

MAXIMAL OPERATOR FOR THE HIGHER ORDER CALDERÓN COMMUTATOR

XUDONG LAI

ABSTRACT. In this paper, we investigate the weighted multilinear boundedness properties of the maximal higher order Calderón commutator for the dimensions larger than two. We establish all weighted multilinear estimates on the product of the $L^p(\mathbb{R}^d, w)$ space, including some peculiar endpoint estimates of the higher dimensional Calderón commutator.

1. INTRODUCTION

In the recent work [21], the author studied the multilinear boundedness of the higher order Calderón commutator. The purpose of this paper is to further generalize those results to the weighted space for its maximal type operator. Before stating our main results, let us give some notation and background. Define the truncated higher (n -th) order Calderón commutator by

$$\mathcal{C}_\varepsilon[\nabla A_1, \dots, \nabla A_n, f](x) = \int_{|x-y| \geq \varepsilon} K(x-y) \left(\prod_{i=1}^n \frac{A_i(x) - A_i(y)}{|x-y|} \right) \cdot f(y) dy,$$

where n is a positive integer and K is the Calderón-Zygmund convolution kernel on $\mathbb{R}^d \setminus \{0\}$ ($d \geq 2$) which means that K satisfies the following three conditions:

$$(1.1) \quad |K(x)| \lesssim |x|^{-d},$$

$$(1.2) \quad \int_{r < |x| < R} K(x) (x/|x|)^\alpha dx = 0, \quad \forall 0 < r < R < \infty \text{ and } \forall \alpha \in \mathbb{Z}_+^d \text{ with } |\alpha| = n,$$

$$(1.3) \quad |K(x-y) - K(x)| \lesssim |y|^\delta / |x|^{d+\delta} \text{ for some } 0 < \delta \leq 1 \text{ if } |x| > 2|y|.$$

Then we define the higher order Calderón commutator and its maximal operator by

$$(1.4) \quad \begin{aligned} \mathcal{C}[\nabla A_1, \dots, \nabla A_n, f](x) &= \lim_{\varepsilon \rightarrow 0} \mathcal{C}_\varepsilon[\nabla A_1, \dots, \nabla A_n, f](x), \\ \mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f](x) &= \sup_{\varepsilon > 0} |\mathcal{C}_\varepsilon[\nabla A_1, \dots, \nabla A_n, f](x)|. \end{aligned}$$

It is the standard context to check that these functions $\mathcal{C}[\nabla A_1, \dots, \nabla A_n, f](x)$ and $\mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f](x)$ are well defined for $A_1, \dots, A_n, f \in C_c^\infty(\mathbb{R}^d)$ (see *e.g.* [17]). This kind of commutator was first introduced by A. P. Calderón [2] when $n = 1$ and

2010 *Mathematics Subject Classification.* Primary 42B20, 42B25, Secondary 46B70, 47G30.

Key words and phrases. Multilinear, Calderón commutator, maximal operator, weighted space.

This work was supported by China Postdoctoral Science Foundation (No. 2017M621253, No. 2018T110279), the National Natural Science Foundation of China (No. 11801118) and the Fundamental Research Funds for the Central Universities.

$K(x)$ is a homogeneous kernel and later [3] [4] for the higher order one (see also [6], [7]). One can easily see that the first order Calderón commutator $\mathcal{C}[\nabla A, f](x)$ is a generalization of

$$[A, S]f(x) = A(x)S(f)(x) - S(Af)(x) = -\text{p.v.} \frac{1}{\pi} \int_{\mathbb{R}} \frac{1}{x-y} \frac{A(x) - A(y)}{x-y} f(y) dy$$

where $S = \frac{d}{dx} \circ H$ and H denotes the Hilbert transform. It is well known that the commutator $[A, S]$ and its generalization are elementary operators in harmonic analysis, which play an important role in the theory of the Cauchy integral along Lipschitz curve in \mathbb{C} , the boundary value problem of elliptic equation on non-smooth domain, the Kato square root problem on \mathbb{R} and the mixing flow problem (see *e.g.* [2], [4], [13], [24], [10], [18], [8], [25], [19], [23] for the details).

Many classical known results about the higher order Calderón commutator take place in the setting of the dimension $d = 1$. For example, the endpoint estimate that the n -th order Calderón commutator \mathcal{C} maps $L^1(\mathbb{R}) \times \cdots \times L^1(\mathbb{R}) \times L^1(\mathbb{R})$ to $L^{\frac{1}{1+n}, \infty}(\mathbb{R})$ was proved by C. P. Calderón [5] when $n = 1$, Coifman and Meyer [6] when $n = 1, 2$ and Duong, Grafakos and Yan [11] when $n \geq 1$. Here we point out that one important fact used by Coifman and Meyer [6], Duong, Grafakos and Yan [11] is that the one dimensional higher order Calderón commutator can be reduced to the multilinear Calderón Zygmund operator (see the very nice exposition [18, Chapter 7] and the reference therein). However when the dimension $d \geq 2$, things become complicated since Calderón commutator is a non standard multilinear Calderón-Zygmund operator. If we consider the Calderón-Zygmund kernel $K(x) = |x|^{-d}$, then the sharp bilinear estimates (except some endpoint estimates) of the first order Calderón commutator in this case has been established by Fong [14] via the time-frequency analysis method. For the more general Calderón-Zygmund kernel or even rough homogeneous kernel, the author [21] established all multilinear boundedness of the higher order Calderón commutator for the higher dimensions, especially the endpoint estimate that the n -th order Calderón commutator \mathcal{C} maps the product of Lorentz space $L^{d,1}(\mathbb{R}^d) \times \cdots \times L^{d,1}(\mathbb{R}^d) \times L^1(\mathbb{R}^d)$ to $L^{\frac{d}{d+n}, \infty}(\mathbb{R}^d)$.

The weighted results related to the Calderón commutator is also only known for the case $d = 1$. Duong, Gong, Grafakos, Li and Yan [12, Theorem 4.3] proved that \mathcal{C}_* maps $L^{q_1}(\mathbb{R}, w) \times \cdots \times L^{q_n}(\mathbb{R}, w) \times L^p(\mathbb{R}^d, w)$ to $L^r(\mathbb{R}, w)$ if $\frac{1}{r} = (\sum_{i=1}^n \frac{1}{q_i}) + \frac{1}{p}$ with $\frac{1}{n+1} < r < \infty$, $1 < q_1, \dots, q_n \leq \infty$, $1 < p < \infty$ and $w \in \cap_{i=1}^n A_{q_i}(\mathbb{R}) \cap A_p(\mathbb{R})$. For the endpoint estimate, Grafakos, Liu and Yang [16, Corollary 1.7] showed that \mathcal{C}_* maps $L^1(\mathbb{R}, w) \times \cdots \times L^1(\mathbb{R}, w) \times L^1(\mathbb{R}, w)$ to $L^{\frac{1}{1+n}, \infty}(\mathbb{R}, w)$ under the assumption $w \in A_1(\mathbb{R})$. The method used in Duong et al. [12] and Grafakos et al. [16] is both that by establishing the weighted theory for a class of multilinear Calderón-Zygmund operators with non-smooth kernel and then applying it to the Calderón commutator for the dimension $d = 1$. For the higher dimensional case of the Calderón commutator, no proper weighted multilinear Calderón-Zygmund theory can be applied directly.

In this paper, we are interested in the following weighted strong type multilinear estimate (or weighted weak type estimate) for the maximal operator of the higher

order Calderón commutator

$$(1.5) \quad \|\mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f]\|_{L^r(\mathbb{R}^d, w)} \lesssim \left(\prod_{i=1}^n \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d, w)} \right) \|f\|_{L^p(\mathbb{R}^d, w)}$$

where $\frac{1}{r} = \left(\sum_{i=1}^n \frac{1}{q_i}\right) + \frac{1}{p}$ with $1 \leq q_i \leq \infty$, $(i = 1, \dots, n)$, and $1 \leq p \leq \infty$. However, it is unknown whether those kind of estimates hold for the maximal Calderón commutator \mathcal{C}_* even in the unweighted case. In this paper, we will work directly on the weighted space and state our main results as follows.

Theorem 1.1. *Let $d \geq 2$ and n be a positive integer. Suppose K satisfies (1.1), (1.2) and (1.3). Assume that $\frac{1}{r} = \left(\sum_{i=1}^n \frac{1}{q_i}\right) + \frac{1}{p}$ with $1 \leq q_i \leq \infty$ ($i = 1, \dots, n$), and $1 \leq p \leq \infty$. Suppose $w \in \left(\bigcap_{i=1}^n A_{\max\{\frac{q_i}{d}, 1\}}(\mathbb{R}^d)\right) \cap A_p(\mathbb{R}^d)$. We may have the following conclusions:*

- (i). *If $\frac{d}{d+n} < r < \infty$, $1 < q_i \leq \infty$ ($i = 1, \dots, n$) and $1 < p \leq \infty$, then (1.5) holds.*
- (ii). *If $\frac{d}{d+n} \leq r < \infty$ with $q_i = 1$ for some $i = 1, \dots, n$; or $p = 1$; or $r = \frac{d}{d+n}$, then the following multilinear estimate holds*

$$(1.6) \quad \|\mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f]\|_{L^{r, \infty}(\mathbb{R}^d, w)} \lesssim \left(\prod_{i=1}^n \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d, w)} \right) \|f\|_{L^p(\mathbb{R}^d, w)}$$

and in this case, if $q_i = d$ for some $i = 1, \dots, n$, $L^{q_i}(\mathbb{R}^d, w)$ in the above inequality should be replaced by $L^{d,1}(\mathbb{R}^d, w)$, the weighted Lorentz space. Specially, we have the following endpoint estimate

$$(1.7) \quad \|\mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f]\|_{L^{\frac{d}{d+n}, \infty}(\mathbb{R}^d, w)} \lesssim \left(\prod_{i=1}^n \|\nabla A_i\|_{L^{d,1}(\mathbb{R}^d, w)} \right) \|f\|_{L^1(\mathbb{R}^d, w)}.$$

Remark 1.2. (i). These results in Theorem 1.1 are new even in the unweighted case when the dimension $d \geq 2$.

- (ii). When $0 < r < \frac{d}{d+n}$, these multilinear strong type estimates (1.5) (or weak type estimates (1.6)) do not hold for the maximal Calderón operator \mathcal{C}_* . In fact, some counterexamples has been constructed in [21, Theorem 1.1] to show that those multilinear strong type estimates (or weak type estimates) fail even for the operator \mathcal{C} in the case $0 < r < \frac{d}{d+n}$. Thus our results in Theorem 1.1 are optimal in this sense.

- (iii). The condition of the weight $w \in \left(\bigcap_{i=1}^n A_{\max\{\frac{q_i}{d}, 1\}}(\mathbb{R}^d)\right) \cap A_p(\mathbb{R}^d)$ seems to be unnatural at the first sight, since it doesn't appear previously. However, this kind of condition is just appropriate for the higher dimensional Calderón commutator as we will see in our later proof. In fact $w \in A_{\max\{\frac{q_i}{d}, 1\}}(\mathbb{R}^d)$ comes from $\nabla A_i \in L^{q_i}(\mathbb{R}^d, w)$ and $w \in A_p(\mathbb{R}^d)$ comes from $f \in L^p(\mathbb{R}^d, w)$. When the dimension $d = 1$, (1.5) turns out to be that \mathcal{C}_* maps $L^{q_1}(\mathbb{R}, w) \times \dots \times L^{q_n}(\mathbb{R}, w) \times L^p(\mathbb{R}^d, w)$ to $L^r(\mathbb{R}, w)$ if $\frac{1}{n+1} < r < \infty$, $1 < q_1, \dots, q_n \leq \infty$, $1 < p \leq \infty$ and $w \in \bigcap_{i=1}^n A_{q_i}(\mathbb{R}) \cap A_p(\mathbb{R})$, which has been proved by Duong, Gong, Grafakos, Li and Yan [12, Theorem 4.3] except the endpoint case $q_i = \infty$ for some i or $p = \infty$. Therefore even in the one dimensional case (1.5) is new at the endpoint case $q_i = \infty$ for some i or $p = \infty$. To the best knowledge of the author, (1.5) is new when $d \geq 2$.

- (iv). Notice that $L^{1,1}(\mathbb{R}, w) = L^1(\mathbb{R}, w)$. Therefore when the dimension $d = 1$, (1.7) is just that the maximal n -th order Calderón commutator maps $L^1(\mathbb{R}, w) \times \cdots \times L^1(\mathbb{R}, w) \times L^1(\mathbb{R}, w)$ to $L^{\frac{1}{1+n}, \infty}(\mathbb{R}, w)$ under the assumption $w \in A_1(\mathbb{R})$, which has been proved by Grafakos, Liu and Yang [16, Corollary 1.7]. To the best knowledge of the author, (1.7) is new when $d \geq 2$. Although we assume that $d \geq 2$ in our main results, the proof presented in this paper is also valid for $d = 1$. Therefore even when $d = 1$, the proof of (1.5) and (1.7) here are quite different from that by Duong, Gong, Grafakos, Li and Yan [12], Grafakos, Liu and Yang [16], thus we give new proofs of (1.5) and (1.7) for $d = 1$.
- (v). Currently, there are extensively research on seeking the optimal quantitative weighted bound for singular integral. We do not pursue this topic in this paper but hope to work on it in the future work.

Notice first that if $q_i = \infty$ with $i = 1, \dots, n$, i.e. A_i is a Lipschitz function, then $\mathcal{C}[\nabla A_1, \dots, \nabla A_n, \cdot]$ is a standard Calderón Zygmund operator. By the standard weighted theory of the Calderón-Zygmund operator, we may easily get that \mathcal{C}_* maps $L^\infty(\mathbb{R}^d, w) \times \cdots \times L^\infty(\mathbb{R}^d, w) \times L^p(\mathbb{R}^d, w)$ to $L^p(\mathbb{R}^d, w)$ for $1 < p < \infty$ and $L^\infty(\mathbb{R}^d, w) \times \cdots \times L^\infty(\mathbb{R}^d, w) \times L^1(\mathbb{R}^d, w)$ to $L^{1, +\infty}(\mathbb{R}^d, w)$. Recall the method used in [12] or [16], by establishing the Cotlar inequality for the multilinear Calderón-Zygmund operator, the authors in [12] or [16] proved the weighted multilinear estimates for the Calderón-Zygmund operator and then applies them to the one-dimensional Calderón commutator. There are also variants of the Cotlar inequality for the higher dimensional Calderón commutator, which is available only for the multilinear estimates (1.5) in the case that all $q_i > d, i = 1, \dots, n, r > 1$ (see Proposition 3.3). To deal with the remainder case, our strategy is as follows. We straightforward establish the endpoint estimates in (ii) of Theorem 1.1, which means that we need to give some weak type estimates. Note that A_i belongs to the Sobolev space $W^{1, q_i}(\mathbb{R}^d, w)$. We will construct an *exceptional set* which satisfies the required weighted weak type estimate. And on the complementary set of *exceptional set*, the function A_i is a Lipschitz function with a bound $\lambda^{\frac{r}{q_i}}$. Then, roughly speaking, the strong type estimate and the weak type $L^{1, \infty}(\mathbb{R}^d, w)$ boundedness (with $q_i = \infty, i = 1, \dots, n$) of $\mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f](x)$ could be applied on the complementary set of *exceptional set*. To construct the *exceptional set*, we will make use of the Marry Weiss maximal operator and the weighted Sobolev inequality.

This paper is organized as follows. Firstly some preliminary lemmas are presented in Section 2. In Section 3, we give the proof of Theorem 1.1. The proof is divided into several case. In Subsection 3.1, we prove some strong type estimates of (i) in Theorem 1.1. The proofs of (ii) in Theorem 1.1 are given in Subsections 3.2 and 3.3. In Subsection 3.4, we shall use the linear Marcinkiewicz interpolation with some strong type estimates of (i) and full weak type estimates of (ii) to show the rest of (i) in Theorem 1.1.

Notation. Throughout this paper, we only consider the dimension $d \geq 2$ and the letter C stands for a positive finite constant which is independent of the essential variables, not necessarily the same one in each occurrence. $A \lesssim B$ means $A \leq CB$ for some constant C . By the notation C_ε means that the constant depends on the

parameter ε . $A \approx B$ means that $A \lesssim B$ and $B \lesssim A$. n represents the order of Calderón commutator. The indexes r, q_1, \dots, q_n and p satisfy $\frac{1}{r} = \left(\sum_{i=1}^n \frac{1}{q_i}\right) + \frac{1}{p}$ with $1 \leq q_i \leq \infty$ ($i = 1, \dots, n$) and $1 \leq p \leq \infty$ in the whole paper. For a set $E \subset \mathbb{R}^d$, we denote by $w(E) = \int_E w(x)dx$. ∇A will stand for the vector $(\partial_1 A, \dots, \partial_d A)$ where $\partial_i A(x) = \partial A(x)/\partial x_i$. Define $\mathbb{N}_i^j = \{i, i+1, \dots, j\}$. Set

$$\|\nabla A\|_X = \left\| \left(\sum_{i=1}^d |\partial_i A|^2 \right)^{\frac{1}{2}} \right\|_X$$

for $X = L^p(\mathbb{R}^d, w)$ or $X = L^{d,1}(\mathbb{R}^d, w)$. \mathbb{Z}_+ denotes the set of all nonnegative integers and $\mathbb{Z}_+^d = \underbrace{\mathbb{Z}_+ \times \dots \times \mathbb{Z}_+}_d$. For $\alpha \in \mathbb{Z}_+^d$ and $x \in \mathbb{R}^d$, we define $x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_d^{\alpha_d}$.

2. SOME PRELIMINARY LEMMAS

In this section, we will introduce the weighted properties of some operators which are useful in the proof of Theorem 1.1. Those operators include the Hardy-Littlewood maximal operator with order δ , the maximal sharp function operator, the Marry Weiss maximal operator and many others. And also a weighted Sobolev inequality is needed.

Definition 2.1 ($A_p(\mathbb{R}^d)$ weight). A nonnegative locally integrable function w on \mathbb{R}^d is called to be an $A_p(\mathbb{R}^d)$ weight if there exists a constant $C > 0$ such that

$$(2.1) \quad \sup_Q \left(\frac{1}{|Q|} \int_Q w(x)dx \right) \left(\frac{1}{|Q|} \int_B w(x)^{-\frac{1}{p-1}} dx \right)^{p-1} \leq C < \infty,$$

where the supremum is taken all cube Q in \mathbb{R}^d . The smallest constant C for (2.1) holds is called the A_p bound of w and is denoted by $[w]_{A_p}$. We call w an $A_1(\mathbb{R}^d)$ weight if there exists a constant C independent of Q such that

$$(2.2) \quad \frac{1}{|Q|} \int_Q w(z)dz \leq Cw(y), \quad a.e. \quad y \in Q.$$

And we set the smallest constant C in (2.2) as $[w]_{A_1}$, which is called the A_1 bound of w . We also set $A_\infty(\mathbb{R}^d) = \bigcup_{1 \leq q < \infty} A_q(\mathbb{R}^d)$.

It is easy to see that an equivalent definition of $A_1(\mathbb{R}^d)$ weight is that $M(w) \leq Cw(x)$, where M is the Hardy-Littlewood maximal operator. Recall the following basic fact about $A_p(\mathbb{R}^d)$ weight (see [17]):

$$A_p(\mathbb{R}^d) \subsetneq A_q(\mathbb{R}^d), \quad \text{if } 1 \leq p < q \leq \infty.$$

Lemma 2.2 (see [20] or [22]). Suppose that $w \in A_\infty(\mathbb{R}^d)$. Let $0 < \delta, q < \infty$. Then there exists a constant C depends only on w, δ, q such that

$$\int_{\mathbb{R}^d} [M_\delta f(x)]^q w(x)dx \leq C \int_{\mathbb{R}^d} [M_\delta^\sharp f(x)]^q w(x)dx$$

holds for any function provided that the left side integral is finite. Here M_δ , M_δ^\sharp are the Hardy-Littlewood maximal operator with order δ and the maximal sharp function operator, which are defined as

$$M_\delta(f)(x) = \sup_{r>0} \left(\frac{1}{|Q(x,r)|} \int_{Q(x,r)} |f(y)|^\delta dy \right)^{\frac{1}{\delta}},$$

$$M_\delta^\sharp(f)(x) = \sup_{Q \ni x} \inf_c \left(\frac{1}{|Q|} \int_Q |f(z) - c|^\delta dz \right)^{\frac{1}{\delta}},$$

where Q is a cube in \mathbb{R}^d and $Q(x,r)$ is a cube with center x and sidelength r .

Next we state some properties of a special maximal function introduced firstly by Mary Weiss (see [5]), which is defined as

$$\mathcal{M}(\nabla A)(x) = \sup_{h \in \mathbb{R}^d \setminus \{0\}} \frac{|A(x+h) - A(x)|}{|h|}.$$

Lemma 2.3. *Suppose that $w \in A_{p/d}(\mathbb{R}^d)$ with $p > d$. Let $\nabla A \in L^p(\mathbb{R}^d, w)$. Then \mathcal{M} is bounded on $L^p(\mathbb{R}^d, w)$, that is*

$$\|\mathcal{M}(\nabla A)\|_{L^p(\mathbb{R}^d, w)} \lesssim \|\nabla A\|_{L^p(\mathbb{R}^d, w)}.$$

Proof. By using the dense argument, it is sufficient to consider A as a C^∞ function with compact support. Then by the result of [5, Lemma 1.4], we get that for any $q > d$,

$$\frac{|A(x) - A(y)|}{|x - y|} \lesssim \left(\frac{1}{|x - y|^d} \int_{|x-z| \leq 2|x-y|} |\nabla A(z)|^q dz \right)^{\frac{1}{q}}.$$

Since $w \in A_{p/d}(\mathbb{R}^d)$, by the reverse Hölder inequality of $A_{p/d}(\mathbb{R}^d)$ weight (see [17]) and its definition, there exist $\varepsilon > 0$ such that $w \in A_{p/d-\varepsilon}(\mathbb{R}^d)$ and $p/d - \varepsilon \geq 1$. Therefore we may choose q in the above inequality such that $p/d - \varepsilon = p/q$ and $d < q < p$. Applying the fact that the Hardy-Littlewood maximal operator M maps $L^s(\mathbb{R}^d, w)$ to itself if $1 < s \leq \infty$ and $w \in A_s(\mathbb{R}^d)$, we may get

$$\|\mathcal{M}(\nabla A)\|_{L^p(\mathbb{R}^d, w)} \lesssim \|M_q(\nabla A)\|_{L^p(\mathbb{R}^d, w)} = \|M(|\nabla A|^q)\|_{L^{p/q}(\mathbb{R}^d, w)}^{\frac{1}{q}} \lesssim \|\nabla A\|_{L^p(\mathbb{R}^d, w)},$$

which completes the proof. \square

Lemma 2.4. *Let A be a function such that $\nabla A \in L^{d,1}(\mathbb{R}^d, w)$, the Lorentz space with weight $w \in A_1(\mathbb{R}^d)$. Then for any $\lambda > 0$,*

$$w(\{x \in \mathbb{R}^d : \mathcal{M}(\nabla A)(x) > \lambda\}) \lesssim \lambda^{-d} \|\nabla A\|_{L^{d,1}(\mathbb{R}^d, w)}^d.$$

Proof. By the dense argument, it is sufficient to consider A as a smooth function with compact support. Using the formula [26, page 125, (17)], one may write

$$A(x) = C_d \sum_{j=1}^d \int_{\mathbb{R}^d} \frac{x_j - y_j}{|x - y|^d} \partial_j A(y) dy = \mathbb{K} * f(x)$$

where $\mathbb{K}(x) = 1/|x|^{d-1}$, $f = C_d \sum_{j=1}^d R_j(\partial_j A)$ with R_j s the Riesz transforms. Notice that $w \in A_1(\mathbb{R}^d) \subsetneq A_p(\mathbb{R}^d)$ for all $1 < p \leq \infty$. By applying the general form of the

Marcinkiewicz interpolation theorem (see [28, page 197, Theorem 3.15]), we obtain that the Riesz transform R_j maps $L^{d,1}(\mathbb{R}^d, w)$ to itself. Then it is easy to see that

$$\|f\|_{L^{d,1}(\mathbb{R}^d, w)} \lesssim \|\nabla A\|_{L^{d,1}(\mathbb{R}^d, w)}.$$

Hence to finish the proof, it is sufficient to prove that

$$(2.3) \quad w(\{x \in \mathbb{R}^d : \mathcal{M}(\nabla A)(x) > \lambda\}) \lesssim \lambda^{-d} \|f\|_{L^{d,1}(\mathbb{R}^d, w)}^d$$

with $A = \mathbb{K} * f$. Below we shall show that for any $x \in \mathbb{R}^d$, the following estimate

$$|A(x+h) - A(x)| \lesssim |h|T(f)(x)$$

holds uniformly for $h \in \mathbb{R}^d \setminus \{0\}$ where T is an operator maps $L^{d,1}(\mathbb{R}^d, w)$ to $L^{d,\infty}(\mathbb{R}^d, w)$. Once we show this, we get (2.3) and hence finish the proof of Lemma 2.4. We write

$$\begin{aligned} & A(x+h) - A(x) \\ &= \int_{|x-y| \leq 2|h|} |x+h-y|^{-d+1} f(y) dy - \int_{|x-y| \leq 2|h|} |x-y|^{-d+1} f(y) dy \\ & \quad + \int_{|x-y| > 2|h|} (|x+h-y|^{-d+1} - |x-y|^{-d+1}) f(y) dy \\ &= I + II + III. \end{aligned}$$

Consider I firstly. Observe that $\mathbb{K} \in L^{d',\infty}(\mathbb{R}^d)$ where $d' = d/(d-1)$. Set $B(x, r) = \{y \in \mathbb{R}^d : |x-y| \leq r\}$. Applying the rearrangement inequality (see [17, page 74, Exercise 1.4.1]), we obtain that

$$\begin{aligned} |I| &\leq \int_{\mathbb{R}^d} \mathbb{K}(x+h-y) |f \chi_{B(x, 2|h|)}(y)| dy \leq \int_0^\infty \mathbb{K}^*(s) (f \chi_{B(x, 2|h|)})^*(s) ds \\ &\leq \left(\int_0^\infty (f \chi_{B(x, 2|h|)})^*(s) s^{\frac{1}{d}} \frac{ds}{s} \right) \cdot \sup_{s>0} \left(\mathbb{K}^*(s) s^{\frac{1}{d'}} \right) \\ &\lesssim \|f \chi_{B(x, 2|h|)}\|_{L^{d,1}(\mathbb{R}^d)} \|\mathbb{K}\|_{L^{d',\infty}(\mathbb{R}^d)}, \end{aligned}$$

where f^* stands for the decreasing rearrangement of f . Applying the definition of Lorentz space, we may get that $\|\chi_E\|_{L^{d,1}(\mathbb{R}^d)} = \|\chi_E\|_{L^d(\mathbb{R}^d)}$ holds for any characteristic function χ_E of set E of finite measure, thus $\|\chi_{B(x, 2|h|)}\|_{L^{d,1}(\mathbb{R}^d)} = C_d |h|$. Then we obtain that

$$|I| \lesssim |h| \Lambda(f)(x), \quad \text{where } \Lambda(f)(x) = \sup_{r>0} \frac{\|f \chi_{B(x, r)}\|_{L^{d,1}(\mathbb{R}^d)}}{\|\chi_{B(x, r)}\|_{L^{d,1}(\mathbb{R}^d)}}.$$

In the following it is sufficient to show that the operator Λ maps $L^{d,1}(\mathbb{R}^d, w)$ to $L^{d,\infty}(\mathbb{R}^d, w)$. Note that $L^{d,1}(\mathbb{R}^d, w)$ is a Banach space (see *e.g.* [28, page 204, Theorem 3.22]), it suffices to show that Λ is restricted of type (d, d) , thus is $\|\Lambda(\chi_E)\|_{L^{d,\infty}(\mathbb{R}^d, w)} \lesssim w(E)^{\frac{1}{d}}$ (see *e.g.* [17, page 62, Lemma 1.4.20]). However in this case, the proof is equivalent to show that

$$w(\{x \in \mathbb{R}^d : M(\chi_E)(x) > \lambda\}) \lesssim \lambda^{-1} \|\chi_E\|_{L^1(\mathbb{R}^d, w)},$$

where M is the Hardy-Littlewood maximal operator. Since M is weighted weak type $(1,1)$ if $w \in A_1(\mathbb{R}^d)$, hence we prove that Λ maps $L^{d,1}(\mathbb{R}^d, w)$ to $L^{d,\infty}(\mathbb{R}^d, w)$.

Next consider *II*. Observe that the kernel $k(x) := \varepsilon^{-1}|x|^{-d+1}\chi_{\{|x|\leq\varepsilon\}}$ is radial non-increasing and L^1 integrable in \mathbb{R}^d , we get

$$|II| \lesssim \|k\|_{L^1(\mathbb{R}^d)} |h| M(f)(x).$$

Notice that $L^{p,1}(\mathbb{R}^d, w) \subset L^p(\mathbb{R}^d, w)$ and M maps $L^p(\mathbb{R}^d, w)$ to itself for $1 < p < \infty$. Hence we get that M maps $L^{d,1}(\mathbb{R}^d, w)$ to $L^{d,\infty}(\mathbb{R}^d, w)$.

Finally consider *III*. Notice that it suffices to consider $|x-y| > 2|h|$. Then applying the Taylor expansion of $|x-y+h|^{-d+1}$, we get

$$(2.4) \quad \frac{1}{|x-y+h|^{d-1}} - \frac{1}{|x-y|^{d-1}} = (-d+1) \sum_{j=1}^d h_j \frac{x_j - y_j}{|x-y|^{d+1}} + R(x, y, h)$$

where the remainder term $R(x, y, h)$ in the Taylor expansion satisfies

$$|R(x, y, h)| \leq C|h|^2|x-y|^{-d-1} \text{ if } |x-y| > 2|h|.$$

Inserting (2.4) into the term *III* with the above estimate of $R(x, y, h)$, we conclude that

$$|III| \lesssim |h| \sum_{j=1}^d R_j^*(f)(x) + |h|^2 \int_{|x-y|>2|h|} |x-y|^{-d-1} |f(y)| dy$$

where R_j^* is the maximal Riesz transform defined by

$$R_j^*(f)(x) = \sup_{\varepsilon>0} \left| \int_{|x-y|>\varepsilon} \frac{x_j - y_j}{|x-y|^{d+1}} f(y) dy \right|.$$

Since R_j^* is bounded on $L^p(\mathbb{R}^d, w)$ for $1 < p < \infty$, we immediately obtain that R_j^* maps $L^{d,1}(\mathbb{R}^d, w)$ to $L^{d,\infty}(\mathbb{R}^d, w)$. The second term which controls *III* can be dealt similar to that of the estimate of *II* once we observe that $\varepsilon|x|^{-d-1}\chi_{\{|x|>\varepsilon\}}$ is radial non-increasing and L^1 integrable. \square

In the following, we introduce a weighted Sobolev inequality and a key weighted weak type estimate for $\nabla A \in L^p(\mathbb{R}^d, w)$ with $1 \leq p < d$. Define the weighted Hardy-Littlewood maximal operator of order p $M_{w,p}$ and the weighted maximal operator $\mathfrak{M}_{w,s}$ by

$$M_{w,p}(f)(x) = \sup_{r>0} \left(\frac{1}{w(Q(x,r))} \int_{Q(x,r)} |f(y)|^p w(y) dy \right)^{1/p},$$

$$\mathfrak{M}_{w,s}(\nabla A)(x) = \sup_{r>0} \left(\frac{1}{w(Q(x,r))} \int_{Q(x,r)} \left| \frac{A(x) - A(y)}{r} \right|^s w(y) dy \right)^{1/s},$$

where $Q(x, r)$ is a cube with center x and sidelength r .

Lemma 2.5 (see [9]). *If w is an $A_1(\mathbb{R}^d)$ weight, then the following weighted Sobolev inequality*

$$\left(\int_{\mathbb{R}^d} |g(x)|^s w(y) dy \right)^{\frac{1}{s}} \leq C \left(\int_{\mathbb{R}^d} [w(x)^{-\frac{1}{d}} |\nabla g(x)|]^p w(x) dx \right)^{\frac{1}{p}}$$

holds for $1 \leq p < d$ and $1/s = 1/p - 1/d$, where g is a C^1 function with compact support and the constant C does not depend on g .

Lemma 2.6. *Let $w \in A_1(\mathbb{R}^d)$ and $\nabla A \in L^p(\mathbb{R}^d, w)$ with $1 \leq p < d$. Set $1/s = 1/p - 1/d$. Then we have*

$$w(\{x \in \mathbb{R}^d : \mathfrak{M}_{w,s}(\nabla A)(x) > \lambda\}) \lesssim \lambda^{-p} \|\nabla A\|_{L^p(\mathbb{R}^d, w)}^p.$$

Proof. By using a standard limiting argument, we only need to consider A as a C^∞ function with compact support. Fix a cube $Q(x, r)$. Choose a C_c^∞ function ϕ such that $\phi(y) \equiv 1$ if $y \in Q(x, r)$, $\text{supp}\phi \subset Q(x, 2r)$ and $\|\nabla\phi\|_{L^\infty(\mathbb{R}^d)} \lesssim r^{-1}$. Consider the auxiliary function $\phi(y)(A(x) - A(y))$ where x is fixed and y is the variable. Using the weighted Sobolev inequality in Lemma 2.5 and the property (2.2) of $A_1(\mathbb{R}^d)$ weight, one may get that

$$\begin{aligned} \left(\int_{Q(x,r)} |A(x) - A(y)|^s w(y) dy \right)^{\frac{1}{s}} &\lesssim \left[\int_{\mathbb{R}^d} [\nabla_y(\phi(y)(A(x) - A(y)))]^p w(y)^{1-\frac{p}{d}} dy \right]^{\frac{1}{p}} \\ &\lesssim \left(\int_{Q(x,2r)} |\nabla A(y)|^p w(y)^{1-\frac{p}{d}} dy \right)^{\frac{1}{p}} \\ &\quad + \left(\int_{Q(x,2r) \setminus Q(x,r)} \left| \frac{A(x) - A(y)}{r} \right|^p w(y)^{1-\frac{p}{d}} dy \right)^{\frac{1}{p}} \\ &\lesssim \left[\frac{w(Q(x, 2r))}{|Q(x, 2r)|} \right]^{-\frac{1}{d}} \left[\left(\int_{Q(x,2r)} |\nabla A(y)|^p w(y) dy \right)^{\frac{1}{p}} \right. \\ &\quad \left. + \left(\int_{Q(x,2r) \setminus Q(x,r)} \left| \frac{A(x) - A(y)}{r} \right|^p w(y) dy \right)^{\frac{1}{p}} \right]. \end{aligned}$$

The above estimate, via the doubling property of $w(x)dx$ (i.e. $w(2Q) \lesssim w(Q)$, see [17]) and $\frac{1}{s} = \frac{1}{p} - \frac{1}{d}$, yields that

$$\left(\frac{1}{w(Q(x, r))} \int_{Q(x,r)} \left| \frac{A(x) - A(y)}{r} \right|^s w(y) dy \right)^{\frac{1}{s}} \lesssim M_{w,p}(\nabla A)(x) + S_p(\nabla A)(x),$$

where

$$S_p(\nabla A)(x) := \left[\frac{1}{w(Q(x, 2r))} \int_{Q(x,2r) \setminus Q(x,r)} \left| \frac{A(x) - A(y)}{r} \right|^p w(y) dy \right]^{\frac{1}{p}}.$$

Again using the fact that $w(x)dx$ satisfies the doubling property, one may see that the Hardy-Littlewood maximal operator $M_{w,1}$ with the weight w is of weak type $(1,1)$, thus is $M_{w,1}$ maps $L^1(\mathbb{R}^d, w)$ to $L^{1,\infty}(\mathbb{R}^d, w)$. Then we get that $M_{w,p}$ maps $L^p(\mathbb{R}^d, w)$ to $L^{p,\infty}(\mathbb{R}^d, w)$. Therefore to complete the proof, it is enough to show that $S_p(\nabla A)(x) \lesssim T(\nabla A)(x)$ with T mapping $L^p(\mathbb{R}^d, w)$ to $L^{p,\infty}(\mathbb{R}^d, w)$.

Below we give some explicit estimates of $A(x) - A(y)$ similar to that in the proof of Lemma 2.4. By the formula given in [26, page 125, (17)], we may write

$$A(x) = C_d \sum_{j=1}^d \int_{\mathbb{R}^d} \frac{x_j - y_j}{|x - y|^d} \partial_j A(y) dy.$$

Split $A(x) - A(y)$ into three terms as follows,

$$\begin{aligned}
& A(x) - A(y) \\
(2.5) \quad &= C_d \sum_{j=1}^d \left[\int_{|x-z| \leq 2|x-y|} \frac{x_j - z_j}{|x-z|^d} \partial_j A(z) dz - \int_{|x-z| \leq 2|x-y|} \frac{y_j - z_j}{|y-z|^d} \partial_j A(z) dz \right. \\
&\quad \left. + \int_{|x-z| > 2|x-y|} \left(\frac{x_j - z_j}{|x-z|^d} - \frac{y_j - z_j}{|y-z|^d} \right) \partial_j A(z) dz \right] \\
&= I(x) + II(x) + III(x).
\end{aligned}$$

Plug the above three terms back into $S_p(\nabla A)(x)$ and define these three terms as $S_{p,1}(\nabla A)(x)$, $S_{p,2}(\nabla A)(x)$ and $S_{p,3}(\nabla A)(x)$ respectively.

Let us first consider $S_{p,1}(\nabla A)(x)$. By applying the Hölder inequality,

$$|I(x)|^p \lesssim |x-y|^{p-1} \left(\int_{|x-z| \leq 2|x-y|} \frac{|\nabla A(z)|^p}{|x-z|^{d-1}} dz \right).$$

Plugging the above inequality into $S_{p,1}(\nabla A)(x)$ with $|x-y| \approx r$, and then using the kernel $k(x) = \varepsilon^{-1}|x|^{-d+1}\chi_{\{|x| \leq C\varepsilon\}}$ is a radial non-increasing function and L^1 integrable in \mathbb{R}^d , we get that

$$(2.6) \quad S_{p,1}(\nabla A)(x) \lesssim \left(r^{-1} \int_{|x-z| \lesssim r} \frac{|\nabla A(z)|^p}{|x-z|^{d-1}} dz \right)^{\frac{1}{p}} \lesssim M_p(\nabla A)(x).$$

It is easy to see that M_p maps $L^p(\mathbb{R}^d, w)$ to $L^{p,\infty}(\mathbb{R}^d, w)$ with w an $A_1(\mathbb{R}^d)$ weight is equivalent to that the Hardy-Littlewood maximal operator M maps $L^1(\mathbb{R}^d, w)$ to $L^{1,\infty}(\mathbb{R}^d, w)$, which is however well known.

Next we consider $S_{p,2}(\nabla A)(x)$. By using the Hölder inequality to deal with $II(x)$ as those of $I(x)$, then applying $|x-y| \approx r$ and the Fubini theorem, we get

$$\begin{aligned}
(2.7) \quad S_{p,2}(\nabla A)(x) &\lesssim \left[\frac{1}{w(Q(x, 2r))} \int_{Q(x, 2r) \setminus Q(x, r)} \frac{1}{r} \left(\int_{|x-z| \lesssim r} \frac{|\nabla A(z)|^p}{|y-z|^{d-1}} dz \right) w(y) dy \right]^{\frac{1}{p}} \\
&\lesssim \left[\frac{1}{w(Q(x, 2r))} \int_{|x-z| \lesssim r} r^{-1} \left(\int_{|y-z| \lesssim r} \frac{w(y)}{|y-z|^{d-1}} dy \right) |\nabla A(z)|^p dz \right]^{\frac{1}{p}} \\
&\lesssim \left[\frac{1}{w(Q(x, 2r))} \int_{|x-z| \lesssim r} M(w)(z) |\nabla A(z)|^p dz \right]^{\frac{1}{p}} \\
&\lesssim \left[\frac{1}{w(Q(x, 2r))} \int_{|x-z| \lesssim r} |\nabla A(z)|^p w(z) dz \right]^{\frac{1}{p}} \lesssim M_{w,p}(\nabla A)(x),
\end{aligned}$$

where in the third inequality we use again the fact that the kernel function $k(x) = \varepsilon^{-1}|x|^{-d+1}\chi_{\{|x| \leq C\varepsilon\}}$ is a radial non-increasing function and L^1 integrable in \mathbb{R}^d , the last second inequality follows from (2.2). As showed previously, $M_{w,p}$ maps $L^p(\mathbb{R}^d, w)$ to $L^{p,\infty}(\mathbb{R}^d, w)$.

We consider $S_{p,3}(\nabla A)(x)$. Set $\mathcal{K}_j(x) = \frac{x_j}{|x|^d}$. Notice that $|x - z| > 2|x - y|$. Applying the Taylor expansion of $\mathcal{K}_j(x - z)$, we may get

$$\mathcal{K}_j(x - z) - \mathcal{K}_j(y - z) = \sum_{i=1}^d (x_i - y_i) \partial_i \mathcal{K}_j(x - z) + R(x, y, z)$$

where the Taylor expansion's remainder term $R(x, y, z)$ satisfies

$$|R(x, y, z)| \lesssim |x - y|^2 |x - z|^{-d-1}, \quad \forall |x - z| > 2|x - y|.$$

Plunge the Taylor expansion's main term and reminder term into $S_{p,3}(\nabla A)(x)$ and split $S_{p,3}(\nabla A)(x)$ as two terms $S_{p,3,m}(\nabla A)(x)$ (related to main term) and $S_{p,3,r}(\nabla A)(x)$ (related to reminder term), respectively. Then by $|x - y| \approx r$, we have the following estimate of $S_{p,3,m}(\nabla A)(x)$,

$$\begin{aligned} & S_{p,3,m}(\nabla A)(x) \\ & \lesssim \left[\frac{1}{w(Q(x, 2r))} \int_{Q(x, 2r)} \left(\sum_{j=1}^d \sum_{i=1}^d \left| \int_{|x-z|>2|x-y|} \partial_i \mathcal{K}_j(x - z) \partial_j A(z) dz \right| \right)^p w(y) dy \right]^{\frac{1}{p}} \\ & \lesssim \sum_{j=1}^d \sum_{i=1}^d T_{i,j}^*(\partial_j A)(x), \end{aligned}$$

where the maximal singular integral operator $T_{i,j}^*(f)(x)$ is defined as follows

$$(2.8) \quad T_{i,j}^*(f)(x) = \sup_{\varepsilon > 0} \left| \int_{|x-y|>\varepsilon} \partial_i \mathcal{K}_j(x - y) f(y) dy \right|.$$

One can easily check that the kernel $\partial_i \mathcal{K}_j(x - y)$ is a standard Calderón-Zygmund convolution kernel which satisfies (1.1), (1.3) and has mean value zero on \mathbf{S}^{d-1} . Then by the standard weighted Calderón-Zygmund theory (see [17]), $T_{i,j}^*$ is bounded on $L^p(\mathbb{R}^d, w)$. So $T_{i,j}^*$ maps $L^p(\mathbb{R}^d, w)$ to $L^{p,\infty}(\mathbb{R}^d, w)$.

Finally one may apply the method similar to that of I to handle the reminder term $S_{p,3,r}(\nabla A)(x)$. Indeed, by the Hölder inequality and $|x - y| \approx r$, we get

$$\begin{aligned} S_{p,3,r}(\nabla A)(x) & \lesssim \left[\frac{1}{w(Q(x, 2r))} \int_{Q(x, 2r) \setminus Q(x, r)} r \left(\int_{r \lesssim |x-z|} \frac{|\nabla A(z)|^p}{|x - z|^{d+1}} dz \right) w(y) dy \right]^{\frac{1}{p}} \\ & \lesssim M_p(\nabla A)(x), \end{aligned}$$

where in the last inequality we use that the function $\varepsilon|x|^{-d-1} \chi_{\{|x|>\varepsilon\}}$ is radial non-increasing and L^1 integrable. As showed in the estimate of I , we get that M_p maps $L^p(\mathbb{R}^d, w)$ to $L^{p,\infty}(\mathbb{R}^d, w)$. Hence we complete the proof. \square

Remark 2.7. When giving an estimate in (2.5), we in fact prove that the following inequality

$$(2.9) \quad \frac{|A(x) - A(y)|}{|x - y|} \lesssim M(\nabla A)(x) + M(\nabla A)(y) + \sum_{i=1}^d \sum_{j=1}^d T_{i,j}^*(\partial_j A)(x)$$

holds for almost every $x, y \in \mathbb{R}^d$ if A is a C_c^∞ function, where $T_{i,j}^*$ is defined in (2.8).

Lemma 2.8. *Let $\{Q_k\}_k$ be the disjoint cubes in \mathbb{R}^d . Denote by $l(Q_k)$ the side length of Q_k . Define the operator T_s as*

$$T_s(f)(x) = \sum_k \int_{Q_k} \frac{l(Q_k)^s}{[l(Q_k) + |x - y|]^{d+s}} |f(y)| dy.$$

Suppose that $1 \leq q \leq \infty$, $w \in A_q(\mathbb{R}^d)$ and $f \in L^q(\mathbb{R}^d, w)$. Then for any $s > 0$, we get that

$$\|T_s(f)\|_{L^q(\mathbb{R}^d, w)} \lesssim \|f\|_{L^q(\mathbb{R}^d, w)}.$$

Proof. If $q = 1$, Lemma 2.8 just follows from the property (2.2) of $A_1(\mathbb{R}^d)$ weight and the Fubini theorem. In fact, we have

$$\begin{aligned} \|T_s(f)\|_{L^1(\mathbb{R}^d, w)} &\leq \sum_{Q_k} \int_{Q_k} \left[\int_{\mathbb{R}^d} \frac{w(x) \cdot l(Q_k)^s}{[l(Q_k) + |x - y|]^{d+s}} dx \right] \cdot |f(y)| dy \\ &\lesssim \sum_{Q_k} \int_{Q_k} M(w)(y) \cdot |f(y)| dy \\ &\lesssim [w]_{A_1} \sum_{Q_k} \int_{Q_k} |f(y)| w(y) dy \lesssim [w]_{A_1} \|f\|_{L^1(\mathbb{R}^d, w)}, \end{aligned}$$

where the second inequality follows from that splitting the kernel $\frac{l(Q_k)^s}{[l(Q_k) + |x - y|]^{d+s}}$ into two parts according whether $|x - y| \leq l(Q_k)$ or $|x - y| > l(Q_k)$, the third inequality follows from the property (2.2) and in the last inequality we use that Q_k s are cubes disjoint each other. After we establish T_s is bounded on $L^1(\mathbb{R}^d, w)$ with bound $[w]_{A_1}$, the proof of the case $1 < q < \infty$ just follows from the famous extrapolation theorem (see *e.g.* Theorem 7.5.3 in [17]). If $q = \infty$, apply the Fubini theorem,

$$|T_s(f)(x)| \leq \sum_{Q_k} \|f\|_{L^\infty(Q_k)} \sup_{x \in \mathbb{R}^d} \int_{Q_k} \frac{l(Q_k)^s}{[l(Q_k) + |x - y|]^{d+s}} dy \lesssim \|f\|_{L^\infty(\mathbb{R}^d)}.$$

Then T_s is bounded on $L^\infty(\mathbb{R}^d, w)$ is just a consequence of the chain of inequalities:

$$(2.10) \quad \|T_s(f)\|_{L^\infty(\mathbb{R}^d, w)} \lesssim \|T_s(f)\|_{L^\infty(\mathbb{R}^d)} \lesssim \|f\|_{L^\infty(\mathbb{R}^d)} \lesssim \|f\|_{L^\infty(\mathbb{R}^d, w)},$$

which can be proved as follows. Notice that we have the equivalent definition of $L^\infty(\mathbb{R}^d, \mu)$: $\|f\|_{L^\infty(\mathbb{R}^d, \mu)} = \sup\{\alpha : \mu(\{x \in \mathbb{R}^d : |f(x)| > \alpha\}) > 0\}$. The first inequality in (2.10) follows from the fact that $w(E) > 0$ implies $|E| > 0$. Likewise, the last inequality in (2.10) follows from the fact that $|E| > 0$ implies $w(E) > 0$, because $w(x) = 0$ only for the points in a set of measure zero by the definition of $A_\infty(\mathbb{R}^d)$ weight. Hence we complete the proof. \square

Remark 2.9. By the last argument above, for any $w \in A_\infty(\mathbb{R}^d)$, the follow equality

$$\|f\|_{L^\infty(\mathbb{R}^d)} \approx \|f\|_{L^\infty(\mathbb{R}^d, w)}$$

holds. We will straightforward apply this equivalence many times later.

3. PROOF OF THEOREM 1.1

3.1. Some basic strong type multilinear estimates.

In the following, we begin to give the proof of Theorem 1.1. In this subsection, we will first show our theorem in the case $q_1 = \cdots = q_n = \infty$, $r = p \in [1, \infty)$ which is not quite complicated and the case $d < q_1, \dots, q_n \leq \infty$, $1 < r < \infty$, $p = \infty$.

Proposition 3.1. Let $q_i = \infty$ with $i = 1, \dots, n$, $1 \leq r = p < \infty$. We have the following conclusions:

(i). If $p = r \in (1, \infty)$, $w \in A_p(\mathbb{R}^d)$, then

$$\|\mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f]\|_{L^p(\mathbb{R}^d, w)} \lesssim \left(\prod_{i=1}^n \|\nabla A_i\|_{L^\infty(\mathbb{R}^d, w)} \right) \|f\|_{L^p(\mathbb{R}^d, w)}.$$

(ii). If $p = r = 1$, $w \in A_1(\mathbb{R}^d)$, then

$$\|\mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f]\|_{L^1(\mathbb{R}^d, w)} \lesssim \left(\prod_{i=1}^n \|\nabla A_i\|_{L^\infty(\mathbb{R}^d, w)} \right) \|f\|_{L^1(\mathbb{R}^d, w)}.$$

Proof. The proof of this lemma is quite standard, so we just give some key steps. When $q_1 = \cdots = q_n = \infty$, A_i is a Lipschitz function for $i = 1, \dots, n$. Fix all A_i . Observe that the kernel

$$(3.1) \quad \mathfrak{K}(x, y) := K(x - y) \left(\prod_{i=1}^n \frac{A_i(x) - A_i(y)}{|x - y|} \right)$$

is a standard Calderón-Zygmund kernel satisfying the boundedness condition and regularity condition with bound $\prod_{i=1}^n \|\nabla A_i\|_{L^\infty(\mathbb{R}^d)}$ (see *e.g.* [18, Definition 4.1.2]). Then we may have the following L^2 boundedness

$$(3.2) \quad \|\mathcal{C}[\nabla A_1, \dots, \nabla A_n, f]\|_{L^2(\mathbb{R}^d)} \lesssim \left(\prod_{i=1}^n \|\nabla A_i\|_{L^\infty(\mathbb{R}^d)} \right) \|f\|_{L^2(\mathbb{R}^d)}.$$

This in fact can be seen by using the famous $T1$ theorem (see [18]) or by applying the mean value formula

$$\frac{A_i(x) - A_i(y)}{|x - y|} = \int_0^1 \left\langle \frac{x - y}{|x - y|}, \nabla A_i(sx + (1 - s)y) \right\rangle ds$$

to reduce the operator \mathcal{C} to the following operator introduced by Christ and Journé [8]

$$\mathcal{C}_{CJ}[a_1, \dots, a_n, f](x) = \text{p.v.} \int_{\mathbb{R}^d} k(x - y) \left(\prod_{i=1}^n m_{x,y} a_i \right) f(y) dy$$

which maps $L^\infty(\mathbb{R}^d) \times \cdots \times L^\infty(\mathbb{R}^d) \times L^2(\mathbb{R}^d)$ to $L^2(\mathbb{R}^d)$ (see [8]). Here in the above operator $k(x - y)$ is a standard Calderón-Zygmund kernel and $m_{x,y} a = \int_0^1 a(sx + (1 - s)y) dy$. Then the rest of the proof just follows from the standard weighted Calderón-Zygmund theory (see [17, Theorem 7.4.6.]). \square

Proposition 3.2. Suppose that $1 < r < \infty$, $d < q_1, \dots, q_n \leq \infty$ and $\frac{1}{r} = \sum_{i=1}^n \frac{1}{q_i}$. Let $w \in \bigcap_{i=1}^n A_{q_i/d}(\mathbb{R}^d)$. Assume that $\nabla A_i \in L^{q_i}(\mathbb{R}^d, w)$, $i = 1, \dots, n$ and $f \in L^\infty(\mathbb{R}^d, w)$. Then we get

$$\|\mathcal{C}[\nabla A_1, \dots, \nabla A_n, f]\|_{L^r(\mathbb{R}^d, w)} \lesssim \left(\prod_{i=1}^n \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d, w)} \right) \|f\|_{L^\infty(\mathbb{R}^d, w)}.$$

Proof. By the standard limiting arguments, it is enough to consider that each A_i are C_c^∞ functions and f is bounded compact function. Then one can easily check that $\int_{\mathbb{R}^d} [M_\delta(\mathcal{C}[\nabla A_1, \dots, \nabla A_n, f])(x)]^r w(x) dx$ is finite (for example one may use the method in [22, page 1248] to show this). Therefore using the Fefferman-Stein inequality in Lemma 2.2, we may get that for any $\delta > 0$,

$$\begin{aligned} \|\mathcal{C}[\nabla A_1, \dots, \nabla A_n, f]\|_{L^r(\mathbb{R}^d, w)} &\leq \|M_\delta(\mathcal{C}[\nabla A_1, \dots, \nabla A_n, f])\|_{L^r(\mathbb{R}^d, w)} \\ &\lesssim \|M_\delta^\sharp(\mathcal{C}[\nabla A_1, \dots, \nabla A_n, f])\|_{L^r(\mathbb{R}^d, w)}. \end{aligned}$$

In the following, we need to give an estimate of the maximal sharp function. Fix x and a cube $Q \ni x$. Define $f_1 = f\chi_{3Q}$ and $f_2 = f - f_1$. Then write

$$\mathcal{C}[\nabla A_1, \dots, \nabla A_n, f](z) = \mathcal{C}[\nabla A_1, \dots, \nabla A_n, f_1](z) + \mathcal{C}[\nabla A_1, \dots, \nabla A_n, f_2](z).$$

Choose a constant $c = \mathcal{C}[\nabla A_1, \dots, \nabla A_n, f_2](x)$ in the maximal sharp function. Then we see that this maximal sharp function $M_\delta^\sharp(\mathcal{C}[\nabla A_1, \dots, \nabla A_n, f])(x)$ is bounded by the following two functions

$$\begin{aligned} I(x) + II(x) &:= \sup_{Q \ni x} \left(\frac{1}{|Q|} \int_Q |\mathcal{C}[\nabla A_1, \dots, \nabla A_n, f_1](z)|^\delta dz \right)^{\frac{1}{\delta}} \\ &\quad + \sup_{Q \ni x} \left(\frac{1}{|Q|} \int_Q |\mathcal{C}[\nabla A_1, \dots, \nabla A_n, f_2](z) - \mathcal{C}[\nabla A_1, \dots, \nabla A_n, f_2](x)|^\delta dz \right)^{\frac{1}{\delta}}. \end{aligned}$$

We first consider the above first function $I(x)$. Define $\tilde{A}_i = A_i\chi_{3Q}$. Then for any $z \in Q$, we may write

$$\mathcal{C}[\nabla A_1, \dots, \nabla A_n, f_1](z) = \mathcal{C}[\nabla \tilde{A}_1, \dots, \nabla \tilde{A}_n, f_1](z).$$

Choose $\delta \leq d/n$. Applying the Hölder inequality, strong type multilinear estimate (see [21, Theorem 1.1]) and the definition of \tilde{A}_i , we may get

$$\begin{aligned} \left(\frac{1}{|Q|} \int_Q |\mathcal{C}[\nabla A_1, \dots, \nabla A_n, f_1](z)|^\delta dz \right)^{\frac{1}{\delta}} &\lesssim \|\mathcal{C}[\nabla \tilde{A}_1, \dots, \nabla \tilde{A}_n, f_1]\|_{L^{\frac{d}{\delta}}(Q, \frac{dx}{|Q|})} \\ &\lesssim \|f_1\|_{L^\infty(\mathbb{R}^d)} \prod_{i=1}^n \|\nabla \tilde{A}_i\|_{L^d(\mathbb{R}^d, \frac{dx}{|Q|})} \\ &\lesssim \|f\|_{L^\infty(\mathbb{R}^d)} \prod_{i=1}^d M_d(\nabla A_i)(x), \end{aligned}$$

where M_d is the Hardy-Littlewood maximal operator of order d . Notice that $w \in \bigcap_{i=1}^n A_{q_i/d}(\mathbb{R}^d)$, by using the weighted boundedness of the Hardy-Littlewood maximal

operator, one may easily get that $\|M_d(\nabla A_i)\|_{L^{q_i}(\mathbb{R}^d, w)} \lesssim \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d, w)}$. Therefore

$$\|I\|_{L^r(\mathbb{R}^d, w)} \lesssim \left(\prod_{i=1}^n \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d, w)} \right) \|f\|_{L^\infty(\mathbb{R}^d, w)}.$$

Next we turn to $II(x)$. Write

$$\mathcal{C}[\nabla A_1, \dots, \nabla A_n, f_2](z) - \mathcal{C}[\nabla A_1, \dots, \nabla A_n, f_2](x) = \int_{(3Q)^c} [\mathfrak{K}(z, y) - \mathfrak{K}(x, y)] f(y) dy$$

where $\mathfrak{K}(x, y) := K(x - y) \prod_{i=1}^n \frac{A_i(x) - A_i(y)}{|x - y|}$. Then write

$$\begin{aligned} & \mathfrak{K}(z, y) - \mathfrak{K}(x, y) \\ &= \left(\frac{K(z - y)}{|z - y|^n} - \frac{K(x - y)}{|x - y|^n} \right) \prod_{i=1}^n (A_i(z) - A_i(y)) \\ & \quad + \frac{K(x - y)}{|x - y|^n} \left(\prod_{i=1}^n (A_i(z) - A_i(y)) - \prod_{i=1}^n (A_i(x) - A_i(y)) \right) \\ &=: \mathfrak{K}_1(z, x, y) + \mathfrak{K}_2(z, x, y). \end{aligned}$$

We consider the term $\mathfrak{K}_1(z, x, y)$. Notice that $x, z \in Q$ and $y \in (3Q)^c$, then $|z - y| \approx |x - y|$. By the regularity condition (1.3) and the formula (2.9), we get that

$$|\mathfrak{K}_1(z, x, y)| \lesssim \frac{(l(Q))^\delta}{|x - y|^{d+\delta}} \prod_{i=1}^d [M(\nabla A_i)(z) + T(\nabla A_i)(y)],$$

where here and in the following, T is the sum of combination of the Hardy-Littlewood maximal operator and maximal singular integral $T_{i,j}^*$ defined in (2.8), which both map $L^q(\mathbb{R}^d, w)$ to itself for $1 < q < \infty$.

Next we consider the term $\mathfrak{K}_2(z, x, y)$. We may split $\mathfrak{K}_2(z, x, y)$ into n terms and apply (2.9),

$$\begin{aligned} \mathfrak{K}_2(z, x, y) &\lesssim \frac{1}{|x - y|^{d+n}} \left| \sum_{i=1}^n [A_i(z) - A_i(x)] \prod_{k=1}^{i-1} [A_k(x) - A_i(y)] \prod_{k=i+1}^n [A_k(z) - A_k(y)] \right| \\ &\lesssim \frac{l(Q)}{|x - y|^{d+1}} \prod_{i=1}^n [M(\nabla A_i)(z) + T(\nabla A_i)(x) + T(\nabla A_i)(y)]. \end{aligned}$$

Combining these estimates of \mathfrak{K}_1 and \mathfrak{K}_2 , we get

$$\begin{aligned} & |\mathfrak{K}(z, y) - \mathfrak{K}(x, y)| \\ &\lesssim \frac{(l(Q))^\delta}{|x - y|^{d+\delta}} \prod_{i=1}^n [M(\nabla A_i)(z) + T(\nabla A_i)(x) + T(\nabla A_i)(y)] \\ &\lesssim \frac{(l(Q))^\delta}{|x - y|^{d+\delta}} \sum_{\mathbb{N}_1^n} \prod_{i \in N_1} M(\nabla A_i)(z) \prod_{i \in N_2} T(\nabla A_i)(x) \prod_{i \in N_3} T(\nabla A_i)(y), \end{aligned}$$

where in the last inequality we divide $\mathbb{N}_1^n = N_1 \cup N_2 \cup N_3$ with $\mathbb{N}_1^n = \{1, \dots, n\}$ and N_1, N_2, N_3 non intersecting each other. Plugging the above estimates into $II(x)$ and

applying the Hölder inequality, we get that

$$\begin{aligned}
II(x) &\lesssim \sum_{\mathbb{N}_1^n} \left[\prod_{i \in \mathbb{N}_2} T(\nabla A_i)(x) \right] \frac{1}{|Q|} \int_Q \int_{(3Q)^c} \frac{(l(Q))^\delta}{|x-y|^{d+\delta}} \left[\prod_{i \in \mathbb{N}_3} T(\nabla A_i)(y) \right] f(y) dy \\
&\quad \times \left[\prod_{i \in \mathbb{N}_1} T(\nabla A_i)(z) \right] dz \\
&\lesssim \|f\|_{L^\infty(\mathbb{R}^d)} \sum_{\mathbb{N}_1^n} \left[\prod_{i \in \mathbb{N}_2} T(\nabla A_i)(x) \right] M \left[\prod_{i \in \mathbb{N}_3} T(\nabla A_i) \right](x) \cdot M \left[\prod_{i \in \mathbb{N}_1} T(\nabla A_i) \right](x).
\end{aligned}$$

Now using the Hölder inequality and the fact that M, T are bounded on $L^q(\mathbb{R}^d, w)$ for $1 < q < \infty$, we get

$$\|II\|_{L^r(\mathbb{R}^d, w)} \lesssim \left(\prod_{i=1}^n \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d, w)} \right) \|f\|_{L^\infty(\mathbb{R}^d, w)},$$

which completes the proof. \square

Proposition 3.3. Suppose that $1 < r < \infty$, $d < q_1, \dots, q_n \leq \infty$, $p = \infty$ and $\frac{1}{r} = \sum_{i=1}^n \frac{1}{q_i}$. Let $w \in \bigcap_{i=1}^n A_{q_i/d}(\mathbb{R}^d)$. Then we get

$$\|\mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f]\|_{L^r(\mathbb{R}^d, w)} \lesssim \left(\prod_{i=1}^n \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d, w)} \right) \|f\|_{L^\infty(\mathbb{R}^d)}.$$

Proof. Let φ be a C_c^∞ function which is supported in $\{x \in \mathbb{R}^d : |x| < 1/4\}$, $\varphi(x) = 1$ if $|x| < 1/8$ and $\int_{\mathbb{R}^d} \varphi(x) dx = 1$. Set $\varphi_\varepsilon(x) = \varepsilon^{-d} \varphi(\varepsilon^{-1}x)$. It is easy to see that $\varphi_\varepsilon * \mathcal{C}[\nabla A_1, \dots, \nabla A_n, f](x)$ is bounded by $M(\mathcal{C}[\nabla A_1, \dots, \nabla A_n, f])(x)$. By the weighted boundedness of the Hardy-Littlewood maximal operator M and Proposition 3.2, we may get that

$$\|M(\mathcal{C}[\nabla A_1, \dots, \nabla A_n, f])\|_{L^r(\mathbb{R}^d, w)} \lesssim \left(\prod_{i=1}^n \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d, w)} \right) \|f\|_{L^\infty(\mathbb{R}^d, w)}.$$

So, to complete the proof, it is suffice to show that the following difference

$$\mathcal{C}_\varepsilon[\nabla A_1, \dots, \nabla A_n, f](x) - \varphi_\varepsilon * \mathcal{C}[\nabla A_1, \dots, \nabla A_n, f](x)$$

is controlled uniformly in ε by a function which is bounded from $L^{q_1}(\mathbb{R}^d, w) \times \dots \times L^{q_n}(\mathbb{R}^d, w) \times L^\infty(\mathbb{R}^d, w)$ to $L^r(\mathbb{R}^d, w)$. We write the difference in the above equality as follows

$$\begin{aligned}
&\int_{\mathbb{R}^d} \varphi_\varepsilon(z) \left[\int_{|x-y|>\varepsilon} \left(\mathfrak{K}(x, y) - \mathfrak{K}(x-z, y) \right) f(y) dy \right] dz \\
&\quad + \int_{\mathbb{R}^d} \varphi_\varepsilon(z) \left[\text{p.v.} \int_{|x-y|<\varepsilon} \mathfrak{K}(x-z, y) f(y) dy \right] dz =: P_\varepsilon(x) + Q_\varepsilon(x),
\end{aligned}$$

where $\mathfrak{K}(x, y) := K(x-y) \prod_{i=1}^n \frac{A_i(x) - A_i(y)}{|x-y|}$. Now we first give an estimate of $Q_\varepsilon(x)$. By the Fubini theorem,

$$\begin{aligned}
Q_\varepsilon(x) &= \int_{\mathbb{R}^d} \varphi_\varepsilon(x-z) \left[\text{p.v.} \int_{|x-y|<\varepsilon} \mathfrak{K}(z,y) f(y) dy \right] dz \\
&= \int_{|x-y|<\varepsilon} \left[\text{p.v.} \int_{\mathbb{R}^d} \varphi_\varepsilon(x-z) \mathfrak{K}(z,y) dz \right] f(y) dy.
\end{aligned}$$

Notice that $|x-y| < \varepsilon$ and $|x-z| < \frac{\varepsilon}{4}$. For each $i = 1, \dots, n$, define $\tilde{A}_i(\cdot) = A_i(\cdot) \chi_{\{|\cdot-x|<\varepsilon\}}$. The all A_i in the above inequality can be replaced by \tilde{A}_i . Choose $\frac{1}{\tilde{r}} = (\sum_{i=1}^n \frac{1}{\tilde{q}_i}) + \frac{1}{\tilde{p}}$ such that $1 < \tilde{r} < +\infty$, $1 \leq \tilde{q}_i < q_i < \infty$ for all $i = 1, \dots, n$, $1 < \tilde{p} < \infty$. Then by the multilinear boundedness properties of $\mathcal{C}[\nabla A_1, \dots, \nabla A_n, f](x)$, we may continue to give an estimate of $Q_\varepsilon(x)$ as follows

$$\begin{aligned}
|Q_\varepsilon(x)| &\lesssim \|f\|_{L^\infty(\mathbb{R}^d)} \varepsilon^{\frac{d}{\tilde{r}}} \prod_{i=1}^n \|\nabla \tilde{A}_i\|_{L^{\tilde{q}_i}(\mathbb{R}^d)} \|\varphi_\varepsilon\|_{L^{\tilde{p}}(\mathbb{R}^d)} \\
&\leq \|f\|_{L^\infty(\mathbb{R}^d)} \left(\prod_{i=1}^n \frac{1}{\varepsilon^d} \int_{|x-z|<\varepsilon} |\nabla A_i(z)|^{\tilde{q}_i} dz \right)^{\frac{1}{\tilde{q}_i}} \varepsilon^{\frac{d}{\tilde{r}} + \sum_{i=1}^n \frac{d}{\tilde{q}_i} - d \frac{\tilde{p}-1}{\tilde{p}}} \|\varphi\|_{L^{\tilde{p}}(\mathbb{R}^d)} \\
&\lesssim \|f\|_{L^\infty(\mathbb{R}^d, w)} \prod_{i=1}^n M_{\tilde{q}_i}(\nabla A_i)(x).
\end{aligned}$$

As we have done in the proof of Lemma 2.3, we see that $M_{\tilde{q}_i}$ maps $L^{q_i}(\mathbb{R}^d, w)$ to itself for $w \in A_{q_i/d}(\mathbb{R}^d)$. Then by using the Hölder inequality, we may get that

$$\| \sup_\varepsilon |Q_\varepsilon| \|_{L^r(\mathbb{R}^d, w)} \lesssim \left(\prod_{i=1}^n \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d, w)} \right) \|f\|_{L^\infty(\mathbb{R}^d, w)}.$$

Next we turn to $P_\varepsilon(x)$. Write

$$\begin{aligned}
&\mathfrak{K}(x,y) - \mathfrak{K}(x-z,y) \\
&= \left(\frac{K(x-y)}{|x-y|^n} - \frac{K(x-z-y)}{|x-z-y|^n} \right) \prod_{i=1}^n (A_i(x) - A_i(y)) \\
&\quad + \frac{K(x-z-y)}{|x-z-y|^n} \left[\prod_{i=1}^n (A_i(x) - A_i(y)) - \prod_{i=1}^n (A_i(x-z) - A_i(y)) \right] \\
&=: I + II.
\end{aligned}$$

We consider the term I . Notice that $|x-y| > \varepsilon$ and $|z| < \frac{1}{4}\varepsilon$, then $|x-y| \approx |x-z-y|$. By the regularity condition (1.3) and (2.9), we get that

$$|I| \lesssim \frac{\varepsilon^\delta}{|x-y|^{d+\delta}} \prod_{i=1}^d [M(\nabla A_i)(x) + T(\nabla A_i)(y)].$$

Consider the term II . We may split II into n terms and use (2.9),

$$\begin{aligned} |II| &\lesssim \frac{1}{|x-y|^{d+n}} \left| \sum_{i=1}^n [A_i(x) - A_i(x-z)] \prod_{k=1}^{i-1} [A_k(x-z) - A_k(y)] \prod_{k=i+1}^n [A_k(x) - A_k(y)] \right| \\ &\lesssim \frac{\varepsilon}{|x-y|^{d+1}} \prod_{i=1}^n [M(\nabla A_i)(x) + T(\nabla A_i)(x-z) + T(\nabla A_i)(y)]. \end{aligned}$$

Combining the estimates of I and II , we get

$$\begin{aligned} &|\mathfrak{K}(x,y) - \mathfrak{K}(x-z,y)| \\ &\lesssim \frac{\varepsilon^\delta}{|x-y|^{d+\delta}} \prod_{i=1}^n [M(\nabla A_i)(x) + T(\nabla A_i)(x-z) + T(\nabla A_i)(y)] \\ &\lesssim \frac{\varepsilon^\delta}{|x-y|^{d+\delta}} \sum_{\mathbb{N}_1^n} \prod_{i \in N_1} M(\nabla A_i)(x) \prod_{i \in N_2} T(\nabla A_i)(x-z) \prod_{i \in N_3} T(\nabla A_i)(y), \end{aligned}$$

where in the last inequality we divide $\mathbb{N}_1^n = N_1 \cup N_2 \cup N_3$ with N_1, N_2, N_3 non intersecting each other. Plugging the above estimate into $P_\varepsilon(x)$, we get that

$$\begin{aligned} |P_\varepsilon(x)| &\lesssim \sum_{\mathbb{N}_1^n} \left[\prod_{i \in N_1} M(\nabla A_i)(x) \right] \int_{\mathbb{R}^d} \int_{|x-y|>\varepsilon} \frac{\varepsilon^\delta}{|x-y|^{d+\delta}} \left[\prod_{i \in N_3} T(\nabla A_i)(y) \right] f(y) dy \\ &\quad \times \varphi_\varepsilon(z) \left[\prod_{i \in N_2} T(\nabla A_i)(x-z) \right] dz \\ &\lesssim \|f\|_{L^\infty(\mathbb{R}^d)} \sum_{\mathbb{N}_1^n} \left[\prod_{i \in N_1} M(\nabla A_i)(x) \right] M \left[\prod_{i \in N_2} T(\nabla A_i) \right](x) \cdot M \left[\prod_{i \in N_3} T(\nabla A_i) \right](x). \end{aligned}$$

Now using the Hölder inequality and the fact that M, T are bounded on $L^q(\mathbb{R}^d, w)$ for $1 < q < \infty$, we get that

$$\| \sup_{\varepsilon} |P_\varepsilon| \|_{L^r(\mathbb{R}^d, w)} \lesssim \left(\prod_{i=1}^n \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d, w)} \right) \|f\|_{L^\infty(\mathbb{R}^d, w)},$$

which completes the proof. \square

3.2. Case: all q_i s are larger than d .

In this subsection, we consider the case $d/(d+n) \leq r < \infty$ and $d \leq q_1, \dots, q_n \leq \infty$. Without loss of generality, we assume that the first $q_1, \dots, q_l > d$ and $q_{l+1}, \dots, q_n = d$ with $0 \leq l \leq n$. Here and in the following, when $l = 0$, we mean all $q_1 = \dots = q_n = d$. The proof of the case $p = \infty$ is slight different from that of $1 \leq p < \infty$. So we shall give two propositions below. Let us see the case $1 \leq p < \infty$ firstly and we emphasize in the proof where it doesn't work for $p = \infty$.

Proposition 3.4. Let $\frac{1}{r} = \left(\sum_{i=1}^n \frac{1}{q_i} \right) + \frac{1}{p}$, $\frac{d}{d+n} \leq r < \infty$, $d < q_1, \dots, q_l \leq \infty$ and $q_{l+1} = \dots = q_n = d$ with $0 \leq l \leq n$, $1 \leq p < \infty$. Suppose that $w \in \left(\bigcap_{i=1}^n A_{\max\{\frac{q_i}{d}, 1\}}(\mathbb{R}^d) \right) \cap$

$A_p(\mathbb{R}^d)$. Then

$$(3.3) \quad \begin{aligned} & \|\mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f]\|_{L^{r,\infty}(\mathbb{R}^d, w)} \\ & \lesssim \left(\prod_{i=1}^l \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d, w)} \right) \left(\prod_{i=l+1}^n \|\nabla A_i\|_{L^{d,1}(\mathbb{R}^d, w)} \right) \|f\|_{L^p(\mathbb{R}^d, w)}, \end{aligned}$$

where $L^{d,1}(\mathbb{R}^d, w)$ is the weighted Lorentz space.

Proof. By the dense limiting argument and scaling argument, it is sufficient to prove that when A_i ($i = 1, \dots, n$) and f are C^∞ functions with compact supports,

$$\|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d, w)} = \|\nabla A_j\|_{L^{d,1}(\mathbb{R}^d, w)} = \|f\|_{L^p(\mathbb{R}^d, w)} = 1,$$

for $i = 1, \dots, l$ and $j = l+1, \dots, n$, the following inequality

$$w(\{x \in \mathbb{R}^d : \mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f](x) > \lambda\}) \lesssim \lambda^{-r}$$

holds for any $\lambda > 0$. Fix $\lambda > 0$. For convenience we set

$$(3.4) \quad E_\lambda = \{x \in \mathbb{R}^d : \mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f](x) > \lambda\}.$$

Our goal is to show $w(E_\lambda) \lesssim \lambda^{-r}$. First assume that all $q_1, \dots, q_l < \infty$. Once the proof in this situation is well understood, we can modify the proof to the other case that there exist some $q_i = \infty$ for $i = 1, \dots, l$. We shall show how to do this in the last part of the proof. Define the *exceptional set*

$$J_{i,\lambda} = \{x \in \mathbb{R}^d : \mathcal{M}(\nabla A_i)(x) > \lambda^{\frac{r}{q_i}}\}.$$

for $i = 1, \dots, n$. Here it should be pointed out that the above definition is meaningless if $q_i = \infty$. Therefore we need to assume all $q_i < \infty$ firstly. By Lemma 2.3 and Lemma 2.4, \mathcal{M} maps $L^p(\mathbb{R}^d, w)$ to itself for $p > d$ and maps $L^{d,1}(\mathbb{R}^d, w)$ to $L^{d,\infty}(\mathbb{R}^d, w)$, i.e.

$$(3.5) \quad \begin{aligned} & w(J_{i,\lambda}) \lesssim \lambda^{-r} \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d, w)}^{q_i} = \lambda^{-r}, \quad i = 1, \dots, l; \\ & w(J_{j,\lambda}) \lesssim \lambda^{-r} \|\nabla A_j\|_{L^{d,1}(\mathbb{R}^d, w)}^d = \lambda^{-r}, \quad j = l+1, \dots, n. \end{aligned}$$

Set $J_\lambda = \cup_{i=1}^n J_{i,\lambda}$. Since $w(x)dx$ satisfies the doubling property, we may choose an open set G_λ which satisfies the following conditions: (1) $J_\lambda \subset G_\lambda$; (2) $w(G_\lambda) \lesssim w(J_\lambda)$. By the property (3.5) of $J_{i,\lambda}$, we see that $w(G_\lambda) \lesssim \lambda^{-r}$. Next making a Whitney decomposition of G_λ (see e.g. [17]), we may obtain a family of disjoint dyadic cubes $\{Q_k\}_k$ such that

- (i). $G_\lambda = \bigcup_{k=1}^\infty Q_k$;
- (ii). $\sqrt{d} \cdot l(Q_k) \leq \text{dist}(Q_k, (G_\lambda)^c) \leq 4\sqrt{d} \cdot l(Q_k)$.

With those properties (i) and (ii), for each Q_k , we may construct a larger cube Q_k^* so that $Q_k \subset Q_k^*$, Q_k^* is centered at y_k and $y_k \in (G_\lambda)^c$, $l(Q_k^*) \approx l(Q_k)$. By the property (ii) above, the distance between Q_k and $(G_\lambda)^c$ equals to $Cl(Q_k)$. Therefore by the construction of Q_k^* and y_k , one may get

$$(3.6) \quad \text{dist}(y_k, Q_k) \approx l(Q_k), \quad w(Q_k^*) \approx w(Q_k).$$

Now we come back to give an estimate of $w(E_\lambda)$. Split f into two parts $f = f_1 + f_2$ where $f_1(x) = f(x)\chi_{(G_\lambda)^c}(x)$ and $f_2(x) = f(x)\chi_{G_\lambda}(x)$. By the definition of J_λ , when restricted on $(G_\lambda)^c$, A_i is a Lipschitz function with $\|\nabla A_i\|_{L^\infty((G_\lambda)^c)} \leq \lambda^{\frac{r}{q_i}}$ for $i =$

$1, \dots, n$. Let \tilde{A}_i represent the Lipschitz extension of A_i from $(G_\lambda)^c$ to \mathbb{R}^d (see [26, page 174, Theorem 3]) so that for each $i = 1, \dots, n$,

$$\tilde{A}_i(y) = A_i(y) \quad \text{if } y \in (G_\lambda)^c;$$

$$|\tilde{A}_i(x) - \tilde{A}_i(y)| \leq \lambda^{\frac{r}{q_i}} |x - y| \quad \text{for all } x, y \in \mathbb{R}^d.$$

Since the operator $\mathcal{C}_*[\dots, \cdot]$ is sub-multilinear, we split E_λ as three terms and give estimates as follows:

$$\begin{aligned} & w(\{x \in \mathbb{R}^d : \mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f](x) > \lambda\}) \\ (3.7) \quad & \leq w(10G_\lambda) + w(\{x \in (10G_\lambda)^c : \mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f_1](x) > \lambda/2\}) \\ & \quad + w(\{x \in (10G_\lambda)^c : \mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f_2](x) > \lambda/2\}). \end{aligned}$$

The above first term satisfies $w(10G_\lambda) \lesssim \lambda^{-r}$, which is our required estimate. In the following, we only consider the second terms. Notice that we only need to consider $x \in (10G_\lambda)^c$. By the definition of f_1 , it is not difficulty to see that

$$\mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f_1](x) = \mathcal{C}_*[\nabla \tilde{A}_1, \dots, \nabla \tilde{A}_n, f_1](x).$$

With this equality in hand, Proposition 3.1 ($1 \leq p < \infty$) implies

$$\begin{aligned} & w(\{x \in (10G_\lambda)^c : \mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f_1](x) > \lambda/2\}) \\ (3.8) \quad & = w(\{x \in (10G_\lambda)^c : \mathcal{C}_*[\nabla \tilde{A}_1, \dots, \nabla \tilde{A}_n, f_1](x) > \lambda/2\}) \\ & \lesssim \lambda^{-p} \left(\prod_{i=1}^n \|\nabla \tilde{A}_i\|_{L^\infty(\mathbb{R}^d, w)}^p \right) \|f_1\|_{L^p(\mathbb{R}^d, w)}^p \lesssim \lambda^{-p+p \sum_{i=1}^n \frac{r}{q_i}} = \lambda^{-r}. \end{aligned}$$

If $p = \infty$, the above method does not work. We will show how to prove this kind of estimate in the next proposition.

Let us turn to $\mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f_2](x)$. Recall $\mathbb{N}_i^j = \{i, i+1, \dots, j\}$ and our construction of G_λ , y_k , Q_k and Q_k^* above (3.6). Then by the property (i) of $\{Q_k\}_k$, we may write $f_2 = \sum_k f \chi_{Q_k}$. Therefore we may get

$$\mathcal{C}_\varepsilon[\nabla A_1, \dots, \nabla A_n, f_2](x) = \sum_k \mathcal{C}_\varepsilon[\nabla A_1, \dots, \nabla A_n, f \chi_{Q_k}](x).$$

In the following we need to study carefully $\prod_{i=1}^n \frac{A_i(x) - A_i(y)}{|x-y|}$. We will separate it into several terms and then give an estimate for each term. Write

$$\begin{aligned} & \prod_{i=1}^n \frac{A_i(x) - A_i(y)}{|x-y|} \\ & = \prod_{i=1}^n \left(\frac{\tilde{A}_i(x) - \tilde{A}_i(y)}{|x-y|} + \frac{\tilde{A}_i(y) - \tilde{A}_i(y_k)}{|x-y|} + \frac{A_i(y_k) - A_i(y)}{|x-y|} \right) \\ & = \sum \left(\prod_{i \in N_1} \frac{\tilde{A}_i(x) - \tilde{A}_i(y)}{|x-y|} \right) \left(\prod_{i \in N_2} \frac{\tilde{A}_i(y) - \tilde{A}_i(y_k)}{|x-y|} \right) \left(\prod_{i \in N_3} \frac{A_i(y_k) - A_i(y)}{|x-y|} \right) \\ & = I(x, y) + II(x, y, y_k), \end{aligned}$$

where in the third equality we divide $\mathbb{N}_1^n = N_1 \cup N_2 \cup N_3$ with N_1, N_2, N_3 non intersecting each other; and $I(x, y), II(x, y, y_k)$, are defined as follows

$$(3.9) \quad \begin{aligned} I(x, y) &= \prod_{i=1}^n \frac{\tilde{A}_i(x) - \tilde{A}_i(y)}{|x - y|}, \\ II(x, y, y_k) &= \sum_{N_1 \subsetneq \mathbb{N}_1^n} \left[\prod_{i \in N_1} \frac{\tilde{A}_i(x) - \tilde{A}_i(y)}{|x - y|} \right] \\ &\quad \times \left[\prod_{i \in N_2} \frac{\tilde{A}_i(y) - \tilde{A}_i(y_k)}{|x - y|} \right] \left[\prod_{i \in N_3} \frac{A_i(y_k) - A_i(y)}{|x - y|} \right]. \end{aligned}$$

By the above decomposition, we in fact write $\mathcal{C}_\varepsilon[\nabla A_1, \dots, \nabla A_n, f\chi_{Q_k}](x)$ into 3^n terms and separate these terms into two parts according I and II .

Weighted estimate of $\mathcal{C}_[\dots, \cdot]$ related to I .* This estimate is similar to (3.8). In fact, in this case there is only one term $\mathcal{C}_*[\nabla \tilde{A}_1, \dots, \nabla \tilde{A}_n, f_2]$. Then by Proposition 3.1 ($1 \leq p < \infty$), we get

$$\begin{aligned} &w(\{x \in (10G_\lambda)^c : \mathcal{C}_*[\nabla \tilde{A}_1, \dots, \nabla \tilde{A}_n, f_2](x) > \lambda/2\}) \\ &\lesssim \lambda^{-p} \left(\prod_{i=1}^n \|\nabla \tilde{A}_i\|_{L^\infty(\mathbb{R}^d, w)}^p \right) \|f_2\|_{L^p(\mathbb{R}^d, w)}^p \lesssim \lambda^{-p+p \sum_{i=1}^n \frac{r}{q_i}} = \lambda^{-r}. \end{aligned}$$

If $p = \infty$, the above argument may not work again.

Weighted estimate of $\mathcal{C}_[\dots, \cdot]$ related to II .* It is sufficient to consider one term $\mathcal{C}_*[\dots, \cdot]$ related to II in which N_1 is a proper subset of \mathbb{N}_1^n . In such a case, without loss of generality, we may suppose that $N_1 = \{1, \dots, v\}$, $N_2 = \{v+1, \dots, m\}$ and $N_3 = \{m+1, \dots, n\}$ with $0 \leq v \leq m \leq n$ and $v < n$. Here if $v = 0$, it means that $N_1 = \emptyset$; if $v = m$, $N_2 = \emptyset$; if $m = n$, $N_3 = \emptyset$. With these notation, it is easy to see that N_1 is a proper subset of \mathbb{N}_1^n . By a slight abuse of notation, we still utilize $II(x, y, y_k)$ to represent one term related to N_1, N_2 and N_3 in (3.9) and utilize $H_{II}(x)$ to represent $\mathcal{C}_*[\dots, \cdot]$ related to $II(x, y, y_k)$, i.e.

$$H_{II}(x) = \sup_{\varepsilon > 0} \left| \sum_k \int_{|x-y| > \varepsilon} K(x-y) II(x, y, y_k) f(y) \chi_{Q_k}(y) dy \right|.$$

Notice that $y_k \in (G_\lambda)^c$, thus $y_k \in (J_{i,\lambda})^c$. Therefore we obtain that

$$\mathcal{M}(\nabla A_i)(y_k) \leq \lambda^{\frac{r}{q_i}}, \text{ for } i = m+1, \dots, n.$$

With the above fact and \tilde{A}_i is a Lipschitz function with bound λ^{r/q_i} for $i = 1, \dots, m$, we get

$$\begin{aligned} |II(x, y, y_k)| &\lesssim \lambda^{\sum_{i=1}^m \frac{r}{q_i}} \frac{|y - y_k|^{n-v}}{|x - y|^{n-v}} \prod_{i=m+1}^n \mathcal{M}(\nabla A_i)(y_k) \\ &\lesssim \lambda^{\sum_{i=1}^n \frac{r}{q_i}} \frac{|y - y_k|^{n-v}}{|x - y|^{n-v}}. \end{aligned}$$

Since it is sufficient to consider $x \in (10G_\lambda)^c$, then for $y \in Q_k$, $|x-y| \geq 2l(Q_k) \approx |y-y_k|$ by (3.6). Combining the above discussion with (1.1), we obtain

$$(3.10) \quad \begin{aligned} H_{II}(x) &\leq \sum_k \int_{Q_k} |K(x-y)| \cdot |II(x,y,y_k)| \cdot |f(y)| dy \\ &\lesssim \lambda^{\sum_{i=1}^n \frac{r}{q_i}} \sum_k \int_{Q_k} \frac{l(Q_k)^{n-v}}{[l(Q_k) + |x-y|]^{d+n-v}} |f(y)| dy = \lambda^{\sum_{i=1}^n \frac{r}{q_i}} T_{n-v} f(x) \end{aligned}$$

where T_{n-v} is defined in Lemma 2.8. Applying the Chebyshev inequality with the above estimate, and utilizing Lemma 2.8 (notice that $n-v \geq 1$ because N_1 is a proper set of \mathbb{N}_1^n), we finally get

$$w(\{x \in (10G_\lambda)^c : H_{II}(x) > \lambda\}) \leq \lambda^{-p + \sum_{i=1}^n \frac{rp}{q_i}} \|T_{n-v} f\|_{L^p(\mathbb{R}^d, w)}^p \lesssim \lambda^{-r} \|f\|_{L^p(\mathbb{R}^d, w)}^p.$$

Hence we finish the proof of the term II . If $p = \infty$, the above last argument may not work and some different discussion should be involved, see the proof in the next proposition.

Finally, we show how to modify our proof here to the case $q_i = \infty$ for some $i = 1, \dots, l$. We may assume that only $q_1 = \dots = q_u = \infty$ with $1 \leq u \leq l$. Thus A_1, \dots, A_u are Lipschitz functions which in fact are nice functions. Then we just fix A_1, \dots, A_u in the rest of the proof. We only make a construction of *exceptional set* for A_{u+1}, \dots, A_n and study $\prod_{i=u+1}^n \frac{A_i(x) - A_i(y)}{|x-y|}$ by using the same way as we have done previously. After that utilizing A_1, \dots, A_u are Lipschitz functions to deal with all estimates involved with A_1, \dots, A_u , we could obtain our required bound. \square

Proposition 3.5. Let $\frac{1}{r} = (\sum_{i=1}^n \frac{1}{q_i}) + \frac{1}{p}$, $\frac{d}{d+n} \leq r < \infty$, $d < q_1, \dots, q_l \leq \infty$ and $q_{l+1}, \dots, q_n = d$ with $0 \leq l \leq n$, $p = \infty$. Suppose that $w \in \bigcap_{i=1}^n A_{\max\{\frac{q_i}{d}, 1\}}(\mathbb{R}^d)$. Then

$$\begin{aligned} &\|\mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f]\|_{L^{r,\infty}(\mathbb{R}^d, w)} \\ &\lesssim \left(\prod_{i=1}^l \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d, w)} \right) \left(\prod_{i=l+1}^n \|\nabla A_i\|_{L^{d,1}(\mathbb{R}^d, w)} \right) \|f\|_{L^\infty(\mathbb{R}^d, w)}, \end{aligned}$$

where $L^{d,1}(\mathbb{R}^d, w)$ is the weighted Lorentz space.

Proof. The proof here is similar to that of Proposition 3.4. So we shall be brief and only indicate necessary modifications here. Proceeding the proof in Proposition 3.4, there are four different arguments.

The first one is that when we choose the set E_λ , we set

$$E_\lambda = \{x \in \mathbb{R}^d : \mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f](x) > C_0 \lambda\},$$

where C_0 is a constant determined later. Our goal is to show $w(E_\lambda) \lesssim \lambda^{-r}$. We split E_λ as several terms and give estimates as follows:

$$(3.11) \quad \begin{aligned} &w(\{x \in \mathbb{R}^d : \mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f](x) > C_0 \lambda\}) \\ &\leq w(10G_\lambda) + w(\{x \in (10G_\lambda)^c : \mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f_1](x) > C_0 \lambda/2\}) \\ &\quad + w(\{x \in (10G_\lambda)^c : \mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f_2](x) > C_0 \lambda/2\}). \end{aligned}$$

The first term above satisfies $w(10G_\lambda) \lesssim \lambda^{-r}$, so it is sufficient to consider the second and third terms. Thus we only need to consider $x \in (10G_\lambda)^c$.

The second difference is the estimate related to the second term in (3.11). Here we choose $\tilde{r}, \tilde{q}_1, \dots, \tilde{q}_n$, such that $1 < \tilde{r} < \infty$, $q_1 < \tilde{q}_1 < \infty$, \dots , $q_n < \tilde{q}_n < \infty$, $d < \tilde{q}_1, \dots, \tilde{q}_n$ and $\frac{1}{\tilde{r}} = \sum_{i=1}^n \frac{1}{\tilde{q}_i}$. Utilize Proposition 3.3 with those above $\tilde{r}, \tilde{q}_1, \dots, \tilde{q}_n$ and A_i is a Lipschitz function on $(G_\lambda)^c$ with Lipschitz bound $\lambda^{\frac{r}{q_i}}$ for $i = 1, \dots, n$, we may obtain

$$\begin{aligned} & w(\{x \in (10G_\lambda)^c : \mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f_1](x) > C_0\lambda/2\}) \\ & \leq w(\{x \in (G_\lambda)^c : \mathcal{C}_*[\nabla(A_1\chi_{(G_\lambda)^c}), \dots, \nabla(A_n\chi_{(G_\lambda)^c}), f_1](x) > C_0\lambda/2\}) \\ & \lesssim \lambda^{-\tilde{r}} \left(\prod_{i=1}^n \|\nabla(A_i\chi_{(G_\lambda)^c})\|_{L^{\tilde{q}_i}(\mathbb{R}^d, w)} \right) \|f_1\|_{L^\infty(\mathbb{R}^d, w)} \\ & \lesssim \lambda^{-\tilde{r}} \left(\prod_{i=1}^n \|\nabla A_i\|_{L^\infty((G_\lambda)^c)}^{(\tilde{q}_i - q_i) \frac{\tilde{r}}{q_i}} \right) \left(\prod_{i=1}^n \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d, w)}^{\frac{q_i}{\tilde{q}_i} \tilde{r}} \right) \|f_1\|_{L^\infty(\mathbb{R}^d, w)} \\ & \lesssim \lambda^{-\tilde{r} + \tilde{r} \left(\sum_{i=1}^n \frac{r}{q_i} \right) - r \left(\sum_{i=1}^n \frac{\tilde{r}}{q_i} \right)} = \lambda^{-r}. \end{aligned}$$

Next consider the estimate related to the third term in (3.11). As done in the proof of Proposition 3.4, we divide $\mathcal{C}_\varepsilon[\nabla A_1, \dots, \nabla A_n, f_2](x)$ into several terms and then separate these terms into two parts according I and II in (3.9). So we get

$$\begin{aligned} & w(\{x \in (10G_\lambda)^c : \mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f_2](x) > C_0\lambda/2\}) \\ & \leq w(\{x \in (10G_\lambda)^c : \mathcal{C}_*[\nabla \tilde{A}_1, \dots, \nabla \tilde{A}_n, f_2](x) > C_0\lambda/4\}) \\ & \quad + w(\{x \in (10G_\lambda)^c : H_{II}(x) > C_0\lambda/4\}). \end{aligned}$$

The third difference is the *weighted estimate of $\mathcal{C}_*[\dots, \cdot]$ related to I* . Here we utilize Lemma 3.1 and the estimate $\|f_2\|_{L^1(\mathbb{R}^d, w)} \lesssim \|f\|_{L^\infty(\mathbb{R}^d, w)} w(G_\lambda) \lesssim \lambda^{-r}$ to get

$$\begin{aligned} & w(\{x \in (10G_\lambda)^c : \mathcal{C}_*[\nabla \tilde{A}_1, \dots, \nabla \tilde{A}_n, f_2](x) > C_0\lambda/4\}) \\ & \lesssim \lambda^{-1} \left(\prod_{i=1}^n \|\nabla \tilde{A}_i\|_{L^\infty(\mathbb{R}^d)} \right) \|f_2\|_{L^1(\mathbb{R}^d, w)} \lesssim \lambda^{-1 + \left(\sum_{i=1}^n \frac{r}{q_i} \right) - r} = \lambda^{-r}. \end{aligned}$$

The fourth difference is the *weighted estimate of $\mathcal{C}_*[\dots, \cdot]$ related to II* . We shall prove that

$$(3.12) \quad \{x \in (10G_\lambda)^c : H_{II}(x) > C_0\lambda/4\} = \emptyset.$$

In fact, by (3.10) and Lemma 2.8 with $q = \infty$, we get for any $x \in (10G_\lambda)^c$,

$$H_{II}(x) \leq C_d \lambda^{\sum_{i=1}^n \frac{r}{q_i}} \|f\|_{L^\infty(\mathbb{R}^d, w)} = C_d \lambda.$$

If we choose $C_0 > 4C_d$, we get (3.12). So we complete the proof. \square

3.3. Case: some q_i s are smaller than d and some are not.

In this subsection, we consider the case: $d/(d+n) \leq r < \infty$ with at least one $q_i < d$, $1 \leq p \leq \infty$. By our condition, the weight w satisfies $w \in \left(\bigcap_{i=1}^n A_{\max\{\frac{q_i}{d}, 1\}}(\mathbb{R}^d)\right) \cap A_p(\mathbb{R}^d) = A_1(\mathbb{R}^d)$. Without loss of generality, we may suppose that $d \leq q_1, \dots, q_l \leq \infty$ and $1 \leq q_{l+1}, \dots, q_n < d$ with $0 \leq l < n$. If $l = 0$, it means that all $q_1, \dots, q_n \in [1, d]$. Also we suppose that $q_1 = \dots = q_k = d$ and $d < q_{k+1}, \dots, q_l \leq \infty$ with $0 \leq k \leq l$. If $k = 0$, we mean that there is no index in q_1, \dots, q_l equals to d , i.e. $d < q_1, \dots, q_l \leq \infty$; if $k = l$, we mean that $q_1 = \dots = q_l = d$. Since the proof of $p = \infty$ is a little different from that of $1 \leq p < \infty$, we shall give two propositions.

Proposition 3.6. Suppose $w \in A_1(\mathbb{R}^d)$. Let $\frac{1}{r} = \left(\sum_{i=1}^n \frac{1}{q_i}\right) + \frac{1}{p}$, $\frac{d}{d+n} \leq r < \infty$, $q_1 = \dots = q_k = d$, $d < q_{k+1}, \dots, q_l \leq \infty$ and $1 \leq q_{l+1}, \dots, q_n < d$ with $0 \leq k \leq l$ and $0 \leq l < n$, $1 \leq p < \infty$. Then

$$\begin{aligned} & \|\mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f]\|_{L^{r, \infty}(\mathbb{R}^d, w)} \\ & \lesssim \left(\prod_{i=1}^k \|\nabla A_i\|_{L^{d, 1}(\mathbb{R}^d, w)}\right) \left(\prod_{i=k+1}^n \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d, w)}\right) \|f\|_{L^p(\mathbb{R}^d, w)}, \end{aligned}$$

where $L^{d, 1}(\mathbb{R}^d, w)$ is the weighted Lorentz space.

Proof. We need to prove that for any $\lambda > 0$, the following inequality holds

$$\begin{aligned} & w(\{x \in \mathbb{R}^d : \mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f](x) > \lambda\}) \\ (3.13) \quad & \lesssim \lambda^{-r} \left(\prod_{i=1}^k \|\nabla A_i\|_{L^{d, 1}(\mathbb{R}^d, w)}^r\right) \left(\prod_{i=k+1}^n \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d, w)}^r\right) \|f\|_{L^p(\mathbb{R}^d, w)}^r. \end{aligned}$$

By the standard dense and scaling argument, it is sufficient to consider that each A_i ($i = 1, \dots, n$) and f as smooth functions with compact supports and

$$\begin{aligned} & \|\nabla A_1\|_{L^{d, 1}(\mathbb{R}^d, w)} = \dots = \|\nabla A_k\|_{L^{d, 1}(\mathbb{R}^d, w)} = 1; \\ & \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d, w)} = 1 \quad \text{for } i = k+1, \dots, n, \text{ and } \|f\|_{L^p(\mathbb{R}^d, w)} = 1. \end{aligned}$$

As done in the proof of Proposition 3.4, we first suppose that all $q_{k+1}, \dots, q_l < \infty$ since the other case is easy and we will show lastly how to modify the proof to the case that there exist $q_i = \infty$ for some $i = k+1, \dots, l$. Fix $\lambda > 0$ and set

$$E_\lambda = \{x \in \mathbb{R}^d : \mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f](x) > \lambda\}.$$

Our goal is to show $w(E_\lambda) \lesssim \lambda^{-r}$. The main idea is to construct some *exceptional set* such that the w measure of *exceptional set* is bounded by λ^{-r} , which is our required estimate. At the same time on the complementary set of *exceptional set* these functions A_i s should be Lipschitz functions with bound $\lambda^{\frac{r}{q_i}}$ for each $i = 1, \dots, n$. The constructions of *exceptional sets* are different between $d \leq q_i < \infty$ and $1 \leq q_i < d$. Now we begin our constructions of some *exceptional sets*.

Step 1: Exceptional set related to q_1, \dots, q_l . Define the *exceptional set* for $i = 1, \dots, l$

$$J_{i, \lambda} = \{x \in \mathbb{R}^d : \mathcal{M}(\nabla A_i)(x) > \lambda^{\frac{r}{q_i}}\}; \quad J_\lambda = \bigcup_{i=1}^l J_{i, \lambda}.$$

Since $w \in A_1(\mathbb{R}^d)$, by Lemma 2.3 and Lemma 2.4, \mathcal{M} maps $L^p(\mathbb{R}^d, w)$ to itself for $p > d$ and maps $L^{d,1}(\mathbb{R}^d, w)$ to $L^{d,\infty}(\mathbb{R}^d, w)$, i.e.

$$(3.14) \quad \begin{aligned} w(J_{i,\lambda}) &\lesssim \lambda^{-r} \|\nabla A_i\|_{L^{d,1}(\mathbb{R}^d, w)}^d = \lambda^{-r}, \quad i = 1, \dots, k; \\ w(J_{i,\lambda}) &\lesssim \lambda^{-r} \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d, w)}^{q_i} = \lambda^{-r}, \quad i = k+1, \dots, l. \end{aligned}$$

So we obtain that $w(J_\lambda) \lesssim \lambda^{-r}$.

Step 2: Calderón-Zygmund decomposition. By the formula given in [26, page 125, (17)], for each A_i , $i = l+1, \dots, n$, we may write

$$A_i(x) = \sum_{j=1}^d C_d \int_{\mathbb{R}^d} \frac{x_j - y_j}{|x - y|^d} \partial_j A_i(y) dy =: \sum_{j=1}^d A_{i,j}(x).$$

Notice that $w \in A_1(\mathbb{R}^d)$ in this case. For each $|\partial_j A_i|^{q_i} \in L^1(\mathbb{R}^d, w)$ with $j = 1, \dots, d$ and $i = l+1, \dots, n$, making a Calderón-Zygmund decomposition at level λ^r , we may have the following conclusions (see e.g. [15, page 413, Theorem 3.5]):

- (cz-i) $\partial_j A_i = g_{j,i} + b_{j,i}$, $\|g_{j,i}\|_{L^\infty(\mathbb{R}^d)} \lesssim \lambda^{\frac{r}{q_i}}$, $\|g_{j,i}\|_{L^{q_i}(\mathbb{R}^d, w)} \lesssim \|\partial_j A_i\|_{L^{q_i}(\mathbb{R}^d, w)}$;
- (cz-ii) $b_{j,i} = \sum_{Q \in \mathcal{Q}_{j,i}} b_{j,i,Q}$, $\text{supp } b_{j,i,Q} \subset Q$, where $\mathcal{Q}_{j,i}$ is a countable set of disjoint dyadic cubes;
- (cz-iii) Let $E_{j,i} = \bigcup_{Q \in \mathcal{Q}_{j,i}} Q$, then $w(E_{j,i}) \lesssim \lambda^{-r} \|\partial_j A_i\|_{L^{q_i}(\mathbb{R}^d, w)}^{q_i}$;
- (cz-iv) $\int b_{j,i,Q}(y) dy = 0$ for each $Q \in \mathcal{Q}_{j,i}$, the unweighted estimate $\|b_{j,i,Q}\|_{L^{q_i}(\mathbb{R}^d)} \lesssim \lambda^r |Q|$ and the weighted estimate $\|b_{j,i}\|_{L^{q_i}(\mathbb{R}^d, w)} \lesssim \|\partial_j A_i\|_{L^{q_i}(\mathbb{R}^d, w)}$ hold.

We shall split $A_{i,j}$ into two parts according the above Calderón-Zygmund decomposition (cz-i):

$$\begin{aligned} A_{i,j}^g(x) &= C_d \int_{\mathbb{R}^d} \frac{x_j - y_j}{|x - y|^d} g_{j,i}(y) dy; \\ A_{i,j}^b(x) &= C_d \int_{\mathbb{R}^d} \frac{x_j - y_j}{|x - y|^d} b_{j,i}(y) dy. \end{aligned}$$

Define the *exceptional set* $B_\lambda = \bigcup_{i=l+1}^n \bigcup_{j=1}^d E_{j,i}$. Then by (cz-iii), we obtain $w(B_\lambda) \lesssim \lambda^{-r}$.

Step 3: Exceptional set D_λ . Set $\frac{1}{s_i} = \frac{1}{q_i} - \frac{1}{d}$ for $i = l+1, \dots, n$. Notice that w belongs to $A_1(\mathbb{R}^d)$. Define the following *exceptional set*

$$D_{i,\lambda} = \left\{ x \in \mathbb{R}^d : \mathfrak{M}_{w, s_i}(\nabla A_i)(x) > \lambda^{\frac{r}{q_i}} \right\}$$

where the maximal operator \mathfrak{M}_{w, s_i} is defined in the paragraph above Lemma 2.5. Denote $D_\lambda = \bigcup_{i=l+1}^n D_{i,\lambda}$. Then by Lemma 2.6, we get that

$$w(D_{i,\lambda}) \lesssim \lambda^{-r} \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d, w)}^{q_i} = \lambda^{-r}; \quad w(D_\lambda) \lesssim \lambda^{-r}.$$

Step 4: Exceptional set F_λ . For each $j = 1, \dots, d$, $i = l+1, \dots, n$, define the functions

$$\Delta^{j,i}(x) = \sum_{Q \in \mathcal{Q}_{j,i}} \frac{l(Q)}{[l(Q) + |x - y_Q|]^{d+1}} m(Q)$$

where y_Q is the center of Q . Define another *exceptional set*

$$F_{j,i,\lambda} = \{x \in \mathbb{R}^d : \Delta^{j,i}(x) > 1\}, \quad F_\lambda = \cup_{j=1}^d \cup_{i=l+1}^n F_{j,i,\lambda}.$$

Notice that we have $w \in A_1(\mathbb{R}^d)$. We claim the following property of A_1 weight: For any cube Q and $\alpha > 1$, there exists a constant C independent of Q and α such that

$$(3.15) \quad w(\alpha Q) \leq C\alpha^d w(Q).$$

In fact by (2.2) in Definition 2.1, there exist a constant C independent of α and Q such that

$$\frac{1}{|\alpha Q|} \int_{\alpha Q} w(z) dz \leq C \operatorname{ess\,inf}_{y \in Q} w(y) \leq C \frac{1}{|Q|} \int_Q w(y) dy,$$

which immediately implies (3.15). In the following, by using the Chebyshev inequality in the first inequality, (3.15) with $\alpha = 2^k$ in the last second inequality and (cz-iii) in the last inequality, we get

$$\begin{aligned} w(F_{j,i,\lambda}) &\leq \int_{\mathbb{R}^d} \Delta_{j,i,\lambda}(x) w(x) dx \lesssim \sum_{Q \in \mathcal{Q}_{j,i}} \left[\int_{\mathbb{R}^d} \frac{l(Q)^{d+1}}{[l(Q) + |x - y_Q|]^{d+1}} w(x) dx \right] \\ &\leq \sum_{Q \in \mathcal{Q}_{j,i}} \left[\int_Q w(x) dx + \sum_{k=1}^{\infty} \int_{2^k Q \setminus 2^{k-1} Q} \frac{l(Q)^{d+1}}{[l(Q) + |x - y_Q|]^{d+1}} w(x) dx \right] \\ &\lesssim \sum_{Q \in \mathcal{Q}_{j,i}} \left[w(Q) + \sum_{k=1}^{\infty} \frac{w(2^k Q)}{2^{k(d+1)}} \right] \lesssim \sum_{Q \in \mathcal{Q}_{j,i}} w(Q) \lesssim \lambda^{-r}. \end{aligned}$$

Therefore we obtain that $w(F_\lambda) \lesssim \lambda^{-r}$.

Step 5: Exceptional set H_λ . Define the *exceptional set* for $i = l+1, \dots, n$, $j = 1, \dots, d$,

$$H_{i,j,\lambda} = \{x \in \mathbb{R}^d : \mathcal{M}(\nabla A_{i,j}^g)(x) > \lambda^{r/q_i}\}, \quad H_\lambda = \cup_{i=l+1}^n \cup_{j=1}^d H_{i,j,\lambda}.$$

Notice that by the definition of $A_{i,j}^g$, for each $s = 1, \dots, d$, we get

$$\mathcal{F}(\partial_s A_{i,j}^g)(\xi) = C \frac{\xi_s \xi_j}{|\xi|^2} \mathcal{F}(g_{j,i})(\xi) \Rightarrow \nabla A_{i,j}^g = C R R_j g_{j,i},$$

where \mathcal{F} is the Fourier transform, R_j is the Riesz transform and $R = (R_1, \dots, R_d)$. Recall $w \in A_1(\mathbb{R}^d)$. Since R_j is weighted strong type (q, q) for $1 < q < \infty$, we get $\|\nabla A_{i,j}^g\|_{L^q(\mathbb{R}^d, w)} \lesssim \|g_{j,i}\|_{L^q(\mathbb{R}^d, w)}$. Choose $d < q < \infty$, by the Chebyshev inequality, Lemma 2.3 and (cz-i) in Step 2, we get

$$\begin{aligned} w(H_{i,j,\lambda}) &\lesssim \lambda^{-\frac{qr}{q_i}} \int_{\mathbb{R}^d} [\mathcal{M}(\nabla A_{i,j}^g)(x)]^q w(x) dx \lesssim \lambda^{-\frac{qr}{q_i}} \int_{\mathbb{R}^d} |\nabla A_{i,j}^g(x)|^q w(x) dx \\ &\lesssim \lambda^{-\frac{qr}{q_i}} \int_{\mathbb{R}^d} |g_{j,i}(x)|^q w(x) dx \lesssim \lambda^{-r} \int_{\mathbb{R}^d} |g_{j,i}(x)|^{q_i} w(x) dx \lesssim \lambda^{-r}. \end{aligned}$$

Therefore we get $w(H_\lambda) \lesssim \lambda^{-r}$.

Step 6: Final exceptional set G_λ . Based on the construction of $J_\lambda, B_\lambda, D_\lambda, F_\lambda, H_\lambda$ in Step 1-5 and the fact w satisfies the doubling property, we choose an open set G_λ which satisfies the following conditions:

- (1). $(10J_\lambda \cup 10B_\lambda \cup 10D_\lambda \cup 10F_\lambda \cup 10H_\lambda) \subset G_\lambda$;
- (2). $w(G_\lambda) \lesssim w(J_\lambda) + w(B_\lambda) + w(D_\lambda) + w(F_\lambda) + w(H_\lambda)$.

Applying the previous weighted estimates of $J_\lambda, B_\lambda, D_\lambda, F_\lambda$ and H_λ , we obtain that $w(G_\lambda) \lesssim \lambda^{-r}$. Next making a Whitney decomposition of G_λ (see [17]), we may obtain a family of disjoint dyadic cubes $\{Q_k\}_k$ such that

- (i). $G_\lambda = \bigcup_{k=1}^{\infty} Q_k$;
- (ii). $\sqrt{d} \cdot l(Q_k) \leq \text{dist}(Q_k, (G_\lambda)^c) \leq 4\sqrt{d} \cdot l(Q_k)$.

By the property (ii) above, the distance between Q_k and $(G_\lambda)^c$ equals to $Cl(Q_k)$. For each Q_k above, we could construct a larger cube Q_k^* so that $Q_k \subset Q_k^*$, Q_k^* is centered at y_k and $y_k \in (G_\lambda)^c$, $l(Q_k^*) \approx l(Q_k)$. Therefore by the construction of Q_k^* and y_k , we may get that

$$(3.16) \quad \text{dist}(y_k, Q_k) \approx l(Q_k).$$

Clearly, the *exceptional set* G_λ constructed in Step 6 satisfies that $w(G_\lambda) \lesssim \lambda^{-r}$. Below we will prove that these functions A_i s are Lipschitz functions on $(G_\lambda)^c$.

Step 7: Lipschitz estimates of A_i on $(G_\lambda)^c$. Choose any $x, y \in (G_\lambda)^c$. By the *exceptional set* J_λ constructed in Step 1, we see that for $i = 1, \dots, l$

$$(3.17) \quad |A_i(x) - A_i(y)| \leq \lambda^{\frac{r}{q_i}} |x - y|.$$

In the following we only consider $i = l + 1, \dots, n$. By the Calderón-Zygmund decomposition in Step 2, it is sufficient to prove that $A_{i,j}^g$ and $A_{i,j}^b$ satisfy Lipschitz estimates on $(G_\lambda)^c$ for each $i = l + 1, \dots, n$ and $j = 1, \dots, d$. Firstly, it is easy to see that $A_{i,j}^g$ satisfies Lipschitz estimates by the construction of H_λ in Step 5. In fact, $x, y \in (G_\lambda)^c$ implies that $x, y \in H_\lambda^c$, we obtain that for $i = l + 1, \dots, n$, $j = 1, \dots, d$,

$$(3.18) \quad |A_{i,j}^g(x) - A_{i,j}^g(y)| \leq \lambda^{\frac{r}{q_i}} |x - y|.$$

We devote to proving that $A_{i,j}^b$ is a Lipschitz function on $(G_\lambda)^c$. Recall the Calderón-Zygmund decomposition properties (cz-ii), (cz-iii) and (cz-iv) in Step 2. For each $b_{j,i} = \sum_{Q \in \mathcal{Q}_{j,i}} b_{j,i,Q}$, $\text{supp } b_{j,i,Q} \subset Q$, where $\mathcal{Q}_{j,i}$ is a countable set of disjoint dyadic cubes. Then for each $Q \in \mathcal{Q}_{j,i}$, we define

$$A_{i,j}^{bQ}(x) = C_d \int_{\mathbb{R}^d} \frac{x_j - z_j}{|x - z|^d} b_{j,i,Q}(z) dz.$$

Now we fix a dyadic cube $Q \in \mathcal{Q}_{j,i}$. We are going to give a straight-forward Lipschitz estimate of $A_{i,j}^{bQ}$. By the construction of G_λ , we get that $x, y \in (10B_\lambda)^c$, i.e. $x, y \in (10Q)^c$, therefore we obtain $\text{dist}(x, Q) \geq \frac{9}{2}l(Q)$ and $\text{dist}(y, Q) \geq \frac{9}{2}l(Q)$. Let z_Q be the center of Q . Without loss of generality, assume that $|x - z_Q| \leq |y - z_Q|$. Choose a point $Z \in \mathbb{R}^d$ such that

$$|x - Z| < 100|x - y|; \quad |y - Z| \leq 100|x - y|; \quad |x - z_Q| > \frac{2}{5}|x - z_Q|$$

for any X belongs to the polygonal with vertex x, y, Z . We could draw a figure to show that such a point Z exists under the condition that $\text{dist}(x, Q) > \frac{9}{2}l(Q)$ and $\text{dist}(y, Q) > \frac{9}{2}l(Q)$. Now we write $A_{i,j}^{b_Q}(x) - A_{i,j}^{b_Q}(y) = A_{i,j}^{b_Q}(x) - A_{i,j}^{b_Q}(Z) + A_{i,j}^{b_Q}(Z) - A_{i,j}^{b_Q}(y)$. By using the mean value formula, we have

$$(3.19) \quad \begin{aligned} A_{i,j}^{b_Q}(x) - A_{i,j}^{b_Q}(Z) &= \int_0^1 \langle x - Z, \nabla(A_{i,j}^{b_Q})(tx + (1-t)Z) \rangle dt; \\ A_{i,j}^{b_Q}(Z) - A_{i,j}^{b_Q}(y) &= \int_0^1 \langle Z - y, \nabla(A_{i,j}^{b_Q})(tZ + (1-t)y) \rangle dt. \end{aligned}$$

For any $0 \leq t \leq 1$, the points $tx + (1-t)Z$ and $tZ + (1-t)y$ lie in the polygonal with vertex x, y, Z . Notice that $|x - z_Q| \geq 5l(Q)$. Then by our choice of Z , we get

$$(3.20) \quad \begin{aligned} |tx + (1-t)Z - z_Q| &> \frac{2}{5}|x - z_Q| > 2l(Q), \\ |tZ + (1-t)y - z_Q| &> \frac{2}{5}|x - z_Q| > 2l(Q). \end{aligned}$$

We set $Z(t)$ equals to $tx + (1-t)Z$ or $tZ + (1-t)y$ and $K_j(x) = x_j/|x|^d$. Using the cancelation condition of $b_{j,i,Q}$, (3.20) and the unweighted estimate in (cz-iv) of Step 2, we get that

$$\begin{aligned} |\nabla(A_{i,j}^{b_Q})(Z(t))| &= \left| \int_{\mathbb{R}^d} [(\nabla K_j)(Z(t) - z) - (\nabla K_j)(Z(t) - z_Q)] b_{j,i,Q}(z) dz \right| \\ &\lesssim \frac{l(Q)}{[l(Q) + |x - z_Q|]^{d+1}} \|b_{j,i,Q}\|_{L^1(Q)} \lesssim \lambda^{\frac{r}{q_i}} \frac{l(Q)}{[l(Q) + |x - z_Q|]^{d+1}} |Q|. \end{aligned}$$

Combining the above arguments with (3.19) and the construction of Z , we obtain

$$|A_{i,j}^{b_Q}(x) - A_{i,j}^{b_Q}(y)| \lesssim \lambda^{\frac{r}{q_i}} \frac{l(Q)}{[l(Q) + |x - z_Q|]^{d+1}} |Q| |x - y|.$$

Notice $x \in (G_\lambda)^c$ implies that $x \in (F_\lambda)^c$ in Step 3. Then we get that

$$(3.21) \quad |A_{i,j}^{b_Q}(x) - A_{i,j}^{b_Q}(y)| \lesssim \lambda^{\frac{r}{q_i}} |x - y| \sum_{Q \in \mathcal{Q}_{j,i}} \frac{l(Q)}{[l(Q) + |x - z_Q|]^{d+1}} |Q| \leq \lambda^{\frac{r}{q_i}} |x - y|.$$

Therefore we conclude that the Lipschitz estimates in (3.17) for $i = 1, \dots, l$, good function (3.18) and bad function (3.21) for $i = l+1, \dots, n$, to obtain that for any $i = 1, \dots, n$, $x, y \in (G_\lambda)^c$,

$$(3.22) \quad |A_i(x) - A_i(y)| \leq \lambda^{\frac{r}{q_i}} |x - y|.$$

Step 8: Weighted estimate of E_λ . We come back to give an estimate of E_λ . Split f into two parts $f = f_1 + f_2$ where $f_1(x) = f(x)\chi_{(G_\lambda)^c}(x)$ and $f_2(x) = f(x)\chi_{G_\lambda}(x)$. By the Lipschitz estimate in (3.22), when restricted on $(G_\lambda)^c$, A_i is a Lipschitz function with $\|\nabla A_i\|_{L^\infty((G_\lambda)^c)} \leq \lambda^{\frac{r}{q_i}}$ for $i = 1, \dots, n$. Let \tilde{A}_i represent the Lipschitz extension of A_i from $(G_\lambda)^c$ to \mathbb{R}^d (see [26, page 174, Theorem 3]) so that for each $i = 1, \dots, n$,

$$(3.23) \quad \begin{aligned} \tilde{A}_i(y) &= A_i(y) \quad \text{if } y \in (G_\lambda)^c; \\ |\tilde{A}_i(x) - \tilde{A}_i(y)| &\leq \lambda^{\frac{r}{q_i}} |x - y| \quad \text{for all } x, y \in \mathbb{R}^d. \end{aligned}$$

Since the operator $\mathcal{C}_*[\cdots, \cdot]$ is sub-multilinear, we split E_λ as three terms

$$\begin{aligned} & w(\{x \in \mathbb{R}^d : \mathcal{C}_*[\nabla A_1, \cdots, \nabla A_n, f](x) > \lambda\}) \\ & \leq w(10G_\lambda) + w(\{x \in (10G_\lambda)^c : \mathcal{C}_*[\nabla A_1, \cdots, \nabla A_n, f_1](x) > \lambda/2\}) \\ & \quad + w(\{x \in (10G_\lambda)^c : \mathcal{C}_*[\nabla A_1, \cdots, \nabla A_n, f_2](x) > \lambda/2\}). \end{aligned}$$

The above first term satisfies $w(10G_\lambda) \lesssim \lambda^{-r}$, which is our required estimate. Below we consider the second term. We only consider $x \in (10G_\lambda)^c$. By the definition of f_1 ,

$$\mathcal{C}_*[\nabla A_1, \cdots, \nabla A_n, f_1](x) = \mathcal{C}_*[\nabla \tilde{A}_1, \cdots, \nabla \tilde{A}_n, f_1](x).$$

Notice that $w \in A_1(\mathbb{R}^d) \subsetneq A_p(\mathbb{R}^d)$. Applying the above equality and Proposition 3.1 ($1 \leq p < \infty$), we derive that

$$\begin{aligned} & w(\{x \in (10G_\lambda)^c : \mathcal{C}_*[\nabla A_1, \cdots, \nabla A_n, f_1](x) > \lambda/2\}) \\ & = w(\{x \in (10G_\lambda)^c : \mathcal{C}_*[\nabla \tilde{A}_1, \cdots, \nabla \tilde{A}_n, f_1](x) > \lambda/2\}) \\ & \lesssim \lambda^{-p} \left(\prod_{i=1}^n \|\nabla \tilde{A}_i\|_{L^\infty(\mathbb{R}^d, w)}^p \right) \|f_1\|_{L^p(\mathbb{R}^d, w)}^p \lesssim \lambda^{-p+p \sum_{i=1}^n \frac{r}{q_i}} = \lambda^{-r}. \end{aligned}$$

If $p = \infty$, the above argument does not work.

Step 9: Weighted estimate of $\mathcal{C}_[\nabla A_1, \cdots, \nabla A_n, f_2](x)$.* Recall $\mathbb{N}_i^j = \{i, i+1, \cdots, j\}$ and the construction of G_λ, y_k, Q_k and Q_k^* above (3.16). Then we may write $f_2 = \sum_k f \chi_{Q_k}$. So

$$\mathcal{C}_\varepsilon[\nabla A_1, \cdots, \nabla A_n, f_2](x) = \sum_k \mathcal{C}_\varepsilon[\nabla A_1, \cdots, \nabla A_n, f \chi_{Q_k}](x).$$

In the following we study $\prod_{i=1}^n \frac{A_i(x) - A_i(y)}{|x-y|}$. We shall separate it into several terms and then give an estimate for each term. Write

$$\begin{aligned} & \prod_{i=1}^n \frac{A_i(x) - A_i(y)}{|x-y|} \\ & = \prod_{i=1}^n \left(\frac{\tilde{A}_i(x) - \tilde{A}_i(y)}{|x-y|} + \frac{\tilde{A}_i(y) - \tilde{A}_i(y_k)}{|x-y|} + \frac{A_i(y_k) - A_i(y)}{|x-y|} \right) \\ & = \sum_{\mathbb{N}_1^n} \left(\prod_{i \in \mathbb{N}_1} \frac{\tilde{A}_i(x) - \tilde{A}_i(y)}{|x-y|} \right) \left(\prod_{i \in \mathbb{N}_2} \frac{\tilde{A}_i(y) - \tilde{A}_i(y_k)}{|x-y|} \right) \left(\prod_{i \in \mathbb{N}_3} \frac{A_i(y_k) - A_i(y)}{|x-y|} \right) \\ & = I(x, y) + II(x, y, y_k) + III(x, y, y_k) + IV(x, y, y_k), \end{aligned}$$

where in the third equality we divide $\mathbb{N}_1^n = N_1 \cup N_2 \cup N_3$ with N_1, N_2, N_3 non intersecting each other; and $I(x, y), II(x, y, y_k), III(x, y, y_k)$ and $IV(x, y, y_k)$ are defined as follows

$$\begin{aligned}
I(x, y) &= \prod_{i=1}^n \frac{\tilde{A}_i(x) - \tilde{A}_i(y)}{|x - y|}, \\
II(x, y, y_k) &= \sum_{\substack{N_1 \subsetneq \mathbb{N}_1^n \\ N_3 = \emptyset}} \left(\prod_{i \in N_1} \frac{\tilde{A}_i(x) - \tilde{A}_i(y)}{|x - y|} \right) \left(\prod_{i \in N_2} \frac{\tilde{A}_i(y) - \tilde{A}_i(y_k)}{|x - y|} \right), \\
III(x, y, y_k) &= \sum_{\substack{N_1 \subsetneq \mathbb{N}_1^n \\ N_3 \neq \emptyset, N_3 \subset \{1, \dots, l\}}} \left(\prod_{i \in N_1} \frac{\tilde{A}_i(x) - \tilde{A}_i(y)}{|x - y|} \right) \left(\prod_{i \in N_2} \frac{\tilde{A}_i(y) - \tilde{A}_i(y_k)}{|x - y|} \right) \\
&\quad \times \left(\prod_{i \in N_3} \frac{A_i(y_k) - A_i(y)}{|x - y|} \right), \\
IV(x, y, y_k) &= \sum_{\substack{N_1 \subsetneq \mathbb{N}_1^n \\ N_3 \neq \emptyset, N_3 \cap \{l+1, \dots, n\} \neq \emptyset}} \left(\prod_{i \in N_1} \frac{\tilde{A}_i(x) - \tilde{A}_i(y)}{|x - y|} \right) \left(\prod_{i \in N_2} \frac{\tilde{A}_i(y) - \tilde{A}_i(y_k)}{|x - y|} \right) \\
&\quad \times \left(\prod_{i \in N_3} \frac{A_i(y_k) - A_i(y)}{|x - y|} \right).
\end{aligned} \tag{3.24}$$

In the above decomposition, we in fact divide $\mathcal{C}_\varepsilon[\nabla A_1, \dots, \nabla A_n, f\chi_{Q_k}](x)$ into 3^n terms and separate these terms into four parts according I, II, III and IV .

Step 10: Weighted estimate of $\mathcal{C}_[\dots, \cdot]$ related to I .* In this case there is only one term, i.e. $\mathcal{C}_*[\nabla \tilde{A}_1, \dots, \nabla \tilde{A}_n, f_2]$. Then by Proposition 3.1 ($1 \leq p < \infty$), we get

$$\begin{aligned}
&w(\{x \in (10G_\lambda)^c : \mathcal{C}_*[\nabla \tilde{A}_1, \dots, \nabla \tilde{A}_n, f_2](x) > \lambda/2\}) \\
&\lesssim \lambda^{-p} \left(\prod_{i=1}^n \|\nabla \tilde{A}_i\|_{L^\infty(\mathbb{R}^d, w)}^p \right) \|f_2\|_{L^p(\mathbb{R}^d, w)}^p \lesssim \lambda^{-p+p \sum_{i=1}^n \frac{r}{q_i}} = \lambda^{-r}.
\end{aligned}$$

If $p = \infty$, the above argument may not work.

Step 11: Weighted estimate of $\mathcal{C}_[\dots, \cdot]$ related to II .* It is sufficient to consider one term $\mathcal{C}_*[\dots, \cdot]$ related to II in which N_1 is a proper subset of \mathbb{N}_1^n and $N_3 = \emptyset$. In such a case, without loss of generality, we may suppose $N_1 = \{1, \dots, v\}$, $N_2 = \{v+1, \dots, n\}$ with $0 \leq v < n$. Here if $v = 0$, it means that $N_1 = \emptyset$. With these notation, we see that N_1 is a proper subset of \mathbb{N}_1^n . By a slight abuse of notation, we still utilize $II(x, y, y_k)$ to represent one term related to N_1, N_2 and N_3 in (3.24) and utilize $H_{II}(x)$ to represent $\mathcal{C}_*[\dots, \cdot]$ related to $II(x, y, y_k)$, i.e.

$$H_{II}(x) = \sup_{\varepsilon > 0} \left| \sum_k \int_{|x-y| > \varepsilon} K(x-y) II(x, y, y_k) f\chi_{Q_k}(y) dy \right|.$$

Notice that \tilde{A}_i is a Lipschitz function with bound $\lambda^{\frac{r}{q_i}}$ for $i = 1, \dots, n$ by (3.23). Then we obtain that

$$|II(x, y, y_k)| \lesssim \lambda^{\sum_{i=1}^n \frac{r}{q_i}} \frac{|y - y_k|^{n-v}}{|x - y|^{n-v}}.$$

Since we only need to consider $x \in (10G_\lambda)^c$, then by (3.16), we obtain that

$$(3.25) \quad |x - y| \geq 2l(Q_k) \approx |y - y_k| \quad \text{for any } y \in Q_k.$$

Now combining with (1.1), the above estimate of $II(x, y, y_k)$ and (3.25), we obtain that

$$\begin{aligned} H_{II}(x) &\leq \sum_k \int_{Q_k} |K(x - y)| \cdot |II(x, y, y_k)| \cdot |f(y)| dy \\ &\lesssim \lambda^{\sum_{i=1}^n \frac{r}{q_i}} \sum_k \int_{Q_k} \frac{l(Q_k)^{n-v}}{[l(Q_k) + |x - y|]^{d+n-v}} |f(y)| dy. \end{aligned}$$

Utilizing the Chebyshev inequality, the above estimate of H_{II} and Lemma 2.8 (since $n - v \geq 1$), we finally obtain that

$$w(\{x \in (10G_\lambda)^c : H_{II}(x) > \lambda\}) \leq \lambda^{-p + \sum_{i=1}^n \frac{rp}{q_i}} \|T_{n-v} f\|_{L^p(\mathbb{R}^d, w)}^p \lesssim \lambda^{-r} \|f\|_{L^p(\mathbb{R}^d, w)}^p.$$

Hence we finish the proof related to II .

Step 12: Weighted estimate of $\mathcal{C}_[\dots, \cdot]$ related to III .* It is sufficient to consider one term $\mathcal{C}_*[\dots, \cdot]$ related to III in which N_1 is a proper subset of \mathbb{N}_1^n and N_3 is a nonempty subset of $\{1, \dots, l\}$. By the condition in this proposition, for any $i \in N_3$, $d \leq q_i < \infty$. Thus $\nabla A_i \in L^{q_i}(\mathbb{R}^d, w)$ (or $L^{d,1}(\mathbb{R}^d, w)$ if $q_i = d$) with $d \leq q_i < \infty$. Then by using the fact y_k lies in the $(G_\lambda)^c$, i.e. $y_k \in (J_{i,\lambda})^c$, we give the estimates in N_3 as follows

$$(3.26) \quad \frac{|A_i(y_k) - A_i(y)|}{|y - y_k|} \leq \mathcal{M}(\nabla A_i)(y_k) \leq \lambda^{\frac{r}{q_i}}, \quad \text{for } i \in N_3.$$

Define $v = \text{card}(N_1)$. Then we see that $0 \leq v < n$. By a slight abuse of notation, we still utilize $III(x, y, y_k)$ to stand for one term related to N_1, N_2 and N_3 in (3.24) and utilize $H_{III}(x)$ to represent $\mathcal{C}_*[\dots, \cdot]$ related to $III(x, y, y_k)$, i.e.

$$H_{III}(x) = \sup_{\varepsilon > 0} \left| \sum_k \int_{|x-y|>\varepsilon} K(x-y) III(x, y, y_k) f \chi_{Q_k}(y) dy \right|.$$

From the fact \tilde{A}_i s are Lipschitz functions with bounds λ^{r/q_i} for $i \in N_1 \cup N_2$ and (3.26), we obtain that

$$|III(x, y, y_k)| \lesssim \lambda^{\sum_{i \in N_1 \cup N_2} \frac{r}{q_i}} \frac{|y - y_k|^{n-v}}{|x - y|^{n-v}} \prod_{i \in N_3} \mathcal{M}(\nabla A_i)(y_k) \lesssim \lambda^{\sum_{i=1}^n \frac{r}{q_i}} \frac{|y - y_k|^{n-v}}{|x - y|^{n-v}}.$$

Inserting this estimate of $III(x, y, y_k)$ into H_{III} , combining with (1.1) and (3.25), and next utilizing the Chebyshev inequality and Lemma 2.8 (since $n - v \geq 1$), we finally obtain that

$$w(\{x \in (10G_\lambda)^c : H_{III}(x) > \lambda\}) \leq \lambda^{-p+p\left(\sum_{i=1}^n \frac{r}{q_i}\right)} \|T_{n-v} f\|_{L^p(\mathbb{R}^d, w)}^p \lesssim \lambda^{-r} \|f\|_{L^p(\mathbb{R}^d, w)}^p.$$

Hence we finish the proof of this part.

Step 13: Weighted estimate of $\mathcal{C}_[\cdots, \cdot]$ related to IV .* It is sufficient to consider one term $\mathcal{C}_*[\cdots, \cdot]$ related to IV in which $N_1 \subsetneq \mathbb{N}_1^n$ and $N_3 \neq \emptyset$ with $N_3 \cap \{l+1, \cdots, n\} \neq \emptyset$. In such a case, without loss of generality, we may suppose $l+1, \cdots, v \in N_3$ with $l+1 \leq v \leq n$ and $v+1, \cdots, n$ belongs to N_1 or N_2 . So we may assume that $N_3 = \{\iota, \cdots, w, l+1, \cdots, v\}$ with $0 \leq \iota \leq w \leq l$. Define $u = \text{card}(N_1)$. Then $n-u \geq 1$. With these notation, we can easily see that N_3 is a nonempty set with $N_3 \cap \{l+1, \cdots, n\} \neq \emptyset$. Note that $w \in A_1(\mathbb{R}^d)$. By a slight abuse of notation, we still use $IV(x, y, y_k)$ to stand for one term related to N_1, N_2 and N_3 in (3.24) and use $H_{IV}(x)$ to stand for $\mathcal{C}_*[\cdots, \cdot]$ related to $IV(x, y, y_k)$, i.e.

$$H_{IV}(x) = \sup_{\varepsilon > 0} \left| \sum_k \int_{|x-y| > \varepsilon} K(x-y) IV(x, y, y_k) f \chi_{Q_k}(y) dy \right|.$$

Note that $d \leq q_1, \cdots, q_l \leq \infty$ and $1 \leq q_{l+1}, \cdots, q_n < d$. Recall in Step 3, we set $\frac{1}{s_i} = \frac{1}{q_i} - \frac{1}{d}$ for $i = l+1, \cdots, n$. We also set $\frac{1}{q} = \left(\sum_{i=l+1}^v \frac{1}{s_i} \right) + \frac{1}{p}$. Since $r \geq \frac{d}{d+n}$ and $\frac{1}{r} = \left(\sum_{i=1}^n \frac{1}{q_i} \right) + \frac{1}{p}$, we could obtain $1 \leq q \leq \infty$ which will be crucial when we use Lemma 2.8. With (3.16) and \tilde{A}_i is a Lipschitz function with bound λ^{r/q_i} for $i \in N_1 \cup N_2$, we have

$$\begin{aligned} |IV(x, y, y_k)| &\lesssim \lambda^{(\sum_{i \in N_1 \cup N_2} \frac{r}{q_i})} \frac{(l(Q_k))^{n-u}}{|x-y|^{n-u}} \left[\prod_{i=\iota}^w \mathcal{M}(\nabla A_i)(y_k) \right] \prod_{i=l+1}^v \frac{|A_i(y_k) - A_i(y)|}{l(Q_k)} \\ &\lesssim \lambda^{(\sum_{i=1}^l + \sum_{i=v+1}^n) \frac{r}{q_i}} \frac{(l(Q_k))^{n-u}}{|x-y|^{n-u}} \prod_{i=l+1}^v \frac{|A_i(y_k) - A_i(y)|}{l(Q_k)}. \end{aligned}$$

Then inserting the above estimate of IV into H_{IV} with (1.1) and (3.25), we get

$$\begin{aligned} H_{IV}(x) &\leq \sum_k \int_{Q_k} |K(x-y)| \cdot |IV(x, y, y_k)| \cdot |f(y)| dy \\ &\lesssim \lambda^{(\sum_{i=1}^l + \sum_{i=v+1}^n) \frac{r}{q_i}} \sum_k \int_{Q_k} \frac{l(Q_k)^{n-u}}{[l(Q_k) + |x-y|]^{d+n-u}} h_{l,v}(y) dy \\ &= \lambda^{(\sum_{i=1}^l + \sum_{i=v+1}^n) \frac{r}{q_i}} T_{n-u}(h_{l,v})(x), \end{aligned}$$

where the operator T_{n-u} is defined in Lemma 2.8 and the function $h_{l,v}(y)$ is defined as

$$h_{l,v}(y) = \sum_{Q_k} \prod_{i=l+1}^v \left(\frac{|A_i(y_k) - A_i(y)|}{l(Q_k)} \right) \chi_{Q_k} |f|(y).$$

Utilizing the Chebyshev inequality and the above estimate of H_{IV} , applying Lemma 2.8 (note that $1 \leq q \leq \infty$ and $n-u \geq 1$), we finally obtain that

$$\begin{aligned} (3.27) \quad &w(\{x \in (10G_\lambda)^c : H_{IV}(x) > \lambda\}) \\ &\leq \lambda^{-q + (\sum_{i=1}^l + \sum_{i=v+1}^n) \frac{rq}{q_i}} \int_{(10G_\lambda)^c} [T_{n-u}(h_{l,v})(x)]^q w(x) dx \\ &\lesssim \lambda^{-q + (\sum_{i=1}^l + \sum_{i=v+1}^n) \frac{rq}{q_i}} \|h_{l,v}\|_{L^q(\mathbb{R}^d, w)}^q. \end{aligned}$$

In the following we give an estimate of $\|h_{l,v}\|_{L^q(\mathbb{R}^d,w)}^q$. We may write

$$\begin{aligned} \|h_{l,v}\|_{L^q(\mathbb{R}^d,w)}^q &= \sum_{Q_k} \int_{Q_k} \left[\prod_{i=l+1}^v \left(\frac{|A_i(y_k) - A_i(y)|}{l(Q_k)} \right)^q \right] |f(y)|^q w(y) dy \\ &\leq \sum_{Q_k} \prod_{i=l+1}^v \left[\int_{Q_k} \left(\frac{|A_i(y_k) - A_i(y)|}{l(Q_k)} \right)^{s_i} w(y) dy \right]^{\frac{q}{s_i}} \left[\int_{Q_k} |f(y)|^p w(y) dy \right]^{\frac{q}{p}} \\ &\lesssim \sum_{Q_k} \prod_{i=l+1}^v \left[\int_{Q_k^*} \left(\frac{|A_i(y_k) - A_i(y)|}{l(Q_k^*)} \right)^{s_i} w(y) dy \right]^{\frac{q}{s_i}} \left[\int_{Q_k} |f(y)|^p w(y) dy \right]^{\frac{q}{p}} \\ &\lesssim \sum_{Q_k} \left[\prod_{i=l+1}^v \mathfrak{M}_{w,s_i}(\nabla A_i)(y_k)^q w(Q_k)^{\frac{q}{s_i}} \right] \left[\int_{Q_k} |f(y)|^p w(y) dy \right]^{\frac{q}{p}}, \end{aligned}$$

where in the second inequality we use the Hölder inequality and the third inequality follows from the fact $Q_k \subset Q_k^*$, y_k is the center of Q_k^* and $l(Q_k^*) \approx l(Q_k)$. Notice that y_k lies in the $(G_\lambda)^c$, i.e. $y_k \in (D_{i,\lambda})^c$ (see Step 3). Then we obtain that

$$\mathfrak{M}_{w,s_i}(\nabla A_i)(y_k) \leq \lambda^{\frac{r}{q_i}}, \text{ for } i = l+1, \dots, v.$$

Utilizing the above inequality, the Hölder inequality again and (cz-iii) in Step 2, we get

$$\begin{aligned} \|h_{l,v}\|_{L^q(\mathbb{R}^d,w)}^q &\lesssim \lambda^{\sum_{i=l+1}^v \frac{qr}{q_i}} \left[\sum_{Q_k} w(Q_k) \right]^{\sum_{i=l+1}^v \frac{q}{s_i}} \|f\|_{L^p(\mathbb{R}^d,w)}^q \\ &\lesssim \lambda^{\sum_{i=l+1}^v \frac{qr}{q_i}} [w(G_\lambda)]^{\sum_{i=l+1}^v \frac{q}{s_i}} \lesssim \lambda^{\sum_{i=l+1}^v \left(\frac{qr}{q_i} - \frac{qr}{s_i} \right)}. \end{aligned}$$

Plunge the above estimate into (3.27) with some elementary calculations, we finally obtain that

$$w(\{x \in (10G_\lambda)^c : H_{IV}(x) > \lambda\}) \lesssim \lambda^{-q + \left(\sum_{i=1}^n \frac{rq}{q_i} \right) - \left(\sum_{i=l+1}^v \frac{qr}{s_i} \right)} \lesssim \lambda^{-r},$$

hence we finish the proof of the term *IV*.

Finally, we show how to modify the above argument to the case $q_i = \infty$ for some $i = k+1, \dots, l$. Notice that only in Step 1 the construction of *exceptional set* is involved with A_{k+1}, \dots, A_l . We may assume that only $q_{k+1} = \dots = q_u = \infty$ with $k+1 \leq u \leq l$. Therefore A_{k+1}, \dots, A_u are Lipschitz functions. Then we just fix A_{k+1}, \dots, A_u in the rest of the proof. In Step 1 we modify the argument that we only make a construction of *exceptional set* for A_1, \dots, A_k and A_{u+1}, \dots, A_l . These proofs in Steps 2-8 are the same. Later when studying

$$\left(\prod_{i=1}^n \frac{A_i(x) - A_i(y)}{|x-y|} \right) = \left(\prod_{i=k+1}^u \prod_{i=1}^k \prod_{i=u+1}^n \right) \frac{A_i(x) - A_i(y)}{|x-y|},$$

we just use the same way as in Steps 9-13 to deal with the terms from $\prod_{i=1}^k \prod_{i=u+1}^n$ since the term $\prod_{i=k+1}^u \frac{A_i(x) - A_i(y)}{|x-y|}$ could be absorbed by the kernel $K(x-y)$ if we observe that $K(x-y) \prod_{i=k+1}^u \frac{A_i(x) - A_i(y)}{|x-y|}$ is a standard Calderón-Zygmund kernel. \square

Proposition 3.7. Let $\frac{d}{d+n} \leq r \leq 1$, $q_1 = \cdots = q_k = d$, $d < q_{k+1}, \dots, q_l \leq \infty$ and $1 \leq q_{l+1}, \dots, q_n < d$ with $0 \leq k \leq l$ and $1 \leq l < n$, $p = \infty$. Suppose that $w \in A_1(\mathbb{R}^d)$. Then

$$\begin{aligned} & \|\mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f]\|_{L^{r,\infty}(\mathbb{R}^d, w)} \\ & \lesssim \left(\prod_{i=1}^k \|\nabla A_i\|_{L^{d,1}(\mathbb{R}^d, w)} \right) \left(\prod_{i=k+1}^n \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d, w)} \right) \|f\|_{L^\infty(\mathbb{R}^d, w)}, \end{aligned}$$

where $L^{d,1}(\mathbb{R}^d, w)$ is the standard Lorentz space.

Proof. The proof is similar to that of Proposition 3.6 and one could follow the idea in the proof of Proposition 3.5, so the details of the proof is omitted. \square

3.4. Interpolation.

Notice that we have already proven all the cases (ii) in Theorem 1.1 by Propositions 3.4, 3.5, 3.6 and 3.7. And only part strong type multilinear estimates of (i) in Theorem 1.1 has been established by Proposition 3.1 and Proposition 3.3. The rest part of (i) in Theorem 1.1 just follow from the linear Marcinkiewicz interpolation (see [28] or [1]). In the following we show how to do this.

Since the maximal Calderón commutator \mathcal{C}_* is $(n+1)$ th submultilinear, when using the Marcinkiewicz interpolation, our main strategy is that we consider \mathcal{C}_* as a sublinear operator if we fix part of n variables.

Let $\nabla A_i \in L^{q_i}(\mathbb{R}^d, w)$ and $f \in L^p(\mathbb{R}^d, w)$ with $\frac{1}{r} = \left(\sum_{i=1}^n \frac{1}{q_i}\right) + \frac{1}{p}$, $\frac{d}{d+n} < r < \infty$, $1 < q_i \leq \infty$ ($i = 1, \dots, n$) and $1 < p \leq \infty$. Let $w \in \left(\bigcap_{i=1}^n A_{\max\{\frac{q_i}{d}, 1\}}(\mathbb{R}^d)\right) \cap A_p(\mathbb{R}^d)$. Our goal is to show the follow strong type estimate

$$(3.28) \quad \|\mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, f]\|_{L^r(\mathbb{R}^d, w)} \lesssim \left(\prod_{i=1}^n \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d, w)} \right) \|f\|_{L^p(\mathbb{R}^d, w)}.$$

We divide the proof into several cases. We first consider the case all $q_i \neq d$ for $i = 1, \dots, n$. Therefore by (ii) of Theorem 1.1, the multilinear estimates (1.6) are not involved with $L^{d,1}(\mathbb{R}^d, w)$ spaces. We further divide this case into two cases: $1 < p < \infty$ and $p = \infty$. Consider firstly the case $1 < p < \infty$. We fix all ∇A_i , q_i and $w \in \left(\bigcap_{i=1}^n A_{\max\{\frac{q_i}{d}, 1\}}(\mathbb{R}^d)\right) \cap A_p(\mathbb{R}^d)$. By the basic property of $A_p(\mathbb{R}^d)$ weight, $w \in A_{p_1}(\mathbb{R}^d)$ for all $p_1 > p$. If we choose p_1, r_1 such that $p < p_1 < \infty$ and $\frac{1}{r_1} = \left(\sum_{i=1}^n \frac{1}{q_i}\right) + \frac{1}{p_1}$, then by (ii) of Theorem 1.1

$$(3.29) \quad \mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, \cdot] : L^{p_1}(\mathbb{R}^d, w) \mapsto L^{r_1, \infty}(\mathbb{R}^d, w).$$

Since $w \in A_p(\mathbb{R}^d)$, by the revers Hölder inequality of $A_p(\mathbb{R}^d)$ weight (see [17]) and its definition, there exist $\varepsilon > 0$ such that $w \in A_{p-\varepsilon}$ and $p - \varepsilon \geq 1$. Then we may choose p_0, r_0 such that $p - \varepsilon \leq p_0 < p$, $\frac{d}{d+n} < r_0 < \infty$ and $\frac{1}{r_0} = \left(\sum_{i=1}^n \frac{1}{q_i}\right) + \frac{1}{p_0}$. Hence we obtain $w \in \left(\bigcap_{i=1}^n A_{\max\{\frac{q_i}{d}, 1\}}(\mathbb{R}^d)\right) \cap A_{p_0}(\mathbb{R}^d)$. By using (ii) of Theorem 1.1, we get

$$(3.30) \quad \mathcal{C}_*[\nabla A_1, \dots, \nabla A_n, \cdot] : L^{p_0}(\mathbb{R}^d, w) \mapsto L^{r_0, \infty}(\mathbb{R}^d, w).$$

Applying the Marcinkiewicz interpolation with (3.29) and (3.30), we establish the strong type estimate (3.28) provided that all $q_i \neq d$ and $1 < p < \infty$. Next we consider

another case all $q_i \neq d$ and $p = \infty$. By our condition $r < \infty$, there is at least one $q_i < \infty$. Without loss of generality, we may suppose that $q_1 < \infty$. If $d < q_1 < \infty$, then the rest of proof is similar to the case $q_i \neq d$ and $1 < p < \infty$ once fixing $\nabla A_i, q_i$ for $i = 2, \dots, n, f \in L^\infty(\mathbb{R}^d, w)$ and $w \in \bigcap_{i=2}^n A_{\max\{\frac{q_i}{d}, 1\}}(\mathbb{R}^d)$. If $1 < q_1 < d$, then $w \in A_1(\mathbb{R}^d)$ by our condition. Therefore it is easy to show (3.28) using (ii) of Theorem 1.1 once we fix $\nabla A_i, q_i$ for $i = 2, \dots, n, f \in L^\infty(\mathbb{R}^d, w)$ and $w \in A_1(\mathbb{R}^d)$.

Secondly let us consider the case there is only one q_i which equals to d . Without loss of generality, we may suppose $q_1 = d$. Then by our condition $w \in A_1(\mathbb{R}^d)$ in this case. Fix $\nabla A_i, q_i$ for $i = 2, \dots, n, f \in L^p(\mathbb{R}^d, w)$ and $w \in A_1(\mathbb{R}^d)$. Then we may choose $r_0, r_1, q_{1,0}, q_{1,1}$ such that $\frac{d}{d+n} < r_0, r_1 < \infty, 1 < q_{1,0} < d < q_{1,1}, \frac{1}{r_0} = \frac{1}{q_{1,0}} + \left(\sum_{i=2}^n \frac{1}{q_i}\right) + \frac{1}{p}, \frac{1}{r_1} = \frac{1}{q_{1,1}} + \left(\sum_{i=2}^n \frac{1}{q_i}\right) + \frac{1}{p}$. Then by (ii) of Theorem 1.1, we get

$$\mathcal{C}_*[\cdot, \nabla A_2, \dots, \nabla A_n, f] : L^{q_1, j}(\mathbb{R}^d, w) \mapsto L^{r_j, \infty}(\mathbb{R}^d, w) \quad j = 0, 1.$$

Using the Marcinkiewicz interpolation with the above two estimate, we get (3.28) in the case $q_1 = d$ and all $q_2, \dots, q_n \neq d$.

Finally we consider the general case there are m numbers of q_i s which equal to d . We only need to show $m = 2$, the general case just follows from the induction. Without loss of generality, we suppose that $q_1 = q_2 = d$. In this case, $w \in A_1(\mathbb{R}^d)$. Fix $\nabla A_i, q_i$ for $i = 2, \dots, n, f \in L^p(\mathbb{R}^d, w)$ and $w \in A_1(\mathbb{R}^d)$. Then we may choose $r_0, r_1, q_{1,0}, q_{1,1}$ such that $\frac{d}{d+n} < r_0, r_1 < \infty, 1 < q_{1,0} < d < q_{1,1}, \frac{1}{r_0} = \frac{1}{q_{1,0}} + \left(\sum_{i=2}^n \frac{1}{q_i}\right) + \frac{1}{p}, \frac{1}{r_1} = \frac{1}{q_{1,1}} + \left(\sum_{i=2}^n \frac{1}{q_i}\right) + \frac{1}{p}$. Since $q_2 = d$, by the result of the case there is only one $q_i = d$ we discussed above, we get the strong type estimate

$$\mathcal{C}_*[\cdot, \nabla A_2, \dots, \nabla A_n, f] : L^{q_1, j}(\mathbb{R}^d, w) \mapsto L^{r_j}(\mathbb{R}^d, w) \quad j = 0, 1.$$

Using the Marcinkiewicz interpolation with the above two estimate, we get (3.28) in the case $q_1 = q_2 = d$ and all $q_3, \dots, q_n \neq d$. Applying the induction of m , we finish the proof.

Remark 3.8. Instead of using the linear Marcinkiewicz interpolation in this proof, another possible more straightforward method is the multilinear interpolation with change of measures. To the best knowledge of the author, such kind of multilinear interpolation with change of measures is currently unknown. Therefore it will be interesting to establish the multilinear version of Stein-Weiss interpolation with change of measures (see [27]).

Acknowledgement. The author would like to thank the referees for their very careful reading and valuable suggestions.

REFERENCES

1. J. Bergh and J. Löfström, *Interpolation spaces. An introduction*. Grundlehren der Mathematischen Wissenschaften, No. 223. Springer-Verlag, Berlin-New York, 1976. x+207 pp. [34](#)
2. A. P. Calderón, *Commutators of singular integral operators*, Proc. Nat. Acad. Sci. USA. **53** (1965), 1092-1099. [1, 2](#)
3. A. P. Calderón, *Cauchy integrals on Lipschitz curves and related operators*, Proc. Nat. Acad. Sci. USA, **74** (1977), 1324-1327. [2](#)
4. A. P. Calderón, *Commutators, singular integrals on Lipschitz curves and application*, Proc. Inter. Con. Math. Helsinki, 1978, 85-96, Acad. Sci. Fennica, Helsinki, 1980. [2](#)

5. C. P. Calderón, *On commutators of singular integrals*, *Studia Math.*, **53** (1975), 139-174. [2](#), [6](#)
6. R. Coifman and Y. Meyer, *On commutators of singular integral and bilinear singular integrals*, *Trans. Amer. Math. Soc.*, 212 (1975), 315-331. [2](#)
7. R. Coifman and Y. Meyer, *Au delà des opérateurs pseudo-diffrentiels*, *Astérisque*, 57 (1978). [2](#)
8. M. Christ and J. L. Journé, *Polynomial growth estimates for multilinear singular integral operators*, *Acta Math.* **159** (1987), 51-80. [2](#), [13](#)
9. G. David and S. Semmes, *Strong A_∞ weights, Sobolev inequalities and quasiconformal mappings*. *Analysis and partial differential equations*, 101-111, *Lecture Notes in Pure and Appl. Math.*, 122, Dekker, New York, 1990. [8](#)
10. Y. Ding and X. Lai, *Weak type (1,1) bound criterion for singular integral with rough kernel and its applications*. *Trans. Amer. Math. Soc.* 371(2019), No. 3, 1649-1675. [2](#)
11. X. Duong, L. Grafakos and L. Yan, *Multilinear operators with non-smooth kernels and commutators of singular integrals*. *Trans. Amer. Math. Soc.* 362 (2010), no. 4, 2089-2113. [2](#)
12. X. Duong, R. Gong, L. Grafakos, J. Li and L. Yan, *Maximal operator for multilinear singular integrals with non-smooth kernels*. *Indiana Univ. Math. J.* 58 (2009), no. 6, 2517-2541. [2](#), [3](#), [4](#)
13. C. Fefferman, *Recent Progress in Classical Fourier Analysis*, *Proc. Inter. Con. Math.*, Vancouver, 1974, 95-118. [2](#)
14. P. W. Fong, *Smoothness properties of symbols, Calderón commutators and generalizations*. Thesis (Ph.D.)-Cornell University. 2016. 60 pp. [2](#)
15. J. García-Cuerva and J. Rubio de Francia, *Weighted norm inequalities and related topics*, *North-Holland Math. Studies* **116**, North-Holland, Amsterdam, 1985. [25](#)
16. L. Grafakos, L. Liu and D. Yang, *Multiple-weighted norm inequalities for maximal multi-linear singular integrals with non-smooth kernels*. *Proc. Roy. Soc. Edinburgh Sect. A* 141 (2011), no. 4, 755-775. [2](#), [4](#)
17. L. Grafakos, *Classic Fourier Analysis*, *Graduate Texts in Mathematics*, Vol. **249** (Third edition), Springer, New York, 2014. [1](#), [5](#), [6](#), [7](#), [9](#), [11](#), [12](#), [13](#), [19](#), [27](#), [34](#)
18. L. Grafakos, *Modern Fourier Analysis*, *Graduate Texts in Mathematics*, Vol. **250** (Third edition), Springer, New York, 2014. [2](#), [13](#)
19. M. Hadžić, A. Seeger, C. K. Smart, B. Street, *Singular integrals and a problem on mixing flows*. *Ann. Inst. H. Poincaré Anal. Non Linéaire.* 35 (2018), no. 4, 921-43. [2](#)
20. J. L. Journé, *Calderón-Zygmund operators, pseudodifferential operators and the Cauchy integral of Calderón*. *Lecture Notes in Mathematics*, 994. Springer-Verlag, Berlin, 1983. vi+128 pp. [5](#)
21. X. Lai, *Multilinear estimates for Calderón commutators*. *Int. Math. Res. Not. IMRN.* 2018. Doi: 10.1093/imrn/rny197. See also arXiv:1712.09020. [1](#), [2](#), [3](#), [14](#)
22. A. K. Lerner, S. Ombrosi, C. Pérez, R. H. Torres, and R. Trujillo-Gonzlez, *New maximal functions and multiple weights for the multilinear Calderón-Zygmund theory*. *Adv. Math.* 220 (2009), no. 4, 1222-1264. [5](#), [14](#)
23. F. Léger, *A new approach to bounds on mixing*. *Math. Models Methods Appl. Sci.* **28** (2018), no. 5, 829-849. [2](#)
24. Y. Meyer and R. Coifman, *Wavelets. Calderón-Zygmund and multilinear operators*. Translated from the 1990 and 1991 French originals by David Salinger. *Cambridge Studies in Advanced Mathematics*, 48. Cambridge University Press, Cambridge, 1997. [2](#)
25. A. Seeger, C. K. Smart and B. Street, *Multilinear singular integral forms of Christ-Journé type*. *Mem. Amer. Math. Soc.* vol. 257, No. 1231 (2019). [2](#)
26. E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*. *Princeton Mathematical Series*, No. 30, Princeton University Press, Princeton, N.J. 1970. [6](#), [9](#), [20](#), [25](#), [28](#)
27. E. M. Stein and G. Weiss, *Interpolation of operators with change of measures*. *Trans. Amer. Math. Soc.* 87 (1958), 159-172. [35](#)
28. E. M. Stein and G. Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*. *Princeton Mathematical Series*, No. 32, Princeton University Press, Princeton, N.J. 1971. [7](#), [34](#)

XUDONG LAI: INSTITUTE FOR ADVANCED STUDY IN MATHEMATICS, HARBIN INSTITUTE OF TECHNOLOGY, HARBIN, 150001, PEOPLE'S REPUBLIC OF CHINA

E-mail address: xudonglai@hit.edu.cn xudonglai@mail.bnu.edu.cn