MULTILINEAR ESTIMATES FOR CALDERÓN COMMUTATORS

XUDONG LAI

ABSTRACT. In this paper, we investigate the multilinear boundedness properties of the higher (n-th) order Calderón commutator for dimensions larger than two. We establish all multilinear endpoint estimates for the target space $L^{\frac{d}{d+n},\infty}(\mathbb{R}^d)$, including that Calderón commutator maps the product of Lorentz spaces $L^{d,1}(\mathbb{R}^d) \times \cdots \times L^{d,1}(\mathbb{R}^d) \times L^1(\mathbb{R}^d)$ to $L^{\frac{d}{d+n},\infty}(\mathbb{R}^d)$, which is the higher dimensional nontrivial generalization of the endpoint estimate that the n-th order Calderón commutator maps $L^1(\mathbb{R}) \times \cdots \times L^1(\mathbb{R}) \times L^1(\mathbb{R})$ to $L^{\frac{1}{1+n},\infty}(\mathbb{R})$. When considering the target space $L^r(\mathbb{R}^d)$ with $r < \frac{d}{d+n}$, some counterexamples are given to show that these multilinear estimates may not hold. The method in the present paper seems to have a wide range of applications and it can be applied to establish the similar results for Calderón commutator with a rough homogeneous kernel.

1. INTRODUCTION

The study of multilinear Calderón-Zygmund operators was initiated by Coifman and Meyer (see [8], [9], [20]). One of their motivations is to study the second order Calderón commutator (see [8]). Now a fruitful theory has grown around the multilinear Calderón-Zygmund operator and there are still many works on going, we refer to see the very nice exposition [18, Chapter 7] and the references therein. Despite of the intensive research of the multilinear Calderón-Zygmund theory, there are still some open problems related to Calderón commutators, the original model of multilinear Calderón-Zygmund operators. For example, there are no appropriate multilinear endpoint estimates of the higher order Calderón commutator for higher dimensions.

In this paper, we investigate the multilinear boundedness properties of the higher (n-th) order Calderón commutator for dimensions larger than two. We establish all multilinear endpoint estimates for the target space $L^{\frac{d}{d+n},\infty}(\mathbb{R}^d)$, in which the endpoint estimates exist on a plane $\frac{1}{q_1} + \cdots + \frac{1}{q_n} + \frac{1}{p} = \frac{d+n}{d}$ with $(\frac{1}{q_1}, \cdots, \frac{1}{q_n}, \frac{1}{p}) \in \mathbb{R}^{n+1}$ intersects $[0,1]^{n+1}$. Specially, these endpoint estimates include that Calderón commutator maps the product of Lorentz spaces $L^{d,1}(\mathbb{R}^d) \times \cdots \times L^{d,1}(\mathbb{R}^d) \times L^1(\mathbb{R}^d)$ to $L^{\frac{d}{d+n},\infty}(\mathbb{R}^d)$, which is the higher dimensional nontrivial generalization of the endpoint estimate that the n-th order Calderón commutator maps $L^1(\mathbb{R}) \times \cdots \times L^1(\mathbb{R}) \times L^1(\mathbb{R})$ to $L^{\frac{1}{1+n},\infty}(\mathbb{R})$. If the dimension $d = 1$, the above endpoint estimates for the n-th order commutators on products of $L^1(\mathbb{R})$ spaces have been obtained by C. P. Calderón [6] when $n = 1$,

by Coifman and Meyer [8] when $n = 1, 2$ and by Duong, Grafakos and Yan [14] when $n \geq 1$. However when the dimension $d \geq 2$, things become more complicated since Calderón commutator in this case is a non standard multilinear Calderón-Zygmund operator. No appropriate multilinear Calderón-Zygmund theory can be applied to it directly. Therefore it is interesting to establish the multilinear estimates of Caderón commutator for $d \geq 2$ and the purpose of the present paper is to develop the theory in this respect.

Before stating our results, we give some notation and the background. Define the higher (n-th) order Calderón commutator by

$$(1.1) \qquad \mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f](x) = \text{p.v.} \int_{\mathbb{R}^d} K(x-y) \Big(\prod_{i=1}^{n} \frac{A_i(x) - A_i(y)}{|x-y|} \Big) \cdot f(y) dy,$$

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:

$$(1.2) \qquad\qquad\qquad\qquad |K(x)| \lesssim |x|^{-d},$$

$$(1.3) \qquad \int_{R < |x| < 2R} K(x) \Big(\frac{x}{|x|} \Big)^{\alpha} dx = 0, \ \forall R > 0, \ \forall \alpha \in \mathbb{Z}_+^d \text{ with } |\alpha| = n,$$

$$(1.4) \qquad |K(x-y) - K(x)| \lesssim \frac{|y|^{\delta}}{|x|^{d+\delta}} \ \text{ for some } 0 < \delta \leq 1 \text{ if } |x| > 2|y|.$$

Such kind of commutator was first introduced by A. P. Calderón in [3] for the first order with $K(x)$ a homogeneous kernel and also later in [4] [5] for the higher order one (see also [8], [9]). It is easy to see that $\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f](x)$ is well defined for $A_1, \cdots, A_n, f \in C_c^{\infty}(\mathbb{R}^d)$. For its applications, let us look at the first order Calderón commutator (1.1). Indeed $\mathcal{C}[\nabla A, f](x)$ is a generalization of

$$(1.5) \quad [A, S]f(x) = A(x)S(f)(x) - S(Af)(x) = -\text{p.v.} \frac{1}{\pi} \int_{\mathbb{R}} \frac{1}{x-y} \frac{A(x) - A(y)}{x-y} f(y) dy$$

where $S = \frac{d}{dx} \circ H$ and H denotes the Hilbert transform (one can deduce $S = \frac{d}{dx} \circ H$ just by taking a derivation into the kernel $\frac{1}{\pi x}$ or utilizing the Fourier transform for both sides). It is well known that the commutator $[A, S]$ is a fundamental operator in harmonic analysis and plays 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, and the Kato square root problem on $\mathbb{R}$ (see *e.g.* [3], [5], [15], [20], [18] for the details). Recently, there has been a renewed interest into the commutator $[A, S]$ and *d-commutator* introduced by M. Christ and J. Journé (see [10]) since they have applications in the mixing flow problem (see *e.g.* [21], [19]).

In this paper, we are interested in the following strong type multilinear estimate (or weak type estimate)

$$(1.6) \qquad \|\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f]\|_{L^r(\mathbb{R}^d)} \lesssim \Big(\prod_{i=1}^{n} \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d)} \Big) \|f\|_{L^p(\mathbb{R}^d)}$$

where $\frac{1}{r} = \big(\sum_{i=1}^{n} \frac{1}{q_i} \big) + \frac{1}{p}$ with $1 \leq q_i \leq \infty$, $(i = 1, \cdots, n)$ and $1 \leq p \leq \infty$. Our main results are as follows.

Theorem 1.1. *Let $d \geq 2$ and n be a positive integer. Suppose K satisfies* (1.2), (1.3) *and* (1.4). *Assume that $\frac{1}{r} = \left(\sum_{i=1}^{n} \frac{1}{q_i}\right) + \frac{1}{p}$ with $1 \leq q_i \leq \infty$ $(i = 1, \cdots, n)$, and $1 \leq p \leq \infty$. We have the following conclusions:*

(i). If $\frac{d}{d+n} < r < \infty$, $1 < q_i \leq \infty$ $(i = 1, \cdots, n)$ and $1 < p \leq \infty$, then the multilinear estimate (1.6) *holds.*

(ii). If $\frac{d}{d+n} \leq r < \infty$ with $q_i = 1$ for some $i = 1, \cdots, n$; or $p = 1$; or $r = \frac{d}{d+n}$, then there exists a constant $C > 0$ such that

$$(1.7) \qquad \|\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f]\|_{L^{r,\infty}(\mathbb{R}^d)} \leq C \Big(\prod_{i=1}^{n} \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d)} \Big) \|f\|_{L^p(\mathbb{R}^d)}$$

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

$$(1.8) \qquad \|\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f]\|_{L^{\frac{d}{d+n},\infty}(\mathbb{R}^d)} \leq C \Big(\prod_{i=1}^{n} \|\nabla A_i\|_{L^{d,1}(\mathbb{R}^d)} \Big) \|f\|_{L^1(\mathbb{R}^d)}.$$

(iii). If $0 < r < \frac{d}{d+n}$, $1 \leq q_i \leq \infty$ $(i = 1, \cdots, n)$ and $1 \leq p \leq \infty$, there exist functions A_i for $i = 1, \cdots, n$, and f such that $\|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d)} < \infty$ for $i = 1, \cdots, n$, and $\|f\|_{L^p(\mathbb{R}^d)} < \infty$. But

$$\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f](x) = \infty \quad \text{in a ball in } \mathbb{R}^d.$$

Remark 1.2. Notice that (i) gives strong type estimates (1.6) for $\frac{d}{d+n} < r < \infty$. (ii) gives all endpoint estimates for $\frac{d}{d+n} \leq r \leq 1$, especially the case $r = \frac{d}{d+n}$ where the endpoints $\left(\frac{1}{q_1}, \cdots, \frac{1}{q_n}, \frac{1}{p}\right)$ exist in the intersection between the plane $\left(\sum_{i=1}^{n} \frac{1}{q_i}\right) + \frac{1}{p} = \frac{d+n}{d}$ and $[0,1]^{n+1}$, which is the most difficult part in our proof. Here we point out that the condition $r \geq \frac{d}{d+n}$ is crucial in the proof of (i) and (ii) in Theorem 1.1, which will be emphasized further in the proof where we use this condition. Our basic strategy is first to show (1.6) for $1 \leq r < \infty$ and (ii), then use the multilinear interpolation between (1.6) for $1 \leq r < \infty$ and the result of (ii), to justify the rest part of (1.6) for $\frac{d}{d+n} < r < 1$. All those will be clear in our proof. Obviously, the conclusion of (iii) indicates that the requirement $r \geq \frac{d}{d+n}$ is a necessary condition to guarantee the strong type estimates (or weak type estimates) (1.6) hold, thus our results in Theorem 1.1 are optimal in this sense. Some counterexamples will be constructed to prove conclusion (iii).

Remark 1.3. Notice that $L^{1,1}(\mathbb{R}) = L^1(\mathbb{R})$. Therefore when the dimension $d = 1$, (1.8) turns out to be the n-th Calderón commutator mapping $L^1(\mathbb{R}) \times \cdots \times L^1(\mathbb{R}) \times L^1(\mathbb{R})$ to $L^{\frac{1}{1+n},\infty}(\mathbb{R})$, which has been previously proved by Duong, Grafakos and Yan [14]. To the best knowledge of the author, (1.8) is new when $d \geq 2$. Currently we still do not know whether $L^{d,1}(\mathbb{R}^d)$ in (1.8) could be replaced by $L^{d,1+\varepsilon}(\mathbb{R}^d)$ for some $\varepsilon > 0$ when $d \geq 2$ and we will further explore this problem in our future research.

We next briefly introduce the methods employed and the main procedures in the proof of Theorem 1.1. We first establish the assertion (i) of Theorem 1.1 in the case $1 \leq r < \infty$ based on the recent deep result of A. Seeger, C. K. Smart and B. Street

in [21]. Next we show that if $q_i = \infty$ with $i = 1, \cdots, n$ and $p = 1$, i.e. A_i is a Lipschitz function, then the weak type $L^{1,\infty}(\mathbb{R}^d)$ boundedness holds by the standard Calderón-Zygmund theory. We will devote to proving (ii), i.e. we need to give a weak type estimate. In the case of (ii), by our condition, A_i satisfies $\nabla A_i \in L^{q_i}(\mathbb{R}^d)$. We will construct an *exceptional set* which satisfies the required weak type estimate. And on the complementary set of *exceptional set*, the function A_i is a Lipschitz function. Then, roughly speaking, the strong type estimate in (i) and the weak type $L^{1,\infty}(\mathbb{R}^d)$ boundedness of $\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f](x)$ could be applied on the complementary set of *exceptional set*. The idea partly comes from C. P. Calderón [6], [7]. However we develop further more here. Our argument works once we establish the strong type estimate (1.6) when $1 < r < \infty$, $1 < q_1, \cdots, q_n \leq \infty$, $1 \leq p \leq \infty$ and weak type $L^{1,\infty}(\mathbb{R}^d)$ boundedness when $r = 1$, $q_1 = \cdots = q_n = \infty$, $p = 1$.

The strategy to construct the *exceptional set* is as follows. Notice that the estimate $\|\nabla A\|_{L^q(\mathbb{R}^d)}$ is related to the Sobolev space $W^{1,q}(\mathbb{R}^d)$. When $1 \leq q < d$, it is well known that Sobolev space $W^{1,q}(\mathbb{R}^d)$ is embedded into $L^{q^*}(\mathbb{R}^d)$ with $\frac{1}{q^*} = \frac{1}{q} - \frac{1}{d}$. This property is crucial to help us establish a boundedness property of maximal operator (see Lemma 2.4). When $q > d$, *exceptional set* can be constructed by using the Mary Weiss maximal operator $\mathcal{M}$ (see Subsection 2.1 for its definition), which maps $L^q(\mathbb{R}^d)$ to $L^q(\mathbb{R}^d)$ (or $L^{q,\infty}(\mathbb{R}^d)$) only when $q > d$. But when $q = d$, the critical Sobolev $W^{1,d}(\mathbb{R}^d)$ is imbedded into an Orlicz space (see [1]) which may be not useful to us. This forces us to study the Mary Weiss maximal operator on $L^d(\mathbb{R}^d)$, which is quite challenging. Fortunately, we find a substitute that $\mathcal{M}$ maps the Lorentz space $L^{d,1}(\mathbb{R}^d)$ to $L^{d,\infty}(\mathbb{R}^d)$ which is enough to construct an *exceptional set*. Base on this, we can establish the multilinear endpoint estimate that $\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f](x)$ maps $L^{d,1}(\mathbb{R}^d) \times \cdots \times L^{d,1}(\mathbb{R}^d) \times L^1(\mathbb{R}^d)$ to $L^{\frac{d}{d+n},\infty}(\mathbb{R}^d)$. 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 (1.8) here is quite different from that by Duong, Grafakos and Yan [14], thus we give a new proof of (1.8) for $d = 1$.

As aforementioned, the above method built in this paper works as long as we establish the strong type estimate (1.6) when $1 < r < \infty$ and weak type $L^{1,\infty}(\mathbb{R}^d)$ boundedness when $r = 1$, $q_1 = \cdots = q_n = \infty$, $p = 1$. Therefore we can use the method here to establish the similar multilinear estimates of Calderón commutator with a homogeneous rough kernel. Define the higher order Calderón commutator with a rough kernel by

$$\mathcal{C}_\Omega[\nabla A_1, \cdots, \nabla A_n, f](x) = \text{p.v.} \int_{\mathbb{R}^d} \frac{\Omega(x-y)}{|x-y|^d} \Big(\prod_{i=1}^n \frac{A_i(x) - A_i(y)}{|x-y|} \Big) \cdot f(y) dy,$$

here Ω is a function defined on $\mathbb{R}^d \setminus \{0\}$ which satisfies:

(1.9) $\Omega(r\theta) = \Omega(\theta)$ for $r > 0$, $\theta \in \mathbf{S}^{d-1}$; $\Omega(-\theta) = (-1)^{n+1}\Omega(\theta)$[1]

and $\Omega \in L^1(\mathbf{S}^{d-1})$. $\mathbf{S}^{d-1}$ is the unit sphere in $\mathbb{R}^d$. Similar to those in Theorem 1.1, we have the following result.

---

[1]One may also consider the case $\Omega(-\theta) = (-1)^n\Omega(\theta)$ with some other moment cancelation conditions, we refer to see Remark 3.2 for further discussion.

Theorem 1.4. *Suppose Ω satisfies* (1.9) *and* $\Omega \in L \log^+ L(\mathbf{S}^{d-1})$ *for $d \geq 2$. Then all the results in Theorem 1.1 also hold for* $\mathcal{C}_\Omega[\nabla A_1, \cdots, \nabla A_n, f](x)$.

When $n = 1$, part of results in Theorem 1.4 have been established by A. P. Calderón [3] and C. P. Calderón [6] [7]. We summarize their results [3], [6], [7] in Figure 1. More precisely, A. P. Calderón [3] showed that if $\frac{1}{r} = \frac{1}{q} + \frac{1}{p}$ with $1 < r < \infty$, $1 < q \leq \infty$, $1 < p < \infty$, then (1.6) holds when $\Omega \in L \log^+ L(\mathbf{S}^{d-1})$ (see the region with diagonal lines in Figure 1). Later C. P. Calderón [6] extended these results to the boundary of the region with diagonal lines where he proved (1.6) is still true in the case $1 < r = q < \infty$, $p = \infty$ and in the case $r = 1, q > 1, p > 1$. C. P. Calderón [6] also showed that if Ω satisfies the Hörmander condition, then (1.6) holds when $d/(d+1) < r < 1$, $q > d$, $p > 1$ (see the region with vertical lines in Figure 1). In [7], C. P. Calderón showed that if $\frac{d}{d+1} \leq r \leq 1$, $1 \leq q < d$, $1 < p \leq \infty$, then the weak type estimate (1.7) holds when $\Omega \in L \log^+ L(\mathbf{S}^{d-1})$ (see the region with horizontal lines in Figure 1). With the above results in hand, by using the interpolation arguments, one may easily get the strong type estimate (1.6) holds for $\frac{d}{d+n} < r < 1$, $1 < q < \infty$, $1 < p < \infty$ if $\Omega \in L \log^+ L(\mathbf{S}^{d-1})$. Recently Fong [16] considered the special case $\Omega \equiv 1$ and used the time-frequency analysis method to show (1.6) holds for $\frac{d}{d+n} < r < \infty$, $1 < q < \infty$, $1 < p < \infty$.
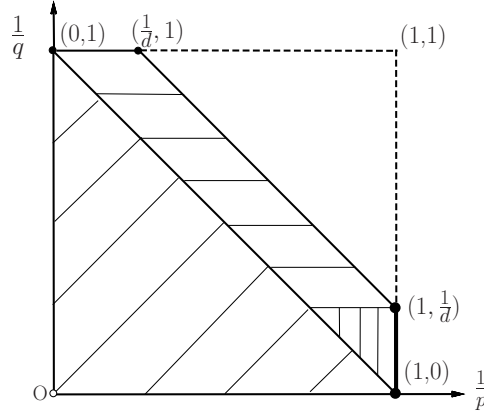

FIGURE 1. In the case $n = 1$, our main results in Theorem 1.4 are new when $0 < \frac{1}{q} \leq \frac{1}{d}$ and $\frac{1}{p} = 1$, see the **bold** line including the endpoint $(\frac{1}{p}, \frac{1}{q}) = (1, \frac{1}{d})$.

For the endpoint $(\frac{1}{p}, \frac{1}{q}) = (1, 0)$, the weak type $L^{1,\infty}(\mathbb{R}^d)$ boundedness of $\mathcal{C}_\Omega[\nabla A, f]$ with $\Omega \in L \log^+ L(\mathbf{S}^{d-1})$ has been recently derived by Ding and the author [11]. The contribution of Theorem 1.4 in the case $n = 1$ is the estimates with $p = 1$, $0 < \frac{1}{q} \leq \frac{1}{d}$ (see the bold line including the endpoint $(\frac{1}{p}, \frac{1}{q}) = (1, \frac{1}{d})$ in Figure 1), which complements the aforementioned works for $n = 1$. To the best knowledge of the author, Theorem 1.4 is new when $n \geq 2$.

An immediate consequence of Theorem 1.1 or 1.4 is the following n-th order commutator of the Riesz transform with n-th derivation which may have potential applications in partial differential equations.

Corollary 1.5. *Let R_j be the Riesz transform. Then all the results in Theorem 1.1 also hold for the following operator*

$$[A_1, \cdots, [A_n, \partial^\alpha \circ R_j] \cdots] f(x)$$
$$= \text{p.v.} \int_{\mathbb{R}^d} \partial_x^\alpha \Big(\frac{x_j - y_j}{|x - y|^{d+1}} \Big) \cdot \Big(\prod_{i=1}^n [A_i(x) - A_i(y)] \Big) \cdot f(y) dy$$

where $\alpha \in \mathbb{Z}_+^d$ is a multi-indice with $|\alpha| = n$.

This paper is organized as follows. In Section 2, we give the proof of Theorem 1.1, which will be divided into several cases. First some preliminary lemmas are presented in Subsection 2.1. Subsection 2.2 is devoted to proving (i) of Theorem 1.1 in the case $1 \le r < \infty$ and weak type $L^{1,\infty}(\mathbb{R}^d)$ boundedness on $L^\infty(\mathbb{R}^d) \times \cdots \times L^\infty(\mathbb{R}^d) \times L^1(\mathbb{R}^d)$. The proofs of (ii) in Theorem 1.1 are given in Subsections 2.3, 2.4 and 2.5. In Subsection 2.6, we proceed to proving the rest part of (i) in Theorem 1.1 by the multilinear interpolation theorem. Finally some counterexamples are given in Subsection 2.7 to prove (iii) in Theorem 1.1. The proof of Theorem 1.4 is similar to that of Theorem 1.1. So in Section 3, we outline the proof of Theorem 1.4.

Notation. Throughout this paper, we only consider the dimension $d \ge 2$ and the letter C stands for a positive finite constant which is independent of the essential variables and not necessarily the same one in each occurrence. $A \lesssim B$ means $A \le CB$ 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 indices $r, q_1, \cdots, q_n$ and p satisfy $\frac{1}{r} = \Big(\sum_{i=1}^n \frac{1}{q_i} \Big) + \frac{1}{p}$ with $1 \le q_i \le \infty$ $(i = 1, \cdots, n)$ and $1 \le p \le \infty$ in the whole paper. For a set $E \subset \mathbb{R}^d$, we denote by $|E|$ or $m(E)$ the Lebesgue measure of E. $\mathbf{S}^{d-1}$ is the unit sphere in $\mathbb{R}^d$. $d\sigma$ denotes the spherical measure on $\mathbf{S}^{d-1}$. ∇A will stand for the vector $(\partial_1 A, \cdots, \partial_d A)$ where $\partial_i A(x) = \partial A(x)/\partial x_i$. Define

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

for $X = L^p(\mathbb{R}^d)$ or $X = L^{d,1}(\mathbb{R}^d)$. $\mathbb{Z}_+$ denotes the set of all nonnegative integers and $\mathbb{Z}_+^d = \underbrace{\mathbb{Z}_+ \times \cdots \times \mathbb{Z}_+}_{d}$.

2. Proof of Theorem 1.1

2.1. **Some preliminary lemmas.**

Before giving the proof of Theorem 1.1, we introduce some lemmas which play a key role in the proof of Theorem 1.1. For those readers who are not familiar with the theory of the Lorentz space $L^{p,q}(\mathbb{R}^d)$, we refer to see [23, Chapter V.3]. We will use the theory of the Lorentz space $L^{p,q}(\mathbb{R}^d)$ in Lemma 2.2. Now we begin by some properties of a special maximal function which was introduced by Mary Weiss (see [6]). It is defined

as

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

Lemma 2.1. *Let $\nabla A \in L^p(\mathbb{R}^d)$ with $p > d$. Then $\mathcal{M}$ is bounded on $L^p(\mathbb{R}^d)$, that is*

$$\|\mathcal{M}(\nabla A)\|_{L^p(\mathbb{R}^d)} \leq C\|\nabla A\|_{L^p(\mathbb{R}^d)},$$

where the constant C is independent of A.

Proof. By using a standard limiting argument, we only need to consider A as a C^∞ function with compact support. Then the lemma just follows from the inequality

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

which holds for any $q > d$ (see [6, Lemma 1.4]) and the fact that the Hardy-Littlewood maximal operator is of strong type (p, p) for $p > 1$. $\square$

Lemma 2.2. *Let $\nabla A \in L^{d,1}(\mathbb{R}^d)$, the standard Lorentz space. Then for any $\lambda > 0$, there exist a finite constant C independent of A such that*

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

Proof. It suffices to consider A as a smooth function with compact support. By the formula given in [22, page 125, (17)], we may write

$$A(x) = C_d \sum_{i=1}^d \int_{\mathbb{R}^d} \frac{x_i - y_i}{|x - y|^d} \partial_i A(y) dy = \mathbb{K} * f(x)$$

where $\mathbb{K}(x) = 1/|x|^{d-1}$, $f = C_d \sum_{j=1}^d R_j(\partial_j A)$ with R_j the Riesz transforms. By using the fact the Riesz transform R_j maps $L^{d,1}(\mathbb{R}^d)$ to itself which follows from the general form of the Marcinkiewicz interpolation theorem (see [23, Theorem 3.15 in page 197]), one can easily get that

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

Hence to prove the lemma, it is enough to show that

$$(2.1) \qquad \lambda^d |\{x \in \mathbb{R}^d : \mathcal{M}(\nabla A)(x) > \lambda\}| \lesssim \|f\|_{L^{d,1}(\mathbb{R}^d)}^d$$

with $A = \mathbb{K} * f$. In the following our goal is to prove that for any $x \in \mathbb{R}^d$, the estimate

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

holds uniformly for $h \in \mathbb{R}^d \backslash \{0\}$ with T an operator maps $L^{d,1}(\mathbb{R}^d)$ to $L^{d,\infty}(\mathbb{R}^d)$. Once we prove this, we get (2.1) and hence complete the proof of Lemma 2.2. We write

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

Let us first consider I. By an elementary calculation, one may get $\mathbb{K} \in L^{d',\infty}(\mathbb{R}^d)$ where $d' = d/(d-1)$. Set $B(x,r) = \{y \in \mathbb{R}^d : |x-y| \leq r\}$. Using the rearrangement inequality (see [17, page 74, Exercise 1.4.1]), we have

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

here f^* represents the decreasing rearrangement of f. Using the definition of Lorentz space, one may get $\|\chi_E\|_{L^{d,1}(\mathbb{R}^d)} = \|\chi_E\|_{L^d(\mathbb{R}^d)}$ holds for any characteristic function χ_E of set E of finite Lebesgue measure, thus $\|\chi_{B(x,2|h|)}\|_{L^{d,1}(\mathbb{R}^d)} = C_d|h|$. Therefore we get

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

Below we need to show that the operator Λ maps $L^{d,1}(\mathbb{R}^d)$ to $L^{d,\infty}(\mathbb{R}^d)$, which can be found in [24]. Since the proof is short, for completeness, we also give a proof here. Note that $L^{d,1}(\mathbb{R}^d)$ is a Banach space (see [23, page 204, Theorem 3.22]), it is sufficient to show that Λ maps the characteristic function $\chi_E \in L^{d,1}(\mathbb{R}^d)$ to $L^{d,\infty}(\mathbb{R}^d)$ (see [17, page 62, Lemma 1.4.20]). However in this case, it is equivalent to show that

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

where M is the Hardy-Littlewood maximal operator. It is well known that M is of weak type $(1,1)$, hence we have shown that Λ maps $L^{d,1}(\mathbb{R}^d)$ to $L^{d,\infty}(\mathbb{R}^d)$.

Next we consider II. This estimate is quite simple. Since the kernel $k(x) = \varepsilon^{-1}|x|^{-d+1}\chi_{\{|x|\leq\varepsilon\}}$ is a radial non-increasing function and L^1 integrable in $\mathbb{R}^d$, we get

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

Notice that $L^{p,1}(\mathbb{R}^d) \subset L^p(\mathbb{R}^d)$ and M is of strong type (p,p), $1 < p < \infty$, of course those imply that M maps $L^{d,1}(\mathbb{R}^d)$ to $L^{d,\infty}(\mathbb{R}^d)$.

Finally we give an estimate of III. Notice that we only consider $|x-y| > 2|h|$. Then by the Taylor expansion of $|x-y+h|^{-d+1}$, one may have

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

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

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

Inserting (2.2) 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 which is defined by

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

Since R_j^* is bounded on $L^p(\mathbb{R}^d)$, $1 < p < \infty$, one immediately gets that R_j^* maps $L^{d,1}(\mathbb{R}^d)$ to $L^{d,\infty}(\mathbb{R}^d)$. The second term which controls III can be dealt with the same way as we do in the estimate of II once we notice that the function $\varepsilon |x|^{-d-1} \chi_{\{|x| > \varepsilon\}}$ is radial non-increasing and L^1 integrable. $\qquad\square$

Remark 2.3. Here it should be pointed out that some idea in this proof is similar to that in [24], where E. M. Stein proved that for a function F defined in $\mathbb{R}^d$ with $\nabla F \in L_{loc}^{d,1}(\mathbb{R}^d)$, then F is equivalent with a continuous function and

(2.3) $\qquad F(x+h) - F(x) - h(\nabla F)(x) = o(|h|)$ for almost every x,

as $|h| \to 0$. The method of proving (2.3) in [24] is just giving a direct estimate of (2.3). See also another proof by using elementary principle in [12], [13]. The property of the maximal operator $\mathcal{M}$ that maps $L^{d,1}(\mathbb{R}^d)$ to $L^{d,\infty}(\mathbb{R}^d)$ seems to be more powerful since it implies (2.3) immediately. In fact, using the dense limiting arguments and Lemma 2.2, we get for any function F defined in $\mathbb{R}^d$ with $\nabla F \in L^{d,1}(\mathbb{R}^d)$,

$$\lim_{s \to 0} \frac{F(x + s\theta) - F(x)}{s} = (\nabla F)(x) \cdot \theta, \text{ for any } \theta \in \mathbf{S}^{d-1}, \text{ a.e. } x \in \mathbb{R}^d,$$

which is inequivalent to (2.3).

Lemma 2.4. *Let $\nabla A \in L^p(\mathbb{R}^d)$ with $1 \le p < d$. Set $1/s = 1/p - 1/d$. Define the maximal operator $\mathfrak{M}_s$ and the Hardy-Littlewood maximal operator of order p M_p by*

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

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

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

$$\mathfrak{M}_s(\nabla A)(x) \lesssim M_p(\nabla A)(x).$$

Proof. We refer to see [7, Lemma 3.2] and its proof there from line (3.2.2) to (3.2.7). $\qquad\square$

Lemma 2.5. *Let $\{Q_k\}_k$ be the disjoint cubes in $\mathbb{R}^d$. Denote by $l(Q_k)$ the side length of Q_k. Suppose Ω satisfies (1.9). Define the operator T_s as*

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

Then for any $s > 0$ and $1 \le q \le \infty$, we get that,

$$\|T_s(f)\|_{L^q(\mathbb{R}^d)} \lesssim \|\Omega\|_{L^1(\mathbf{S}^{d-1})} \|f\|_{L^q(\mathbb{R}^d)}.$$

Proof. If $q = 1$, Lemma 2.5 just follows from the Fubini theorem. In fact, we have

$$\|T_s(f)\|_{L^1(\mathbb{R}^d)} \leq \sum_{Q_k} \int_{Q_k} \Big| \int_{\mathbb{R}^d} \frac{|\Omega(x-y)| \cdot l(Q_k)^s}{[l(Q_k) + |x-y|]^{d+s}} dx \Big| \cdot |f(y)| dy \lesssim \|\Omega\|_{L^1(\mathbf{S}^{d-1})} \|f\|_{L^1(\mathbb{R}^d)},$$

here we use that Q_ks are cubes disjoint each other. If $q = \infty$, applying the Fubini theorem again,

$$|T_s(f)(x)| \leq \sum_{Q_k} \|f\|_{L^\infty(Q_k)} \sup_{x \in \mathbb{R}^d} \int_{Q_k} \frac{|\Omega(x-y)| \cdot l(Q_k)^s}{[l(Q_k) + |x-y|]^{d+s}} dy \lesssim \|\Omega\|_{L^1(\mathbf{S}^{d-1})} \|f\|_{L^\infty(\mathbb{R}^d)}.$$

Now using the Marcinkiewicz interpolation theorem (see *e.g.* [23]), one may get T_s maps $L^q(\mathbb{R}^d)$ to $L^q(\mathbb{R}^d)$ for any $1 < q < \infty$. Hence we complete the proof. $\square$

In the following, we begin to give the proof of Theorem 1.1. We will first show our theorem for $r \geq 1$ which is not quite complicated. Define the multi-indice set

$$\mathrm{MI} = \Big\{ (\frac{1}{q_1}, \cdots, \frac{1}{q_n}, \frac{1}{p}) : \frac{1}{r} = \big(\sum_{i=1}^n \frac{1}{q_i} \big) + \frac{1}{p}, \frac{d}{d+n} \leq r < \infty, 1 \leq q_1, \cdots, q_n, p \leq \infty \Big\}.$$

If $\frac{d}{d+n} \leq r \leq 1$, we will divide the proof into several cases according whether q_i is bigger than d or smaller than d. And in this case, we will establish the weak type estimate at all boundary points of MI. Although we don't take a rigorous classification, we will cover all cases for $\frac{d}{d+n} \leq r < \infty$, $1 \leq q_1, \cdots, q_n \leq \infty$ and $1 \leq p \leq \infty$. Next we will use the multilinear interpolation to establish the strong type estimate in the interior of MI between $r > 1$ and $r = \frac{d}{d+n}$. Finally, we give some examples to show that if $0 < r < \frac{d}{d+n}$, there are no multilinear strong type estimates like (1.6) (or weak type estimates).

2.2. **Case: $1 \leq r < \infty$.**

Proposition 2.6. Let $1 \leq r < +\infty$, $1 < q_i \leq \infty$, $i = 1, \cdots, n$, $1 < p \leq \infty$. Then the strong type estimate (1.6) holds.

Proof. We do not plan to give a direct proof here. The proof relies on the recent deep results in [21]. In fact, by using the mean value formula, one may get

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

For each $i = 1, \cdots, n$, plunge the above equality into $\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f](x)$ and write it as follows:

$$\mathrm{p.v.} \int_{\mathbb{R}^d} K(x-y) \Big(\prod_{i=1}^n \Big[\sum_{j=1}^d \frac{x_j - y_j}{|x-y|} \int_0^1 \partial_j A_i(sx + (1-s)y) ds \Big] \Big) \cdot f(y) dy.$$

Then by the moment cancelation condition (1.3), the bound condition (1.2) and the regularity condition (1.4), for any multi-indice $\alpha \in \mathbb{Z}_+^d$ with $|\alpha| = n$, $K(x)(x/|x|)^\alpha$ is a standard Calderón-Zygmund kernel. Therefore the proof reduces to show that the

following operator

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

maps $L^{q_1}(\mathbb{R}^d) \times \cdots \times L^{q_n}(\mathbb{R}^d) \times L^p(\mathbb{R}^d)$ to $L^r(\mathbb{R}^d)$, where k is a standard Calderón-Zygmund kernel and $m_{x,y}a = \int_0^1 a(sx + (1-s)y)dy$. However, this estimate has been proved by A. Seeger, C. K. Smart and B. Street in [21]. $\qquad\square$

Proposition 2.7. Let $r = 1$, $q_1 = \cdots = q_n = \infty$, $p = 1$. Then

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

Proof. When $q_1 = \cdots = q_n = \infty$, A_i is a Lipschitz function for $i = 1, \cdots, n$. Fix all A_i. We may regard $\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f](x)$ as a linear function of f. Then the kernel

$$K(x,y) =: K(x-y)\Big(\prod_{i=1}^n \frac{A_i(x) - A_i(y)}{|x-y|}\Big)$$

is a standard Calderón-Zygmund kernel (see *e.g.* [18, Page 211, Definition 4.1.2])which in fact satisfies the boundedness condition $|K(x,y)| \lesssim (\prod_{i=1}^n \|\nabla A_i\|_{L^\infty(\mathbb{R}^d)})|x-y|^{-d}$ and the following regularity conditions

$$|K(x_1,y) - K(x_2,y)| \lesssim (\prod_{i=1}^n \|\nabla A_i\|_{L^\infty(\mathbb{R}^d)})\frac{|x_1-x_2|^\delta}{|x_1-y|^{d+\delta}} \quad \text{for } |x_1 - y| > 2|x_1 - x_2|,$$

$$|K(x,y_1) - K(x,y_2)| \lesssim (\prod_{i=1}^n \|\nabla A_i\|_{L^\infty(\mathbb{R}^d)})\frac{|y_1-y_2|^\delta}{|x-y_1|^{d+\delta}} \quad \text{for } |x - y_1| > 2|y_1 - y_2|.$$

Therefore by Proposition 2.6 with $q_1 = \cdots = q_n = \infty, p = 2$ and the standard Calderón-Zygmund theory (see *e.g.* [18, Page 226, Theorem 4.2.2]), we may get that the operator $\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, \cdot]$ is of weak type (1,1) with bound $\prod_{i=1}^n \|\nabla A_i\|_{L^\infty(\mathbb{R}^d)}$, thus we complete the proof. $\qquad\square$

2.3. **Case:** $d/(d+n) \leq r \leq 1$ **and** $d \leq q_1, \cdots, q_n \leq \infty$.

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

Proposition 2.8. Let $d/(d+n) \leq r \leq 1$, $d < q_1, \cdots, q_l \leq \infty$ and $q_{l+1}, \cdots, q_n = d$ with $0 \leq l \leq n$, $1 \leq p < \infty$. Then

$$\|\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f]\|_{L^{r,\infty}(\mathbb{R}^d)}$$

(2.4)
$$\lesssim \Big(\prod_{i=1}^l \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d)}\Big)\Big(\prod_{i=l+1}^n \|\nabla A_i\|_{L^{d,1}(\mathbb{R}^d)}\Big)\|f\|_{L^p(\mathbb{R}^d)},$$

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

Proof. By using a standard limiting argument, we only need to show that when A_i $(i = 1, \cdots, n)$ and f are C^∞ functions with compact supports, the following inequality

$$m(\{x \in \mathbb{R}^d : |\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f](x)| > \lambda\})$$
$$\lesssim \lambda^{-r} \Big(\prod_{i=1}^{l} \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d)}^r \Big) \Big(\prod_{i=l+1}^{n} \|\nabla A_i\|_{L^{d,1}(\mathbb{R}^d)}^r \Big) \|f\|_{L^p(\mathbb{R}^d)}^r,$$

holds for any $\lambda > 0$. By a simple scaling argument, we may assume that

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

for $i = 1, \cdots, l$ and $j = l+1, \cdots, n$. Fix $\lambda > 0$. For convenience we set

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

We need to show $|E_\lambda| \lesssim \lambda^{-r}$. First suppose that all $q_1, \cdots, q_l < \infty$. Once we have understood the proof in this situation, we can modify the proof to the other case that there exist some $q_i = \infty$ for $i = 1, \cdots, l$. We shall show how to do this in the last part of the proof. Define the *exceptional set*

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

for $i = 1, \cdots, n$. Here it should be pointed out that if $q_i = \infty$, the above definition is meaningless. Therefore we need to assume all $q_i < \infty$ firstly. From Lemma 2.1 and Lemma 2.2, $\mathcal{M}$ maps $L^p(\mathbb{R}^d)$ to itself for $p > d$ and maps $L^{d,1}(\mathbb{R}^d)$ to $L^{d,\infty}(\mathbb{R}^d)$, i.e.

(2.6)
$$|J_{i,\lambda}| \lesssim \lambda^{-r} \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d)}^{q_i} = \lambda^{-r}, \quad i = 1, \cdots, l;$$
$$|J_{j,\lambda}| \lesssim \lambda^{-r} \|\nabla A_j\|_{L^{d,1}(\mathbb{R}^d)}^d = \lambda^{-r}, \quad j = l+1, \cdots, n.$$

Set $J_\lambda = \cup_{i=1}^n J_{i,\lambda}$. Choose an open set G_λ which satisfies the following conditions: (1) $J_\lambda \subset G_\lambda$; (2) $m(G_\lambda) \leq 2|J_\lambda|$. By the property (2.6) of $J_{i,\lambda}$, we see that $m(G_\lambda) \lesssim \lambda^{-r}$. Next making a Whitney decomposition of G_λ (see *e.g.* [17]), one may get a family of disjoint dyadic cubes $\{Q_k\}_k$ such that

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

With those properties (i) and (ii), for each Q_k, we could construct a larger cube Q_k^* so that $Q_k \subset Q_k^*$, Q_k^* is centered at y_k and $y_k \in (G_\lambda)^c$, $|Q_k^*| \approx |Q_k|$. 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

(2.7) $$dist(y_k, Q_k) \approx l(Q_k).$$

Now we return 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 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 $i = 1, \cdots, n$. Let $\tilde{A}_i$ stand for the Lipschitz extension of A_i from $(G_\lambda)^c$ to $\mathbb{R}^d$ (see [22, page 174, Theorem 3]) so that for each $i = 1, \cdots, n$,

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

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

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

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

The first term above satisfies $|10G_\lambda| \lesssim \lambda^{-r}$, which is the required bound. In the following, we only consider $x \in (10G_\lambda)^c$. By the definition of f_1, one may see that

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

With this equality in hand, Proposition 2.6 ($1 < p < \infty$) and Proposition 2.7 ($p = 1$) imply

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

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 $\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f_2](x)$. Define $\mathbb{N}_i^j = \{i, i+1, \cdots, j\}$. Recall our construction of G_λ, y_k, Q_k and Q_k^* in the paragraph above (2.7). Then we can write $f_2 = \sum_k f\chi_{Q_k}$. Therefore we may get

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

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

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

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

$$I(x,y) = \prod_{i=1}^{n} \frac{\tilde{A}_i(x) - \tilde{A}_i(y)}{|x-y|},$$

(2.10)
$$II(x,y,y_k) = \sum_{N_1 \subsetneq \mathbb{N}_1^n} \Big(\prod_{i \in N_1} \frac{\tilde{A}_i(x) - \tilde{A}_i(y)}{|x-y|} \Big) \times$$
$$\times \Big(\prod_{i \in N_2} \frac{\tilde{A}_i(y) - \tilde{A}_i(y_k)}{|x-y|} \Big) \Big(\prod_{i \in N_3} \frac{A_i(y_k) - A_i(y)}{|x-y|} \Big).$$

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

Estimate of $\mathcal{C}[\cdots, \cdot]$ *related to* I. This estimate is similar to (2.9). In fact, in this case there is only one term $\mathcal{C}[\nabla \tilde{A}_1, \cdots, \nabla \tilde{A}_n, f_2]$. Then by Proposition 2.6 ($1 < p < \infty$) and Proposition 2.7 ($p = 1$), we get

$$m\big(\big\{x \in (10G_\lambda)^c : |\mathcal{C}[\nabla \tilde{A}_1, \cdots, \nabla \tilde{A}_n, f_2](x)| > \lambda/2\big\}\big)$$
$$\lesssim \lambda^{-p} \Big(\prod_{i=1}^{n} \|\nabla \tilde{A}_i\|_{L^\infty(\mathbb{R}^d)}^p \Big) \|f_2\|_{L^p(\mathbb{R}^d)}^p \lesssim \lambda^{-p+p\sum_{i=1}^n \frac{r}{q_i}} = \lambda^{-r}.$$

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

Estimate of $\mathcal{C}[\cdots, \cdot]$ *related to* II. It suffices to consider one term $\mathcal{C}[\cdots, \cdot]$ related to II in which N_1 is a proper subset of $\mathbb{N}_1^n$. In this case, without loss of generality, we may assume $N_1 = \{1, \cdots, v\}$, $N_2 = \{v+1, \cdots, m\}$ and $N_3 = \{m+1, \cdots, n\}$ with $0 \le v \le m \le n$ and $v < n$. Here when $v = 0$, it means that $N_1 = \emptyset$; when $v = m$, $N_2 = \emptyset$; when $m = n$, $N_3 = \emptyset$. With these notation, one can easily see that N_1 is a proper subset of $\mathbb{N}_1^n$. By a slight abuse of notation, we still use $II(x,y,y_k)$ to represent one term related to N_1, N_2 and N_3 in (2.10) and use $H_{II}(x)$ to represent $\mathcal{C}[\cdots, \cdot]$ related to $II(x,y,y_k)$, i.e.

$$H_{II}(x) = \sum_k \int_{Q_k} K(x-y) II(x,y,y_k) f(y) dy.$$

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

(2.11)
$$\mathcal{M}(\nabla A_i)(y_k) \le \lambda^{\frac{r}{q_i}}, \text{ for } i = m+1, \cdots, n.$$

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

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

Notice that we only consider $x \in (10G_\lambda)^c$, then for $y \in Q_k$, $|x - y| \geq 2l(Q_k) \approx |y - y_k|$ by (2.7). Combining the above discussion with (1.2), we get

$$
(2.12)
\begin{aligned}
|H_{II}(x)| &\leq \sum_k \int_{Q_k} |K(x - y)| \cdot |II(x, y, y_k)| \cdot |f(y)| dy \\
&\lesssim \lambda^{\sum_{i=1}^n \frac{r}{q_i}} \sum_k \int_{Q_k} \frac{l(Q_k)^{n-v}}{[l(Q_k) + |x - y|]^{d+n-v}} |f(y)| dy.
\end{aligned}
$$

Applying the Chebyshev inequality with the above estimate, and utilizing Lemma 2.5 with $|\Omega| \equiv 1$ (note that $n - v \geq 1$), we finally get

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

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

Finally, we add some word about how to modify the above proof to the case $q_i = \infty$ for some $i = 1, \cdots, l$. We may suppose only $q_1 = \cdots = q_u = \infty$ with $1 \leq u \leq l$. Thus $A_1, \cdots, A_u$ are Lipschitz functions which in fact are nice functions. Then we just fix $A_1, \cdots, A_u$ in the rest of the proof. We only make a construction of *exceptional set* for $A_{u+1}, \cdots, A_n$ 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 $A_1, \cdots, A_u$ are Lipschitz functions to deal with all estimates involved with $A_1, \cdots, A_u$, we may get the required bound. $\qquad \square$

Proposition 2.9. Let $d/(d + n) \leq r \leq 1$, $d < q_1, \cdots, q_l \leq \infty$ and $q_{l+1}, \cdots, q_n = d$ with $0 \leq l \leq n$, $p = \infty$. Then the weak type estimate (2.4) holds.

Proof. The proof is quite similar to that of Proposition 2.8. So we shall be brief and only indicate necessary modifications here. Proceeding the proof as we do that in Proposition 2.8, there are four different arguments involved. We will point out below one by one.

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

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

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

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

The first term above satisfies $|10G_\lambda| \lesssim \lambda^{-r}$, so it suffices to consider the second and third term. We only consider $x \in (10G_\lambda)^c$.

The second difference is the estimate related to f_1. Here we need to choose $\tilde{r}$, $\tilde{q}_1$, $\cdots$, $\tilde{q}_n$, such that $1 < \tilde{r} < \infty$, $q_1 < \tilde{q}_1 < \infty$, $\cdots$, $q_n < \tilde{q}_n < \infty$ and $\frac{1}{\tilde{r}} = \sum_{i=1}^n \frac{1}{\tilde{q}_i}$. Apply

Lemma 2.6 with those above $\tilde{r}$, $\tilde{q}_1$, $\cdots$, $\tilde{q}_n$,

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

where in the last second inequality we use A_i is a Lipschitz function on $(G_\lambda)^c$ with Lipschitz bound $\lambda^{\frac{r}{q_i}}$ for $i = 1, \cdots, n$ and $L^{d,1}(\mathbb{R}^d) \subsetneq L^d(\mathbb{R}^d)$ if $q_i = d$.

Next consider the estimate related to f_2. As we have done in the proof of Proposition 2.8, we divide $\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f_2](x)$ into several terms and separate these terms into two part according I and II in (2.10). Then we get

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

The third difference is the *estimate of* $\mathcal{C}[\cdots, \cdot]$ *related to* I. Here we apply Lemma 2.7 and the estimate $\|f_2\|_{L^1(\mathbb{R}^d)} \lesssim \|f\|_{L^\infty(\mathbb{R}^d)}|G_\lambda| \lesssim \lambda^{-r}$ to get

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

The fourth difference is the *estimate of* $\mathcal{C}[\cdots, \cdot]$ *related to* II. We will show that

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

In fact, by (2.12) and Lemma 2.5 with $q = \infty$, we get for any $x \in (10G_\lambda)^c$,

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

So if we choose $C_0 > 4C_d$, we get (2.13). Thus we finish the proof. $\qquad\square$

2.4. **Case:** $d/(d+n) \leq r \leq 1$ **and** $1 \leq q_1, \cdots, q_n < d$.

In this subsection, we consider the case $d/(d+n) \leq r \leq 1$ and $1 \leq q_1, \cdots, q_n < d$. Again here the proof of $p = \infty$ is a little different from that of $1 \leq p < \infty$. We first consider $1 \leq p < \infty$ and point out in the proof where it doesn't work for $p = \infty$.

Proposition 2.10. Let $d/(d+n) \leq r \leq 1$, $1 \leq q_1, \cdots, q_n < d$, $1 \leq p < \infty$. Then the weak type estimate (1.7) holds.

Proof. Our main goal is to prove that for any $\lambda > 0$, the following inequality holds

$$m(\{x \in \mathbb{R}^d : |\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f](x)| > \lambda\})$$

$$\lesssim \lambda^{-r} \Big(\prod_{i=1}^{n} \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d)}^{r} \Big) \|f\|_{L^p(\mathbb{R}^d)}^{r}.$$

Now we fix $\lambda > 0$. Recall E_λ defined in (2.5). By rescaling as showed in the proof of Proposition 2.8, we only need to show $|E_\lambda| \lesssim \lambda^{-r}$. The main idea is to construct some *exceptional set* such that the 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 should be Lipschitz functions with bound $\lambda^{\frac{r}{q_i}}$ for each $i = 1, \cdots, n$. Below we begin our constructions of some *exceptional set* which will be involved with several steps.

Step 1: Calderón-Zygmund decomposition. By the formula given in [22, page 125, (17)], for each A_i, $i = 1, \cdots, n$, one may write

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

For each $|\partial_j A_i|^{q_i} \in L^1(\mathbb{R}^d)$ with $j = 1, \cdots, d$ and $i = 1, \cdots, n$, making a Calderón-Zygmund decomposition at level λ^r, one may have the following conclusions (see *e.g.* [17]):

(cz-i) $\partial_j A_i = g_{j,i} + b_{j,i}$, $\|g_{j,i}\|_{L^\infty(\mathbb{R}^d)} \lesssim \lambda^{\frac{r}{q_i}}$, $\|g_{j,i}\|_{L^{q_i}(\mathbb{R}^d)} \lesssim \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d)}$;

(cz-ii) $b_{j,i} = \sum_{Q \in \mathcal{Q}_{j,i}} b_{j,i,Q}$, 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 $m(E_{j,i}) \lesssim \lambda^{-r} \|\partial_j A_i\|_{L^{q_i}(\mathbb{R}^d)}^{q_i}$;

(cz-iv) $\int b_{j,i,Q}(y) dy = 0$ for each $Q \in \mathcal{Q}_{j,i}$ and $\|b_{j,i,Q}\|_{L^{q_i}(\mathbb{R}^d)}^{q_i} \lesssim \lambda^r |Q|$, so we get $\|b_{j,i}\|_{L^{q_i}(\mathbb{R}^d)} \lesssim \|\partial_j A_i\|_{L^{q_i}(\mathbb{R}^d)}$ by (cz-ii) and (cz-iii).

We are going to separate $A_{i,j}$ into two parts according the above Calderón-Zygmund decomposition property (cz-i):

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

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

Set the *exceptional set* $B_\lambda = \cup_{i=1}^{n} \cup_{j=1}^{d} E_{j,i}$. Then by (cz-iii), we get $m(B_\lambda) \lesssim \lambda^{-r}$.

Step 2: Exceptional set D_λ. Set $\frac{1}{s_i} = \frac{1}{q_i} - \frac{1}{d}$ for $i = 1, \cdots, n$. Define the following *exceptional set*

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

where the maximal operator $\mathfrak{M}_{s_i}$ is defined in Lemma 2.4. We denote $D_\lambda = \cup_{i=1}^{n} D_{i,\lambda}$. Then by Lemma 2.4 and the weak type (1,1) bound for the Hardy-Littlewood maximal operator, we see

$$m(D_{i,\lambda}) \leq m(\{x \in \mathbb{R}^d : M_{q_i}(\nabla A_i)(x) > C\lambda^{\frac{r}{q_i}}\}) \lesssim \lambda^{-r} \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d)}^{q_i} = \lambda^{-r}.$$

So does $m(D_\lambda) \lesssim \lambda^{-r}$.

Step 3: Exceptional set F_λ. For each $j = 1, \cdots, d$, $i = 1, \cdots, n$, we 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. We define another *exceptional set*

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

Then by the Chebyshev inequality and (cz-iii), we get

$$|F_{j,i,\lambda}| \le \int_{\mathbb{R}^d} \Delta_{j,i,\lambda}(x) dx \le \Big[\int_{\mathbb{R}^d} \frac{1}{(1 + |y|)^{d+1}} dy \Big] \Big[\sum_{Q \in \mathcal{Q}_{j,i}} |Q| \Big] \lesssim \lambda^{-r}.$$

So does $m(F_\lambda) \lesssim \lambda^{-r}$.

Step 4: Exceptional set H_λ. We define the *exceptional set*

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

Notice that by the definition of $A_{i,j}^g$, for each $z = 1, \cdots, d$, we have

$$\mathcal{F}(\partial_z A_{i,j}^g)(\xi) = C \frac{\xi_z \xi_j}{|\xi|^2} \mathcal{F}(g_{j,i})(\xi) \Rightarrow \nabla A_{i,j}^g = CRR_j g_{j,i},$$

where $\mathcal{F}$ is the Fourier transform, R_j is the Riesz transform and $R = (R_1, \cdots, R_d)$. Since R_j is of strong (q, q) type for $1 < q < \infty$, we get $\|\nabla A_{i,j}^g\|_{L^q(\mathbb{R}^d)} \lesssim \|g_{j,i}\|_{L^q(\mathbb{R}^d)}$. By the Chebyshev inequality, Lemma 2.1 and (cz-i), we get for $d < q < \infty$,

$$m(H_{i,j,\lambda}) \lesssim \lambda^{-\frac{qr}{q_i}} \int_{\mathbb{R}^d} [\mathcal{M}(\nabla A_{i,j}^g)(x)]^q dx \lesssim \lambda^{-\frac{qr}{q_i}} \int_{\mathbb{R}^d} |\nabla A_{i,j}^g(x)|^q dx$$

$$\lesssim \lambda^{-\frac{qr}{q_i}} \int_{\mathbb{R}^d} |g_{j,i}(x)|^q dx \lesssim \lambda^{-r} \int_{\mathbb{R}^d} |g_{j,i}(x)|^{q_i} dx \lesssim \lambda^{-r}.$$

So does $m(H_\lambda) \lesssim \lambda^{-r}$.

Step 5: Final exceptional set G_λ. Based on the construction of $B_\lambda, D_\lambda, F_\lambda, H_\lambda$ in Step 1-4, we choose an open set G_λ which satisfies the following conditions:

(1). $(10B_\lambda \cup 10D_\lambda \cup 10F_\lambda \cup 10H_\lambda) \subset G_\lambda$;

(2). $m(G_\lambda) \le (20)^d(|B_\lambda| + |D_\lambda| + |F_\lambda| + |H_\lambda|)$.

By the property of B_λ, D_λ, F_λ and H_λ, we see that $m(G_\lambda) \lesssim \lambda^{-r}$. Next making a Whitney decomposition of G_λ (see [17]), we may get a family of disjoint dyadic cubes $\{Q_k\}_k$ such that

(i). $G_\lambda = \bigcup_{k=1}^\infty Q_k$;

(ii). $\sqrt{d} \cdot l(Q_k) \le dist(Q_k, (G_\lambda)^c) \le 4\sqrt{d} \cdot l(Q_k)$.

With those properties (i) and (ii), for each Q_k, we could construct a larger cube Q_k^* so that $Q_k \subset Q_k^*$, Q_k^* is centered at y_k and $y_k \in (G_\lambda)^c$, $|Q_k^*| \approx |Q_k|$. 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, we get

(2.14) $dist(y_k, Q_k) \approx l(Q_k).$

Clearly, the *exceptional set* G_λ constructed in Step 5 satisfies that the measure is bounded by λ^{-r}. In the following we will show that these functions A_i are Lipschitz functions on the complementary set of G_λ.

Step 6: Lipschitz estimates of A_i on $(G_\lambda)^c$. By the Calderón-Zygmund decomposition in Step 1, it suffices to show that $A_{i,j}^g$ and $A_{i,j}^b$ satisfy Lipschitz estimates on $(G_\lambda)^c$ for each $i = 1, \cdots, n$ and $j = 1, \cdots, d$. Firstly, it is easy to see that $A_{i,j}^g$ satisfies Lipschitz estimates by the construction of H_λ in Step 4. In fact, for any $x, y \in H_\lambda^c$, we get

$$(2.15) \qquad |A_{i,j}^g(x) - A_{i,j}^g(y)| \le \lambda^{\frac{r}{q_i}} |x - y|.$$

We give effort to showing $A_{i,j}^b$ is a Lipschitz function on $(G_\lambda)^c$. Recall the Calderón-Zygmund decomposition property (cz-ii), (cz-iii) and (cz-iv) in Step 1. For each $b_{j,i} = \sum_{Q \in \mathcal{Q}_{j,i}} b_{j,i,Q}$, 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

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

Now we choose $x, y \in (G_\lambda)^c$ and fix a dyadic cube $Q \in \mathcal{Q}_{j,i}$. Then by the construction of G_λ, $x, y \in (10B_\lambda)^c$, i.e. $x, y \in (10Q)^c$, therefore we get $dist(x, Q) \ge \frac{9}{2}l(Q)$ and $dist(y, Q) \ge \frac{9}{2}l(Q)$. We will give a straight-forward Lipschitz estimate of $A_{i,j}^{b_Q}$. Let z_Q be the center of Q. Without loss of generality, suppose that $|x - z_Q| \le |y - z_Q|$. Choose a point $Z \in \mathbb{R}^d$ such that

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

for any X belongs to the polygonal with vertex x, y, Z. One may draw a figure to check that such a point Z always exists, provided that $dist(x, Q) > \frac{9}{2}l(Q)$ and $dist(y, Q) > \frac{9}{2}l(Q)$. Now we split $A_{i,j}^{b_Q}(x) - A_{i,j}^{b_Q}(y) = A_{i,j}^{b_Q}(x) - A_{i,j}^{b_Q}(Z) + A_{i,j}^{b_Q}(Z) - A_{i,j}^{b_Q}(y)$. By using the mean value formula, we see that

$$(2.16) \qquad \begin{aligned} A_{i,j}^{b_Q}(x) - A_{i,j}^{b_Q}(Z) &= \int_0^1 \langle x - Z, \nabla(A_{i,j}^{b_Q})(tx + (1-t)Z) \rangle dt; \\ A_{i,j}^{b_Q}(Z) - A_{i,j}^{b_Q}(y) &= \int_0^1 \langle Z - y, \nabla(A_{i,j}^{b_Q})(tZ + (1-t)y) \rangle dt. \end{aligned}$$

For any $0 \le t \le 1$, $tx + (1-t)Z$ and $tZ + (1-t)y$ lie in the polygonal with vertex x, y, Z. Notice that $|x - z_Q| \ge 5l(Q)$. By our choice of Z, we have

$$(2.17) \qquad \begin{aligned} |tx + (1-t)Z - z_Q| &> \frac{2}{5}|x - z_Q| > 2l(Q), \\ |tZ + (1-t)y - z_Q| &> \frac{2}{5}|x - z_Q| > 2l(Q). \end{aligned}$$

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 $b_{j,i,Q}$, (2.17) and (cz-iv) in Step 1, we see that

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

Combining the above arguments with (2.16) and the construction of Z, we get

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

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

$$(2.18) \quad \big|A_{i,j}^{b}(x) - A_{i,j}^{b}(y)\big| \lesssim \lambda^{\frac{r}{q_i}} |x - y| \sum_{Q \in \mathcal{Q}_{j,i}} \frac{l(Q)}{[l(Q) + |x - z_Q|]^{d+1}} |Q| \leq \lambda^{\frac{r}{q_i}} |x - y|.$$

Now we conclude the Lipschitz estimates related good functions (2.15) and bad functions (2.18) to get that for any $i = 1, \cdots, n$, $x, y \in (G_\lambda)^c$,

$$(2.19) \qquad\qquad |A_i(x) - A_i(y)| \lesssim \lambda^{\frac{r}{q_i}} |x - y|.$$

Step 7: Estimate of E_λ. We return 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 (2.19), 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 $i = 1, \cdots, n$. Let $\tilde{A}_i$ stand for the Lipschitz extension of A_i from $(G_\lambda)^c$ to $\mathbb{R}^d$ (see [22, page 174, Theorem 3]) so that for each $i = 1, \cdots, n$,

$$(2.20) \qquad\qquad \tilde{A}_i(y) = A_i(y) \text{ if } y \in (G_\lambda)^c;$$
$$\big|\tilde{A}_i(x) - \tilde{A}_i(y)\big| \leq \lambda^{\frac{r}{q_i}} |x - y| \text{ for all } x, y \in \mathbb{R}^d.$$

As we have done in (2.8) and (2.9) in the proof of Proposition 2.8, we may reduce the proof to the following estimate

$$m\big(\{x \in (10G_\lambda)^c : |\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f_2](x)| > \lambda/2\}\big) \lesssim \lambda^{-r}.$$

Step 8: 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 our construction of G_λ, y_k, Q_k and Q_k^* in the paragraph above (2.14). Then we can write $f_2 = \sum_k f\chi_{Q_k}$. Therefore we may get

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

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

$$\prod_{i=1}^{n} \frac{A_i(x) - A_i(y)}{|x-y|}$$

$$= \prod_{i=1}^{n} \Big(\frac{\tilde{A}_i(x) - \tilde{A}_i(y)}{|x-y|} + \frac{\tilde{A}_i(y) - \tilde{A}_i(y_k)}{|x-y|} + \frac{A_i(y_k) - A_i(y)}{|x-y|} \Big)$$

$$= \sum \Big(\prod_{i \in N_1} \frac{\tilde{A}_i(x) - \tilde{A}_i(y)}{|x-y|} \Big) \Big(\prod_{i \in N_2} \frac{\tilde{A}_i(y) - \tilde{A}_i(y_k)}{|x-y|} \Big) \Big(\prod_{i \in N_3} \frac{A_i(y_k) - A_i(y)}{|x-y|} \Big)$$

$$= I(x,y) + II(x,y,y_k) + III(x,y,y_k),$$

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 $I(x,y)$, $II(x,y,y_k)$, $III(x,y,y_k)$ are defined as follows

(2.21)

$$I = \prod_{i=1}^{n} \frac{\tilde{A}_i(x) - \tilde{A}_i(y)}{|x-y|},$$

$$II = \sum_{\substack{N_1 \subsetneq \mathbb{N}_1^n \\ N_3 = \emptyset}} \Big(\prod_{i \in N_1} \frac{\tilde{A}_i(x) - \tilde{A}_i(y)}{|x-y|} \Big) \Big(\prod_{i \in N_2} \frac{\tilde{A}_i(y) - \tilde{A}_i(y_k)}{|x-y|} \Big),$$

$$III = \sum_{\substack{N_1 \subsetneq \mathbb{N}_1^n \\ N_3 \neq \emptyset}} \Big(\prod_{i \in N_1} \frac{\tilde{A}_i(x) - \tilde{A}_i(y)}{|x-y|} \Big) \Big(\prod_{i \in N_2} \frac{\tilde{A}_i(y) - \tilde{A}_i(y_k)}{|x-y|} \Big)$$

$$\times \Big(\prod_{i \in N_3} \frac{A_i(y_k) - A_i(y)}{|x-y|} \Big).$$

By the above decomposition, we in fact divide $\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f\chi_{Q_k}](x)$ into 3^n terms. We separate these terms into three parts according I, II and III.

Step 9: Estimate of $\mathcal{C}[\cdots, \cdot]$ related to I. This estimate is similar to the term related to I in the proof of Proposition 2.8. In fact, in this case there is only one term $\mathcal{C}[\nabla \tilde{A}_1, \cdots, \nabla \tilde{A}_n, f_2]$. Then by Proposition 2.6 ($1 < p < \infty$) and Proposition 2.7 ($p = 1$), we get

$$m\big(\{ x \in (10G_\lambda)^c : |\mathcal{C}[\nabla \tilde{A}_1, \cdots, \nabla \tilde{A}_n, f_2](x)| > \lambda/2 \} \big)$$

$$\lesssim \lambda^{-p} \Big(\prod_{i=1}^{n} \|\nabla \tilde{A}_i\|_{L^\infty(\mathbb{R}^d)}^p \Big) \|f_2\|_{L^p(\mathbb{R}^d)}^p \lesssim \lambda^{-p + p \sum_{i=1}^{n} \frac{r}{q_i}} = \lambda^{-r}.$$

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

Step 10: Estimate of $\mathcal{C}[\cdots, \cdot]$ related to II. The proof of this part is similar to the estimate related to II in Proposition 2.8. It suffices to consider one term $\mathcal{C}[\cdots, \cdot]$ related to II in which N_1 is a proper subset of $\mathbb{N}_1^n$. In this case, without loss of generality, we may assume $N_1 = \{1, \cdots, l\}$, $N_2 = \{l+1, \cdots, n\}$ with $0 \le l < n$. Here when $l = 0$, it means that $N_1 = \emptyset$. 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 use $II(x, y, y_k)$ to represent

one term related to N_1 and N_2 in (2.21) and use $H_{II}(x)$ to represent $\mathcal{C}[\cdots,\cdot]$ related to $II(x,y,y_k)$, i.e.

$$H_{II}(x) = \sum_k \int_{Q_k} K(x-y)II(x,y,y_k)f(y)dy.$$

Notice that $\tilde{A}_i$ is a Lipschitz function with bound $\lambda^{\frac{r}{q_i}}$ by (2.20) for all $i = 1,\cdots,n$. Then we get

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

Since we only consider $x \in (10G_\lambda)^c$, then by (2.14), we get

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

Therefore utilizing (1.2) and the above estimate, we get

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

where the operator T_{n-l} is defined in Lemma 2.5 with $|\Omega| \equiv 1$. Now applying the Chebyshev inequality and the above estimate, and using Lemma 2.5 since $n-l \geq 1$, we finally get

$$\begin{aligned}
m(\{x \in (G_\lambda)^c : |H_{II}(x)| > \lambda\}) &\leq \lambda^{-p+\sum_{i=1}^n \frac{rp}{q_i}} \int_{(10G_\lambda)^c} |T_{n-l}f(x)|^p dx \\
&\lesssim \lambda^{-r} \|f\|_{L^p(\mathbb{R}^d)}^p.
\end{aligned}$$

Hence we complete the proof related to II.

Step 11: Estimate of $\mathcal{C}[\cdots,\cdot]$ related to III. It suffices to consider one term $\mathcal{C}[\cdots,\cdot]$ related to III in which N_1 is a proper subset of $\mathbb{N}_1^n$ and N_3 is a nonempty set. In this case, without loss of generality, we may assume $N_1 = \{1,\cdots,l\}$, $N_2 = \{l+1,\cdots,m\}$ and $N_3 = \{m+1,\cdots,n\}$ with $0 \leq l \leq m < n$. Here when $l = 0$, it means that $N_1 = \emptyset$; when $l = m$, $N_2 = \emptyset$. With these notation, one can easily see that N_1 is a proper subset of $\mathbb{N}_1^n$ and N_3 is a nonempty set. By a slight abuse of notation, we still use $III(x,y,y_k)$ to represent one term related to N_1, N_2 and N_3 in (2.21) and use $H_{III}(x)$ to represent $\mathcal{C}[\cdots,\cdot]$ related to $III(x,y,y_k)$, i.e.

$$H_{III}(x) = \sum_k \int_{Q_k} K(x-y)III(x,y,y_k)f(y)dy.$$

Recall in Step 2, we set $\frac{1}{s_i} = \frac{1}{q_i} - \frac{1}{d}$ for all $i = 1,\cdots,n$. We also set $\frac{1}{q} = \left(\sum_{i=m+1}^n \frac{1}{s_i}\right) + \frac{1}{p}$. Since $r \geq d/(d+n)$ and $\frac{1}{r} = \left(\sum_{i=1}^n \frac{1}{q_i}\right) + \frac{1}{p}$, by some elementary calculation, one may get $1 \leq q \leq \infty$, this will be crucial when we use Lemma 2.5.

This is the place where we use the condition $r \geq d/(d + n)$. With the above fact and $\tilde{A}_i$ is a Lipschitz function with bound λ^{r/q_i} for $i = 1, \cdots, m$, we get

$$|III(x, y, y_k)| \lesssim \lambda^{\sum_{i=1}^{m} \frac{r}{q_i}} \frac{(l(Q_k))^{n-l}}{|x - y|^{n-l}} \prod_{i=m+1}^{n} \frac{|A_i(y_k) - A_i(y)|}{l(Q_k)}.$$

Applying (1.2) and the above estimate with (2.22), we get

$$|H_{III}(x)| \leq \sum_k \int_{Q_k} |K(x - y)| \cdot |III(x, y, y_k)| \cdot |f(y)| dy$$

$$\lesssim \lambda^{\sum_{i=1}^{m} \frac{r}{q_i}} \sum_k \int_{Q_k} \frac{l(Q_k)^{n-l}}{[l(Q_k) + |x - y|]^{d+n-l}} h_{m,n}(y) dy$$

$$\lesssim \lambda^{\sum_{i=1}^{m} \frac{r}{q_i}} T_{n-l}(h_{m,n})(x)$$

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

Applying the Chebyshev inequality and the above estimate of H_{III}, utilizing Lemma 2.5 with $|\Omega| \equiv 1$, we then get

$$m(\{x \in (10G_\lambda)^c : |H_{III}(x)| > \lambda\})$$

(2.23)
$$\leq \lambda^{-q+\left(\sum_{i=1}^{m} \frac{rq}{q_i}\right)} \int_{(10G_\lambda)^c} |T_{n-l}(h_{m,n})(x)|^q dx$$

$$\lesssim \lambda^{-q+\left(\sum_{i=1}^{m} \frac{rq}{q_i}\right)} \|h_{m,n}\|^q_{L^q(\mathbb{R}^d)},$$

since $1 \leq q \leq \infty$ and $n - l \geq 1$. Below we give an estimate of $\|h_{m,n}\|^q_{L^q(\mathbb{R}^d)}$. We write

$$\|h_{m,n}\|^q_{L^q(\mathbb{R}^d)} = \sum_{Q_k} \int_{Q_k} \prod_{i=m+1}^{n} \left[\frac{|A_i(y_k) - A_i(y)|}{l(Q_k)} \right]^q |f(y)|^q dy$$

(2.24)
$$\leq \sum_{Q_k} \prod_{i=m+1}^{n} \left[\int_{Q_k} \left(\frac{|A_i(y_k) - A_i(y)|}{l(Q_k)} \right)^{s_i} dy \right]^{\frac{q}{s_i}} \left(\int_{Q_k} |f(y)|^p dy \right)^{\frac{q}{p}}$$

$$\lesssim \sum_{Q_k} \prod_{i=m+1}^{n} \left[\int_{Q_k^*} \left(\frac{|A_i(y_k) - A_i(y)|}{l(Q_k^*)} \right)^{s_i} dy \right]^{\frac{q}{s_i}} \left(\int_{Q_k} |f(y)|^p dy \right)^{\frac{q}{p}}$$

$$\lesssim \sum_{Q_k} \left[\prod_{i=m+1}^{n} \mathfrak{M}_{s_i}(\nabla A_i)(y_k)^q |Q_k|^{\frac{q}{s_i}} \right] \left(\int_{Q_k} |f(y)|^p dy \right)^{\frac{q}{p}},$$

where the second inequality just follows from the Hölder inequality and in the third inequality we use 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. $y_k \in (D_{i,\lambda})^c$. Therefore by the construction of $D_{i,\lambda}$ in Step 2, we get

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

Applying the above inequality, the Hölder inequality again and (cz-iii) in Step 1, we get

$$\|h_{m,n}\|_{L^q(\mathbb{R}^d)}^q \lesssim \lambda^{\sum_{i=m+1}^n \frac{qr}{q_i}} \Big(\sum_{Q_k} |Q_k| \Big)^{\sum_{i=m+1}^n \frac{q}{s_i}} \|f\|_{L^p(\mathbb{R}^d)}^q$$

$$\lesssim \lambda^{\sum_{i=m+1}^n \frac{qr}{q_i}} |G_\lambda|^{\sum_{i=m+1}^n \frac{q}{s_i}} \lesssim \lambda^{\sum_{i=m+1}^n \left(\frac{qr}{q_i} - \frac{qr}{s_i} \right)}.$$

Submitting the above estimate into (2.23) with some elementary calculations, we finally get

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

which is the required bound. Hence we complete the proof. Notice that this argument for the term III also works in the case $p = \infty$. $\qquad\square$

Proposition 2.11. Let $d/(d+n) \le r \le 1$, $1 \le q_1, \cdots, q_n < d$, $p = \infty$. Then the weak type estimate (1.7) holds.

Proof. The proof is quite similar to that of Proposition 2.10. One may follow the four different arguments that we deal with E_λ, f_1, I and II in the proof of Proposition 2.9. The proof of the term III is similar to Step 11 in the proof of Proposition 2.10. We omit the details of the proof here. $\qquad\square$

2.5. **Case:** $d/(d+n) \le r \le 1$, **some** $q_i \ge d$ **and some** $q_i < d$.

In this subsection, we consider the most complicated case: $d/(d+n) \le r \le 1$ with some $q_i \ge d$ and some $q_i < d$. After the warm-up of the case all $q_i \ge d$ in Subsection 2.3 and all $q_i < d$ in Subsection 2.4, the strategy here is quite clear that we will put the two arguments in Subsections 2.3 and 2.4 together. Without loss of generality, we may suppose that $d \le q_1, \cdots, q_l \le \infty$ and $1 \le q_{l+1}, \cdots, q_n < d$ with $1 \le l < n$. Also we may assume that $q_1 = \cdots = q_k = d$ and $d < q_{k+1}, \cdots, q_l \le \infty$ with $0 \le k \le l$. When $k = 0$, we mean that there is no indice in $q_1, \cdots, q_l$ equals to d, i.e. $d < q_1, \cdots, q_l \le \infty$; when $k = l$, we mean that $q_1 = \cdots = q_l = d$. Since the proof of $p = \infty$ is a little different from that of $1 \le p < \infty$, we will give two propositions here. We first consider $1 \le p < \infty$.

Proposition 2.12. Let $d/(d+n) \le r \le 1$, $q_1 = \cdots = q_k = d$, $d < q_{k+1}, \cdots, q_l \le \infty$ and $1 \le q_{l+1}, \cdots, q_n < d$ with $0 \le k \le l$ and $1 \le l < n$, $1 \le p < \infty$. Then

$$\|\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f]\|_{L^{r,\infty}(\mathbb{R}^d)}$$

(2.25)
$$\lesssim \Big(\prod_{i=1}^k \|\nabla A_i\|_{L^{d,1}(\mathbb{R}^d)} \Big) \Big(\prod_{i=k+1}^n \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d)} \Big) \|f\|_{L^p(\mathbb{R}^d)}.$$

Proof. The proof of this proposition is involved with the idea that we have done in the proof of Proposition 2.8 for $q_i \ge d$ and Proposition 2.10 for $1 \le q_i < d$. We will combine these two arguments in the proof of Proposition 2.8 and Proposition 2.10. One will see below that part of discussions have been appeared in the previous proposition. So we shall be brief and only indicate necessary differences.

Now we start our proof. Our main goal is to prove that for any $\lambda > 0$, the following inequality holds

$$\lambda^r m(\{x \in \mathbb{R}^d : |\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f](x)| > \lambda\})$$

$$\lesssim \Big(\prod_{i=1}^{k} \|\nabla A_i\|_{L^{d,1}(\mathbb{R}^d)}^r \Big) \Big(\prod_{i=k+1}^{n} \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d)}^r \Big) \|f\|_{L^p(\mathbb{R}^d)}^r.$$

Fix $\lambda > 0$. Recall E_λ defined in (2.5). By rescaling as showed in the proof of Proposition 2.8 or Proposition 2.10, it suffices to show $|E_\lambda| \lesssim \lambda^{-r}$. The main idea is to construct some *exceptional set* such that the 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 should be Lipschitz functions with bound $\lambda^{\frac{r}{q_i}}$ for each $i = 1, \cdots, n$. If $d \leq q_i < \infty$, the construction of *exceptional set* is similar to that of Proposition 2.8. And if $1 \leq q_i < d$, the construction of *exceptional set* is similar to that of Proposition 2.10. As we have done in Proposition 2.8, we only need to consider that all $q_{k+1}, \cdots, q_l < \infty$. Below we begin our constructions of some *exceptional set*.

Step 1: Exceptional set J_λ. Define the *exceptional set* for $i = 1, \cdots, l$

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

Step 2: Calderón-Zygmund decomposition. For each $|\partial_j A_i|^{q_i} \in L^1(\mathbb{R}^d)$ with $j = 1, \cdots, d$ and $i = l+1, \cdots, n$, making a Calderón-Zygmund decomposition at level λ^r as we have done in the proof of Proposition 2.10, one may get the properties of $g_{j,i}, b_{j,i}, E_{j,i}, A_{i,j}^g, A_{i,j}^b$ similarly. Set the *exceptional set* $B_\lambda = \cup_{i=l+1}^{n} \cup_{j=1}^{d} E_{j,i}$.

Step 3: Exceptional set D_λ. Set $\frac{1}{s_i} = \frac{1}{q_i} - \frac{1}{d}$ for $i = l+1, \cdots, n$. Define the following *exceptional set*

$$D_{i,\lambda} = \Big\{ x \in \mathbb{R}^d : \mathfrak{M}_{s_i}(\nabla A_i)(x) > \lambda^{\frac{r}{q_i}} \Big\}, \quad D_\lambda = \cup_{i=l+1}^{n} D_{i,\lambda}.$$

Step 4: Exceptional set F_λ. For each $j = 1, \cdots, d$, $i = l+1, \cdots, n$, we define the functions $\Delta^{j,i}(x)$ as those in the proof of Proposition 2.10. Define another *exceptional set*

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

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

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

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

(1). $\big(10J_\lambda \cup 10B_\lambda \cup 10D_\lambda \cup 10F_\lambda \cup 10H_\lambda \big) \subset G_\lambda$;

(2). $m(G_\lambda) \leq (20)^d(|J_\lambda| + |B_\lambda| + |D_\lambda| + |F_\lambda| + |H_\lambda|)$.

As showed in the proof of Proposition 2.8 and Proposition 2.10, one may get that the measures of J_λ, B_λ, D_λ, F_λ and H_λ are bounded by λ^{-r}. So we see that $m(G_\lambda) \lesssim \lambda^{-r}$. Next making a Whitney decomposition of G_λ, we may get a family of disjoint dyadic cubes $\{Q_k\}_k$ and then we 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 construction of Q_k^* and y_k, one may get

$$(2.26) \qquad\qquad dist(y_k, Q_k) \approx l(Q_k).$$

In the following we will show that these functions A_i are Lipschitz functions on the complementary set of G_λ.

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 $i = 1, \cdots, l$

$$(2.27) \qquad\qquad |A_i(x) - A_i(y)| \leq \lambda^{\frac{r}{q_i}} |x - y|.$$

Below we consider $i = l + 1, \cdots, n$. By the Calderón-Zygmund decomposition in Step 2, it suffices to show that $A_{i,j}^g$ and $A_{i,j}^b$ satisfy Lipschitz estimates on $(G_\lambda)^c$ for each $i = l+1, \cdots, n$ and $j = 1, \cdots, d$. Firstly, one may easily 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 get for $i = l + 1, \cdots, n$, $j = 1, \cdots, d$,

$$(2.28) \qquad\qquad |A_{i,j}^g(x) - A_{i,j}^g(y)| \leq \lambda^{\frac{r}{q_i}} |x - y|.$$

While considering $A_{i,j}^b$, we see that by using the similar method that we prove (2.18) in Proposition 2.10, we get for $i = l + 1, \cdots, n$, $j = 1, \cdots, d$,

$$(2.29) \qquad\qquad \left|A_{i,j}^b(x) - A_{i,j}^b(y)\right| \lesssim \lambda^{\frac{r}{q_i}} |x - y|.$$

Therefore we conclude the Lipschitz estimates in (2.27) for $i = 1, \cdots, l$, good function (2.28) and bad function (2.29) for $i = l + 1, \cdots, n$, to get that for any $i = 1, \cdots, n$, $x, y \in (G_\lambda)^c$,

$$(2.30) \qquad\qquad |A_i(x) - A_i(y)| \leq \lambda^{\frac{r}{q_i}} |x - y|.$$

Step 8: Estimate of E_λ. As we have done in the proof of Proposition 2.8, we may split $f = f_1 + f_2$. Following (2.8) and (2.9), we may reduce the estimate of E_λ to the following inequality

$$m\big(\{x \in (10G_\lambda)^c : |\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f_2](x)| > \lambda/2\}\big) \lesssim \lambda^{-r}.$$

Step 9: 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 our construction of G_λ, y_k, Q_k and Q_k^* in the paragraph above (2.26). Then we can write $f_2 = \sum_k f\chi_{Q_k}$. Therefore we may get

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

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

$$\prod_{i=1}^n \frac{A_i(x) - A_i(y)}{|x - y|} = I(x, y) + II(x, y, y_k) + III(x, y, y_k) + IV(x, y, y_k),$$

where $I(x, y)$, $II(x, y, y_k)$, $III(x, y, y_k)$ and $IV(x, y, y_k)$ are defined as follows

$$I = \prod_{i=1}^{n} \frac{\tilde{A}_i(x) - \tilde{A}_i(y)}{|x - y|},$$

$$II = \sum_{\substack{N_1 \subsetneq \mathbb{N}_1^n \\ N_3 = \emptyset}} \Big(\prod_{i \in N_1} \frac{\tilde{A}_i(x) - \tilde{A}_i(y)}{|x - y|} \Big) \Big(\prod_{i \in N_2} \frac{\tilde{A}_i(y) - \tilde{A}_i(y_k)}{|x - y|} \Big),$$

$$III = \sum_{\substack{N_1 \subsetneq \mathbb{N}_1^n \\ N_3 \neq \emptyset, N_3 \subset \{1, \cdots, l\}}} \Big(\prod_{i \in N_1} \frac{\tilde{A}_i(x) - \tilde{A}_i(y)}{|x - y|} \Big) \Big(\prod_{i \in N_2} \frac{\tilde{A}_i(y) - \tilde{A}_i(y_k)}{|x - y|} \Big)$$

(2.31)
$$\times \Big(\prod_{i \in N_3} \frac{A_i(y_k) - A_i(y)}{|x - y|} \Big),$$

$$IV = \sum_{\substack{N_1 \subsetneq \mathbb{N}_1^n \\ N_3 \neq \emptyset, N_3 \cap \{l+1, \cdots, n\} \neq \emptyset}} \Big(\prod_{i \in N_1} \frac{\tilde{A}_i(x) - \tilde{A}_i(y)}{|x - y|} \Big) \Big(\prod_{i \in N_2} \frac{\tilde{A}_i(y) - \tilde{A}_i(y_k)}{|x - y|} \Big)$$

$$\times \Big(\prod_{i \in N_3} \frac{A_i(y_k) - A_i(y)}{|x - y|} \Big),$$

here $\mathbb{N}_1^n = N_1 \cup N_2 \cup N_3$ with N_1, N_2, N_3 non intersecting each other. By the above decomposition, we in fact divide $\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f\chi_{Q_k}](x)$ into 3^n terms. We separate these terms into four parts according I, II, III and IV.

Step 10: Estimate of $\mathcal{C}[\cdots, \cdot]$ *related to* I. Since I is the same as I term in the proof of Proposition 2.8, so this estimate is similar to that there. We omit the proof here.

Step 11: Estimate of $\mathcal{C}[\cdots, \cdot]$ *related to* II. This estimate is similar to the term related to II in the proof of Proposition 2.10. So we omit the proof.

Step 12: Estimate of $\mathcal{C}[\cdots, \cdot]$ *related to* III. It suffices to consider one term $\mathcal{C}[\cdots, \cdot]$ related to III in which N_1 is a proper subset of $\mathbb{N}_1^n$ and N_3 is a nonempty subset of $\{1, \cdots, l\}$. By the condition in this proposition, for any $i \in N_3$, $d \leq q_i < \infty$. Thus $\nabla A_i \in L^{q_i}(\mathbb{R}^d)$ (or $L^{d,1}(\mathbb{R}^d)$ if $q_i = d$) with $d \leq q_i < \infty$. Therefore the estimates in N_3 will be straightforward since

(2.32)
$$\frac{|A_i(y_k) - A_i(y)|}{|y - y_k|} \leq \mathcal{M}(\nabla A_i)(y_k) \leq \lambda^{\frac{r}{q_i}}, \text{ for } i \in N_3.$$

Once we give the above estimate in N_3, the rest terms related to N_1 and N_2 can be dealt as the same way to those related to II. For the rest of the proof, one can follow the term related to II in the proof of Proposition 2.10.

Step 13: Estimate of $\mathcal{C}[\cdots, \cdot]$ *related to* IV. It suffices to consider one term $\mathcal{C}[\cdots, \cdot]$ related to IV in which N_1 is a proper subset of $\mathbb{N}_1^n$ and N_3 is a nonempty set with $N_3 \cap \{l+1, \cdots, n\} \neq \emptyset$. In this case, without loss of generality, we may assume $l+1, \cdots, v \in N_3$ with $l+1 \leq v \leq n$ and $v+1, \cdots, n$ belongs to N_1 or N_2. So we may suppose that $N_3 = \{\iota, \cdots, w, l+1, \cdots, v\}$ with $0 \leq \iota \leq w \leq l$ and $N_1 \neq \emptyset$. Set $u = \text{card}(N_1)$. Then $n - u \geq 1$. With these notation, it is easy to see that N_3 is a

nonempty set with $N_3 \cap \{l+1, \cdots, n\} \neq \emptyset$. We use $IV(x, y, y_k)$ to represent one term related to N_1, N_2 and N_3 in (2.31) and use $H_{IV}(x)$ to represent $\mathcal{C}[\cdots, \cdot]$ related to $IV(x, y, y_k)$, i.e.

$$H_{IV}(x) = \sum_k \int_{Q_k} K(x-y) IV(x, y, y_k) f(y) dy.$$

By our condition, $d \leq q_l, \cdots, q_l \leq \infty$ 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 $r \geq d/(d+n)$ and $\frac{1}{r} = \left(\sum_{i=1}^{n} \frac{1}{q_i} \right) + \frac{1}{p}$, by some elementary calculation, one may get $1 \leq q \leq \infty$. With the above fact and $\tilde{A}_i$ is a Lipschitz function with bound λ^{r/q_i} for $i \in N_1 \cup N_2$, we get

$$|IV(x, y, y_k)| \lesssim \lambda^{(\sum_{i \in N_1 \cup N_2}) \frac{r}{q_i}} \frac{(l(Q_k))^{n-u}}{|x-y|^{n-u}} \Big(\prod_{i=\iota}^{w} \mathcal{M}(\nabla A_i)(y_k) \Big)$$

$$\times \prod_{i=l+1}^{v} \frac{|A_i(y_k) - A_i(y)|}{l(Q_k)}$$

$$\lesssim \lambda^{(\sum_{i=1}^{l} + \sum_{i=v+1}^{n}) \frac{r}{q_i}} \frac{(l(Q_k))^{n-u}}{|x-y|^{n-u}} \prod_{i=l+1}^{v} \frac{|A_i(y_k) - A_i(y)|}{l(Q_k)}.$$

Then inserting the above estimate of IV into H_{IV} with (1.2), we get

$$|H_{IV}(x)| \leq \sum_k \int_{Q_k} |K(x-y)| \cdot |IV(x, y, y_k)| \cdot |f(y)| dy$$

$$\lesssim \lambda^{(\sum_{i=1}^{l} + \sum_{i=v+1}^{n}) \frac{r}{q_i}} \sum_k \int_{Q_k} \frac{l(Q_k)^{n-u}}{[l(Q_k) + |x-y|]^{d+n-u}} h_{l,v}(y) dy$$

$$= \lambda^{(\sum_{i=1}^{l} + \sum_{i=v+1}^{n}) \frac{r}{q_i}} T_{n-u}(h_{l,v})(x).$$

Now the rest of the proof is similar to (2.23) in the proof of Proposition 2.10. We omit the details here. $\square$

Proposition 2.13. Let $d/(d+n) \leq r \leq 1$, $q_1 = \cdots = q_k = d$, $d < q_{k+1}, \cdots, q_l \leq \infty$ and $1 \leq q_{l+1}, \cdots, q_n < d$ with $0 \leq k \leq l$ and $1 \leq l < n$, $p = \infty$. Then the weak type estimate (2.25) holds.

Proof. The proof is similar to that of Proposition 2.12, one may follow the idea in the proof of Proposition 2.9 and Proposition 2.11. We omit the details here. $\square$

2.6. **Multilinear interpolation arguments.**

Notice that we have already proven all cases (ii) in Theorem 1.1 by Propositions 2.7, 2.8, 2.9, 2.10, 2.11, 2.12, 2.13. And we also prove the case $1 \leq r < \infty$ of (i) in Theorem 1.1 by Proposition 2.6. The rest part of (i) in Theorem 1.1 follows from the standard multilinear interpolation. In fact, in the case $1 \leq r < +\infty$, for all point $(\frac{1}{q_1}, \cdots, \frac{1}{q_n}, \frac{1}{p})$ in the polyhedron $\left(\sum_{i=1}^{n} \frac{1}{q_i} \right) + \frac{1}{p} = \frac{1}{r}$, we have the follow strong type estimate

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

In the case $r = \frac{d}{d+n}$, for all points $(\frac{1}{q_1}, \cdots, \frac{1}{q_n}, \frac{1}{p})$ in the plane $\left(\sum_{i=1}^{n} \frac{1}{q_i}\right) + \frac{1}{p} = \frac{d+n}{d}$, we have the weak type estimate

$$\|\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f]\|_{L^{\frac{d}{d+n},\infty}(\mathbb{R}^d)} \leq C\Big(\prod_{i=1}^{n} \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d)}\Big) \|f\|_{L^p(\mathbb{R}^d)}$$

where $L^{q_i}(\mathbb{R}^d)$ in the above inequality should be replaced by $L^{d,1}(\mathbb{R}^d)$ if $q_i = d$ for some $i = 1, \cdots, n$. Notice that in the rest part of (i) in Theorem 1.1, we consider $\frac{d}{d+n} < r < 1$, $1 < q_i \leq \infty$ $(i = 1, \cdots, n)$ and $1 < p \leq \infty$, thus the point $(\frac{1}{q_1}, \cdots, \frac{1}{q_n}, \frac{1}{p})$ lies in the interior of the polyhedron between the polyhedron $\left(\sum_{i=1}^{n} \frac{1}{q_i}\right) + \frac{1}{p} = \frac{1}{r}$ with $1 \leq r < \infty$ and plane $\left(\sum_{i=1}^{n} \frac{1}{q_i}\right) + \frac{1}{p} = \frac{d+n}{d}$. Then if we choose $n + 1$ points in the polyhedron $\left(\sum_{i=1}^{n} \frac{1}{q_i}\right) + \frac{1}{p} = \frac{1}{r}$ with $1 \leq r < \infty$ and one point in the plane $\left(\sum_{i=1}^{n} \frac{1}{q_i}\right) + \frac{1}{p} = \frac{d+n}{d}$, using the multilinear interpolation theorem (see [18, Theorem 7.2.2]), we get all strong type estimate in (i). Therefore we complete the proof of (i) and (ii) in Theorem 1.1. Proof of (iii) in Theorem 1.1 will be given in the next subsection.

2.7. Examples.

Proposition 2.14. If $0 < r < \frac{d}{d+n}$, $1 \leq q_i \leq \infty$ $(i = 1, \cdots, n)$ and $1 \leq p \leq \infty$, there exist functions K, A_i $(i = 1, \cdots, n)$ and f such that K satisfies (1.2), (1.3), (1.4); $\|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d)} < \infty$ for $i = 1, \cdots, n$ and $\|f\|_{L^p(\mathbb{R}^d)} < \infty$. But

$$\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f](x) = \infty \quad \text{in a ball in } \mathbb{R}^d.$$

Proof. We may suppose that n is an odd integer. Then we may choose K as

$$K(x) = \frac{1}{|x|^d}, \quad \text{for } x \in \mathbb{R}^d \setminus \{0\}.$$

It is easy to see that K satisfies (1.2), (1.3) and (1.4).

Next we choose α_i $(i = 1, \cdots, n)$ and β such that $-1 \leq \alpha_i < 1$ for $i = 1, \cdots, n$, $\beta \geq 0$ and $\left(\sum_{i=1}^{n} \alpha_i\right) + \beta = d$. For each $i = 1, \cdots, n$, we choose A_i as a C^∞ function in $\mathbb{R}^d \setminus \{0\}$ such that

$$A_i(x) = \begin{cases} |x|^{-\alpha_i} & \text{if } x \in Cone_\rho; \\ 0 & \text{if } x \notin Cone_{\rho,\varepsilon}, \end{cases}$$

where $Cone_\rho$ and $Cone_{\rho,\varepsilon}$ are defined as follows

$$Cone_\rho = \Big\{x \in \mathbb{R}^d : \kappa \sum_{j=2}^{d} x_j^2 < x_1^2, \ 0 < x_1 < \rho\Big\},$$

$$Cone_{\rho,\varepsilon} = \Big\{x \in \mathbb{R}^d : \kappa \sum_{j=2}^{d} x_j^2 < (x_1 + \varepsilon)^2, \ -\varepsilon < x_1 < \rho + \varepsilon\Big\},$$

here ρ, ε are fixed positive constants and κ is a large positive constant. Below we choose

$$f(x) = \begin{cases} |x|^{-\beta} & \text{if } x \in Cone_\rho, \\ 0 & \text{if } x \notin Cone_\rho. \end{cases}$$

By some elementary calculation, one can easily get that for $1 \le q_i < \infty$ and $1 \le p < \infty$

$$\|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d)} < \infty \ \text{ if } \ 1 \le q_i < \frac{d}{1+\alpha_i}; \quad \|f\|_{L^p(\mathbb{R}^d)} < \infty \ \text{ if } \ 1 \le p < \frac{d}{\beta}.$$

If $q_i = \infty$, we may choose $\alpha_i = -1$, then we get $\|\nabla A_i\|_{L^\infty(\mathbb{R}^d)} < \infty$. If $p = \infty$, we may choose $\beta = 0$, then $\|f\|_{L^\infty(\mathbb{R}^d)} < \infty$. Since $(\sum_{i=1}^n \alpha_i) + \beta = d$, it is impossible that all q_i and p equal to ∞. Then by this choice of q_i and p, we see that $0 < r < d/(d+n)$.

Set $z_0 = (-2\varepsilon, 0, \cdots, 0)$. Let x be a point in the small neighborhood of z_0 such that

$$|x - y| \le C|\rho + 4\varepsilon| \ \text{ for all } y \in Cone_\rho.$$

Then combining the choice of A_i $(i = 1, \cdots, n)$ and f, and noticing that n is a odd integer, we finally get

$$-\mathcal{C}[\nabla A_1, \cdots, \nabla A_n, f](x) = \int_{Cone_\rho} K(x-y) \Big[\prod_{i=1}^n \frac{A_i(y)}{|x-y|} \Big] f(y) dy$$

$$\ge \frac{C}{|\rho + 4\varepsilon|^{d+n}} \int_{Cone_\rho} \frac{dy}{|y|^d} = +\infty.$$

$$\square$$

3. Proof of Theorem 1.4

In this section, we just outline the proof of Theorem 1.4 since it is similar to that of Theorem 1.1.

Proposition 3.1. Let $1 \le r < +\infty$, $1 < q_i \le \infty$, $i = 1, \cdots, n$, $1 < p \le \infty$. Then

$$\|\mathcal{C}_\Omega[\nabla A_1, \cdots, \nabla A_n, f]\|_{L^r(\mathbb{R}^d)} \lesssim C_\Omega \Big(\prod_{i=1}^n \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d)} \Big) \|f\|_{L^p(\mathbb{R}^d)}.$$

Proof. We use the method of rotation to prove our main result. The method is standard, so we will be brief. Applying the condition (1.9) and making a change of variable $x - y = r\theta$, we get that

$$\mathcal{C}_{\Omega,\varepsilon}[\nabla A_1, \cdots, \nabla A_n, f](x) = \int_{|x-y|>\varepsilon} \frac{\Omega(x-y)}{|x-y|^d} \Big(\prod_{i=1}^n \frac{A_i(x) - A_i(y)}{|x-y|} \Big) \cdot f(y) dy$$

$$= \int_{\mathbf{S}^{d-1}} \Omega(\theta) \Big[\int_{r>\varepsilon} \frac{1}{r} \Big(\prod_{i=1}^n \frac{A_i(x) - A_i(x - r\theta)}{r} \Big) \cdot f(x - r\theta) dr \Big] d\sigma(\theta)$$

$$= \frac{1}{2} \int_{\mathbf{S}^{d-1}} \Omega(\theta) \Big[\int_{|r|>\varepsilon} \frac{1}{r} \Big(\prod_{i=1}^n \frac{A_i(x) - A_i(x - r\theta)}{r} \Big) \cdot f(x - r\theta) dr \Big] d\sigma(\theta).$$

For convenience, we set the integral in the square bracket as $\mathcal{C}^1_{\varepsilon,\theta}[\nabla A_1, \cdots, \nabla A_n, f](x)$. Now applying the Minkowski inequality, making a change of variable $x = s\theta + z$ with

$s \in \mathbb{R}$ and $z \in L(\theta)$ ($L(\theta)$ is the hyperplane which is perpendicular to θ), utilizing Proposition 2.6 with dimension one and the Hölder inequality, we finally get that

$$\|\mathcal{C}_\Omega[\nabla A_1, \cdots, \nabla A_n, f]\|_{L^r(\mathbb{R}^d)}$$

$$\lesssim \int_{\mathbf{S}^{d-1}} |\Omega(\theta)| \Big[\int_{L(\theta)} \int_{\mathbb{R}} |\mathcal{C}^1_{\varepsilon,\theta}[\nabla A_1, \cdots, \nabla A_n, f](s\theta + z)|^r ds dz \Big]^{\frac{1}{r}} d\sigma(\theta)$$

$$\lesssim \int_{\mathbf{S}^{d-1}} |\Omega(\theta)| \Big[\int_{L(\theta)} \Big(\prod_{i=1}^n \|\nabla A_i(\cdot \theta + z)\|^r_{L^{q_i}(\mathbb{R})} \Big) \|f(\cdot \theta + z)\|^r_{L^p(\mathbb{R})} dz \Big]^{\frac{1}{r}} d\sigma(\theta)$$

$$\lesssim \|\Omega\|_{L^1(\mathbf{S}^{d-1})} \Big(\prod_{i=1}^n \|\nabla A_i\|_{L^{q_i}(\mathbb{R}^d)} \Big) \|f\|_{L^p(\mathbb{R}^d)},$$

which ends the proof. $\qquad\qquad\square$

Remark 3.2. Proposition 3.1 may also hold if we consider that Ω is homogenous of order zero and satisfies

$$(3.1) \quad \Omega(-\theta) = (-1)^n \Omega(\theta); \quad \int_{\mathbf{S}^{d-1}} \Omega(\theta) \theta^\alpha d\theta = 0, \quad \text{for all } \alpha \in \mathbb{Z}_+^d \text{ with } |\alpha| = n;$$

and $\Omega \in L \log^+ L(\mathbf{S}^{d-1})$. In such a case, then method of rotation can not be applies directly. One may need to insert $I = \sum_{j=1}^d R_j^2$ into the kernel of Calderón commutator where R_j is the Riesz transform. Since the commutator is a non convolution operator, there are some tail terms that should be dealt carefully. We don't pursue these matters in the present paper. For those reader who are interested in this case, we refer to see the very detailed discussion by B. Bajsanski and R. Coifman [2], which may also work here.

Proposition 3.3. Let $r = 1$, $q_1 = \cdots = q_n = \infty$, $p = 1$. Suppose $\Omega \in L \log^+ L(\mathbf{S}^{d-1})$. Then

$$\|\mathcal{C}_\Omega[\nabla A_1, \cdots, \nabla A_n, f]\|_{L^{1,\infty}(\mathbb{R}^d)} \lesssim C_\Omega \Big(\prod_{i=1}^n \|\nabla A_i\|_{L^\infty(\mathbb{R}^d)} \Big) \|f\|_{L^1(\mathbb{R}^d)}.$$

Proof. When $q_1 = \cdots = q_n = \infty$, A_i is a Lipschitz function for $i = 1, \cdots, n$. Fix all A_i. We may regard $\mathcal{C}_\Omega[\nabla A_1, \cdots, \nabla A_n, f](x)$ as a linear function of f. Then the kernel

$$\frac{1}{|x-y|^d} \Big(\prod_{i=1}^n \frac{A_i(x) - A_i(y)}{|x-y|} \Big)$$

is a standard Calderón-Zygmund kernel (see *e.g.* [18, Page 211, Definition 4.1.2]). Then by the recent result of Ding and the author [11, Theorem 1.1], one immediately finish the proof of Proposition 3.3. $\qquad\square$

After establishing Proposition 3.1 and Proposition 3.3, one can get the rest of the proof of Theorem 1.4 by using the similar way in the proof of Theorem 1.1. One can check the proof step-by-step in which the applications of Proposition 2.6 and Proposition 2.7 in Propositions 2.9, 2.10, 2.11, 2.12, 2.13 are just replaced by Proposition 3.1 and Proposition 3.3. There is only one thing that we should be careful. When giving a explicit estimate of $K(x-y)$ in Propositions 2.9, 2.10, 2.11, 2.12, 2.13, we just use the boundedness condition (1.2): $|K(x)| \lesssim |x|^{-d}$, and then apply Lemma 2.5 in a special

case $|\Omega| \equiv 1$ to get the required bound. While in the rest of the proof of Theorem 1.4, the above arguments are replaced by $|K(x)| \leq \frac{|\Omega(x)|}{|x|^d}$ and apply Lemma 2.5 with a rough kernel Ω. The verification of the details of this proof is omitted. Finally, it is easy to see that those examples in Proposition 2.14 also work here.

References

1. R. A. Adams and J. F. Fournier, *Sobolev spaces.* Second edition. Pure and Applied Mathematics (Amsterdam), 140. Elsevier/Academic Press, Amsterdam, 2003.
2. B. Bajsanski and R. Coifman, *On singular integrals*, Proc. Sympos. Pure Math., **10**, 1-17, Amer. Math. Soc., Providence, R.I. 1967.
3. A. P. Calderón, *Commutators of singular integral operators*, Proc. Nat. Acd. Sci. USA. **53** (1965), 1092-1099.
4. A. P. Calderón, *Cauchy integrals on Lipschitz curves and related operators,* Proc. Nat. Acad. Sci. USA, **74** (1977), 1324-1327.
5. A. P. Calderón, *Commutators, singular integrals on Lipschitz curves and application*, Proc. Inter. Con. Math. Helsinki, 1978, 85-96, Acad. Sci. Fennica, Helsinki, 1980.
6. C. P. Calderón, *On commutators of singular integrals*, Studia Math., **53** (1975), 139-174.
7. C. P. Calderón, *On a singular integral*, Studia Math., **65** (1979), no.3, 313-335.
8. R. Coifman and Y. Meyer, *On commutators of singular integral and bilinear singular integrals*, Trans. Amer. Math. Soc., 212 (1975), 315-331.
9. R. Coifman and Y. Meyer, *Au delà des opérateurs pseudo-diffrentiels*, Astérisque, 57 (1978).
10. M. Christ and J. Journé, *Polynomial growth estimates for multilinear singular integral operators*, Acta Math. **159** (1987), 51-80.
11. Y. Ding and X. Lai, *Weak type (1,1) bound criterion for singular integral with rough kernel and its applications.* Trans. Amer. Math. Soc., to appear. arXiv:1509.03685.
12. R. A. DeVore and R. C. Sharpley, *On the differentiability of functions in $\mathbb{R}^n$.* Proc. Amer. Math. Soc. 91 (1984), no. 2, 326-328.
13. R. A. DeVore and R. C. Sharpley, *Maximal functions measuring smoothness.* Mem. Amer. Math. Soc. 47 (1984), no. 293, viii+115 pp.
14. 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.
15. C. Fefferman, *Recent Progress in Classical Fourier Analysis*, Proc. Inter. Con. Math., Vancouver, 1974, 95-118.
16. P. W. Fong, *Smoothness properties of symbols, Calderón commutators and generalizations.* Thesis (Ph.D.)-Cornell University. 2016. 60 pp.
17. L. Grafakos, *Classic Fourier Analysis*, Graduate Texts in Mathematics, Vol. **249** (Third edition), Springer, New York, 2014.
18. L. Grafakos, *Modern Fourier Analysis*, Graduate Texts in Mathematics, Vol. **250** (Third edition), Springer, New York, 2014.
19. M. Hadžić, A. Seeger, C. K. Smart, B. Street, *Singular integrals and a problem on mixing flows.* Annales de l'Institut Henri Poincare/Analyse non lineaire, to appear. arXiv:1612.03431
20. 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.

21. A. Seeger, C. K. Smart and B. Street, *Multilinear singular integral forms of Christ-Journé type.* Memoirs of the American Mathematical Society, to appear. arXiv:1510.06990.

22. E. M. Stein, *Singular Integrals and Differentiability Properties of Functions.* Princeton Mathematical Series, No. 30, Princeton University Press, Princeton, N.J. 1970.

23. 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.

24. E. M. Stein, *Editor's note: The differentiability of functions in $\mathbb{R}^n$.* Ann. of Math. (2) 113 (1981), 383-385.

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