)||f_2^2(y_2)|\mathrm{d}y_1\mathrm{d}y_2\\
&\lesssim \sum_{k=1}^{\infty}\int_{2^k\ell(Q)<|z-y_1|+|z-y_2|<2^{k+1}\ell(Q)}\frac{\ell(Q)^\gamma\ |(b-\lambda)f_1^2(y_1)|}{(|z-y_1|+|z-y_2|)^{2n+\gamma}}|f_2^2(y_2)|\mathrm{d}y_1\mathrm{d}y_2\\
&\lesssim \sum_{k=1}^{\infty}\frac{\ell(Q)^\gamma}{(2^k\ell(Q))^{2n+\gamma}}\left(\int_{2^{k+2}Q^*}|(b-\lambda)f_1^2(y_1)|\mathrm{d}y_1\right)\left(\int_{2^{k+2}Q^*}|f_2^2(y_2)|\mathrm{d}y_2\right)\\
&\lesssim \sum_{k=1}^{\infty}2^{-k\gamma}\left(\frac{1}{|2^{k+2}Q^*|}\int_{2^{k+2}Q^*}|(b-\lambda)f_1^2(y_1)|\mathrm{d}y_1\right)\left(\frac{1}{|2^{k+2}Q^*|}\int_{2^{k+2}Q^*}|f_2^2(y_2)|\mathrm{d}y_2\right)\\
&\lesssim \|b\|_{\mathrm{BMO}(\mathbf{R}^n)}\mathcal{M}_{L(\log L)}(\vec{f})(x).
\end{aligned}
$$

Therefore,

$$I_2^4\lesssim \|b\|_{\mathrm{BMO}(\mathbf{R}^n)}\mathcal{M}_{L(\log L)}(\vec{f})(x).$$

Combining the estimates for I_2^1, I_2^2, I_2^3 and I_2^4, lead to that

$$I_2\lesssim \|b\|_{\mathrm{BMO}(\mathbf{R}^n)}(\mathcal{M}_{L(\log L)}(\vec{f})(x)+\sum_{i=1}^{2}\{\mathcal{M}_\beta^i(f_1^r,f_2^r)(x)\}^{1/r}).$$

The proof is completed. □

Now, we are ready to prove Theorem 1.2.

Proof. For the sake of brevity, we only write out the proof of the boundedness of T_b^1, and the other can be got in the same method. By [22, Lemma 6.1], we know that for every $\vec{w}\in A_{\vec{p}}(\mathbf{R}^{2n})$, there exists a finite constant $1<r_0<\min\{p_1,p_2\}$ such that $\vec{w}\in A_{\vec{p}/r_0}(\mathbf{R}^{2n})$. From Lemma 2.3, for $\vec{w}\in A_{\vec{p}/r_0}(\mathbf{R}^{2n})$, there exists a $\beta_0>0$ satisfies that $\sum_{i=1}^{2}\mathcal{M}_{\beta_0}^i(f_1^{r_0},f_2^{r_0})(x)$ is bounded from $L^{p_1/r_0}(w_1) \times L^{p_2/r_0}(w_2)$ to $L^{p/r_0}(\nu_{\vec{w}})$. Hence,

$$\begin{aligned}\sum_{i=1}^{2} \|\{\mathcal{M}^i_{\beta_0}(f_1^{r_0}, f_2^{r_0})(x)\}^{\frac{1}{r_0}}\|_{L^p(\nu_{\vec{w}})} &= \sum_{i=1}^{2} \|\{\mathcal{M}^i_{\beta_0}(f_1^{r_0}, f_2^{r_0})(x)\}\|^{1/r_0}_{L^{p/r_0}(\nu_{\vec{w}})} \\ &\le C\|f_1^{r_0}\|^{1/r_0}_{L^{p_1/r_0}(w_1)}\|f_2^{r_0}\|^{1/r_0}_{L^{p_2/r_0}(w_2)} \\ &= C\|f_1\|_{L^{p_1}(w_1)}\|f_2\|_{L^{p_2}(w_2)}.\end{aligned}$$

Because $\nu_{\vec{w}} \in A_{2p}(\mathbf{R}^n) \subset A_\infty(\mathbf{R}^n)$, using inequality (9) and Lemma 3.2, we deduce that

$$\begin{aligned}&\|T_b^1(\vec{f})\|_{L^p(\nu_{\vec{w}})} \le \|M_\delta(T_b^1(\vec{f}))\|_{L^p(\nu_{\vec{w}})} \\ &\le C\|M_\delta^{\#}(T_b^1(\vec{f}))\|_{L^p(\nu_{\vec{w}})} \\ &\le C\|b\|_{\mathrm{BMO}(\mathbf{R}^n)}\|\mathcal{M}_{L(\log L)}(\vec{f})(x) + M_\epsilon(T(\vec{f}))(x) + \sum_{i=1}^{2}\{\mathcal{M}^i_{\beta_0}(f_1^{r_0}, f_2^{r_0})(x)\}^{\frac{1}{r_0}}\|_{L^p(\nu_{\vec{w}})} \\ &\le C\|b\|_{\mathrm{BMO}(\mathbf{R}^n)}\Big(\|\mathcal{M}_{L(\log L)}(\vec{f})(x)\|_{L^p(\nu_{\vec{w}})} + \|M_\epsilon(T(\vec{f}))(x)\|_{L^p(\nu_{\vec{w}})} \\ &\qquad + \|\sum_{i=1}^{2}\{\mathcal{M}^i_{\beta_0}(f_1^{r_0}, f_2^{r_0})(x)\}^{1/r_0}\|_{L^p(\nu_{\vec{w}})}\Big).\end{aligned}$$

By Lemma 3.1, we have

$$\begin{aligned}\|M_\epsilon(T(\vec{f}))\|_{L^p(\nu_{\vec{w}})} &\le \|M_\epsilon^{\#}(T(\vec{f}))\|_{L^p(\nu_{\vec{w}})} \\ &\le C\|\mathcal{M}(\vec{f})\|_{L^p(\nu_{\vec{w}})} + C\|\sum_{i=1}^{2}\mathcal{M}_{2,i}(\vec{f})(x)\|_{L^p(\nu_{\vec{w}})} \\ &\le C\|\mathcal{M}_{L(\log L)}(\vec{f})\|_{L^p(\nu_{\vec{w}})} + C\|\sum_{i=1}^{2}\mathcal{M}_{2,i}(\vec{f})(x)\|_{L^p(\nu_{\vec{w}})}.\end{aligned}$$

Now the desired result follows from Lemma 2.1, Lemma 2.2 and Lemma 2.3 directly.

In the above proof, we note that when using the inequality (9) we need to explain that $\|M_\epsilon(T(\vec{f}))\|_{L^p(\nu_{\vec{w}})}$ and $\|M_\delta(T_b^1(\vec{f}))\|_{L^p(\nu_{\vec{w}})}$ are finite. A detailed proof was given in page 33 of [22], and the proof can also be applied here owing to the boundedness of T which was given in [20, Theorem 2]. □

§4 Proof of Theorem 1.1

The idea of using truncated operators to prove compactness results in the linear setting can trace back to [21], and this method was adopted in [7]. Recently, Bényi et al. (see [1]) introduced a new smooth truncation to simplify the computations. We will use this technique to prove Theorem 1.1.

Let $\varphi = \varphi(x, y_1, y_2)$ be a non-negative function in $C_c^\infty(\mathbf{R}^{3n})$, and it satisfy

$$\operatorname{supp}\varphi \subset \{(x, y_1, y_2) : \max(|x|, |y_1|, |y_2|) < 1\},$$

$$\int_{\mathbf{R}^{3n}} \varphi(u)du = 1.$$

For $\delta > 0$, let $\chi^\delta = \chi^\delta(x, y_1, y_2)$ be the characteristic function of the set $\{(x, y_1, y_2) : \max(|x -$

$y_1|, |x-y_2|) \geq 3\delta/2\}$, and let
$$\psi^\delta = \varphi_\delta * \chi^\delta,$$
where
$$\varphi_\delta(x, y_1, y_2) = (\delta/4)^{-3n}\varphi(4x/\delta, 4y_1/\delta, 4y_2/\delta).$$
By an easy calculation, we get that $\psi^\delta \in C^\infty(\mathbf{R}^{3n})$, $\|\psi^\delta\|_{L^\infty(\mathbf{R}^{3n})} \leq 1$,
$$\operatorname{supp}\psi^\delta \subset \{(x, y_1, y_2) : \max(|x-y_1|, |x-y_2|) \geq \delta\},$$
and $\psi^\delta(x, y_1, y_2) = 1$ if $\max(|x-y_1|, |x-y_2|) \geq 2\delta$.

We define the truncated kernel
$$K^\delta(x, y_1, y_2) = \psi^\delta(x, y_1, y_2)K(x, y_1, y_2),$$
where $K(x, y_1, y_2)$ is the kernel of the bilinear singular integral operator T considered in Theorem 1.1. It's easy to verify that K^δ also satisfies condition (2) and (8). Let T^δ be the bilinear operator that associated with kernel K^δ in the sense of (1). The following Lemma was proved in [1]:

Lemma 4.1. *For all $x \in \mathbf{R}^n$, $b, b_1, b_2 \in C_c^\infty(\mathbf{R}^n)$, if $\vec{w} = (w_1, w_2) \in A_{\vec{p}}(\mathbf{R}^{2n})$, then*
$$\lim_{\delta\to 0} \|[b, T^\delta]_1 - [b, T]_1\|_{L^{p_1}(w_1)\times L^{p_2}(w_2)\to L^p(\nu_{\vec{w}})} = 0,$$
$$\lim_{\delta\to 0} \|[b, T^\delta]_2 - [b, T]_2\|_{L^{p_1}(w_1)\times L^{p_2}(w_2)\to L^p(\nu_{\vec{w}})} = 0,$$
$$\lim_{\delta\to 0} \|[b_2, [b_1, T^\delta]_1]_2 - [b_2, [b_1, T]_1]_2\|_{L^{p_1}(w_1)\times L^{p_2}(w_2)\to L^p(\nu_{\vec{w}})} = 0.$$

By condition (2), Lemma 4.1 can be proved by the argument used in [1].

Lemma 4.2. *Suppose that T is as in Theorem 1.1. Then, for all $\zeta > 0$, there exists a positive constant C such that for all $\vec{f}$ in the product of $L^{p_j}(\mathbf{R}^n)$ with $1 < p_j < \infty$ and all $x \in \mathbf{R}^n$,*
$$T^*(\vec{f})(x) \leq C\big(M_\zeta(T(\vec{f}))(x) + \sum_{i=1}^{2}\mathcal{M}_{2,i}(\vec{f})(x) + \mathcal{M}(\vec{f})(x)\big),$$
where $T^*(\vec{f})$ is the maximal truncated bilinear singular integral operator defined as
$$T^*(f_1, f_2)(x) = \sup_{\eta>0}\left|\int\int_{\max(|x-y_1|,|x-y_2|)>\eta} K(x, y_1, y_2)f_1(y_1)f_2(y_2)\mathrm{d}y_1\mathrm{d}y_2\right|.$$

The proof of the Lemma 4.2 is similar to the proof of [15, Theorem 1], so we leave it to the interested reader.

Lemma 4.3. *Let $1 < p < \infty$, $w \in A_p(\mathbf{R}^n)$ and $\mathcal{H} \subset L^p(w)$. If*

(i) *$\mathcal{H}$ is bounded in $L^p(w)$;*

(ii) $\lim_{A\to\infty} \int_{|x|>A} |f(x)|^p w(x)\mathrm{d}x = 0$ *uniformly for $f \in \mathcal{H}$;*

(iii) $\lim_{t\to 0} \|f(\cdot + t) - f(\cdot)\|_{L^p(w)} = 0$ *uniformly for $f \in \mathcal{H}$.*

Then $\mathcal{H}$ is precompact in $L^p(w)$.

This Lemma was given in [7].

Now, we are ready to prove Theorem 1.1.

Proof. We will work with the commutator $[b,T]_1$ first, and the proof of the commutator $[b,T]_2$ can be get similarly. From Lemma 4.1, we only need to prove the compactness for $[b,T^\delta]_1$ for any fixed $\delta \le 1/8$. By Theorem 1.2, it suffices to show the result for $b \in C_c^\infty(\mathbf{R}^n)$. Suppose f_1, f_2 belong to

$$B_1(L^{p_1}(w_1)) \times B_1(L^{p_2}(w_2)) = \{(f_1,f_2) : \|f_1\|_{L^{p_1}(w_1)}, \|f_2\|_{L^{p_2}(w_2)} \le 1\},$$

where $\vec{w} \in A_{\vec{p}}(\mathbf{R}^n)$. We need to prove that the following three conditions hold:

(i) $[b,T^\delta]_1(B_1(L^{p_1}(w_1)) \times B_1(L^{p_2}(w_2)))$ is bounded in $L^p(\nu_{\vec{w}})$;

(ii) $\lim\limits_{A\to\infty} \int_{|x|>A} |[b,T^\delta]_1(f_1,f_2)(x)|^p \nu_{\vec{w}}(x)\mathrm{d}x = 0$;

(iii) Given $0 < \xi < 1/8$, there exists a sufficiently small $t_0 (t_0 = t_0(\xi))$ such that for all $0 < |t| < t_0$, we have

$$\|[b,T^\delta]_1(f_1,f_2)(\cdot) - [b,T^\delta]_1(f_1,f_2)(\cdot + t)\|_{L^p(\nu_{\vec{w}})} \le C\xi. \tag{11}$$

It is easy to find that the condition (i) holds because of the boundedness of $[b,T]_1$ in Theorem 1.2 and Lemma 4.1. Now, we prove the condition (ii) using some ideas in [17]. Let $R > 0$ be large enough such that $\operatorname{supp} b \subset B(0,R)$ and let $A \ge \max(2R,1)$, l be a nonnegative integer. For any $|x| > A$, denote

$$V_R^0(x) = \int_{|y_2|\le|x|}\int_{|y_1|\le R} |K^\delta(x,y_1,y_2)| \prod_{j=1}^2 |f_j(y_j)|\mathrm{d}y_1\mathrm{d}y_2,$$

and

$$V_R^l(x) = \int_{2^{l-1}|x|\le|y_2|\le 2^l|x|}\int_{|y_1|\le R} |K^\delta(x,y_1,y_2)| \prod_{j=1}^2 |f_j(y_j)|\mathrm{d}y_1\mathrm{d}y_2,$$

where $l > 0$. From condition (2), we deduce that

$$\begin{aligned}
V_R^l(x) &\le C\int_{2^{l-1}|x|\le|y_2|\le 2^l|x|}\int_{|y_1|\le R} \frac{1}{(|x-y_1|+|x-y_2|)^{2n}}|f_1(y_1)||f_2(y_2)|\mathrm{d}y_1\mathrm{d}y_2\\
&\le C\int_{2^{l-1}|x|\le|y_2|\le 2^l|x|}\int_{|y_1|\le R} \frac{|f_1(y_1)||f_2(y_2)|}{(|x|+|x-y_2|)^{2n}}\mathrm{d}y_1\mathrm{d}y_2\\
&\le C\frac{1}{(2^{l-1}|x|)^{2n}}\int_{2^{l-1}|x|\le|y_2|\le 2^l|x|}\int_{|y_1|\le R} |f_1(y_1)||f_2(y_2)|\mathrm{d}y_1\mathrm{d}y_2\\
&\le C\frac{1}{(2^{l-1}|x|)^{2n}}\left(\int_{B(0,R)} w_1^{-\frac{1}{p_1-1}}(y_1)dy_1\right)^{1-\frac{1}{p_1}}\left(\int_{B(0,2^l|x|)} w_2^{-\frac{1}{p_2-1}}(y_2)dy_2\right)^{1-\frac{1}{p_2}}.
\end{aligned}$$

The same estimate can be got for $V_R^0(x)$. Note that $w_1^{-\frac{1}{p_1-1}} \in A_\infty(\mathbf{R}^n)$, so there exists a constant $\theta_1 \in (0,1)$ such that

$$\int_{B(0,R)} w_1^{-\frac{1}{p_1-1}}(y_1)\mathrm{d}y_1 \le C(2^{-(j+l)}RA^{-1})^{n\theta_1}\int_{B(0,2^{l+j}A)} w_1^{-\frac{1}{p_1-1}}(y_1)\mathrm{d}y_1.$$

Since $p > 1$, it follows that

$$\begin{aligned}
&\left(\int_{2^{j-1}A\le|x|\le 2^jA}|[b,T^\delta]_1(f_1,f_2)(x)|^p\nu_{\vec{w}}(x)\mathrm{d}x\right)^{1/p}\\
&\le C\sum_{l=0}^{\infty}\left(\int_{2^{j-1}A\le|x|\le 2^jA}|V_R^l(x)|^p\nu_{\vec{w}}(x)\mathrm{d}x\right)^{1/p}\\
&\le C\sum_{l=0}^{\infty}\left(\int_{2^{j-1}A\le|x|\le 2^jA}\frac{1}{(2^{l-1}|x|)^{2np}}\nu_{\vec{w}}(x)\mathrm{d}x\right)^{1/p}\\
&\quad\times\left(\int_{B(0,R)}w_1^{-\frac{1}{p_1-1}}(y_1)\mathrm{d}y_1\right)^{1-1/p_1}\left(\int_{B(0,2^{l+j}A)}w_2^{-\frac{1}{p_2-1}}(y_2)\mathrm{d}y_2\right)^{1-1/p_2}\\
&\le C\sum_{l=0}^{\infty}(2^{l+j-2}A)^{-2n}(2^{-(j+l)}RA^{-1})^{n\theta_1(1-1/p_1)}\left(\int_{B(0,2^jA)}\nu_{\vec{w}}(x)\mathrm{d}x\right)^{1/p}\\
&\quad\times\left(\int_{B(0,2^{l+j}A)}w_1^{-\frac{1}{p_1-1}}(y_1)\mathrm{d}y_1\right)^{1-1/p_1}\left(\int_{B(0,2^{l+j}A)}w_2^{-\frac{1}{p_2-1}}(y_2)\mathrm{d}y_2\right)^{1-1/p_2}\\
&\le C\sum_{l=0}^{\infty}(2^{l+j}A)^{-2n}(2^{-(j+l)}RA^{-1})^{n\theta_1(1-1/p_1)}(2^{j+l}A)^{2n}\\
&\le C\sum_{l=0}^{\infty}2^{l(-n\theta_1(1-1/p_1))}2^{j(-n\theta_1(1-1/p_1))}(R/A)^{n\theta_1(1-1/p_1)}\\
&\le C2^{j(-n\theta_1(1-1/p_1))}(R/A)^{n\theta_1(1-1/p_1)}.
\end{aligned}$$

Thus,

$$\left(\int_{|x|>A}|[b,T^\delta]_1(f_1,f_2)(x)|^p\nu_{\vec{w}}(x)\mathrm{d}x\right)^{1/p}\le C(R/A)^{n\theta_1(1-1/p_1)}\to 0,$$

as $A\to\infty$.

So, it suffices to verify condition (iii). We denote

$$\begin{aligned}
E&=\{(x,y_1,y_2):\min(|x-y_1|,|x-y_2|)>\eta\},\\
F&=\{(x,y_1,y_2):\max(|x-y_1|,|x-y_2|)>2\delta\},\\
G&=\{(x,y_1,y_2):\max(|x-y_1|,|x-y_2|)>\eta\},\\
H&=\{(x,y_1,y_2):\delta<\max(|x-y_1|,|x-y_2|)<2\delta\}.
\end{aligned}$$

To prove (11), we decompose the expression inside the $L^p(\nu_{\vec{w}})$ norm as follows:

$$\begin{aligned}
&[b,T^\delta]_1(f_1,f_2)(x)-[b,T^\delta]_1(f_1,f_2)(x+t)\\
&=\int\int_E K^\delta(x,y_1,y_2)(b(x)-b(x+t))\prod_{j=1}^{2}f_j(y_j)\mathrm{d}y_1\mathrm{d}y_2\\
&\quad+\int\int_E(K^\delta(x,y_1,y_2)-K^\delta(x+t,y_1,y_2))(b(x+t)-b(y_1))\prod_{j=1}^{2}f_j(y_j)\mathrm{d}y_1\mathrm{d}y_2\\
&\quad+\int\int_{E^c}K^\delta(x,y_1,y_2)(b(x)-b(y_1))\prod_{j=1}^{2}f_j(y_j)\mathrm{d}y_1\mathrm{d}y_2
\end{aligned}$$

$$+\int\int_{E^c} K^\delta(x+t,y_1,y_2)(b(y_1)-b(x+t))\prod_{j=1}^{2} f_j(y_j)\mathrm{d}y_1\mathrm{d}y_2$$
$$=A(x)+B(x)+C(x)+D(x),$$

where $0<\eta<1$ and the choice of η will be specified later. It is obvious that $K^\delta(x,y_1,y_2)=K(x,y_1,y_2)$ on F. Consequently,

$$\Big|\int\int_E K^\delta(x,y_1,y_2)f_1(y_1)f_2(y_2)\mathrm{d}y_1\mathrm{d}y_2-\int\int_G K(x,y_1,y_2)f_1(y_1)f_2(y_2)\mathrm{d}y_1\mathrm{d}y_2\Big|$$
$$=\Big|\int\int_{(E\cap F)\cup(E\cap H)} K^\delta(x,y_1,y_2)f_1(y_1)f_2(y_2)\mathrm{d}y_1\mathrm{d}y_2$$
$$-\int\int_{(E\cap F)\cup(G\setminus(E\cap F))} K(x,y_1,y_2)f_1(y_1)f_2(y_2)\mathrm{d}y_1\mathrm{d}y_2\Big|$$
$$\le\Big|\int\int_{E\cap H} K^\delta(x,y_1,y_2)f_1(y_1)f_2(y_2)\mathrm{d}y_1\mathrm{d}y_2\Big|$$
$$+\int\int_{G\cap E^c}|K(x,y_1,y_2)f_1(y_1)f_2(y_2)|\mathrm{d}y_1\mathrm{d}y_2$$
$$+\int\int_{G\cap F^c\cap E}|K(x,y_1,y_2)f_1(y_1)f_2(y_2)|\mathrm{d}y_1\mathrm{d}y_2$$
$$+\int\int_{G\cap F^c\cap E^c}|K(x,y_1,y_2)f_1(y_1)f_2(y_2)|\mathrm{d}y_1\mathrm{d}y_2.$$

Now, we estimate the above four parts. From condition (2), we have

$$\Big|\int\int_{E\cap H} K^\delta(x,y_1,y_2)f_1(y_1)f_2(y_2)\mathrm{d}y_1\mathrm{d}y_2\Big|\le\int\int_H\frac{|f_1(y_1)||f_2(y_2)|}{(|x-y_1|+|x-y_2|)^{2n}}\mathrm{d}y_1\mathrm{d}y_2$$
$$\le C\mathcal{M}(f_1,f_2)(x),$$

and

$$\int\int_{G\cap E^c}|K(x,y_1,y_2)f_1(y_1)f_2(y_2)|\mathrm{d}y_1\mathrm{d}y_2$$
$$\le\int_{|x-y_1|<\eta}\int_{|x-y_2|>\eta}\frac{|f_1(y_1)||f_2(y_2)|}{(|x-y_1|+|x-y_2|)^{2n}}\mathrm{d}y_1\mathrm{d}y_2$$
$$\le\int_{|x-y_1|<\eta}|f_1(y_1)|\mathrm{d}y_1\sum_{k=1}^{\infty}\int_{2^{k-1}\eta<|x-y_2|<2^k\eta}\frac{|f_2(y_2)|}{|x-y_2|^{2n}}\mathrm{d}y_2$$
$$\le C\sum_{k=1}^{\infty}2^{-kn}\frac{1}{|B(x,\eta)|}\int_{B(x,\eta)}|f_1(y_1)|\mathrm{d}y_1\frac{1}{|B(x,2^k\eta)|}\int_{B(x,2^k\eta)}|f_2(y_2)|\mathrm{d}y_2$$
$$\le C\sum_{i=1}^{2}\mathcal{M}_{2,i}(f_1,f_2)(x),$$

where the set $G\cap E^c$ includes $\{(x,y_1,y_2):|x-y_1|<\eta,|x-y_2|>\eta\}$ and $\{(x,y_1,y_2):|x-y_1|>\eta,|x-y_2|<\eta\}$. Since the estimates on these two regions are similar, we omit the latter. This method will be used several times in the following.

Because $\eta < |x-y_1| < 2\delta$, $\eta < |x-y_2| < 2\delta$ when $(x, y_1, y_2) \in G \cap F^c \cap E$. Hence,

$$\int\int_{G\cap F^c\cap E} |K(x,y_1,y_2)f_1(y_1)f_2(y_2)|\mathrm{d}\vec{y} \le 4\delta \int\int_G \frac{|f_1(y_1)||f_2(y_2)|}{(|x-y_1|+|x-y_2|)^{2n+1}}\mathrm{d}\vec{y}$$
$$\le C\frac{\delta}{\eta}\mathcal{M}(f_1,f_2)(x),$$

and

$$\int\int_{G\cap F^c\cap E^c} |K(x,y_1,y_2)f_1(y_1)f_2(y_2)|\mathrm{d}\vec{y} \le \int_{|x-y_1|<\eta}\int_{|x-y_2|>\eta} \frac{|f_1(y_1)||f_2(y_2)|}{(|x-y_1|+|x-y_2|)^{2n}}\mathrm{d}\vec{y}$$
$$\le C\sum_{i=1}^{2}\mathcal{M}_{2,i}(f_1,f_2)(x).$$

In summary, we get

$$|A(x)| \le C|t|\|\nabla b\|_{L^\infty(\mathbf{R}^n)}\left|\int\int_E K^\delta(x,y_1,y_2)f_1(y_1)f_2(y_2)\mathrm{d}\vec{y}\right|$$
$$\le C|t|\left|\int\int_E K^\delta(x,y_1,y_2)f_1(y_1)f_2(y_2)\mathrm{d}\vec{y} - \int\int_G K(x,y_1,y_2)f_1(y_1)f_2(y_2)\mathrm{d}\vec{y}\right|$$
$$+ C|t|T^*(f_1,f_2)(x)$$
$$\le C|t|\left(T^*(f_1,f_2)(x) + \frac{1}{\eta}\mathcal{M}(f_1,f_2)(x) + \sum_{i=1}^{2}\mathcal{M}_{2,i}(f_1,f_2)(x)\right).$$

This, along with Lemma 2.2, Lemma 3.1, Lemma 4.2 and [22, Theorem 3.7], leads to that

$$\|A(x)\|_{L^p(\nu_{\vec{w}})} \le C|t|(1+1/\eta). \tag{12}$$

In order to estimate $B(x)$, by a consequence of condition (8), we have

$$|K(x,y_1,y_2) - K(x',y_1,y_2)| \le \frac{D|x-x'|^\gamma}{(|x-y_1|+|x-y_2|)^{2n+\gamma}}$$

when $|x-x'| \le \frac{1}{8}\min\{|x-y_1|,|x-y_2|\}$. Then

$$|B(x)| \le C\|b\|_{L^\infty(\mathbf{R}^n)}\int\int_E |K^\delta(x,y_1,y_2) - K^\delta(x+t,y_1,y_2)||f_1(y_1)||f_2(y_2)|\mathrm{d}\vec{y}$$
$$\le C|t|^\gamma\int\int_G \frac{|f_1(y_1)||f_2(y_2)|}{(|x-y_1|+|x-y_2|)^{2n+\gamma}}\mathrm{d}\vec{y}$$
$$\le C\frac{|t|^\gamma}{\eta^\gamma}\mathcal{M}(f_1,f_2)(x).$$

Therefore,

$$\|B(x)\|_{L^p(\nu_{\vec{w}})} \le C\frac{|t|^\gamma}{\eta^\gamma}. \tag{13}$$

For any $0 < \beta < 1$, we have $|b(x) - b(y_1)| \le |x-y_1|^\beta$. Thus, from the size condition (2) and the property of the support of $K^\delta(x,y_1,y_2)$, we can estimate $C(x)$:

$$\begin{aligned}|C(x)| &\le C\|\nabla b\|_{L^\infty(\mathbf{R}^n)}\eta\int_{|x-y_1|<\eta}\int_{|x-y_2|>\eta}\frac{|f_1(y_1)||f_2(y_2)|}{(|x-y_1|+|x-y_2|)^{2n}}\mathrm{d}y_1\mathrm{d}y_2\\ &\quad + C\int_{|x-y_1|>\eta}\int_{|x-y_2|<\eta}\frac{|f_1(y_1)||f_2(y_2)|}{(|x-y_1|+|x-y_2|)^{2n-\beta}}\mathrm{d}y_1\mathrm{d}y_2\\ &\le C\eta\mathcal{M}_{2,1}(\vec{f})(x) + C\int_{|x-y_2|<\eta}|f_2(y_2)|\mathrm{d}y_2\sum_{k=1}^{\infty}\int_{2^{k-1}\eta<|x-y_1|<2^k\eta}\frac{|f_1(y_1)|}{|x-y_1|^{2n-\beta}}\mathrm{d}y_1\\ &\le C\eta\mathcal{M}_{2,1}(\vec{f})(x) + C\eta^\beta\mathcal{M}_\beta^2(\vec{f})(x),\end{aligned}$$

provided $\eta < \delta$. From Lemma 2.2 and Lemma 2.3, we deduce that

$$\|C(x)\|_{L^p(\nu_{\vec{w}})} \le C\eta, \tag{14}$$

when we take sufficiently small β.

Finally, for the last part $D(x)$ we proceed in a similar way, by replacing x with $x+t$ and the region of integration E^c with a larger one $\{(x, y_1, y_2) : \min(|x+t-y_1|, |x+t-y_2|) < \eta + |t|\}$. By the fact that $x \in B(x+t, \eta+|t|)$, where $B(x+t, \eta+|t|)$ denote the ball centered at $x+t$ and with radius $\eta + |t|$, we obtain

$$\|D(x)\|_{L^p(\nu_{\vec{w}})} \le C(|t| + \eta). \tag{15}$$

Now, let us define $t_0 = \xi^2$ and for each $0 < |t| < t_0$, choose $\eta = |t|/\xi$. Then inequalities (12)-(15) imply (11), and in this way, we can conclude that $[b, T]_1$ is compact. By symmetry, $[b, T]_2$ is also compact. □

References

[1] Á Bényi, W Damián, K Moen, R H Torres. *Compact bilinear commutators: the weighted case*, Michigan Math J, 2015, 64: 39-51.

[2] Á Bényi, R H Torres. *Compact bilinear operators and commutators*, Proc Amer Math Soc, 2013, 141: 3609-3621.

[3] A P Calderón. *Intermediate spaces and interpolation, the complex method*, Studia Math, 1964, 24: 113-190.

[4] J C Chen, G E Hu. *Compact commutators of rough singular integral operators*, Canad Math Bull, 2015, 58: 19-29.

[5] Y P Chen, Y Ding, X X Wang. *Compactness of commutators for singular integrals on Morrey spaces*, Canad J Math, 2012, 64: 257-281.

[6] Y P Chen, Y Ding, X X Wang. *Compactness of commutators of Riesz potential on Morrey spaces*, Potential Anal, 2009, 30: 301-313.

[7] A Clop, V Cruz. *Weighted estimates for Beltrami equations*, Ann Acad Sci Fenn Math, 2013, 38: 91-113.

[8] R R Coifman, Y Meyer. *Nonlinear harmonic analysis, operator theory and PDE*, In: Beijing Lectures in Harmonic Analysis, Ann Math Stud 112, Princeton Univ Press, Princeton, NJ 1986, 3-45.

[9] R R Coifman, Y Meyer. *On commutators of singular integrals and bilinear singular integrals*, Trans Amer Math Soc, 1975, 212: 315-331.

[10] X T Duong, R Gong, L Grafakos, J Li, L X Yan. *Maximal operator for multilinear singular integrals with non-smooth kernels*, Indiana Univ Math J, 2009, 58: 2517-2541.

[11] X T Duong, L Grafakos, L X Yan. *Multilinear operators with non-smooth kernels and commutators of singular integrals*, Trans Amer Math Soc, 2010, 362: 2089-2113.

[12] X T Duong, A MaIntosh. *Singular integral operators with non-smooth kernels on irregular domains*, Rev Mat Iberoamericana, 1999, 15: 233-265.

[13] C Fefferman, E M Stein. *H^p spaces of several variables*, Acta Math, 1972, 129: 137-193.

[14] L Grafakos, L G Liu, D C Yang. *Multiple-weighted norm inequalities for maximal multi-linear singular integrals with non-smooth kernels*, Proc Roy Soc Edinburgh Sect A, 2011, 141: 755-775.

[15] L Grafakos, R H Torres. *Maximal operator and weighted norm inequalities for multilinear singular integrals*, Indiana Univ Math J, 2002, 51: 1261-1276.

[16] L Grafakos, R H Torres. *Multilinear Calderón-Zygmund theory*, Adv Math, 2002, 165: 124-164.

[17] G E Hu. *Weighted compact commutator of bilinear fourier multiplier operator*, Chin Ann Math Ser B, 2017, 38: 795-814.

[18] G E Hu, S Z Lu. *Weighted estimates for the multilinear singular integral operators with non-smooth kernels*, Sci China Math, 2011, 54: 587-602.

[19] G E Hu, D C Yang. *Weighted estimates for singular integral operators with nonsmooth kernels and applications*, J Aust Math Soc, 2008, 85: 377-417.

[20] G E Hu, Y P Zhu. *Weighted norm inequality with general weights for the commutator of Calderón*, Acta Math Sin (Engl Ser), 2013, 29: 505-514.

[21] S G Krantz, S Y Li. *Boundedness and compactness of integral operators on spaces of homogenoeus type and applications, II*, J Math Anal Appl, 2001, 258: 642-657.

[22] A Lerner, S Ombrosi, C Pérez, R H Torres, R Trujillo-González. *New maximal functions and multiple weights for the multilinear Calderón-Zygmund theory*, Adv Math, 2009, 220: 1222-1264.

[23] F Liu. *Continuity and approximate differentiability of multisublinear fractional maximal functions*, Math Inequal Appl, 2018, 21: 25-40.

[24] F Liu, T Chen, H X Wu. *A note on the endpoint regularity of the Hardy-Littlewood maximal functions*, Bull Aust Math Soc, 2016, 94: 121-130.

[25] F Liu, H X Wu. *Endpoint regularity of multisublinear fractional maximal functions*, Canad Math Bull, 2017, 60: 586-603.

[26] F Liu, Q Y Xue, K Yabuta. *Rough singular integral and maximal operators supported by subvarieties on Triebel-Lizorkin spaces*, Nonlinear Anal, 2018, 171: 41-72.

[27] C Pérez. *Endpoint estimates for commutators of singular integral operators*, J Funct Anal, 1995, 128: 163-185.

[28] C Pérez, G Pradolini, R H Torres, R Trujillo-González. *End-points estimates for iterated commutators of multilinear singular integrals*, Bull London Math Soc, 2014, 46: 26-42.

[29] M M Rao, Z D Ren. *Theory of Orlicz Space*, Marcel Dekker, New York, 1991.

[30] A Uchiyama. *On the compactness of operators of Hankel type*, Tôhoku Math J, 1978, 30: 163-171.

[31] K Yosida. *Functional Analysis*, Springer-Verlag, Berlin, 1995.

[32] X Yu, J C Chen, Y D Zhang. A_∞ *weight estimates for vector-valued commutators of multilinear fractional integral*, Appl Math J Chinese Univ Ser B, 2013, 28(3): 335-357.

[1] Department of Mathematics, Qingdao University of Science and Technology, Qingdao 266061, China

Email: burui0@163.com

[2] Department of Mathematics, Zhejiang Normal University, Jinhua 321004, China

Email: jcchen@zjnu.edu.cn