Some jump and variational inequalities for the Calderón commutators and related operators

Yanping Chen
Department of Applied Mathematics, School of Mathematics and Physics,
University of Science and Technology Beijing,
Beijing 100083, The People’s Republic of China
E-mail: yanpingch@ustb.edu.cn

Yong Ding
School of Mathematical Sciences, Beijing Normal University,
Laboratory of Mathematics and Complex Systems (BNU), Ministry of Education,
Beijing 100875, The People’s Republic of China
E-mail: dingy@bnu.edu.cn

Guixiang Hong
School of Mathematics and Statistics, Wuhan University,
Wuhan 430072, The People’s Republic of China
E-mail: guixiang.hong@whu.edu.cn

and

Jie Xiao
Department of Mathematics and Statistics, Memorial University of Newfoundland,
St. Johns, Newfoundland and Labrador, Canada A1C 5S7
E-mail: jxiao@mun.ca

ABSTRACT In this paper, we establish jump and variational inequalities for the Calderón commutators, which are typical examples of non-convolution Calderón-Zygmund operators. For this purpose, we also show jump and variational inequalities for para-products and commutators from pseudo-differential calculus, which are of independent interest. New ingredients in the proofs involve identifying Carleson measures constructed from sequences of stopping times, in addition to many Littlewood-Paley type estimates with gradient.

MR(2000) Subject Classification: 42B20, 42B25

Keywords: Jump and variational inequalities; Calderón commutators; pseudo-differential calculus; stopping times; Carleson measure

*The research was supported by NSF of China (Grant: 11471033, 11371057, 11571160, 11601396), NCET of China (Grant: NCET-11-0574), Thousand Youth Talents Plan of China (Grant: 429900018-101150(2016)), Funds for Talents of China (Grant: 413100002), the Fundamental Research Funds for the Central Universities (FRF-TP-12-006B, 2014KJJC10) and SRFDP of China (Grant: 20130003110003).

1Corresponding author.
1 Introduction

Motivated by the modulus of continuity of Brownian motion, Lépingle [41] established the first variational inequality for general martingales among many other interesting results. In [53], Pisier and Xu established implicitly the jump inequality (explicitly stated in Lemma 6.7 of [34]), and then by real interpolation provided another proof of Lépingle’s variational inequality. The advantage of Pisier and Xu’s approach is that it works also for vector-valued martingales.

Bourgain [4] is the first one who exploited Lépingle’s result to obtain corresponding variational estimates for the Birkhoff ergodic averages along subsequences of natural numbers and then directly deduce point-wise convergence results without previous knowledge that point-wise convergence holds for a dense subclass of functions, which are not available in some ergodic models. In particular, Bourgain’s work [4] has initiated a new research direction in ergodic theory and harmonic analysis. In [34, 36, 35, 5, 6], Jones and his collaborators systematically studied variational inequalities for ergodic averages and truncated singular integrals of homogeneous type. Since then many other publications came to enrich the literature on this subject (cf. e.g. [25, 40, 19, 37, 33, 52, 32]). Recently, several works on weighted as well as vector-valued variational inequalities in ergodic theory and harmonic analysis have also appeared (cf. e.g. [42, 39, 33]); and several results on $\ell^p(\mathbb{Z}^d)$-estimates of $q$-variations for discrete operators of Radon type have also been established (cf. e.g. [38, 46, 47, 50, 61]).

All the operators considered in the previous cited papers have nice symmetry properties, for instance, semigroup property or dilation invariance properties. So far as we know, it is still unknown whether jump and variational inequalities hold for all singular integrals of convolution type (see [47]), let alone for all standard Calderón-Zygmund operators. However, in this paper, we manage to show jump and variational inequalities for the Calderón commutators, which are typical examples of non-convolution Calderón-Zygmund operators. For this purpose, we first show jump and variational inequalities for para-products and commutators from pseudo-differential calculus, which are of independent interest.

The Calderón commutators (see [8, 9]) originate from a representation of linear differential operators by means of singular integral operators, which is an approach to the uniqueness of the Cauchy problem for partial differential equations (see [7]). Given a positive integer $m$, every linear partial differential operator $L$ of homogeneous order $m$ with bounded variable coefficients on Euclidean space $\mathbb{R}^n$ can be expressed as

$$Lf = T\Lambda^m f,$$

where $\hat{\Lambda}f = \varphi(\xi)\hat{f}(\xi)$, $\varphi(\xi)$ is a positive infinitely differentiable function such that $\varphi(\xi) = |\xi|$ if $|\xi| \geq 1$, and $T$ is a singular integral operator

$$T f = \int |\xi|^{-m} \sum_{|\gamma| = m} b_\gamma(x)(-i\xi)^\gamma e^{ix\cdot \xi} \hat{f}(\xi) d\xi + \int r(x,\xi)e^{ix\cdot \xi} \hat{f}(\xi) d\xi$$

with $\gamma$ being an multi-indices of non-negative integers and $|\gamma| = \gamma_1 + \cdots + \gamma_n$. Let $B$ be the operator given by the multiplication of the Lipschitzian function $b(x)$. For simplicity, let us
consider the case $n = 1$, let $H$ be the Hilbert transform, as it is well known, this transform can be expressed as follows

$$Hf(x) = -\frac{i}{2\pi} \int_{-\infty}^{\infty} \text{sgn}\xi e^{ix\xi} \hat{f}(\xi) \, d\xi$$

and this makes it clear that $B, H$ and $BH$ are operators of the type of the generalized $T$ and the simplest of their kind. In order to show that $HB$ is of the same type, since $HB = BH - [b, H]$, it would suffice to show that $[b, H] \frac{d}{dx}$ is bounded in $L^p(\mathbb{R})$, $1 < p < \infty$. Calderón [9] introduced the first Calderón commutator which is defined by

$$[b, H] \frac{d}{dx} := p.v. \int_{-\infty}^{\infty} \frac{(-1)^n}{x - y} \frac{b(x) - b(y)}{x - y} f(y) \, dy.$$ 

The integral on the right, which in the case $b(x) = x$ reduces to the Hilbert transform, is the one studied in [8]. Note that

$$[b, H] \frac{d}{dx} = [b, H] \frac{d}{dx} - H[b, \frac{d}{dx}]$$

and since the operator $[b, \frac{d}{dx}]$ is multiplication by $b'(x)$, which is a bounded function if $b(x)$ is Lipschitzian, $H[b, \frac{d}{dx}]$ is bounded in $L^p(\mathbb{R})$ and the continuity of $[b, H] \frac{d}{dx}$ is equivalent to that of $[b, H] \frac{d}{dx}$. Thus, the role of the first Calderón commutator in the theory of partial differential equations becomes apparent. Commutator $[b, H] \frac{d}{dx}$ also plays an important role in the theory of Cauchy integral along Lipschitz curve in $C$ and the Kato square root problem on $\mathbb{R}$ (see [7, 22, 44, 45] for the details).

As Calderón did in [8], there are large classes of commutators which are of independent interest in harmonic analysis. For $\varepsilon > 0$, suppose that $C_\varepsilon f$ is the truncated Calderón commutator which is defined by

$$C_\varepsilon f(x) = \int_{|x-y| > \varepsilon} \frac{\Omega(x-y)}{|x-y|^{n+1}} (b(x) - b(y)) f(y) \, dy,$$

where $\Omega$ is homogeneous of degree zero, integrable on $S^{n-1}$ (the unit sphere in $\mathbb{R}^n$) and satisfies

$$\int_{S^{n-1}} \Omega(x')(x_k')^N \, d\sigma(x') = 0, k = 1, \ldots, n,$$

for all integers $0 \leq N \leq 1$. Then for $f \in C_0^\infty(\mathbb{R}^n)$, the Calderón commutator

$$C f(x) = \lim_{\varepsilon \to 0^+} C_\varepsilon f(x), \text{ a.e. } x \in \mathbb{R}^n.$$

Using the method of rotation, Calderón [8] has shown the boundedness of the commutator $C$ and as a consequence obtained the boundedness of the operators $[b, T] \nabla$ and $\nabla [b, T]$, where $T$ is a homogeneous singular integral operator with some symbol $K$ which can be defined similarly as $C$: For $f \in C_0^\infty(\mathbb{R}^n)$

$$T f(x) = \lim_{\varepsilon \to 0^+} T_\varepsilon f(x), \text{ a.e. } x \in \mathbb{R}^n.$$
with
\[(1.5)\]
\[T_\varepsilon f(x) = \int_{|y|>\varepsilon} K(y) f(x-y)dy,\]
where \(K\) is homogeneous of degree \(-n\), belongs to \(L^1_{\mathrm{loc}}(\mathbb{R}^n)\) and satisfies the cancelation condition
\[(1.6)\]
\[\int_{S^{n-1}} K(y')d\sigma(y') = 0.\]
Later on, many authors made important progress on the Calderón commutators, one can consult [15, 17, 16, 51, 29, 30, 60, 59, 49, 50, 28, 13], among numerous references, for its development and applications.

That the point-wise principle value (1.3) exists for all \(f \in L^p(\mathbb{R}^n)\) follows from the maximal inequality which was established in [50]. In the present paper, the variational inequality that we will show implies the maximal inequality due to (1.10) below. Moreover our result provide quantitative information of the convergence.

In order to present our results in a precise way, let us fix some notations. Given a family of complex numbers \(a = \{a_t: t \in \mathbb{R}_+\}\) and \(0 < q < \infty\), the strong and the weak \(q\)-variation norm of the family \(a\) is defined respectively by
\[(1.7)\]
\[\|a\|_{V_q} = \sup \| (a_{t_k} - a_{t_{k-1}})_{k \geq 1} \|_{\ell^q},\]
and
\[(1.8)\]
\[\|a\|_{V_{q,\infty}} = \sup \| (a_{t_k} - a_{t_{k-1}})_{k \geq 1} \|_{\ell^{q,\infty}},\]
where the supremum runs over all increasing sequences \(\{t_k: k \geq 0\}\). Here \(\ell^q\) (resp. \(\ell^{q,\infty}\)) denote the Lebesgue \(L^q\) (resp. weak \(L^q\)) norm on the set of integers. From the definition, it is quite clear that the following inequalities hold: For any \(0 < r < q < \infty\),
\[(1.9)\]
\[\|a\|_{V_{q,\infty}} \leq \|a\|_{V_q} \leq \|a\|_{V_{r,\infty}}.\]
On the other hand, it is also trivial that
\[(1.10)\]
\[\|a\|_{L^\infty} := \sup_{t \in \mathbb{R}_+} |a_t| \leq \|a\|_{V_q} + |a_{t_0}| \quad \text{for} \quad 0 < q < \infty,\]
for some fixed \(t_0\).

Via the definition of the strong and weak \(q\)-variation norm of a family of numbers, one may define the strong and the weak \(q\)-variation function \(V_q(\mathcal{F})\) and \(V_{q,\infty}(\mathcal{F})\) of a family \(\mathcal{F}\) of functions. Given a family of Lebesgue measurable functions \(\mathcal{F} = \{F_t: t \in \mathbb{R}_+\}\) defined on \(\mathbb{R}^n\), for fixed \(x\) in \(\mathbb{R}^n\), the value of the strong \(q\)-variation function \(V_q(\mathcal{F})\) of the family \(\mathcal{F}\) at \(x\) is defined by
\[(1.11)\]
\[V_q(\mathcal{F})(x) = \|\{F_t(x)\}_{t \in \mathbb{R}}\|_{V_q}, \quad q > 0;\]
while the value of the weak $q$-variation function $V_{q,\infty}(\mathcal{F})$ of the family $\mathcal{F}$ at $x$ is defined by

$$V_{q,\infty}(\mathcal{F})(x) = \|\{F_t(x)\}_{t \in \mathbb{R}}\|_{V_{q,\infty}}, \quad q > 0;$$

Suppose $\mathcal{A} = \{A_t\}_{t > 0}$ is a family of operators on $L^p(\mathbb{R}^n)$ ($1 \leq p \leq \infty$). The related strong and weak $q$-variation operator are simply defined respectively as

$$V_q(\mathcal{A}f)(x) = \|\{A_t(f)(x)\}_{t > 0}\|_{V_q}, \quad \forall f \in L^p(\mathbb{R}^n)$$

and

$$V_{q,\infty}(\mathcal{A}f)(x) = \|\{A_t(f)(x)\}_{t > 0}\|_{V_{q,\infty}}, \quad \forall f \in L^p(\mathbb{R}^n).$$

It is easy to observe from the definition of $q$-variation norm that for any $x$ if $V_{q,\infty}(\mathcal{A}f)(x) < \infty$ for some $q < \infty$, then $\{A_t(f)(x)\}_{t > 0}$ converges when $t \to 0$ or $t \to \infty$. In particular, if $V_{q,\infty}(\mathcal{A}f)$ belongs to some function spaces such as $L^p(\mathbb{R}^n)$ or $L^{p,\infty}(\mathbb{R}^n)$, then the sequence converges almost everywhere without any additional condition. This is why mapping property of strong or weak $q$-variation operator is so interesting in ergodic theory and harmonic analysis. On the other hand, from (1.10), variational inequality is much stronger than corresponding maximal inequality. Namely, for any $f \in L^p(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$, we have

$$A^*(f)(x) \leq V_q(\mathcal{A}f)(x) \quad \text{for } q \geq 1,$$

where $A^*$ is the maximal operator defined by

$$A^*(f)(x) := \sup_{t > 0} |A_t(f)(x)|$$

and thus is more interesting.

As we know for a family of Lebesgue measurable functions $\mathcal{F} = \{F_t(x) : t \in \mathbb{R}_+\}$, there is another related notion called $\lambda$-jump function $N_\lambda(\mathcal{F})$ whose value at $x$ is defined as the supremum over all $N$ such that there exist $t_0 < t_1 < t_2 < \ldots < t_N$ with

$$|F_{t_k}(x) - F_{t_{k-1}}(x)| > \lambda$$

for all $k = 1, \ldots, N$. It is easy to check that this function is related with the weak $q$-variation norm as follows

$$V_{q,\infty}(\mathcal{F})(x) = \sup_{\lambda > 0} \lambda (N_\lambda(\mathcal{F})(x))^{1/q}.$$

We refer the reader to [37] for more information on $\lambda$-jump functions.

Now, we can formulate our main result as follows.

**Theorem 1.1.** Let $b \in \text{Lip}(\mathbb{R}^n)$ and $\mathcal{C} = \{C_\varepsilon\}_{\varepsilon > 0}$ where $C_\varepsilon(f)$ are as in (1.11) with $\Omega$ satisfying (1.2). If $\Omega \in L(\log^+L)^2(\mathbb{S}^{n-1})$, then the following jump inequality holds for $1 < p < \infty$, namely,

$$\sup_{\lambda > 0} \|\lambda \sqrt{N_\lambda(\mathcal{C}f)}\|_{L^p} \leq C_{p,n,\Omega} \|\nabla b\|_{L^\infty} \|f\|_{L^p}.$$
Quite remarkably, on the one hand, as in [53, 41, 37], the jump inequality (1.14) implies all the strong $q$-variational inequality ($2 < q < \infty$) by a real interpolation argument (see for instance Lemma 2.1 of [37]), and thus implies all the weak $q$-variational inequality and maximal inequality by (1.9) and (1.10). For this reason, we will not state explicitly any $q$-variational inequality in the present paper. On the other hand, the strong 2-variational inequality, and thus any $q$-variational inequality ($q < 2$) may fail, see [54] and [1] for related information. However, it is still unknown whether the weak 2-variational inequality holds, that is, whether the estimate (1.14) is still true if the supremum over $\lambda$ can be put inside the norm.

Consequently, we have the following Calderón-type estimates.

**Corollary 1.2.** Let $1 < p < \infty$. Let $b \in \text{Lip}(\mathbb{R}^n)$ and $\mathcal{T}_b = \{[b, T_\varepsilon]\}_{\varepsilon > 0}$ where $T_\varepsilon$ are as in (1.2). Suppose that $K(x)$ have locally integrable first-order derivatives in $|x| > 0$ and suppose that $K(x)$ and the partial derivatives of $K(x)$ belong locally to $L(\log^+ L)^2$ in $|x| > 0$. If $f$ is continuously differentiable and have compact support, then the following jump inequality holds namely,

$$\sup_{\lambda > 0} \| \lambda \sqrt{N_\lambda(\nabla f)} \|_{L^p} \leq C_{p, n, K} \| \nabla b \|_{L^\infty} \| f \|_{L^p}.$$  

Furthermore, if for $\varepsilon > 0$, $[b, T_\varepsilon]f$ has first-order derivatives in $L^p(\mathbb{R}^n)$, then

$$\sup_{\lambda > 0} \| \lambda \sqrt{N_\lambda(\nabla T_\varepsilon f)} \|_{L^p} \leq C_{p, n, K} \| \nabla b \|_{L^\infty} \| f \|_{L^p}.$$  

As in most of the previously cited paper (in particular see Lemma 1.3 in [37]), we shall show estimate (1.14) by showing separately the corresponding inequalities for the long and short variation. That is, we are reduced to prove for $1 < p < \infty$

$$\sup_{\lambda > 0} \| \lambda \sqrt{N_\lambda(\{C_{2^k} f \}_k)} \|_{L^p} \leq C_{p, n, \Omega} \| \nabla b \|_{L^\infty} \| f \|_{L^p}$$ 

and

$$\| S_2(\mathcal{C} f) \|_{L^p} \leq C_{p, n, \Omega} \| \nabla b \|_{L^\infty} \| f \|_{L^p},$$

where

$$S_2(\mathcal{C} f)(x) = \left( \sum_{j \in \mathbb{Z}} |V_{2,j}(\mathcal{C} f)(x)|^2 \right)^{1/2},$$

with

$$V_{2,j}(\mathcal{C} f)(x) = \left( \sup_{2^j \leq t_0 < \ldots < t_N < 2^{j+1}} \sum_{k=0}^{N-1} |C_{t_{k+1}} f(x) - C_{t_k} f(x)|^2 \right)^{1/2}.$$  

Although we encounter some difficulties in proving (1.16) (see for instance some lemmas in Section 5), the novelty of the proof lies in showing (1.15). We need two results which are of independent interest.

The first auxiliary result is variational inequality for some kind of para-product (see [19] for related results), whose formulation is motivated by the results on maximal operators of
Duoandikoetxea and Rubio de Francia [20] and para-products [26]. Let $\mu$ be a compactly supported finite Borel measure on $\mathbb{R}^n$, that is, $\mu$ is absolutely continuous on the Lebesgue measure $dx$, its Radon-Nikodym derivative is a nonnegative Lebesgue measurable function on $\mathbb{R}^n$ with compact support set. We consider dilates $\mu_k$ of $\mu$ defined with respect to a group of dilations $\{2^k\}_{k \in \mathbb{Z}}$ defined by
$$\int f(x) \, d\mu_k(x) := \int_{\mathbb{R}^n} f(2^k x) \, d\mu.$$ Let $\Upsilon_k f(x) = \mu_k * f(x)$ for $k \in \mathbb{Z}$. A well known fact is, if $\mu$ satisfies the Fourier transform:
$$|\hat{\mu}(\xi)| \leq C|\xi|^{-\alpha},$$
for some $\alpha > 0$, then the maximal operator defined as $M_{\mu} f(x) = \sup_{k \in \mathbb{Z}} |\Upsilon_k f(x)|$ is bounded on $L^p(\mathbb{R}^n)$, $1 < p < \infty$. Further, if the Radon-Nikodym derivative $\zeta$ of $\mu$ satisfies the stronger condition:
$$\int_{\mathbb{R}^n} |\zeta(x + y) - \zeta(x)| \, dx \leq C|y|^{\tau}$$
for some $\tau > 0$, then $M_{\mu}$ is of weak type $(1,1)$ (see [55]). Clearly (1.18) implies (1.17). Suppose $\phi(x) \in \mathcal{S}(\mathbb{R}^n)$. Denote $\Phi_k f(x) = \phi_k * f(x)$, where $\phi_k(x) = 2^{-kn} \phi(2^{-k} x)$. Give a function $b$ on $\mathbb{R}^n$, we define the operator as follows:
$$\mathcal{Z}_b f = \{(\Phi_k b)(\Upsilon_k f)\}_{k \in \mathbb{Z}}.$$ for $f \in L^1_{\text{loc}}(\mathbb{R}^n)$. They are not operators of convolution type. We note that the transpose of the family of operators $\mathcal{Z}_b$ is formally given by the identity
$$\mathcal{Z}_b^t f = \{\Upsilon_k (f \Phi_k b)\}_{k \in \mathbb{Z}}.$$ We are now ready to state the first auxiliary result as follows.

**Theorem 1.3.** Suppose that $b \in L^\infty(\mathbb{R}^n)$. Let $\mathcal{Z}_b$ be defined as in (1.19) with $\mu$ being a compactly supported finite Borel measure on $\mathbb{R}^n$ and $\phi(x) \in \mathcal{S}(\mathbb{R}^n)$.

(i) If $\mu$ satisfies (1.17), then for $1 < p < \infty$ and $f \in L^p(\mathbb{R}^n)$, we have
$$\sup_{\lambda > 0} \|\lambda \sqrt{N_\lambda(\mathcal{Z}_b f)}\|_{L^p} \leq C_{p,n} \|b\|_{L^\infty} \|f\|_{L^p}.$$ (ii) In addition for $p = 1$, if $\mu$ satisfies (1.18), we have
$$\sup_{\lambda > 0} \|\lambda \sqrt{N_\lambda(\mathcal{Z}_b f)}\|_{L^{1,\infty}} \leq C_{p,n} \|b\|_{L^\infty} \|f\|_{L^1}.$$ The proof of Theorem [1.3] involves identifying two Carleson measures constructed from sequences of conditional expectations, one of which is in turn constructed from sequences of stopping times, see below Lemma [2.1] [2.2] and [3.1].

The second auxiliary result we need is the variational inequalities for commutators of pseudo differential calculus with Lipschitz functions. Since the 1960s, the theory of pseudo differential operators has played an important role in many exciting and deep investigations into linear PDE (see [23, 17, 3, 31, 56, 57, 58, 2, 44, 45]).
Theorem 1.4. For $k \in \mathbb{Z}$, let $\Phi_k$ be defined as Theorem 1.3. For $b \in \text{Lip}(\mathbb{R}^n)$, set $\mathcal{F}_b = \{[b, \Phi_k]\}_k$. Suppose that $f$ is continuously differentiable and has compact support.

(i) Then for $1 < p < \infty$, we have
\[
\sup_{\lambda > 0} \| \lambda \sqrt{N_\lambda(\mathcal{F}_b(\nabla f))} \|_{L^p} \leq C_{p,n} \| \nabla b \|_{L^\infty} \| f \|_{L^p}
\]
and
\[
\sup_{\lambda > 0} \| \lambda \sqrt{N_\lambda(\nabla \mathcal{F}_b f)} \|_{L^p} \leq C_{p,n} \| \nabla b \|_{L^\infty} \| f \|_{L^p}.
\]

(ii) In addition for $p = 1$ we have
\[
\sup_{\lambda > 0} \| \lambda \sqrt{N_\lambda(\mathcal{F}_b(\nabla f))} \|_{L^1, \infty} \leq C_{p,n} \| \nabla b \|_{L^\infty} \| f \|_{L^1}
\]
and
\[
\sup_{\lambda > 0} \| \lambda \sqrt{N_\lambda(\nabla \mathcal{F}_b f)} \|_{L^1, \infty} \leq C_{p,n} \| \nabla b \|_{L^\infty} \| f \|_{L^1}.
\]

The paper is organized as follows. In Section 2, some key lemmas will be introduced for the proof of 1.3. In Section 3 and Section 4, we give the proof of Theorem 1.3 and Theorem 1.4, respectively. In Section 5, we give some lemmas for the proof of Theorem 1.1. Section 6 and Section 7 are devoted to the proof of Theorem 1.1. In Section 8, we give the proof of Corollary 1.2. For $p \geq 1$, $p'$ denotes the conjugate exponent of $p$, that is, $p' = p/(p - 1)$. Throughout this paper, the letter “$C$” will stand for a positive constant which is independent of the essential variables and not necessarily the same one in each occurrence.

2 Some key lemmas

Let us begin with some lemmas and their proofs, which will play a key role in proving Theorem 1.3. We borrow some notations and results from [37, pp.6724]. For $j \in \mathbb{Z}$ and $\beta = (m_1, \cdots, m_n) \in \mathbb{Z}^n$, we denote the dyadic cube $\prod_{k=1}^n (m_k 2^j, (m_k + 1)2^j]$ in $\mathbb{R}^n$ by $Q_j^\beta$, and the set of all dyadic cubes with side-length $2^j$ by $D_j$. The conditional expectation of a local integrable $f$ with respect to $D_j$ is given by
\[
E_j f(x) = \sum_{Q \in D_j} \frac{1}{|Q|} \int_Q f(y) dy \cdot \chi_Q(x)
\]
for all $j \in \mathbb{Z}$.

Lemma 2.1. Let $\phi \in \mathcal{S}(\mathbb{R}^n)$ and $\widehat{\phi}(0) = 1$. For $k \in \mathbb{Z}$, denote by $\Phi_k f(x) = \phi_k * f(x)$, where $\phi_k(x) = 2^{-kn} \phi(2^{-k} x)$. Let $E_k$ be given above and $b \in \text{BMO}(\mathbb{R}^n)$. Let $\delta_{2^k}(t)$ be Dirac mass at the point $t = 2^k$. Then there is a constant $C > 0$ such that
\[
d\nu(x,t) = \sum_{k \in \mathbb{Z}} |\Phi_k b(x) - E_k b(x)|^2 dx \delta_{2^k}(t)
\]
is a Carleson measure on $\mathbb{R}^{n+1}_+$ with norm at most $C\|b\|_2^2$. 
Proof. For a cube $Q$ in $\mathbb{R}^n$ we let $Q^*$ be the cube with the same center and orientation whose side length is $100\sqrt{n}\ell(Q)$, where $\ell(Q)$ is the side length of $Q$. Fix a cube $Q$ in $\mathbb{R}^n$, split $b$ as

$$b = (b - b_Q)\chi_{Q^*} + (b - b_Q)\chi_{(Q^*)^c} + b_Q.$$ 

Let $T(Q) = Q \times (0, \ell(Q))$. Since $\Phi_k b_Q = b_Q$ and $\mathbb{E}_k b_Q = b_Q$, then

$$\Phi_k b_Q - \mathbb{E}_k b_Q = 0.$$ 

Thus,

$$\nu(T(Q)) = \sum_{2^k \leq \ell(Q)} \int_Q |\Phi_k(b)(x) - \mathbb{E}_k(b)(x)|^2 \, dx \leq 2\Sigma_1 + 2\Sigma_2,$$

where

$$\Sigma_1 = \sum_{k \in \mathbb{Z}} \int_{\mathbb{R}^n} |\Phi_k((b - b_Q)\chi_{Q^*})(x) - \mathbb{E}_k((b - b_Q)\chi_{Q^*})(x)|^2 \, dx$$

and

$$\Sigma_2 = \sum_{2^k \leq \ell(Q)} \int_Q \Phi_k((b - b_Q)\chi_{(Q^*)^c})(x) - \mathbb{E}_k((b - b_Q)\chi_{(Q^*)^c})(x)|^2 \, dx.$$

Then

$$\Sigma_1 \leq C \int_{Q^*} |b(x) - b_Q|^2 \, dx \leq C|Q||b|^2_*,$$

where in the first inequality we have used that

$$\left\| \left( \sum_{k \in \mathbb{Z}} |\Phi_k(g) - \mathbb{E}_k(g)|^2 \right)^{1/2} \right\|_{L^2} \leq C\|g\|_{L^2}$$

(see [37]). For $\Sigma_2$, we have

$$\Sigma_2 = \sum_{2^k \leq \ell(Q)} \int_Q |\Phi_k((b - b_Q)\chi_{(Q^*)^c})(x) - \mathbb{E}_k((b - b_Q)\chi_{(Q^*)^c})(x)|^2 \, dx$$

$$\leq C \sum_{2^k \leq \ell(Q)} \int_Q |\Phi_k((b - b_Q)\chi_{(Q^*)^c})(x)|^2 \, dx + C \sum_{2^k \leq \ell(Q)} \int_Q |\mathbb{E}_k((b - b_Q)\chi_{(Q^*)^c})(x)|^2 \, dx.$$

Since $\phi(x) \leq \frac{1}{(1 + |x|)^{n+\delta}}$ for some $\delta > 1$, then by the same argument of [26, 27], we get

$$\sum_{2^k \leq \ell(Q)} \int_Q |\Phi_k((b - b_Q)\chi_{(Q^*)^c})(x)|^2 \, dx \leq C|Q||b|^2_*.$$

Recall that

$$\mathbb{E}_k f(x) = \sum_{Q \in \mathcal{D}_k} \frac{1}{|Q|} \int_Q f(y) dy \cdot \chi_Q(x).$$
Then we get
\[
\sum_{2^k \leq \ell(Q)} \int_\Omega |\mathbb{E}_k((b - b_Q)\chi_{(Q^*)^c})(x)|^2 \, dx \leq \sum_{2^k \leq \ell(Q)} \sum_{Q \in D_k} \int_{\tilde{Q}} \left| \int_{\tilde{Q}} \frac{1}{|Q|} \int_{\tilde{Q}} (b - b_Q)\chi_{(Q^*)^c}(y) \, dy \right|^2 \chi_{\tilde{Q}}(x) \, dx
\]
\[
= \sum_{2^k \leq \ell(Q)} \sum_{Q \in D_k} \int_{\tilde{Q}} \left| \int_{\tilde{Q} \cap (Q^*)^c} (b(y) - b_Q) \, dy \right|^2 \, dx.
\]

If \( \tilde{Q} \cap (Q^*)^c \neq \emptyset \), since \( \ell(\tilde{Q}) = 2^k \) and \( \ell(Q^*) = 100\sqrt{n\ell(Q)} \geq 100\sqrt{n2^k} = 100\sqrt{n\ell(\tilde{Q})} \), then we get
\[
\tilde{Q} \cap Q = \emptyset.
\]
Therefore, either \( \tilde{Q} \cap (Q^*)^c = \emptyset \) or \( \tilde{Q} \cap (Q^*)^c \neq \emptyset \), we can get
\[
(2.2) \quad \sum_{2^k \leq \ell(Q)} \int_\Omega |\mathbb{E}_k((b - b_Q)\chi_{(Q^*)^c})(x)|^2 \, dx = 0.
\]
Together,
\[
\Sigma_2 \leq C|Q||b||_s^2.
\]
Then combined this with \( \Sigma_1 \), we get
\[
\nu(T(Q)) \leq C|Q||b||_s^2.
\]
This says that
\[
d\nu(x, t) = \sum_{k \in \mathbb{Z}} |\Phi_k b(x) - \mathbb{E}_k b(x)|^2 \, dx \delta_{2^k}(t)
\]
is a Carleson measure on \( \mathbb{R}^{n+1}_+ \) with norm at most \( C||b||_s^2 \).

**Lemma 2.2.** For \( j \in \mathbb{Z} \), let \( \mathbb{E}_j \) be given above and \( b \in BMO(\mathbb{R}^n) \). Let \( \delta_{t_k}(t) \) be Dirac mass at the point \( t = t_k \). Then there is a constant \( C > 0 \) such that
\[
d\nu(x, t) = \sum_{k \geq 0} |\mathbb{E}_{t_{k+1}} b(x) - \mathbb{E}_{t_k} b(x)|^2 \, dx \delta_k(t)
\]
is a Carleson measure on \( \mathbb{R}^{n+1}_+ \) with norm at most \( C||b||_s^2 \), where \( \{t_k\}_{k \geq 0} \) is any sequence of decreasing stopping times and the bound does not depend on the stopping times.

**Proof.** The proof is essentially similar to Lemma 2.1. More precisely, we need to estimate in \( \Sigma_1 \) with \( \Phi_k((b - b_Q)\chi_{Q^*})(x) - \mathbb{E}_k((b - b_Q)\chi_{Q^*})(x) \) replaced by \( \mathbb{E}_{t_{k+1}}((b - b_Q)\chi_{Q^*})(x) - \mathbb{E}_{t_k}((b - b_Q)\chi_{Q^*})(x) \). The desired result follows from
\[
\left\| \sum_{k \geq 0} |\mathbb{E}_{t_{k+1}}(g) - \mathbb{E}_{t_k}(g)|^2 \right\|_{L^2} \leq C\|g\|_{L^2}
\]
due to Burkholder-Gundy inequality since \( \{\mathbb{E}_{t_k}(g)\}_{k \geq 0} \) forms a new martingale (see for instance [53]). In \( \Sigma_2 \), we replace \( \Phi_k((b - b_Q)\chi_{(Q^*)^c})(x) - \mathbb{E}_k((b - b_Q)\chi_{(Q^*)^c})(x) \) with \( \mathbb{E}_{t_{k+1}}((b -
Then we claim the integrand equal zero. Indeed, for any $x < s < t$ satisfies

$$\sum_{k \in \mathbb{Z}} |\mathbb{E}_{t_k} ((b - bq) \chi_{(Q^*)^c})(x)|^2 dx = 0$$

and

$$\int_{Q} \sum_{k \in \ell(Q)} |\mathbb{E}_{t_k} ((b - bq) \chi_{(Q^*)^c})(x)|^2 dx = 0.$$

Let us explain briefly the second identity. The first identity follows similarly. We first write the left hand side as

$$\sum_{k \geq 0} \int_Q \chi_{2^{k} Q} \mathbb{E}_{t_k} ((b - bq) \chi_{(Q^*)^c})(x) |^2 dx.$$  

Then we claim the integrand equal zero. Indeed, for any $x \in Q$,

$$\mathbb{E}_{t_k} ((b - bq) \chi_{(Q^*)^c})(x) = \frac{1}{|Q(t_k(x))|} \int_{Q(t_k(x)) \cap (Q^*)^c} (b(y) - bq) dy$$

where $Q(t_k(x))$ is the unique dyadic cube containing $x$ with side-length equal to $2^{k}(x)$. Then $\ell(Q(t_k(x))) \leq \ell(Q)$ implies $Q(t_k(x)) \cap (Q^*)^c = \emptyset$.

\[ \square \]

### 3 Proof of Theorem 1.3

We may assume $\int d\mu \neq 0$ since otherwise by the easy fact $\ell^2$ embeds into $\ell^{2, \infty}$, $\lambda \sqrt{N \lambda(\mathbb{M} f)}(x)$ is pointwisely dominated by the square function $C\|b\|_{L^{\infty}}(\sum_{k \in \mathbb{Z}} |H_k * f(x)|^2)^{1/2}$, and known bounds from [20] apply. Therefore we may normalized $\mu$ so that $\int d\mu = 1$. Let $\omega$ be a smooth function with compact support such that $\int_{\mathbb{R}^n} \omega(x) dx = 1$ and decomposes $\mu = \omega * \mu + (\delta_0 - \omega) * \mu$ where $\delta_0$ is the Dirac mass at 0. This in turn decompose $\Phi_k b \mathbb{T}_k f$ into low and high frequency families $L = \{L_k\}$ and $H = \{H_k\}$, where

$$L_k f(x) = \Phi_k b(x) (\omega * \mu) * f(x) \quad \text{and} \quad H_k f(x) = \Phi_k b(x)(\mu * (\delta_0 - \omega)) * f(x).$$

By the quasi-triangle inequality, it suffices to bound $\lambda \sqrt{N \lambda(L f)}$ and $\lambda \sqrt{N \lambda(H f)}$ separately. Since $\mu * (\delta_0 - \omega)$ has vanishing mean value and satisfies condition (1.17), we recall from [20] that the square function

$$g(f)(x) = (\sum_{k \in \mathbb{Z}} |H_k f(x)|^2)^{1/2}$$

satisfies

$$\|g(f)\|_{L^p} \leq C \|f\|_{L^p}$$

for $1 < p < \infty$. Furthermore, if $\mu$ satisfies the stronger hypothesis [1.18], we can also get weak type $(1, 1)$ bounds for $g(f)$. The easy fact $\ell^2$ embeds into $\ell^{2, \infty}$ implies

$$\lambda \sqrt{N \lambda(H f)}(x) \leq C \|b\|_{L^{\infty}} g(f)(x),$$

$$\lambda \sqrt{N \lambda(L f)}(x) \leq C \|b\|_{L^{\infty}} g(f)(x),$$
so matters are reduced to bounding $\lambda \sqrt{N_\lambda(\mathcal{L}f)}$. We need to prove that for $1 < p < \infty$,

$$
\|\lambda \sqrt{N_\lambda(\mathcal{L}f)}\|_{L^p} \leq C\|b\|_{L^\infty}\|f\|_{L^p}
$$

and

$$
\alpha\{x \in \mathbb{R}^n : \lambda \sqrt{N_\lambda(\mathcal{L}f)}(x) > \alpha\} \leq C\|b\|_{L^\infty}\|f\|_{L^1}
$$

uniformly in $\lambda > 0$. Denote by $\Gamma_k$ \( f(x) = (\omega * \mu)_k \ast f(x) \). In the following, we will divide the proof into two cases: Case 1, $\Phi_k 1 \neq 0$; Case 2, $\Phi_k 1 = 0$.

**Case 1**, $\Phi_k 1 \neq 0$. By normalization, $\Phi_k 1$ can be assumed to 1. Then write

$$
\Phi_k b \Gamma_k f = \Phi_k b(\Gamma_k f - E_k f) + (\Phi_k b - E_k b)E_k f + E_k b E_k f
$$

$$
:= W^1_k f + W^2_k f + W^3_k f.
$$

By subadditivity,

$$
\lambda \sqrt{N_\lambda(\mathcal{L}f)} \leq C\lambda \sqrt{N_{\lambda/3}(\{W^1_k f\}_k)} + C\lambda \sqrt{N_{\lambda/3}(\{W^2_k f\}_k)} + C\lambda \sqrt{N_{\lambda/3}(\{W^3_k f\}_k)}.
$$

To bound $\lambda \sqrt{N_\lambda(\mathcal{L}f)}$, we first need to prove $L^2$ norm of the above three parts and then weak $(1, 1)$-norm of $\lambda \sqrt{N_\lambda(\mathcal{L}f)}$. For $W^1_k f$, by Lemma 3.2 in [37], we have for $1 < p < \infty$,

$$
\left\| \lambda \sqrt{N_\lambda(\{W^1_k f\}_k)} \right\|_{L^p} \leq C\left\| \left( \sum_{k \in \mathbb{Z}} |\Phi_k b(\Gamma_k f - E_k f)|^2 \right)^{1/2} \right\|_{L^p}
$$

$$
\leq C\|\Phi_k b\|_{L^\infty}\left\| \left( \sum_{k \in \mathbb{Z}} |\Gamma_k f - E_k f|^2 \right)^{1/2} \right\|_{L^p}
$$

$$
\leq C\|b\|_{L^\infty}\|f\|_{L^p}.
$$

For $W^2_k f$. Let $F(x, 2^k) = E_k f(x)$. Define $F^*(x) = \sup_{k \geq 0} \sup_{y \in \mathbb{R}^n, |y - x| < 2^k} |F(y, 2^k)|$. It is easy to see that $|F^*(x)| \leq C M f(x)$, where $M$ is the Hardy-Littlewood maximal operator. Then by Lemma 2.1 and Carleson’s inequality (see 24), we get

$$
\left\| \lambda \sqrt{N_\lambda(\{W^2_k f\}_k)} \right\|_{L^2} \leq C \sum_{k \in \mathbb{Z}} \int_{\mathbb{R}^n} |\Phi_k b(x) - E_k b(x)|^2 |E_k f(x)|^2 \, dx
$$

$$
\leq C\|b\|_{L^2}^2 \|M f\|_{L^2}^2 \leq C\|b\|_{L^\infty}^2 \|f\|_{L^2}^2.
$$

To deal with the third term $W^3_k f$, we need the following lemma.

**Lemma 3.1.** Fix $\lambda > 0$. For a.e. $x \in \mathbb{R}^n$, we can find a sequence of decreasing stopping times $\{t_i\}_{i \geq 0}$ such that

$$
\lambda \sqrt{N_\lambda(\{W^3_k f\}_k \cap \mathbb{Z})}(x) \leq 2\left( \sum_{i \geq 0} |W^3_{t_{i+1}} f(x) - W^3_{t_i} f(x)|^2 \right)^{1/2}.
$$
Proof. Since \( f \in L^p(\mathbb{R}^n) \), \( b \in L^\infty(\mathbb{R}^n) \), by Jensen inequality we have

\[
\sup_{x \in \mathbb{R}^n} |W_k^3 f(x)| \leq \sup_{x \in \mathbb{R}^n} |E_k f(x) E_k b(x)| \\
\leq \sup_{x \in \mathbb{R}^n} (E_k |f|^p)^{\frac{1}{p}} (x) |E_k b(x)| \leq 2 \frac{\|f\|_{L^p} \|b\|_{L^\infty}}{\|f\|_{L^p}}.
\]

Let \( K \) be the smallest integer such that \( 2^{-\frac{\|f\|_{L^p}}{\|b\|_{L^\infty}}} \|f\|_{L^p} \|b\|_{L^\infty} \leq \lambda/4 \). Since \( W_k^3 f \) is \( k \)-th measurable, that is, constant-valued on the atoms of \( D_k \), we can construct a sequence of decreasing stopping times \( \{t_i\}_{i \geq 0} \) as follows. Let \( t_0 = K \). For \( i \geq 1 \), \( t_i \) is constructed inductively

\[
t_i = \sup\{j : |W_j^3 f - W_{t_{i-1}}^3 f| > \frac{\lambda}{2}\}.
\]

From previous estimates, for all \( x \in \mathbb{R}^n \), \( W_k^3 f(x) \) converges to zero as \( k \to \infty \); By standard arguments—maximal inequality and Banach principle, it is also easy to see \( W_k^3 f \) converges a.e. as \( k \to -\infty \). Hence for a.e. \( x \in \mathbb{R}^n \), \( N_\lambda(\{W_k^3 f\}_{k \in \mathbb{Z}})(x) \) is finite. Fix \( x \in \mathbb{R}^n \), assume \( N_\lambda(\{W_k^3 f\}_{k \in \mathbb{Z}})(x) = N \), which means there exists a sequence of integers \( \{k_i\}_{0 \leq i \leq N} \) such that \( |W_{k_{i+1}}^3 f(x) - W_{k_i}^3 f(x)| > \lambda \). Then \( |W_{k_i}^3 f(x) - W_{k_0}^3 f(x)| > \lambda \) implies either \( |W_{k_i}^3 f(x) - W_{k_0}^3 f(x)| > \frac{\lambda}{2} \) or \( |W_{k_0}^3 f(x) - W_{t_i}^3 f(x)| > \frac{\lambda}{2} \). By the definition of \( t_1 \), we have \( t_1(x) \geq k_1 \). Inductively, we have \( t_i(x) \geq k_i \) for all \( 1 \leq i \leq N \). Thus

\[
\sum_{i \geq 0} |W_{t_{i+1}}^3 f(x) - W_{t_i}^3 f(x)|^2 \geq \sum_{0 \leq i \leq N-1} |W_{t_{i+1}}^3 f(x) - W_{t_i}^3 f(x)|^2 \\
\geq N(\lambda/2)^2 = (\lambda/2)^2 N_\lambda(\{W_k^3 f\}_{k \in \mathbb{Z}})(x),
\]

which yields the desired result.

Now we deal with \( W_k^3 f \). By Lemma 3.1 we can find a sequence of stopping times \( \{t_k\}_{k \geq 0} \) such that

\[
\|\lambda \sqrt{N_\lambda(\{W_k^3 f\}_{k \in \mathbb{Z}})}\|_{L^2} \leq 2 \left\| \left( \sum_{k \geq 0} |E_{t_{k+1}}^b E_{t_k} f - E_{t_k} b E_{t_k} f|^2 \right)^{1/2} \right\|_{L^2} \\
\leq 2 \left\| \left( \sum_{k \geq 0} |(E_{t_{k+1}} b - E_{t_k} b) E_{t_k} f|^2 \right)^{1/2} \right\|_{L^2} \\
+ 2 \left\| \left( \sum_{k \geq 0} |(E_{t_{k+1}} f - E_{t_k} f) E_{t_k} b|^2 \right)^{1/2} \right\|_{L^2}.
\]

By Lemma 2.2 and Carleson’s inequality (see [24]), we get

\[
\sum_{k \geq 0} \int_{\mathbb{R}^n} |E_{t_{k+1}}^b b(x) - E_{t_k} b(x)|^2 |E_{t_k} f(x)|^2 \, dx \leq C \|b\|^2 \|Mf\|^2_{L^2} \\
\leq C \|b\|^2_{L^\infty} \|f\|^2_{L^2},
\]

where \( Mf \) is the Hardy-Littlewood maximal function.
Since $\|E_t b\|_{L^\infty} \leq \|b\|_{L^\infty}$ and $\{E_t f\}_{k \geq 0}$ is still a martingale (see for instance [53]), using Burkholder-Gundy inequality, we get for $1 < p < \infty$

\[
\left( \sum_{k \geq 0} |E_{t_{k+1}} f - E_{t_k} f|^2 |E_{t_{k+1}} b|^2 \right)^{1/2} \leq C \|b\|_{L^\infty} \left( \sum_{k \geq 0} |E_{t_{k+1}} f - E_{t_k} f|^2 \right)^{1/2} \leq C \|b\|_{L^\infty} \|f\|_{L^p}.
\]

Combining the estimates of (3.7) and (3.8) for $p = 2$, we get

\[
\|\lambda \sqrt{N_\lambda(W_k^2 f)}\|_{L^2} \leq C \|b\|_{L^\infty} \|f\|_{L^2}.
\]

Combing the estimates of $W_k^i f$, $i = 1, 2, 3$, we get

\[
\|\lambda \sqrt{N_\lambda(Lf)}\|_{L^2} \leq C \|b\|_{L^\infty} \|f\|_{L^2}.
\]

Next we apply (3.9) to establish weak type $(1,1)$ bounds for $\lambda \sqrt{N_\lambda(Lf)}$. To establish (3.2) we perform the Calderón-Zygmund decomposition of $f$ at height $\alpha$, producing a disjoint family of dyadic cubes $Q$ with total measure $\sum |Q| \leq C \alpha \|f\|_1$ and allowing us to write $f = g + h$ with $\|g\|_{L^\infty} \leq C \alpha$, $\|g\|_{L^1} \leq C \|f\|_{L^1}$ and $h = \sum Q h_Q$, where each $h_Q$ is supported in $Q$ and has mean value zero such that $\sum \|h_Q\|_1 \leq C \|f\|_1$. Since we already know that the $L^2$ norm of $\lambda \sqrt{N_\lambda(Lg)}$ is uniformly controlled by the $L^2$ norm of $g$, matters are reduced in the usual way to estimating $\lambda \sqrt{N_\lambda(Lh)}$ away from $\bigcup \hat{Q}$ where $\hat{Q}$ is a fixed large dilate of $Q$. The fact $\ell^1$ embeds into $\ell^2$ implies

\[
\lambda \sqrt{N_\lambda(Lh)}(x) \leq 2 \sum_{k \in \mathbb{Z}} |\Phi_k b(x) \Gamma_k h(x)|,
\]

we see that

\[
\alpha |\{x \notin \bigcup \hat{Q} : \lambda \sqrt{N_\lambda(Lh)}(x) > \alpha\}| \\
\leq 2 \sum_Q \sum_{k \in \mathbb{Z}} \int_{x \notin \hat{Q}} |\Phi_k b(x) \Gamma_k h_Q(x)| \, dx \\
\leq 2 \sum_Q \sum_{k < k(Q)} \int_{x \notin \hat{Q}} |\Phi_k b(x) \Gamma_k h_Q(x)| \, dx + 2 \sum_Q \sum_{k \geq k(Q)} \int_{x \notin \hat{Q}} |\Phi_k b(x) \Gamma_k h_Q(x)| \, dx.
\]

For $k \leq k(Q)$ (here $2^k(Q)$ is roughly the diameter of $Q$ described in Lemma 3.1 in [37]) we
estimate
\[
\sum_{Q} \sum_{k < k(Q)} \int_{x \notin Q} |\Phi_k b(x) \Gamma_k h_Q(x)| \, dx
\leq C \sum_{Q} \sum_{k < k(Q)} \|\Phi_k b\|_{L^\infty} \int_{Q} |h_Q(y)| \int_{x \notin Q} 2^{-kn(2^{-k}|x - y|)^{(n+1)}} \, dx \, dy
\leq C\|b\|_{L^\infty} \sum_{Q} \sum_{k < k(Q)} \int_{Q} |h_Q(y)| \int_{|x - y| \geq C2^{k(Q)}} 2^{-kn(2^{-k}|x - y|)^{(n+1)}} \, dx \, dy
\leq C\|b\|_{L^\infty} \sum_{Q} \sum_{k < k(Q)} 2^{(k-k(Q))} \|h_Q\|_{L^1} \leq C\|b\|_{L^\infty} \|f\|_{L^1}.
\]

Thus, using the vanishing mean value of \(h_Q\), the right side of the above inequality is dominated by
\[
\sum_{Q} \sum_{k \geq k(Q)} \int_{x \notin Q} |\Phi_k b(x) \Gamma_k h_Q(x)| \, dx
\leq \sum_{Q} \sum_{k \geq k(Q)} \int_{Q} |h_Q(y)| \int_{x \notin Q} |(\mu * \omega)_k(x - y) - (\mu * \omega)_k(x - y_Q)| \, dx \, dy,
\]
where \(y_Q\) denotes the ‘center’ of \(Q\) as described in Lemma 3.1 in [37]. This in turn, using condition (1.18), is
\[
\sum_{Q} \sum_{k \geq k(Q)} \int_{x \notin Q} |\Phi_k b(x) \Gamma_k h_Q(x)| \, dx \leq C\|b\|_{L^\infty} \sum_{Q} \sum_{k \geq k(Q)} 2^{-(k-k(Q))} \|h_Q\|_{L^1} \leq C\|b\|_{L^\infty} \|f\|_{L^1},
\]
establishing the uniform weak-type (1,1) bound for \(\lambda \sqrt{N_\lambda(Lf)}\) and therefore finishing the proof of (3.12). By interpolation between (3.9) and (3.12), imply all the \(L^p\) bounds \(\lambda \sqrt{N_\lambda(Lf)}\) of for \(1 < p \leq 2\). So to prove (3.11), it suffices to prove \(L^p\) bounds of \(\lambda \sqrt{N_\lambda(Lf)}\) for \(2 < p < \infty\). Since we have obtained the \(L^p\) bounds of \(\lambda \sqrt{N_\lambda(W_1^{1}f_k)}\) for \(1 < p < \infty\) in (3.3) and the \(L^p\) bounds of \(I_2\) for \(1 < p < \infty\) in (3.8), we need only to prove for \(2 < p < \infty\)
\[
(3.10) \quad \|\left(\sum_{k \in \mathbb{Z}} |(\Phi_k b - E_k b)E_k f|^2\right)^{1/2}\|_{L^p} \leq C\|b\|_{L^\infty} \|f\|_{L^p}.
\]
and
\[
(3.11) \quad \|\left(\sum_{k \geq 0} |(E_{t_{k+1}} b - E_{t_k} b)E_{t_{k+1}} f|^2\right)^{1/2}\|_{L^p} \leq C\|b\|_{L^\infty} \|f\|_{L^p}.
\]
We first prove (3.10). For $2 < p < \infty$, by H"{o}lder's inequality, we have

$$\left\|\left(\sum_{k \in \mathbb{Z}} |[\Phi_k b - \mathbb{E}_k b] \mathbb{E}_k f|^2\right)^{1/2}\right\|_{L^p} \leq \sup_{\|h_k\|_{L^{p'}}(\mathbb{Z})} \left|\int_{\mathbb{R}^n} \sum_{k \in \mathbb{Z}} ([\Phi_k b(x) - \mathbb{E}_k b(x)] \mathbb{E}_k f(x)) h_k(x) dx\right|$$

$$= \sup_{\|h_k\|_{L^{p'}}(\mathbb{Z})} \left|\int_{\mathbb{R}^n} \sum_{k \in \mathbb{Z}} (\Phi_k (\mathbb{E}_k f \cdot h_k)(y) - \mathbb{E}_k (\mathbb{E}_k f \cdot h_k)(y)) b(y) dy\right|$$

$$\leq \sup_{\|h_k\|_{L^{p'}}(\mathbb{Z})} \left\|\sum_{k \in \mathbb{Z}} \Phi_k (\mathbb{E}_k f \cdot h_k) - \mathbb{E}_k (\mathbb{E}_k f \cdot h_k)\right\|_{L^1} \|b\|_{L^\infty}.$$ 

It suffices to show that

$$\left\|\sum_{k \in \mathbb{Z}} \Phi_k (\mathbb{E}_k f \cdot h_k) - \mathbb{E}_k (\mathbb{E}_k f \cdot h_k)\right\|_{L^1} \leq C \|f\|_{L^p} \|\{h_k\}\|_{L^{p'}(\mathbb{Z})}, \quad 1 < p' \leq 2.$$ 

Clearly, using $\left\|\left(\sum_{k \in \mathbb{Z}} |([\Phi_k b - \mathbb{E}_k b] \mathbb{E}_k f)^2\right)^{1/2}\right\|_{L^2} \leq C \|b\|_{L^\infty} \|f\|_{L^2}$ (see (3.5)) by duality,

$$\left\|\sum_{k \in \mathbb{Z}} \Phi_k (\mathbb{E}_k f \cdot h_k) - \mathbb{E}_k (\mathbb{E}_k f \cdot h_k)\right\|_{L^1} \leq C \|f\|_{L^2} \{h_k\}_{L^2(\mathbb{Z})}.$$ 

Applying $\left|\{x \in \mathbb{R}^n : \left|\sum_{k \in \mathbb{Z}} \Phi_k (g_k(x) - \mathbb{E}_k g_k(x))\right| > \alpha\right| \leq \frac{C}{\alpha} \|\{g_k\}\|_{L^1(\mathbb{Z})}$, which was established in [8] and $\|\mathbb{E}_k f(x)\| \leq \|f\|_{L^\infty}$ for any fixed $x \in \mathbb{R}^n$, we get

$$\left|\{x \in \mathbb{R}^n : \left|\sum_{k \in \mathbb{Z}} \Phi_k (\mathbb{E}_k f \cdot h_k)(x) - \mathbb{E}_k (\mathbb{E}_k f \cdot h_k)(x)\right| > \alpha\right| \leq \frac{C}{\alpha} \|\mathbb{E}_k f \cdot h_k\|_{L^1(\mathbb{Z})}.$$ 

(3.15)

where $\alpha > 0$ and $C$ is independent of $\alpha$, $f$ and $\{h_k\}$. Then by interpolation between (3.14) and (3.15), we get (3.13).

Next we prove (3.11). Similar to the proof of (3.12), we get for $2 < p < \infty$,

$$\left\|\left(\sum_{k \geq 0} \|\mathbb{E}_{t_k+1} b - \mathbb{E}_{t_k} b\|_2 \|\mathbb{E}_{t_k+1} f\|_2^2\right)^{1/2}\right\|_{L^p} \leq \sup_{\|h_k\|_{L^p(\mathbb{Z})}} \left\|\sum_{k \geq 0} \mathbb{E}_{t_k+1} (\mathbb{E}_{t_k+1} f \cdot h_k) - \mathbb{E}_{t_k} (\mathbb{E}_{t_k+1} f \cdot h_k)\right\|_{L^1} \|b\|_{L^\infty}.$$ 

It suffices to show that

$$\left\|\sum_{k \geq 0} \mathbb{E}_{t_k+1} (\mathbb{E}_{t_k+1} f \cdot h_k) - \mathbb{E}_{t_k} (\mathbb{E}_{t_k+1} f \cdot h_k)\right\|_{L^1} \leq C \|f\|_{L^p} \|\{h_k\}\|_{L^p(\mathbb{Z})}, \quad 1 < p' \leq 2.$$ 

By $\left\|\left(\sum_{k \geq 0} \|\mathbb{E}_{t_k+1} b - \mathbb{E}_{t_k} b\|_2 \|\mathbb{E}_{t_k+1} f\|_2^2\right)^{1/2}\right\|_{L^2} \leq C \|b\|_{L^\infty} \|f\|_{L^2}$ (see (3.7)) by duality, we get

$$\left\|\sum_{k \geq 0} \mathbb{E}_{t_k+1} (\mathbb{E}_{t_k+1} f \cdot h_k) - \mathbb{E}_{t_k} (\mathbb{E}_{t_k+1} f \cdot h_k)\right\|_{L^1} \leq C \|f\|_{L^2} \|\{h_k\}\|_{L^2(\mathbb{Z})}. $$

(3.17)
if we can prove that for \( \{ \tilde{h}_k \} \in L^1(\ell^2)(\mathbb{R}^n) \),

\[
(3.18) \quad \{ x \in \mathbb{R}^n : \frac{1}{k} \sum_{k \geq 0} \| E_{x+1} (\tilde{h}_k) - E_{x} (\tilde{h}_k) (x) \| > \alpha \} \leq \frac{C}{\alpha} \| \{ \tilde{h}_k \} \|_{L^1(\ell^2)},
\]

then by \( |E_{x+1} f(x)| \leq \| f \|_{L^\infty} \) for any fixed \( x \in \mathbb{R}^n \), we can get

\[
\{ x \in \mathbb{R}^n : \frac{1}{k} \sum_{k \geq 0} \| E_{x+1} (E_{x+1} f \cdot h_k) (x) - E_{x} (E_{x+1} f \cdot h_k) (x) \| > \alpha \} \leq \frac{C}{\alpha} \| \{ E_{x+1} f \cdot h_k \} \|_{L^1(\ell^2)} 
\]

\[
(3.19) \quad \leq \frac{C}{\alpha} \| f \|_{L^\infty} \| \{ h_k \} \|_{L^1(\ell^2)},
\]

where \( \alpha > 0 \) and \( C \) is independent of \( \alpha, f \) and \( \{ h_k \} \). Thus, by interpolation between (3.17) and (3.19), we get (3.16).

Now we prove (3.18). For \( \alpha > 0 \), we perform Calderón-Zygmund decomposition of \( \| \{ \tilde{h}_k \} \|_{\ell^2} \) at height \( \alpha \), then there exists \( \Lambda \subseteq \mathbb{Z} \times \mathbb{Z}^n \) such that the collection of dyadic cubes \( \{ Q_{\beta}^j \}_{(j, \beta) \in \Lambda} \) are disjoint and the following hold:

(i) \( | \bigcup_{(j, \beta) \in \Lambda} Q_{\beta}^j | \leq \alpha^{-1} \| \{ \tilde{h}_k \} \|_{L^1(\ell^2)} \);

(ii) \( \| \{ \tilde{h}_k (x) \} \|_{\ell^2} \leq \alpha \), if \( x \not\in \bigcup_{(j, \beta) \in \Lambda} Q_{\beta}^j \);

(iii) \( \frac{1}{|Q_{\beta}^j|} \int_{Q_{\beta}^j} \| \{ \tilde{h}_k (x) \} \|_{\ell^2} dx \leq 2^n \alpha \) for each \( (j, \beta) \in \Lambda \).

For \( k \in \mathbb{Z} \), we set

\[
g^{(k)} (x) = \begin{cases} 
\tilde{h}_k (x), & \text{if } x \not\in \bigcup_{(j, \beta) \in \Lambda} Q_{\beta}^j, \\
\frac{1}{|Q_{\beta}^j|} \int_{Q_{\beta}^j} \tilde{h}_k (y) dy, & \text{if } x \in Q_{\beta}^j, (j, \beta) \in \Lambda.
\end{cases}
\]

and

\[
e^{(k)} (x) = \sum_{(j, \beta) \in \Lambda} [\tilde{h}_k (x) - E_{j} \tilde{h}_k (x)] \chi_{Q_{\beta}^j} (x) := \sum_{(j, \beta) \in \Lambda} e^{(k)}_{j, \beta} (x).
\]

First we have \( \| \{ g^{(k)} \} \|_{L^2(\ell^2)}^2 \leq 2\alpha \| \{ \tilde{h}_k \} \|_{L^1(\ell^2)} \). In fact, by (ii), (iii) and Minkowski’s inequality,

\[
\| \{ g^{(k)} \} \|_{L^2(\ell^2)}^2 = \int_{(\bigcup_{(j, \beta) \in \Lambda} Q_{\beta}^j)^c} \| \{ \tilde{h}_k \} \|_{\ell^2} dx + \sum_{(j, \beta) \in \Lambda} \int_{Q_{\beta}^j} \sum_{k \in \mathbb{Z}} \frac{1}{|Q_{\beta}^j|} \int_{Q_{\beta}^j} \tilde{h}_k (y) dy \|_{\ell^2}^2 dx
\]

\[
\leq \alpha \int_{(\bigcup_{(j, \beta) \in \Lambda} Q_{\beta}^j)^c} \| \{ \tilde{h}_k \} \|_{\ell^2} dx + 2^n \alpha \sum_{(j, \beta) \in \Lambda} \int_{Q_{\beta}^j} \| \tilde{h}_k (x) \|_{\ell^2} dx
\]

\[
\leq 2^n \alpha \| \{ \tilde{h}_k \} \|_{L^1(\ell^2)}.
\]

Thus, for above \( \alpha \), by the result in [53] by duality,
\[
\alpha^2 \{ x \in \mathbb{R}^n : \left| \sum_{k \geq 0} [\mathbf{E}_{t_{k+1}} g^{(k)}(x) - \mathbf{E}_{t_k} g^{(k)}(x)] \right| > \alpha \} \\
\leq C \left\| \sum_{k \geq 0} [\mathbf{E}_{t_{k+1}} g^{(k)} - \mathbf{E}_{t_k} g^{(k)}] \right\|_{L^2}^2 \\
\leq C \left\| \{ g^{(k)} \} \right\|_{L^2(\mathcal{E})}^2 \\
\leq C \left\| \{ g^{(k)} \} \right\|_{L^2(\mathcal{E})}^2 \leq C \alpha \left\| \{ \vec{h}_k \} \right\|_{L^1(\mathcal{E})}.
\]

So, we get
\[
\left\{ x \in \mathbb{R}^n : \left| \sum_{k \geq 0} [\mathbf{E}_{t_{k+1}} g^{(k)}(x) - \mathbf{E}_{t_k} g^{(k)}(x)] \right| \leq \frac{C}{\alpha} \left\| \{ \vec{h}_k \} \right\|_{L^1(\mathcal{E})}.
\]

On the other hand, it is easy to see that
\[
\int_{\mathbb{R}^n} e^{(k)}_{j,\beta}(x) dx = 0 \quad \text{for all} \quad k \in \mathbb{Z}, (j, \beta) \in \Lambda.
\]

Let \( \hat{Q}_j^\beta \) be the cube concentric with \( Q_j^\beta \) and with side length 4 times that of \( Q_j^\beta \). It is obvious that
\[
(3.20) \quad \left| \bigcup_{(j, \beta) \in \Lambda} \hat{Q}_j^\beta \right| \leq C \sum_{(j, \beta) \in \Lambda} |Q_j^\beta| \leq \frac{C}{\alpha} \left\| \{ \vec{h}_k \} \right\|_{L^1(\mathcal{E})}.
\]

Note that \( \mathbf{E}_t e^{(k)}_{j,\beta} \) is supported in \( Q_j^\beta \) when \( \ell \leq j \) and \( \mathbf{E}_t e^{(k)}_{j,\beta} \) vanishes everywhere when \( \ell \geq j \).

\[
\alpha \left\{ x \notin \bigcup \hat{Q}_j^\beta : \left| \sum_{k \geq 0} [\mathbf{E}_{t_{k+1}} e^{(k)}(x) - \mathbf{E}_{t_k} e^{(k)}(x)] \right| > \alpha \right\} = 0.
\]

This completes the proof of (3.18).

**Case 2, \( \Phi_k 1 = 0 \).** The argument is very similar to the proof of Case 1 but easier. Since \( \phi \in \mathcal{S}(\mathbb{R}^n) \) and \( \tilde{\phi}(0) = 0 \), then \( \sup_{k \in \mathbb{Z}} \| \Phi_k b \|_{L^\infty} \leq C \| b \|_{L^\infty} \) and \( dv(x, t) = \sum_{k \in \mathbb{Z}} |\Phi_k b(x)|^2 dx \delta_{2k}(t) \) is a Carleson measure on \( \mathbb{R}^{n+1} \) whose norm is controlled by a constant multiple of \( \| b \|_{L^\infty}^2 \) (see [20]). So, we need only a little adjustment in (3.3) with replacing \( \Phi_k b \Gamma_k f = \Phi_k b(\Gamma_k f - E_k f) + (\Phi_k b - E_k b) E_k f + E_k b E_k f \) by \( \Phi_k b \Gamma_k f = \Phi_k b(\Gamma_k f - E_k f) + \Phi_k b E_k f \).

\[\square\]

### 4 Proof of Theorem 1.4

Write
\[
[b, \Phi_k] \nabla = [b, \nabla \Phi_k] - \Phi_k [b, \nabla].
\]

By subadditivity,
\[
\lambda \sqrt{\lambda \Lambda(\mathcal{F}_b \nabla f)} \leq C \lambda \sqrt{N_{\lambda/2}(\{b, \nabla \Phi_k \} f)_{k}} + C \lambda \sqrt{N_{\lambda/2}(\Phi_k [b, \nabla \Phi_k] f)_{k}}.
\]
By Theorem 1.1 in [37], notice that \([b, \nabla] f = -f \nabla b\), we get

\[
\| \lambda \sqrt{N_\lambda(\{ \Phi_k[b, \nabla] f \}_k)} \|_{L^p} \leq C \| [b, \nabla] f \|_{L^p} \leq C \| \nabla b \|_{L^\infty} \| f \|_{L^p}, \quad 1 < p < \infty
\]

and

\[
\alpha \{ x \in \mathbb{R}^n : \lambda \sqrt{N_\lambda(\{ \Phi_k[b, \nabla] f \}_k)}(x) > \alpha \} \leq C \| [b, \nabla] f \|_{L^1} \leq C \| \nabla b \|_{L^\infty} \| f \|_{L^1}
\]

uniformly in \( \lambda > 0 \). So to prove that \( \lambda \sqrt{N_\lambda(\mathcal{K}_b \nabla f)} \) is bounded on \( L^p(\mathbb{R}^n) \) and is of weak \((1, 1)\), it suffices to prove the same properties hold for \( \{ [b, \nabla \Phi_k] \}_k \). Write

\[
[b, \nabla \Phi_k] f = [b, \nabla \Phi_k] f - ([b, \nabla \Phi_k] 1) \Phi_k f + ([b, \nabla \Phi_k] 1) \Phi_k f = P_k f + ([b, \nabla \Phi_k] 1) \Phi_k f.
\]

By subadditivity again,

\[
\lambda \sqrt{N_\lambda(\{ [b, \nabla \Phi_k] f \}_k)} \leq C \lambda \sqrt{N_{\lambda/2}(\{ P_k f \}_k)} + C \lambda \sqrt{N_{\lambda/2}(\{ ([b, \nabla \Phi_k] 1) \Phi_k f \}_k)}.
\]

For \( \{ P_k f \}_{k \in \mathbb{Z}} \), we need the following lemma.

**Lemma 4.1.** ([14, 27, 26]) Denote by \( \Theta_j f(x) := \int_{\mathbb{R}^n} \psi_j(x, y) f(y) \, dy \), where \( \psi_j(x, y) \) satisfies the standard kernel conditions, i.e., for some \( \gamma > 0 \) and \( C > 0 \),

\[
|\psi_j(x, y)| \leq C \frac{2^{j\gamma}}{(2^j + |x - y|)^{n+\gamma}}
\]

and

\[
|\psi_j(x + h, y) - \psi_j(x, y)| + |\psi_j(x, y + h) - \psi_j(x, y)| \leq C \frac{|h|^{\gamma}}{(2^j + |x - y|)^{n+\gamma}}, \quad |h| \leq 2^j,
\]

for all \( x, y \in \mathbb{R}^n \) and \( j \in \mathbb{Z} \). If \( \Theta_j 1 = 0 \), then for \( 1 < p < \infty \),

\[
\left\| \left( \sum_{j \in \mathbb{Z}} |\Theta_j f(x)|^2 \right)^{1/2} \right\|_{L^p} \leq C \| f \|_{L^p}
\]

and

\[
\sup_{\alpha > 0} \alpha \{ x \in \mathbb{R}^n : \left( \sum_{j \in \mathbb{Z}} |\Theta_j f(x)|^2 \right)^{1/2} > \alpha \} \leq C \| f \|_{L^1}.
\]

Denote by \( \tilde{\phi} := \nabla \phi \) and \( \tilde{b} := \nabla b \). Then we can write \( \nabla \Phi_k f = 2^{-k} \tilde{\phi}_k * f \) and \( [b, \nabla \Phi_k] 1 = -\phi_k * \tilde{b} \). Recall that \( P_k f = [b, \nabla \Phi_k] f - ([b, \nabla \Phi_k] 1) \Phi_k f \). Let \( \psi_k(x, y) \) be the kernel of the operator \( P_k \) with

\[
P_k f(x) = \int_{\mathbb{R}^n} \psi_k(x, y) f(y) \, dy.
\]

Then we can write

\[
\psi_k(x, y) = 2^{-k} \tilde{\phi}_k(x - y)(b(x) - b(y)) + (\phi_k * \tilde{b})(x) \phi_k(x - y).
\]
By $|b(x) - b(y)| \leq \|b\|_{L^\infty}|x-y|$ and $|(\phi_k \ast \tilde{b})(x)| \leq \|\tilde{b}\|_{L^\infty}$, we get

$$|\psi_k(x, y)| \leq 2^{-k}|\tilde{b}|_{L^\infty}|\tilde{b}(x-y)|\|\phi_k(x-y)|x-y| + |\tilde{b}|_{L^\infty}|\phi_k(x-y)| \leq C\|\tilde{b}\|_{L^\infty} \frac{2^k}{(2^k + |x-y|)^{n+1}}$$

for all $x, y \in \mathbb{R}^n$ and $k \in \mathbb{Z}$. Also by $\bar{\phi} \in \mathcal{S}(\mathbb{R}^n)$, $|b(x) - b(y)| \leq \|\tilde{b}\|_{L^\infty}|x-y|$ and $|(\phi_k \ast \tilde{b})(x)| \leq \|\tilde{b}\|_{L^\infty}$, we get

$$|\psi_k(x, h) - \psi_k(x, y)| \leq 2^{-k}|\tilde{\phi}_k(x-y-h) - \tilde{\phi}_k(x-y)||b(x) - b(y)|$$

$$+ 2^{-k}|\tilde{\phi}_k(x-y-h)||b(y) - b(y+h)|$$

$$+ |(\phi_k \ast \tilde{b})(x)||\phi_k(x-y-h) - \phi_k(x-y)|$$

$$+ |(\phi_k \ast \tilde{b})(x)| |\phi_k(x-y-h) - \phi_k(x-y)|$$

$$\leq C\|\tilde{b}\|_{L^\infty} \frac{|h|}{(2^k + |x-y|)^{n+1}}, \quad |h| \leq 2^k,$$

for all $x, y \in \mathbb{R}^n$ and $k \in \mathbb{Z}$. This says that the kernel of $P_k$ continues to satisfy (4.1) and (4.2). It is easy to verify that $P_k 1 = 0$ for all $k \in \mathbb{Z}$. Thus by Lemma 4.1, we get for $1 < p < \infty$

$$\left\| \left( \sum_{k \in \mathbb{Z}} |P_k f|^2 \right)^{1/2} \right\|_{L^p} \leq C \|\nabla b\|_{L^\infty} \|f\|_{L^p}$$

and the weak type $(1, 1)$ estimates for $(\sum_{k \in \mathbb{Z}} |P_k f|^2)^{1/2}$. The easy fact $\ell^2$ embeds into $\ell^{2, \infty}$ implies

$$\lambda \sqrt{N}(\{P_k f\}_{k \in \mathbb{Z}})(x) \leq C \left( \sum_{k \in \mathbb{Z}} |P_k f(x)|^2 \right)^{1/2},$$

then gives the desired $L^p$ bounds and weak type $(1, 1)$ bounds for $\lambda \sqrt{N}(\{P_k f\}_{k \in \mathbb{Z}})$. On the other hand, since $[b, \nabla \Phi_k] 1 = -\Phi_k(\nabla b)$, then

$$([b, \nabla \Phi_k] 1) \Phi_k f = -\Phi_k(\nabla b) \Phi_k f.$$

Apply Theorem 1.3, we have

$$\|\lambda \sqrt{N}(\{\Phi_k(\nabla b) \Phi_k f\}_{k})\|_{L^p} \leq C \|\nabla b\|_{L^\infty} \|f\|_{L^p}, \quad 1 < p < \infty$$

(4.3)
and

\[ \alpha \{ x \in \mathbb{R}^n : \lambda \sqrt{\mathcal{N}_\lambda (\{ \Phi_k (\nabla b) \Phi_k \} \Phi_k)}(x) > \alpha \} \leq C \| \nabla b \|_{L^\infty} \| f \|_{L^1} \]

uniformly in \( \lambda > 0 \). Combined these estimates, we get that \( \lambda \sqrt{\mathcal{N}_\lambda (\{ b \nabla \Phi_k \} \Phi_k)} \) is bounded on \( L^p(\mathbb{R}^n) \) and is of weak type \((1, 1)\) if \( b \in \text{Lip}(\mathbb{R}^n) \).

Now we turn to prove that \( \lambda \sqrt{\mathcal{N}_\lambda (\nabla \mathcal{F} b)} \) is bounded on \( L^p(\mathbb{R}^n) \) for \( 1 < p < \infty \) and is of weak type \((1, 1)\) if \( b \in \text{Lip}(\mathbb{R}^n) \). Write

\[ \nabla [b, \Phi_k] f = [b, \nabla \Phi_k] f - [b, \Phi_k] f = [b, \nabla \Phi_k] f + (\nabla b) \Phi_k f. \]

So, we need only to prove

\[ \| \lambda \sqrt{\mathcal{N}_\lambda (\{ \nabla b \Phi_k \} \Phi_k)} \|_{L^p} \leq C \| \nabla b \|_{L^\infty} \| f \|_{L^p}, \quad 1 < p < \infty \]

and

\[ \alpha \{ x \in \mathbb{R}^n : \lambda \sqrt{\mathcal{N}_\lambda (\{ \nabla b \Phi_k \} \Phi_k)}(x) > \alpha \} \leq C \| \nabla b \|_{L^\infty} \| f \|_{L^1} \]

uniformly in \( \lambda > 0 \). Note that \( \nabla b \in L^\infty(\mathbb{R}^n) \), therefore \((4.5)\) and \((4.6)\) can be obtained by the very same argument in \[37]. Therefore we finish the proof of Theorem \[1.4\].

5 Some more lemmas for Theorem \[1.1\]

In this section, we present three more lemmas, which will play a key role in proving Theorem \[1.1\].

**Lemma 5.1.** Let \( \varphi \in \mathcal{S}(\mathbb{R}^n) \) be a radial function such that \( \text{supp} \varphi \subset \{ 1/2 \leq |\xi| \leq 2 \} \) and \( \hat{\Delta_j} f (\xi) = \varphi(2^{-j} \xi) \hat{f}(\xi) \) for \( j \in \mathbb{Z} \). If \( b \in \text{Lip}(\mathbb{R}^n) \), then for \( 1 < p < \infty \) and \( f \in L^p(\mathbb{R}^n) \), we have

\[ \left\| \left( \sum_{l \in \mathbb{Z}} |2^l [b, \Delta_l] f|^2 \right)^{1/2} \right\|_{L^p} \leq C_{n, p} \| \nabla b \|_{L^\infty} \| f \|_{L^p}. \]

**Proof.** Let \( \hat{\Psi} = \varphi \) and \( \Psi_{2^{-j}}(x) = 2^{jn} \Psi(2^j x) \), then \( \Delta_j f = \Psi_{2^{-j}} * f \). Let \( k_j(x, y) = 2^j (b(x) - b(y)) \Psi_{2^{-j}}(x - y) \).

Define the operator \( \mathbb{T} \) by

\[ \mathbb{T} f(x) = \int_{\mathbb{R}^n} \mathbb{K}(x, y) f(y) dy, \]

where \( \mathbb{K} : (x, y) \to \{ k_j(x, y) \}_{j \in \mathbb{Z}} \) with \( \| \mathbb{K}(x, y) \|_{\mathbb{R}^n \times \mathbb{R}^n \to \ell^2} := \left( \sum_{j \in \mathbb{Z}} |k_j(x, y)|^2 \right)^{1/2} \). Lemma 2.3 in \[12\] says that

\[ \| \mathbb{T} f \|_{L^2(\ell^2)} \leq C \| \nabla b \|_{L^\infty} \| f \|_{L^2}. \]

On the other hand, for \( b \in \text{Lip}(\mathbb{R}^n) \), it is easy to verify that for \( 2|h| \leq |x - y| \),

\[ \max \left\{ \left( \sum_{j \in \mathbb{Z}} |k_j(x, y+h) - k_j(x, y)|^2 \right)^{1/2}, \left( \sum_{j \in \mathbb{Z}} |k_j(x+h, y) - k_j(x, y)|^2 \right)^{1/2} \right\} \leq C \| b \|_{\text{Lip}} \frac{|h|}{|x - y|^{n+1}}. \]

Then by the result in \[21, 26, 27\], we get the desired result.
Lemma 5.2. Let $\Omega \in L^1(S^{n-1})$ and satisfy the mean value zero property. For $k \in \mathbb{Z}$, set $\nu_k(x) = \frac{\Omega(x)}{|x|^{n-1}} \chi\{2^k \leq |x| < 2^{k+1}\}(x)$ and $T_k f = \nu_k * f$. Then we have for $1 < p < \infty$,

$$\left\| \left( \sum_{k \in \mathbb{Z}} |T_k f_k|^2 \right)^{1/2} \right\|_{L^p} \leq C_{n, p} \|\Omega\|_{L^1(S^{n-1})} \left\| \left( \sum_{k \in \mathbb{Z}} |\nabla f_k|^2 \right)^{1/2} \right\|_{L^p}. $$

**Proof.** By the mean value zero property of $\Omega$, we have for $t \in \mathbb{R}_+$

$$\left| \int_{S^{n-1}} \Omega(y') f(x - ty')d\sigma(y') \right| = \left| \int_{S^{n-1}} \Omega(y') \left( f(x - ty') - f(x) \right)d\sigma(y') \right| 
\leq \sum_{|\beta| = 1} \int_0^1 \int_{S^{n-1}} |\Omega(y')| |D^\beta f(x + sty')|td\sigma(y')ds. $$

Then, for $\{f_k\}_{k \in \mathbb{Z}}$, by Lemma 2.3 in [11], we have for $1 < p < \infty$

$$\left\| \left( \sum_{k \in \mathbb{Z}} |T_k f_k|^2 \right)^{1/2} \right\|_{L^p} = \left\| \left( \sum_{k \in \mathbb{Z}} \int_{2^{k-1}}^{2^k} \int_{S^{n-1}} |\Omega(y') f_k(-ty')d\sigma(y') \frac{dt}{t^2} \right)^{1/2} \right\|_{L^p} 
\leq C \sum_{|\beta| = 1} \int_0^1 \left\| \left( \sum_{j \in \mathbb{Z}} \int_{2^{k-1}}^{2^k} \int_{S^{n-1}} |\Omega(y')| |D^\beta f_k(x + sty')|d\sigma(y')\frac{dt}{t^2} \right)^{1/2} \right\|_{L^p} ds 
\leq C \sum_{|\beta| = 1} \int_0^1 \left\| \left( \sum_{j \in \mathbb{Z}} \int_{s2^{k-1}}^{s2^k} \frac{|\Omega(y')|}{|y'|^{n+1}} |D^\beta f_k(x + sy)|dy \right)^{1/2} \right\|_{L^p} ds 
= C \sum_{|\beta| = 1} \left\| \left( \sum_{k \in \mathbb{Z}} \int_{s2^{k-1}}^{s2^k} \frac{|\Omega(y')|}{|y'|^{n+1}} |D^\beta f_k(x + y)|dy \right)^{1/2} \right\|_{L^p} ds 
\leq C \sum_{|\beta| = 1} \left\| M_\Omega(D^\beta f_k)(2^{k/2})^2 \right\|_{L^p} 
\leq C \|\Omega\|_{L^1(S^{n-1})} \left\| \left( \sum_{k \in \mathbb{Z}} |\nabla f_k|^2 \right)^{1/2} \right\|_{L^p},$$

where

$$M_\Omega f(x) = \sup_{r > 0} 1 \int_{|x - y| < r} |\Omega(x - y)||f(y)|dy. $$

\[\square\]

Lemma 5.3. Let $\phi \in \mathcal{S}(\mathbb{R}^n)$ and $\Phi_k f(x) = \phi_k * f(x)$, where $\phi_k(x) = 2^{-kn}\phi(2^{-k}x)$. Then for $1 < p < \infty$,

$$\left\| \left( \sum_{k \in \mathbb{Z}} |\nabla [b, \Phi_k] f_k|^2 \right)^{1/2} \right\|_{L^p} \leq C \|\nabla b\|_{L^\infty} \left\| \left( \sum_{k \in \mathbb{Z}} |f_k|^2 \right)^{1/2} \right\|_{L^p}. $$

Write \( \nabla [b, \Phi_k] f = [b, \nabla \Phi_k] f - [b, \nabla] \Phi_k f. \) Then by \([b, \nabla] f = -(\nabla b) f\) and \(\{\Phi_k\}\) is bounded on \(L^p(\ell^2(\mathbb{R}^n))\), we get

\[
\left\| \left( \sum_{k \in \mathbb{Z}} |\nabla [b, \Phi_k] f_k|^2 \right)^{1/2} \right\|_{L^p} \leq C \left\| \left( \sum_{k \in \mathbb{Z}} |[b, \nabla \Phi_k] f_k|^2 \right)^{1/2} \right\|_{L^p} + C \left\| \left( \sum_{k \in \mathbb{Z}} |[b, \nabla] \Phi_k f_k|^2 \right)^{1/2} \right\|_{L^p}.
\]

Note that \(\nabla \phi_k(x) = 2^{-k} (\nabla \phi)_k(x)\) and denote by \(\nabla \phi = \tilde{\phi}\), we get

\[
[b, \nabla \Phi_k] f_k(x) \leq 2^{-k(n+1)} \int_{\mathbb{R}^n} |\tilde{\phi}(2^{-k}(x-y))| |b(x) - b(y)| |f_k(y)| dy.
\]

Therefore by the \(L^p(\ell^2(\mathbb{R}^n))\)-boundedness of \(M\) (see [24]), we get the desired result.

**Lemma 5.4.** Let \(M_{s,\delta,j} \in C_0^\infty(\mathbb{R}^n)(0 < \delta < \infty)\) for any fixed \(s,j \in \mathbb{Z}\), and \(T_{s,\delta,j}\) be the multiplier operator defined by \(T_{s,\delta,j} f(\xi) = M_{s,\delta,j}(\xi)^{1/2} f(\xi)\). Let \(b \in \text{Lip}(\mathbb{R}^n)\) and \([b, T_{s,\delta,j}]\) be the commutator of \(T_{s,\delta,j}\), which is defined by

\[
[b, T_{s,\delta,j}] f(x) = b(x)T_{s,\delta,j} f(x) - T_{s,\delta,j} (bf)(x).
\]

If for some positive constant \(\beta\) and any fixed multi-index \(\alpha\) with \(|\alpha| = 2\),

\[
\|M_{s,\delta,j}\|_{L^\infty} \leq C 2^{-j} \min\{2^{-(1+\beta)s}, 2^s\} \min\{\delta^2, \delta^{-\beta}\}, \quad \|\partial^a M_{s,\delta,j}\|_{L^\infty} \leq C 2^{-j} 2^s,
\]

then there exist some constants \(0 < \lambda, \gamma < 1\) such that

\[
\|[b, T_{s,\delta,j}] f\|_{L^2} \leq C 2^{-j} \min\{2^{-\gamma s}, 2^s\} \min\{\delta^{2\lambda}, \delta^{-\beta\lambda}\} \|b\|_{Lip} \|f\|_{L^2},
\]

where \(C\) is independent of \(s, \delta\) and \(j\).

**Proof.** Taking a \(C_0^\infty(\mathbb{R}^n)\) radial function \(\varphi\) with supp \(\varphi \subset \{1/2 \leq |x| \leq 2\}\) and \(\sum_{l \in \mathbb{Z}} \varphi(2^{-l} x) = 1\) for any \(|x| > 0\). Denote \(\varphi_0(x) = \sum_{l=-\infty}^0 \varphi(2^{-l} x)\) and \(\varphi_l(x) = \varphi(2^{-l} x)\), for positive integer \(l\). Let \(K_{s,\delta,j}(x) = M_{s,\delta,j}^{-1}(x)\), the inverse Fourier transform of \(M_{s,\delta,j}\). Splitting \(K_{s,\delta,j}\) into

\[
K_{s,\delta,j}(x) = K_{s,\delta,j}(x) \varphi_0(x) + \sum_{l=1}^\infty K_{s,\delta,j}(x) \varphi_l(x) =: \sum_{l=0}^\infty K_{s,\delta,j}^l(x).
\]

Write

\[
\tilde{K}_{s,\delta,j}^l(x) = \int_{\mathbb{R}^n} M_{s,\delta,j}(x-y) \tilde{\varphi}_l(y) dy.
\]
Since \( \varphi \) is null in a neighborhood of the origin and a Schwartz function, we have

\[
\int_{\mathbb{R}^n} \hat{\varphi}(\eta) \eta^\theta \, d\eta = 0
\]

for any multi-index \( \vartheta \). Then expanding \( M_{s,\delta,j}(x) \) into a Taylor series around \( x \) and (5.1) gives that

\[
\| K^{l}_{s,\delta,j} \|_{L^\infty} \leq \sum_{|\alpha|=2} \| \partial^\alpha M_{s,\delta,j} \|_{L^\infty} \int_{\mathbb{R}^n} |y|^2 |\hat{\varphi}(y)| \, dy
\]
\[
\leq \sum_{|\alpha|=2} \| \partial^\alpha M_{s,\delta,j} \|_{L^\infty} \int_{\mathbb{R}^n} |2^{-l}y|^2 |\hat{\varphi}(y)| \, dy
\]
\[
\leq C2^{-l}2^{-2l}2^{s} \int_{\mathbb{R}^n} |y|^2 |\hat{\varphi}(y)| \, dy
\]
\[
\leq C2^{-l}2^{-l}2^{s}.
\]

On the other hand, by the Young inequality,

\[
\| \hat{K}^{l}_{s,\delta,j} \|_{L^\infty} = \| \hat{K}_{s,\delta,j} \ast \hat{\varphi} \|_{L^\infty}
\]
\[
\leq \| \hat{K}_{s,\delta,j} \|_{L^\infty} \| \hat{\varphi} \|_{L^1}
\]
\[
\leq C2^{-l} \min\{2^{-(1+\beta)s}, 2^s\} \min\{\delta^2, \delta^{-\beta}\}.
\]

Therefore, interpolating between (5.2) and (5.3), for each \( 1/2 < \theta < \frac{1+\beta}{2+\beta} \),

\[
\| K^{l}_{s,\delta,j} \|_{L^\infty} \leq C2^{-2l}2^{-\theta} \min\{\delta^{2(1-\theta)}, \delta^{-(1-\theta)\beta}\} \min\{2^{\theta s-(1+\beta)(1-\theta)s}, 2^s\}.
\]

Denote by \( \gamma := \theta - (1+\beta)(1-\theta) < 0 \) and \( \lambda := 1 - \theta > 0 \), we get

\[
\| K^{l}_{s,\delta,j} \|_{L^\infty} \leq C2^{-2l}2^{-\gamma} \min\{\delta^{2\gamma}, \delta^{-\lambda \beta}\} \min\{2^{\gamma s}, 2^s\}.
\]

Now we turn our attention to \([b, T^{l}_{s,\delta,j}]\) the commutator of the operator \( T^{l}_{s,\delta,j} \). Decompose \( \mathbb{R}^n \) into a grid of non-overlapping cubes with side length \( 2^l \). That is, \( \mathbb{R}^n = \bigcup_{d=\infty}^{\infty} Q_d \). Set \( f_d = f \chi_{Q_d} \), then

\[
f(x) = \sum_{d=-\infty}^{\infty} f_d(x), \quad a.e. \ x \in \mathbb{R}^n.
\]

It is obvious that \( \text{supp} \ ([b, T^{l}_{s,\delta,j}]f_d) \subset 2^nQ_d \) and that the supports of \( \{[b, T^{l}_{s,\delta,j}]f_d\}_{d=-\infty}^{\infty} \) have bounded overlaps. So we have the following almost orthogonality property

\[
\|[b, T^{l}_{s,\delta,j}]f_d\|_{L^2}^2 \leq C \sum_{d=-\infty}^{\infty} \|[b, T^{l}_{s,\delta,j}]f_d\|_{L^2}^2.
\]
Thus, we may assume that \( \text{supp } f \subset Q \) for some cube with side length \( 2^j \). Choose \( \psi \in C_0^\infty(\mathbb{R}^n) \) with \( 0 \leq \psi \leq 1 \), \( \text{supp } \psi \subset 100nQ \) and \( \psi = 1 \), when \( x \in 30nQ \). Set \( \bar{Q} = 200nQ \) and \( \bar{b} = (\bar{b}(x) - b\bar{Q})\psi(x) \), we can get
\[
\| [b, T_{s, \delta, j}] f \|_{L^2} \leq \sum_{l \geq 0} \| [b, T^l_{s, \delta, j}] f \|_{L^2} \leq \sum_{l \geq 0} \| bT^l_{s, \delta, j} f \|_{L^2} + \sum_{l \geq 0} \| T^l_{s, \delta, j}(\bar{b} f) \|_{L^2}.
\]
By (5.4) with \( \theta > 1/2 \) and \( \| b \|_{L^\infty} \leq 2^l \| b \|_{Lip} \), we have
\[
\sum_{l \geq 0} \| bT^l_{s, \delta, j} f \|_{L^2} \leq \sum_{l \geq 0} \| b \|_{L^\infty} \| T^l_{s, \delta, j} f \|_{L^2}
\leq C \sum_{l \geq 0} 2^{(1-2\theta)l} 2^{-j} \| b \|_{Lip} \min\{\delta^{2\lambda}, \delta^{-\lambda} \} \min\{2^{\gamma s}, 2^s\} \| f \|_{L^2}
\leq C 2^{-j} \| b \|_{Lip} \min\{\delta^{2\lambda}, \delta^{-\lambda} \} \min\{2^{\gamma s}, 2^s\} \| f \|_{L^2}.
\]
Similarly, we can get
\[
\sum_{l \geq 0} \| T^l_{s, \delta, j}(\bar{b} f) \|_{L^2} \leq C 2^{-j} \| b \|_{Lip} \min\{\delta^{2\lambda}, \delta^{-\lambda} \} \min\{2^{\gamma s}, 2^s\} \| f \|_{L^2}.
\]
Thus
\[
\| [b, T_{s, \delta, j}] f \|_{L^2} \leq C 2^{-j} \| b \|_{Lip} \min\{\delta^{2\lambda}, \delta^{-\lambda} \} \min\{2^{\gamma s}, 2^s\} \| f \|_{L^2},
\]
where \( C \) is independent of \( \delta, s \) and \( j \). \( \square \)

6 Proof of Theorem [1.1] (I)

As we have stated in the introduction, to prove Theorem [1.1] it suffices to show (1.15) and (1.16). In this section, we give the proof of (1.15). For \( j \in \mathbb{Z} \), let \( \nu_j(x) = \frac{\Omega(x)}{|x|^{n+1}} \chi_{\{2^j \leq |x| < 2^{j+1}\}}(x) \), then
\[
\nu_j * f(x) = \int_{2^j \leq |y| < 2^{j+1}} \frac{\Omega(y)}{|y|^{n+1}} f(x-y)dy.
\]
Denote by
\[
T^1 f(x) = \text{p.v.} \int_{\mathbb{R}^n} \frac{\Omega(y)}{|y|^{n+1}} f(x-y)dy
\]
and for \( k \in \mathbb{Z} \)
\[
T^1_{2k} f(x) = \int_{|x-y| > 2^k} \frac{\Omega(y)}{|y|^{n+1}} f(x-y)dy.
\]
Let \( \phi \in \mathcal{S}(\mathbb{R}^n) \) be a radial function such that \( \hat{\phi}(\xi) = 1 \) for \( |\xi| \leq 2 \) and \( \hat{\phi}(\xi) = 0 \) for \( |\xi| > 4 \). We have the following decomposition
\[
T^1_{2k} f = \phi_k * T^1 f + \sum_{s \geq 0} (\delta_0 - \phi_k) * \nu_{k+s} * f - \phi_k * \sum_{s < 0} \nu_{k+s} * f,
\]
where \( \phi_k \) satisfies \( \hat{\phi}_k(\xi) = \phi(2^k\xi) \), \( \delta_0 \) is the Dirac measure at 0. Then

\[
C_{2k}f = [b, \phi_k * T^1]f + [b, \sum_{s \geq 0} (\delta_0 - \phi_k) * \nu_{k+s}]f - [b, \phi_k * \sum_{s < 0} \nu_{k+s}]f
\]

\[
: = C_k^1f + C_k^2f - C_k^3f.
\]

Let \( \mathcal{C}^i_f \) denote the family \( \{C^i_kf\}_{k \in \mathbb{Z}} \) for \( i = 1, 2, 3 \). Obviously, to show (1.15) it suffices to prove the following inequalities:

\[
(6.1) \quad \| \lambda \sqrt{N_\lambda(\mathcal{C}^i_f)} \|_{L^p} \leq C\| \nabla b \|_{L^\infty} \| f \|_{L^p}, \quad 1 < p < \infty, \quad i = 1, 2, 3,
\]

uniformly in \( \lambda > 0 \).

**Estimation of (6.1)** for \( i = 1 \). For \( k \in \mathbb{Z} \), denote by \( \Phi_k f(x) = \phi_k * f(x) \) and write

\[
C_k^1f = [b, \Phi_k]T^1f + \Phi_k[b, T^1]f.
\]

Combining Theorem 1.1 in [37] and the \( L^p(1 < p < \infty) \)-boundedness of \( [b, T^1] \) with bounds \( C\| \Omega \|_{L(\log^+ L)(\mathbb{S}^{n-1})} \| \nabla b \|_{L^\infty} \) (see [3]), we can get the following estimate easily for \( 1 < p < \infty \)

\[
(6.2) \quad \| \lambda \sqrt{N_\lambda(\{\Phi_k[b, T^1]f\}^k)} \|_{L^p} \leq C\| [b, T^1] f \|_{L^p} \leq C\| \Omega \|_{L(\log^+ L)(\mathbb{S}^{n-1})} \| \nabla b \|_{L^\infty} \| f \|_{L^p}.
\]

Using \( \sum_{j=1}^n R_j^2 = -I \) (identity operator) and \( R_j = \partial_j I_1, j = 1, \ldots, n \), to get

\[
[b, \Phi_k]T^1f = -[b, \Phi_k] \sum_{j=1}^n R_j^2 T^1f = -\sum_{j=1}^n [b, \Phi_k] \partial_j (R_j I_1 T^1 f),
\]

where \( R_j \) is the \( j \)-th Riesz transform and \( I_1 \) is the Riesz potential operator of order 1. Then by Theorem 1.3 and \( \| R_j f \|_{L^p} \leq C\| f \|_{L^p} \) for \( 1 < p < \infty, j = 1, \ldots, n \), we get

\[
(6.3) \quad \| \lambda \sqrt{N_\lambda(\{[b, \Phi_k]T^1 f\}^k)} \|_{L^p} \leq \sum_{j=1}^n \| \lambda \sqrt{N_\lambda(\{[b, \Phi_k] \partial_j (R_j I_1 T^1 f)\}^k)} \|_{L^p}
\]

\[
\leq C\| \nabla b \|_{L^\infty} \sum_{j=1}^n \| R_j T^1 I_1 f \|_{L^p}
\]

\[
\leq C\| \Omega \|_{L(\log^+ L)(\mathbb{S}^{n-1})} \| \nabla b \|_{L^\infty} \| (-\Delta)^{1/2} I_1 f \|_{L^p}
\]

\[
= C\| \Omega \|_{L(\log^+ L)(\mathbb{S}^{n-1})} \| \nabla b \|_{L^\infty} \| f \|_{L^p},
\]

where in the above inequality, we have used that \( \| T^1 f \|_{L^p} \leq C\| \Omega \|_{L(\log^+ L)(\mathbb{S}^{n-1})} \| (-\Delta)^{1/2} f \|_{L^p} \)

for \( 1 < p < \infty \) (see [10]) and \( (-\Delta)^{1/2} I_1 = I \). Together (6.2) with (6.3), we get for \( 1 < p < \infty \),

\[
\| \lambda \sqrt{N_\lambda(\mathcal{C}^1_f)} \|_{L^p} \leq \| \lambda \sqrt{N_\lambda(\{\Phi_k[b, T^1]f\}^k)} \|_{L^p} + \| \lambda \sqrt{N_\lambda(\{[b, \Phi_k]T^1 f\}^k)} \|_{L^p}
\]

\[
\leq C\| \Omega \|_{L(\log^+ L)(\mathbb{S}^{n-1})} \| \nabla b \|_{L^\infty} \| f \|_{L^p}
\]
uniformly in \( \lambda > 0 \).

**Estimation of (6.1) for** \( i = 2 \). Let \( E_0 = \{ x' \in S^{n-1} : |\Omega(x')| < 2 \} \) and \( E_m = \{ x' \in S^{n-1} : 2^m \leq |\Omega(x')| < 2^{m+1} \} \) for positive integer \( m \). For \( m \geq 0 \), let

\[
\Omega_m(y') = \Omega(y')\chi_{E_m}(y') - \frac{1}{|S^{n-1}|} \int_{E_m} \Omega(x') \, d\sigma(x').
\]

Since \( \Omega \) satisfies (1.6), then

\[
\int_{S^{n-1}} \Omega_m(y') \, d\sigma(y') = 0 \quad \text{for} \quad m \geq 0
\]

and \( \Omega(y') = \sum_{m \geq 0} \Omega_m(y') \). Set \( \nu_{j,m}(x) = \frac{\Omega_m(x)}{|x|^{n+1}} \chi_{\{ |x| \leq 2^j \}}(x) \), then \( \nu_j(x) = \sum_{m \geq 0} \nu_{j,m}(x) \). Thus, by the fact \( \ell^2 \) embeds into \( \ell^{2,\infty} \) and the Minkowski inequality, we get

\[
(6.4) \quad \lambda \sqrt{N_\lambda(\mathcal{G}^2 f)(x)} \leq \sum_{s \geq 0} \left( \sum_{k \in \mathbb{Z}} \left| [b, (\delta_0 - \phi_k) * \nu_{k+s,m}] f(x) \right|^2 \right)^{1/2}
\]

\[
\leq \sum_{s \geq 0} \sum_{m \geq 0} \left( \sum_{k \in \mathbb{Z}} \left| [b, (\delta_0 - \phi_k) * \nu_{k+s,m}] f(x) \right|^2 \right)^{1/2}.
\]

Denote by \( F_{s,k,m} f(x) := (\delta_0 - \phi_k) * \nu_{k+s,m} * f(x) \). Let \( \varphi \in C_0^\infty(\mathbb{R}^n) \) be a radial function such that \( 0 \leq \varphi \leq 1 \), \( \text{supp} \varphi \subset \{ 1/2 \leq |\xi| \leq 2 \} \) and \( \sum_{l \in \mathbb{Z}} \varphi^2(2^{-l} \xi) = 1 \) for \( |\xi| \neq 0 \). Define the multiplier \( \Delta_l \) by \( \Delta_l \varphi(\xi) = \varphi(2^{-l} \xi) \varphi(\xi) \). It is clear that

\[
[b, (\delta_0 - \phi_k) * \nu_{k+s,m}] f(x) = [b, F_{s,k,m}] f(x) = \sum_{l \in \mathbb{Z}} [b, F_{s,k,m} \Delta_l] f(x).
\]

Then by the Minkowski inequality, we get for \( 1 < p < \infty \),

\[
(6.5) \quad \left\| \left( \sum_{k \in \mathbb{Z}} \left| [b, (\delta_0 - \phi_k) * \nu_{k+s,m}] f \right|^2 \right)^{1/2} \right\|_{L^p} \leq \sum_{l \in \mathbb{Z}} \left\| \left( \sum_{k \in \mathbb{Z}} \left| [b, F_{s,k,m} \Delta_l] f \right|^2 \right)^{1/2} \right\|_{L^p}.
\]

If we can prove the following two inequalities: for some \( 0 < \beta < 1 \) and \( 0 < \theta < 1 \),

\[
(6.6) \quad \| G_{s,m; b}^d f \|_{L^2} \leq C 2^{-\beta s} 2^{-\theta |l|} \| \Omega_m \|_{L^\infty(S^{n-1})} \| \nabla b \|_{L^\infty} \| f \|_{L^2}
\]

and

\[
(6.7) \quad \| G_{s,m; b}^d f \|_{L^p} \leq C \| \Omega_m \|_{L^1(S^{n-1})} \| \nabla b \|_{L^\infty} \| f \|_{L^p} \quad \text{for} \quad 1 < p < \infty,
\]

we may get (6.1) for \( i = 2 \). In fact, interpolating between (6.6) and (6.7), we get for \( 0 < \theta_0, \beta_0 < 1 \),

\[
(6.8) \quad \| G_{s,m; b}^d f \|_{L^p} \leq C 2^{-\beta_0 s} 2^{-\theta_0 |l|} \| \nabla b \|_{L^\infty} \| \Omega_m \|_{L^\infty(S^{n-1})} \| f \|_{L^p}, \quad 1 < p < \infty.
\]
Taking a large positive integer $N$, such that $N > \max\{2\theta_0^{-1}, 2\beta_0^{-1}\}$, we have for $1 < p < \infty$,

$$
\|\lambda \sqrt{N\lambda} (\mathcal{E}^2 f)\|_{L^p} \leq \sum_{m \geq 0} \sum_{0 \leq s < Nm} \|G^l_{s,m,b}f\|_{L^p} + \sum_{m \geq 0} \sum_{s \geq Nm} \sum_{l \geq Nm} \|G^l_{s,m,b}f\|_{L^p}
+ \sum_{m \geq 0} \sum_{s > Nm} \sum_{l \geq 0} \|G^l_{s,m,b}f\|_{L^p}.
$$

By (6.7), we get for $1 < p < \infty$,

$$
\sum_{m \geq 0} \sum_{0 \leq s < Nm} \sum_{l \leq Nm} \|G^l_{s,m,b}f\|_{L^p} \leq C \|
abla b\|_{L^\infty} \sum_{m \geq 0} \sum_{0 \leq s < Nm} \sum_{l \leq Nm} 2^m \sigma(E_m) \|f\|_{L^p}
\leq C \|
abla b\|_{L^\infty} \sum_{m \geq 0} m^2 2^m \sigma(E_m) \|f\|_{L^p}
\leq C \|\Omega\|_{L(\log^+L)^2(S^{n-1})} \|
abla b\|_{L^\infty} \|f\|_{L^p}.
$$

Applying (6.8), we get for $1 < p < \infty$,

$$
\sum_{m \geq 0} \sum_{0 \leq s < Nm} \sum_{l \geq Nm} \|G^l_{s,m,b}f\|_{L^p} \leq C \|
abla b\|_{L^\infty} \sum_{m \geq 0} \sum_{0 \leq s < Nm} \sum_{l \geq Nm} 2^{-\beta_0 s} \sum_{m \geq 0} 2^{-\theta_0 l} \|f\|_{L^p}
\leq C \|
abla b\|_{L^\infty} \sum_{m \geq 0} m 2^{(1-\beta_0 N)m} \|f\|_{L^p}
\leq C \|
abla b\|_{L^\infty} \|f\|_{L^p}.
$$

Applying (6.8) again, we get for $1 < p < \infty$,

$$
\sum_{m \geq 0} \sum_{s > Nm} \sum_{l \geq 0} \|G^l_{s,m,b}f\|_{L^p} \leq C \|
abla b\|_{L^\infty} \sum_{m \geq 0} \sum_{s > Nm} \sum_{l \leq Nm} 2^m \sum_{m \geq 0} 2^{-\beta_0 s} \left( \sum_{l \leq Nm} + \sum_{l \geq Nm} 2^{-\theta_0 l} \right) \|f\|_{L^p}
\leq C \|
abla b\|_{L^\infty} \sum_{m \geq 0} (m 2^{(1-\beta_0 N)m} + 2^{(1-\beta_0 N-\theta_0 N)m}) \|f\|_{L^p}
\leq C \|
abla b\|_{L^\infty} \|f\|_{L^p}.
$$

Combining above three estimates, we have for $1 < p < \infty$

$$
\|\lambda \sqrt{N\lambda} (\mathcal{E}^2 f)\|_{L^p} \leq C(1 + \|\Omega\|_{L(\log^+L)^2(S^{n-1})}) \|
abla b\|_{L^\infty} \|f\|_{L^p}.
$$

We therefore finish the estimate of (6.1) for $i = 2$.

Now we are going to give the proof of (6.6) and (6.7). We first prove a rapid decay estimate of $\|G^l_{s,m,b}f\|_{L^2}$ for $l \in \mathbb{Z}$ and $s \in \mathbb{N}$. Set $F^l_{s,k,m}f(x) := F_{s,k,m}\Delta_{l-k}f(x)$. Write

$$
[b, F_{s,k,m}\Delta_{l-k}]f = F^l_{s,k,m} [b, \Delta_{l-k}]f + [b, F^l_{s,k,m}] \Delta_{l-k} f.
$$
Therefore
\[ \|G_{s,m}f\|_{L^2} \leq \left( \sum_{k \in \mathbb{Z}} |F_{s,k,m}^l[b, \Delta_{l-k}]f|^2 \right)^{1/2} + \left( \sum_{k \in \mathbb{Z}} |[b, F_{s,k,m}^l\Delta_{l-k}]f|^2 \right)^{1/2} \]
\[ := I + II. \]

Set
\[ M_{s,k,m}(\xi) = (1 - \hat{\phi}_k(\xi))\hat{\nu}_{k+s,m}(\xi), \quad M_{s,k,m}^l(\xi) = M_{s,k,m}(\xi)\varphi(2^{k-l}\xi). \]

Then write \( F_{s,k,m} \) and \( F_{s,k,m}^l \), respectively by
\[ F_{s,k,m}^l(\xi) = M_{s,k,m}(\xi)\hat{f}(\xi) \]
and \( F_{s,k,m}(\xi) = M_{s,k,m}(\xi)\hat{f}(\xi). \)

Since \( \text{supp} (1 - \hat{\phi}_k)\hat{\nu}_{k+s,m} \subset \{ \xi : |2^k\xi| > 1/2 \} \), by a well-known Fourier transform estimate of Duoandikoetxea and Rubio de Francia (See [20], p.551-552), it is easy to show that there exists some \( \nu \in (0, 1) \) such that
\[ (6.9) \quad |M_{s,k,m}(\xi)| \leq C2^{-k-2-(\nu+1)s} \min\{|2^k\xi|^2, |2^k\xi|^{-\nu}\} \|\Omega_m\|_{L^\infty(S^{n-1})}, \quad s \geq 0. \]

From this and the Plancherel theorem imply the following estimate
\[ (6.10) \quad \|F_{s,k,m}^l\|_{L^2} \leq C2^{-k-2-(\nu+1)s} \min\{2^{2l}, 2^{-\nu l}\} \|\Omega_m\|_{L^\infty(S^{n-1})} \|f\|_{L^2}, \quad \text{for } l \in \mathbb{Z}. \]

Then apply \( (6.10) \) and Lemma 5.1, we have
\[ I \leq C2^{-(\nu+1)s} \min\{2^{-(\nu+1)s}, 2^l\} \|\Omega_m\|_{L^\infty(S^{n-1})} \left( \left( \sum_{k \in \mathbb{Z}} |2^{l-k}[b, \Delta_{l-k}]f|^2 \right)^{1/2} \right) \]
\[ \leq C2^{-(\nu+1)s} \min\{2^{-(\nu+1)s}, 2^l\} \|\Omega_m\|_{L^\infty(S^{n-1})} \|\nabla b\|_{L^\infty} \|f\|_{L^2}. \]

To proceed with the estimate of \( II \), we define multiplier \( \tilde{F}_{s,k,m}^l \), by \( \tilde{F}_{s,k,m}^l(\xi) = M_{s,k,m}^l(2^{-k}\xi)\hat{f}(\xi). \)
As a result of \( (6.9) \), we have the following estimate
\[ (6.11) \quad |M_{s,k,m}^l(2^{-k}\xi)| \leq C2^{-k-2-(\nu+1)s} \min\{2^{2l}, 2^{-\nu l}\} \|\Omega_m\|_{L^\infty(S^{n-1})}. \]

On the other hand, by the trivial computation, we have for any fixed multi-index \( \eta \),
\[ (6.12) \quad |\partial^\eta \hat{\nu}_{k+s,m}(\xi)| \leq C2^{(k+|\eta|-1)} \|\Omega_m\|_{L^1(S^{n-1})}. \]

Then we have for any fixed multi-index \( \alpha \) with \(|\alpha| = 2\),
\[ (6.13) \quad |\partial^\alpha(M_{s,k,m}^l(2^{-k}\xi))| = |\partial^\alpha(\hat{\nu}_{k+s,m}(2^{-k}\xi)(1 - \phi(\xi))\varphi(2^{-l}\xi))| \]
\[ = |\sum_\eta C^\alpha_m \cdots C^\alpha_\eta \partial^\eta(\hat{\nu}_{k+s,m}(2^{-k}\xi))\partial^{\alpha-\eta}[(1 - \phi(\xi))\varphi(2^{-l}\xi)]| \]
\[ \leq C2^{-k-2s} \|\Omega_m\|_{L^1(S^{n-1})}, \]
where the sum is taken over all multi-indices \( \eta \) with \( 0 \leq \eta_j \leq \alpha_j \) for \( 1 \leq j \leq n \). Via Lemma 5.4 to (6.11) and (6.13) with \( \delta = 2^l \) and \( j = k \) says that there exist constants \( \vartheta \in (0,1) \) and \( \gamma \in (0,1) \) such that

\[
\| [b, F^l_{s,k,m}] f \|_{L^2} \leq C 2^{-k} 2^{-\vartheta - \gamma s} \| \Omega_m \|_{L^\infty(S^{n-1})} \| b \|_{Lip} \| f \|_{L^2}, \quad \text{for } l \in \mathbb{Z} \text{ and } s \geq 0.
\]

Further, by \( \| b(2^k \cdot) \|_{Lip} = 2^k \| b \|_{Lip} \), we have

\[
(6.14) \quad \| [b, F^l_{s,k,m}] f \|_{L^2} \leq C 2^{-\vartheta - \gamma s} 2^{-|l|} \| \Omega_m \|_{L^\infty(S^{n-1})} \| b \|_{Lip} \| f \|_{L^2}, \quad \text{for } l \in \mathbb{Z} \text{ and } s \geq 0.
\]

Then by (6.14) and Littlewood-Paley theory, we get

\[
II \leq C 2^{-\vartheta - \gamma s} 2^{-|l|} \| \Omega_m \|_{L^\infty(S^{n-1})} \| \nabla b \|_{L^\infty} \left( \sum_{k \in \mathbb{Z}} \| \Delta_{l-k} f \|_{L^2}^2 \right)^{1/2} \| f \|_{L^2}.
\]

Combining the estimates of \( I \) with \( II \), we establish the proof of (6.6).

Now we give the \( L^p(1 < p < \infty) \) estimate of \( G^l_{s,m,b} f \) for \( l \in \mathbb{Z} \) and \( s \geq 0 \). We write

\[
[b, F_{s,k,m} \Delta^2_{l-k}] f(x) = [b, F_{s,k,m}] \Delta^2_{l-k} f + F_{s,k,m}[b, \Delta^2_{l-k}] f.
\]

By the Minkowski inequality, we get for \( 1 < p < \infty \)

\[
\| G^l_{s,m,b} f \|_{L^p} \leq \left( \sum_{k \in \mathbb{Z}} \| [b, F_{s,k,m}] \Delta^2_{l-k} f \|_{L^2}^2 \right)^{1/2}_{L^p} + \left( \sum_{k \in \mathbb{Z}} \| F_{s,k,m}[b, \Delta^2_{l-k}] f \|_{L^2}^2 \right)^{1/2}_{L^p}.
\]

We estimate each term separately. Firstly, we estimate \( \| (\sum_{k \in \mathbb{Z}} \| [b, F_{s,k,m}] \Delta^2_{l-k} f \|_{L^2}^2)^{1/2} \|_{L^p} \) for \( 1 < p < \infty \). Recall that \( F_{s,k,m} f = (\delta_0 - \varphi_k) \ast \nu_{k+s,m} \ast f \) and \( \Phi_k f = \varphi_k \ast f \). Denote by \( T_{j,m} f = \nu_{j,m} \ast f \) for \( j \in \mathbb{Z} \) and \( m \geq 0 \). Write

\[
[b, F_{s,k,m}] f = [b, T_{k+s,m}] \Phi_k f + T_{k+s,m}[b, \Phi_k] f - [b, T_{k+s,m}] f.
\]

It is well known that for any \( f \in L^p(\mathbb{R}^n) \),

\[
[|b, T_{k+s,m}] f(x) | \leq C \| b \|_{Lip} M_{\Omega_m} f(x).
\]

From this and \( M_{\Omega_m} \) is bounded on \( L^p(\ell^2)(\mathbb{R}^n) \) for \( 1 < p < \infty \) with bounds \( \| \Omega_m \|_{L^1(S^{n-1})} \) (see Lemma 2.3 in [11]), we get for \( 1 < p < \infty \),

\[
(6.15) \quad \left( \sum_{k \in \mathbb{Z}} \| [b, T_{k+s,m}] f_k \|_{L^p}^2 \right)^{1/2} \leq C \| \Omega_m \|_{L^1(S^{n-1})} \| \nabla b \|_{L^\infty} \left( \sum_{k \in \mathbb{Z}} \| f_k \|_{L^p}^2 \right)^{1/2}.
\]
By Lemma 5.2 and Lemma 5.3, we get for $1 < p < \infty$,
\begin{equation}
(6.16) \quad \left\| \left( \sum_{k \in \mathbb{Z}} |T_{k+s,m}[b, \Phi_k]f_k|^2 \right)^{1/2} \right\|_{L^p} \leq C\|\Omega_m\|_{L^1(S^{n-1})} \left\| \left( \sum_{k \in \mathbb{Z}} |\nabla [b, \Phi_k]f_k|^2 \right) \right\|_{L^p}^{1/2}
\end{equation}
\begin{equation*}
\leq C\|\nabla b\|_{L^\infty} \|\Omega_m\|_{L^1(S^{n-1})} \left\| \left( \sum_{k \in \mathbb{Z}} |f_k|^2 \right) \right\|_{L^p}^{1/2}.
\end{equation*}
Together (6.15)-(6.16) with the $L^p(\ell^2)(\mathbb{R}^n)$ $(1 < p < \infty)$ boundedness of \{\Phi_k\} and Littlewood-Paley theory, we get for $1 < p < \infty$,
\begin{equation}
(6.17) \quad \left\| \left( \sum_{k \in \mathbb{Z}} |[b, F_{s,k,m}] \Delta^2_{l-k}f|^2 \right)^{1/2} \right\|_{L^p} \leq C\|\Omega_m\|_{L^1(S^{n-1})} \|\nabla b\|_{L^\infty} \left\| \left( \sum_{k \in \mathbb{Z}} |\Delta^2_{l-k}f|^2 \right) \right\|_{L^p}^{1/2}
\end{equation}
\begin{equation*}
\leq C\|\Omega_m\|_{L^1(S^{n-1})} \|\nabla b\|_{L^\infty} \|f\|_{L^p}.
\end{equation*}
Secondly, we estimate $\|\left( \sum_{k \in \mathbb{Z}} |F_{s,k,m}[b, \Delta^2_{l-k}f]^2 \right)^{1/2}\|_{L^p}$ for $1 < p < \infty$. If the following inequality holds
\begin{equation}
(6.18) \quad \left\| \left( \sum_{k \in \mathbb{Z}} |\nabla [b, \Delta^2_{l-k}]f|^2 \right)^{1/2} \right\|_{L^p} \leq C\|\nabla b\|_{L^\infty} \|f\|_{L^p}, \quad 1 < p < \infty.
\end{equation}
Then apply $L^p(\ell^2)(\mathbb{R}^n)$ $(1 < p < \infty)$ boundedness of \{\Phi_k\} and Lemma 5.2, we can get
\begin{equation}
(6.19) \quad \left\| \left( \sum_{k \in \mathbb{Z}} |F_{s,k,m}[b, \Delta^2_{l-k}]f|^2 \right)^{1/2} \right\|_{L^p} \leq C\|\Omega_m\|_{L^1(S^{n-1})} \left\| \left( \sum_{k \in \mathbb{Z}} |\nabla [b, \Delta^2_{l-k}]f|^2 \right) \right\|_{L^p}^{1/2}
\end{equation}
\begin{equation*}
\leq C\|\Omega_m\|_{L^1(S^{n-1})} \|\nabla b\|_{L^\infty} \|f\|_{L^p}.
\end{equation*}
Combining the estimates of (6.17) and (6.19), we get for $1 < p < \infty$,
\begin{equation}
\|G_{s,m;b}^l f\|_{L^p} \leq C\|\Omega_m\|_{L^1(S^{n-1})} \|\nabla b\|_{L^\infty} \|f\|_{L^p}.
\end{equation}
This gives (6.7). Now we prove (6.18). Since $\nabla \Delta^2_{l-k}f(x) = 2^{l-k}\Delta_{l-k}(x)$ for a.e. $x \in \mathbb{R}^n$, where $\Delta_j$ is the Littlewood-paley operator given on the transform by multiplication with the function $(2^{-j} \xi)\varphi^2(2^{-j} \xi)$ for $j \in \mathbb{Z}$. Then by $\nabla [b, \Delta^2_{l-k}]f = [b, \nabla \Delta^2_{l-k}]f - [b, \nabla] \Delta^2_{l-k}f$ and the Minkowski inequality, we get for $1 < p < \infty$,
\begin{equation}
\left\| \left( \sum_{k \in \mathbb{Z}} |\nabla [b, \Delta^2_{l-k}]f|^2 \right)^{1/2} \right\|_{L^p} \leq \left\| \left( \sum_{k \in \mathbb{Z}} |[b, 2^{-l-k} \Delta_{l-k}]f|^2 \right)^{1/2} \right\|_{L^p} + \left\| \left( \sum_{k \in \mathbb{Z}} |[b, \nabla] \Delta^2_{l-k}f|^2 \right)^{1/2} \right\|_{L^p}.
\end{equation}
By Lemma 5.1 $[b, \nabla]f = -\nabla b)f$ and Littlewood-Paley theory, we get for $1 < p < \infty$,
\begin{equation}
\left\| \left( \sum_{k \in \mathbb{Z}} |\nabla [b, \Delta^2_{l-k}]f|^2 \right)^{1/2} \right\|_{L^p} \leq C\|\nabla b\|_{L^\infty} \|f\|_{L^p}.
\end{equation}
This gives \((6.18)\).

**Estimation of \((6.1)\) for \(i = 3\).** We have the following pointwise estimate

\[(6.20)\]

\[
\lambda \sqrt{\mathcal{N}_\lambda(\mathcal{E}^3 f)}(x) \leq \sum_{s < 0} \left( \sum_{k \in \mathbb{Z}} |[b, \phi_k * \nu_{k+s}] f(x)|^2 \right)^{1/2}. 
\]

The proofs are essentially similar to the proof of \((6.1)\) for \(i = 2\). More precisely, we need to give the estimates on the left hand side of \((6.6)-(6.7)\) with replacing \((\delta_0 - \phi_k) * \nu_{k+s}\) by \(\phi_k * \nu_{k+s}\). Since \(\text{supp} \hat{\phi_k} \nu_{k+s} \subset \{ \xi : |2^k \xi| < 1 \}\) and \(\Omega\) satisfies \((1.2)\), then it is easy to see that

\[
|\hat{\phi_k} \nu_{k+s}(\xi)| \leq C 2^{-k} 2^s \|\Omega\|_{L^1(\mathbb{S}^{n-1})} \min\{|2^k \xi|^2, |2^k \xi|^{-1}\}
\]

and for any fixed multi-index \(\eta\) with \(|\eta| \leq 2\),

\[
|\partial^\eta \hat{\nu}_{k+s}(\xi)| \leq C 2^{(k+s)(|\eta|-1)} \|\Omega\|_{L^1(\mathbb{S}^{n-1})} 2^{k+s} |\eta|^2 \leq C 2^{k(|\eta|-1)} 2^s \|\Omega\|_{L^1(\mathbb{S}^{n-1})} |2^k \xi|^{2-|\eta|}.
\]

Set

\[
R_{s,k}(\xi) = \hat{\phi_k}(\xi) \hat{\nu}_{k+s}(\xi), \quad R_{s,k}^i(\xi) = \hat{R}_{s,k}(\xi) \varphi(2^{k-i} \xi).
\]

Using the two above inequalities, we have the following estimate

\[(6.21)\]

\[
|R_{s,k}^i(2^{-k} \xi)| \leq C 2^{-k} 2^s \min\{2^{|i|}, 2^{-l}\} \|\Omega\|_{L^1(\mathbb{S}^{n-1})},
\]

and for any fixed multi-index \(\alpha\) with \(|\alpha| = 2\),

\[(6.22)\]

\[
|\partial^\alpha (R_{s,k}^i(2^{-k} \xi))| \leq C 2^{-k} 2^s \|\Omega\|_{L^1(\mathbb{S}^{n-1})}.
\]

Then apply Lemma \(5.4\) and the same arguments of the proofs of \((6.1)\) for \(i = 2\), then the right hand side of \((6.6)\) is controlled by \(C 2^{s-3} \|\Omega\|_{L^1(\mathbb{S}^{n-1})} \|\nabla b\|_{L^\infty} \|f\|_{L^2}\) for some \(\theta > 0\). It is also easy to get the same estimates in the right hand side of \((6.7)\) by using \((6.15)\), \((6.16)\) and Lemma \(5.2\). Then we get for \(1 < p < \infty\)

\[
\|\lambda \sqrt{\mathcal{N}_\lambda(\mathcal{E}^3 f)}\|_{L^p} \leq C \|\Omega\|_{L^1(\mathbb{S}^{n-1})} \|\nabla b\|_{L^\infty} \|f\|_{L^p}.
\]

\(\square\)

7. **Proof of Theorem \(1.1\) (II)**

We first prove \((1.16)\) by a key lemma, its proof will be postponed until the end of the section.

**Lemma 7.1.** For \(t \in [1,2)\) and \(j \in \mathbb{Z}\), we define \(\nu_{j,t}\) as

\[
\nu_{j,t}(x) = \frac{\Omega(x')}{|x|^{n+1}} \chi_{[2^{j+1} t \leq |x| \leq 2^{j+1}]}(x).
\]

Denote \(T_{j,t} f(x) = \nu_{j,t} * f(x)\). For \(k \in \mathbb{Z}\), denote by

\[
S_{2,k}(\mathcal{E} f)(x) = \left( \sum_{j \in \mathbb{Z}} \|\{[b, T_{j,t} \Delta_{k-j}^2 f(x)]_{t \in [1,2]}\} 2^{j} \right)^{1/2}.
\]
For \( b \in \text{Lip}(\mathbb{R}^n) \), then the following conclusions hold:

(i) There exists a constant \( \theta_1 \in (0, 1) \) such that

\[
\|S_{2,k}(\mathcal{E}f)\|_{L^2} \leq C 2^{-\theta_1 k} \|\Omega\|_{L^\infty(\mathbb{S}^{n-1})} \|
abla b\|_{L^\infty} \|f\|_{L^2};
\]

(ii) If \( \Omega(x') \) satisfies \( \text{(7.2)} \), there exists a constant \( \theta_2 \in (0, 1) \) such that

\[
\|S_{2,k}(\mathcal{E}f)\|_{L^2} \leq C 2^{\theta_2 k} \|\Omega\|_{L^1(\mathbb{S}^{n-1})} \|
abla b\|_{L^\infty} \|f\|_{L^2};
\]

(iii) For \( 1 < p < \infty \),

\[
\|S_{2,k}(\mathcal{E}f)\|_{L^p} \leq C \|\Omega\|_{L^1(\mathbb{S}^{n-1})} \|
abla b\|_{L^\infty} \|f\|_{L^p}.
\]

The constants \( C \)'s in \( \text{(7.1)} \), \( \text{(7.2)} \) and \( \text{(7.3)} \) are independent of \( k \).

Lemma \( \text{(7.1)} \) will be proved at the end of this section. Let us now finish the proof of \( \text{(1.16)} \) by using Lemma \( \text{(7.1)} \). For \( t \in [1, 2) \), let \( \nu_{j,t} \) and \( T_{j,t} \) be defined in the same way as in Lemma \( \text{(7.1)} \). Observe that \( V_{2,j}(\mathcal{E}f)(x) \) is just the strong 2-variation function of the family \( \{[b, T_{j,t}] f(x)\}_{t \in [1, 2]} \), hence using \( \sum_{k \in \mathbb{Z}} \Delta^2_{k,j} = \mathcal{I} \), we get

\[
S_2(\mathcal{E}f)(x) = \left( \sum_{j \in \mathbb{Z}} \left| V_{2,j}(\mathcal{E}f)(x) \right|^2 \right)^{\frac{1}{2}} = \left( \sum_{j \in \mathbb{Z}} \left\| \{[b, T_{j,t}] f(x)\}_{t \in [1, 2]} \right\|_{L^2}^2 \right)^{\frac{1}{2}}
\leq \sum_{k \in \mathbb{Z}} \left( \sum_{j \in \mathbb{Z}} \left\| \{[b, T_{j,t} \Delta^2_{k,j}] f(x)\}_{t \in [1, 2]} \right\|_{L^2}^2 \right)^{\frac{1}{2}}
:= \sum_{k \in \mathbb{Z}} S_{2,k}(\mathcal{E}f)(x)
= \sum_{k < 0} S_{2,k}(\mathcal{E}f)(x) + \sum_{k \geq 0} S_{2,k}(\mathcal{E}f)(x).
\]

Interpolating between \( \text{(7.2)} \) and \( \text{(7.3)} \), we can get for some constant \( \theta_3 \in (0, 1) \) and \( 1 < p < \infty \),

\[
\|S_{2,k}(\mathcal{E}f)\|_{L^p} \leq C \|\Omega\|_{L^1(\mathbb{S}^{n-1})} 2^{\theta_3 k} \|
abla b\|_{L^\infty} \|f\|_{L^p}, \text{ for } k < 0.
\]

Then for \( 1 < p < \infty \),

\[
\sum_{k < 0} \|S_{2,k}(\mathcal{E}f)\|_{L^p} \leq C \|\Omega\|_{L^1(\mathbb{S}^{n-1})} \|
abla b\|_{L^\infty} \|f\|_{L^p} \sum_{k < 0} 2^{\theta_3 k}
\leq C \|\Omega\|_{L^1(\mathbb{S}^{n-1})} \|
abla b\|_{L^\infty} \|f\|_{L^p}.
\]

Decompose \( \Omega(y') = \sum_{m \geq 0} \Omega_m(y') \) as in Section 3. For \( m \geq 0 \), set

\[
\nu_{j,t,m}(x) = \frac{\Omega_m(x)}{|x|^{n+1}} \chi_{\{2^j \leq |x| < 2^{j+1}\}} (x),
\]
$T_{j,t,m}$ is defined as $T_{j,t}$ by replacing $\nu_{j,t}$ by $\nu_{j,t,m}$.

$$\sum_{k \geq 0} S_{2,k}(\mathcal{E} f)(x) \leq \sum_{m \geq 0} \sum_{k \geq 0} \left( \sum_{j \in \mathbb{Z}} \| [b, T_{j,t,m} \Delta_{k-j}^2] f(x) \|_{L^2} \right)^2 \frac{1}{2}$$

$$:= \sum_{m \geq 0} \sum_{k \geq 0} S_{2,k,m}(\mathcal{E} f)(x).$$

Interpolating between (7.1) and (7.3), we can get for some constant $\theta_4 \in (0, 1/2)$ and $1 < p < \infty$,

$$(7.4) \quad \| S_{2,k,m}(\mathcal{E} f) \|_{L^p} \leq C \| \Omega_m \|_{L^\infty(S^{a-1})} 2^{-\theta_4 k} \| \nabla b \|_{L^\infty} \| f \|_{L^p}, \text{ for } k \geq 0, \ m \geq 0.$$ 

Taking a large positive integer $N$, such that $N > 2 \theta_4^{-1}$. Then, for $1 < p < \infty$,

$$\sum_{k \geq 0} \| S_{2,k}(\mathcal{E} f) \|_{L^p} \leq \sum_{m \geq 0} \sum_{k > Nm} \| S_{2,k,m}(\mathcal{E} f) \|_{L^p} + \sum_{m \geq 0} \sum_{0 \leq k \leq Nm} \| S_{2,k,m}(\mathcal{E} f) \|_{L^p} := J_1 + J_2.$$ 

For $J_1$, using (7.4), we get for $1 < p < \infty$,

$$J_1 \leq C \| \nabla b \|_{L^\infty} \sum_{m \geq 0} 2^m \sum_{k > Nm} 2^{-\theta_4 k} \| f \|_{L^p} \leq C \| \nabla b \|_{L^\infty} \| f \|_{L^p}.$$ 

By (7.3), we get for $1 < p < \infty$,

$$J_2 \leq C \| \nabla b \|_{L^\infty} \sum_{m \geq 0} \sum_{0 \leq k \leq Nm} \| \Omega_m \|_{L^1(S^{a-1})} \| f \|_{L^p} \leq C \| \nabla b \|_{L^\infty} \sum_{m \geq 0} \sum_{0 \leq k \leq Nm} 2^m \sigma(E_m) \| f \|_{L^p} \leq C \| \nabla b \|_{L^\infty} \sum_{m \geq 0} m 2^m \sigma(E_m) \| f \|_{L^p} \leq C \| \Omega \|_{L^{\log^+L}(S^{a-1})} \| \nabla b \|_{L^\infty} \| f \|_{L^p}.$$ 

Combining with the estimates of $J_1$ and $J_2$, we get for $1 < p < \infty$

$$\sum_{k \geq 0} \| S_{2,k}(\mathcal{E} f) \|_{L^p} \leq C (1 + \| \Omega \|_{L^{\log^+L}(S^{a-1})}) \| \nabla b \|_{L^\infty} \| f \|_{L^p}.$$ 

We therefore finish the proof of (1.10).

**Proof of Lemma 7.1.** To deal with (7.1) and (7.2), we borrow the fact $\| a \|_{V^2} \leq \| a \|_{L^2} \| a' \|_{L^2}^{1/2}$, where $a' = \{ \frac{d}{dt} a_t : t \in \mathbb{R} \}$. It is a special case of (39) in [37]. Then,

$$[S_{2,k}(\mathcal{E} f)(x)]^2 \leq \sum_{j \in \mathbb{Z}} \left( \int_1^2 \| [b, T_{j,t} \Delta_{k-j}^2] f(x) \|_{L^2}^2 \right)^{1/2} \left( \int_1^2 \frac{d}{dt} |b, T_{j,t} \Delta_{k-j}^2] f(x)|^2 \right)^{1/2}.$$
By the Cauchy-Schwarz inequality, we have
\[
\|S_{2,k}(\mathcal{Q} f)\|_{L^2}^2 \leq \left\| \left( \int_1^2 \sum_{j \in \mathbb{Z}} |[b, T_{j,t} \Delta_{k-j}^2]f(x)|^2 \frac{dt}{t} \right) \right\|_{L^2} \cdot \left\| \left( \int_1^2 \sum_{j \in \mathbb{Z}} |[b, T_{j,t} \Delta_{k-j}^2]f(x)|^2 \frac{dt}{t} \right) \right\|_{L^2}^{1/2}.
\]
We estimate \(\|I_{1,k} f\|_{L^2}\) and \(\|I_{2,k} f\|_{L^2}\), respectively. To estimate \(\|I_{1,k} f\|_{L^2}\), we need the following estimates
\[
|\hat{\nu}_{j,t}(\xi)| \leq C2^{-j} \|\Omega\|_{L^\infty(S^{n-1})}|2^j \xi|^{-\gamma}, \quad \gamma \in (0, 1)
\]
and for any fixed multi-index \(\eta\) with \(|\eta| \leq 2\),
\[
|\partial^\eta \hat{\nu}_{j,t}(\xi)| \leq C2^{j(|\eta|-1)} \|\Omega\|_{L^1(S^{n-1})}
\]
uniformly in \(t \in [1, 2]\). If \(\Omega\) satisfies (7.5), then
\[
|\hat{\nu}_{j,t}(\xi)| \leq C2^{-j} \|\Omega\|_{L^1(S^{n-1})}|2^j \xi|^2
\]
and for any fixed multi-index \(\eta\) with \(|\eta| \leq 2\),
\[
|\partial^\eta \hat{\nu}_{j,t}(\xi)| \leq C2^{j(|\eta|-1)} \|\Omega\|_{L^1(S^{n-1})}|2^j \xi|^{2-|\eta|}
\]
uniformly in \(t \in [1, 2]\). Set \(M_{j,t}(\xi) = \hat{\nu}_{j,t}(\xi), M_{j,t}^k(\xi) = M_{j,t}(\xi)\varphi(2^{j-k} \xi)\). We define multipliers \(T_{j,t}^k\) and \(\tilde{T}_{j,t}^k\), respectively by
\[
\tilde{T}_{j,t}^k f(\xi) = M_{j,t}^k(\xi) \hat{f}(\xi) \quad \text{and} \quad \tilde{T}_{j,t}^k f(\xi) = M_{j,t}^k(2^{-j} \xi) \hat{f}(\xi).
\]
We use (7.5)–(7.6) to get for \(k \geq 0\),
\[
\|M_{j,t}^k(2^{-j} \cdot)\|_{L^\infty} \leq C2^{-j} 2^{-\gamma k} \|\Omega\|_{L^\infty(S^{n-1})}, \quad |\partial^\alpha [M_{j,t}^k(2^{-j} \cdot)]|_{L^\infty} \leq C2^{-j} \|\Omega\|_{L^1(S^{n-1})},
\]
where \(\alpha\) is a multi-index with \(|\alpha| = 2\). Using (7.7)–(7.8), we get for \(k < 0\),
\[
\|M_{j,t}^k(2^{-j} \cdot)\|_{L^\infty} \leq C2^{-j} 2^{2k} \|\Omega\|_{L^1(S^{n-1})}, \quad |\partial^\alpha [M_{j,t}^k(2^{-j} \cdot)]|_{L^\infty} \leq C2^{-j} \|\Omega\|_{L^1(S^{n-1})},
\]
where \(\alpha\) is a multi-index with \(|\alpha| = 2\). Via Lemma 5.4 to (7.9) and (7.10) with \(\delta = 2^k\), respectively states that for any fixed \(0 < v < 1\),
\[
\|b, \tilde{T}_{j,t}^k f\|_{L^2} \leq C2^{-j} \min\{2^{-\gamma k} \|\Omega\|_{L^\infty(S^{n-1})}, 2^{2k} \|\Omega\|_{L^1(S^{n-1})}\} \|b\|_{lip} \|f\|_{L^2}.
\]
Then by \(\|b(2^\cdot)\|_{lip} = 2^j \|b\|_{lip}\) says that
\[
\|b, T_{j,t}^k f\|_{L^2} \leq C \min\{2^{-\gamma k} \|\Omega\|_{L^\infty(S^{n-1})}, 2^{2k} \|\Omega\|_{L^1(S^{n-1})}\} \|b\|_{lip} \|f\|_{L^2}.
\]
By Plancherel theorem, we also get
\[
\|T_{j,t}^k f\|_{L^2} \leq C2^{-j} \min\{2^{-\gamma k} \|\Omega\|_{L^\infty(S^{n-1})}, 2^{2k} \|\Omega\|_{L^1(S^{n-1})}\} \|f\|_{L^2}.
\]
Write
\[ [b, T_{j,t} \Delta_{k-j}^2]f = [b, T_{j,t}^k \Delta_{k-j}]f = [b, T_{j,t}^k] \Delta_{k-j}f + T_{j,t}^k [b, \Delta_{k-j}]f. \]

Then we get
\[ \|I_{1,k}f\|_{L^2} \leq \left( \int_1^2 \sum_{j \in \mathbb{Z}} \| [b, T_{j,t}^k] \Delta_{k-j}f \|_{L^2}^2 \frac{dt}{t} \right)^{1/2} + \left( \int_1^2 \sum_{j \in \mathbb{Z}} \| T_{j,t}^k [b, \Delta_{k-j}]f \|_{L^2}^2 \frac{dt}{t} \right)^{1/2}. \]

By (7.11) and Littlewood-Paley theory, we get
\begin{align*}
(7.13) \quad & \left( \int_1^2 \sum_{j \in \mathbb{Z}} \| [b, T_{j,t}^k] \Delta_{k-j}f \|_{L^2}^2 \frac{dt}{t} \right)^{1/2} \\
& \leq C \min \{ 2^{-\gamma k} \| \Omega \|_{L^\infty(S^{n-1})}, 2^{2\gamma k} \| \Omega \|_{L^1(S^{n-1})} \} \| \nabla b \|_{L^\infty} \left( \int_1^2 \sum_{j \in \mathbb{Z}} \| \Delta_{k-j}f \|_{L^2}^2 \frac{dt}{t} \right)^{1/2} \\
& \leq C \min \{ 2^{-\gamma k} \| \Omega \|_{L^\infty(S^{n-1})}, 2^{2\gamma k} \| \Omega \|_{L^1(S^{n-1})} \} \| \nabla b \|_{L^\infty} \| f \|_{L^2}.
\end{align*}

By (7.12) and Lemma 5.1 we get
\begin{align*}
(7.14) \quad & \left( \int_1^2 \sum_{j \in \mathbb{Z}} \| T_{j,t}^k [b, \Delta_{k-j}]f \|_{L^2}^2 \frac{dt}{t} \right)^{1/2} \\
& \leq C \min \{ 2^{-(\gamma+1)k} \| \Omega \|_{L^\infty(S^{n-1})}, 2^{k} \| \Omega \|_{L^1(S^{n-1})} \} \left( \int_1^2 \sum_{j \in \mathbb{Z}} \| 2^{-j} [b, \Delta_{k-j}]f \|_{L^2}^2 \frac{dt}{t} \right)^{1/2} \\
& \leq C \min \{ 2^{-(\gamma+1)k} \| \Omega \|_{L^\infty(S^{n-1})}, 2^{k} \| \Omega \|_{L^1(S^{n-1})} \} \| \nabla b \|_{L^\infty} \| f \|_{L^2}.
\end{align*}

Combining the estimates of (7.13) and (7.14), we get
\[ (7.15) \quad \|I_{1,k}f\|_{L^2} \leq C \min \{ 2^{-\gamma k} \| \Omega \|_{L^\infty(S^{n-1})}, 2^{\gamma k} \| \Omega \|_{L^1(S^{n-1})} \} \| \nabla b \|_{L^\infty} \| f \|_{L^2}. \]

Next, we estimate \( \|I_{2,k}f\|_{L^2} \). We write
\[ \frac{d}{dt} [b, T_{j,t} \Delta_{k-j}^2]f = \frac{d}{dt} T_{j,t} [b, \Delta_{k-j}^2]f + [b, \frac{d}{dt} T_{j,t}] \Delta_{k-j}^2 f. \]

Then we get
\begin{align*}
\|I_{2,k}f\|_{L^2} & \leq \left( \int_1^2 \left\| \left( \sum_{j \in \mathbb{Z}} \frac{d}{dt} T_{j,t} [b, \Delta_{k-j}^2]f \right)^2 \right\|_{L^2} \frac{dt}{t} \right)^{1/2} \\
& + \left( \int_1^2 \left\| \left( \sum_{j \in \mathbb{Z}} [b, \frac{d}{dt} T_{j,t}] \Delta_{k-j}^2 f \right)^2 \right\|_{L^2} \frac{dt}{t} \right)^{1/2} \\
& := L_1 + L_2.
\end{align*}
To estimate $L_i$ for $i = 1, 2$, respectively, we need the following elementary fact

$$
(7.16) \quad \frac{d}{dt} T_{j,t} h(x) = \frac{d}{dt} \left[ \int_{2^j t < |y| \leq 2^{j+1}} \frac{\Omega'(y')}{|y|^{n+1}} h(x - y) dy \right]
$$

$$
= \frac{d}{dt} \left[ \int_{S_{n-1}} \Omega(y') \int_{2^j t}^{2^{j+1}} \frac{1}{r^2} h(x - ry') dr d\sigma(y') \right]
$$

$$
= -\frac{1}{2^{j+2}} \int_{S_{n-1}} \Omega(y') h(x - 2^j ty') d\sigma(y')
$$

$$
\leq \sum_{i=1}^{n} \int_{0}^{1} \int_{S_{n-1}} |\Omega(y')| |\partial_i h(x + s2^j ty')| t^{-1} d\sigma(y') ds.
$$

For $t \in [1, 2]$, it is easy to get

$$
(7.17) \quad \| \frac{d}{dt} T_{j,t} h \|_{L^2} \leq C \| \Omega \|_{L^1(S^{n-1})} \| \nabla h \|_{L^2}.
$$

On the other hand, we get

$$
[b, \frac{d}{dt} T_{j,t}] h(x) = -\frac{1}{2^{j+2}} \int_{S_{n-1}} \Omega(y')(b(x) - b(x - 2^j ty')) h(x - 2^j ty') d\sigma(y')
$$

$$
\leq \| \nabla b \|_{L^\infty} \int_{S_{n-1}} |\Omega(y')| |h(x - 2^j ty')| t^{-1} d\sigma(y').
$$

For $t \in [1, 2]$, it is easy to get

$$
(7.18) \quad \|[b, \frac{d}{dt} T_{j,t}] h\|_{L^2} \leq C \| \Omega \|_{L^1(S^{n-1})} \| \nabla b \|_{L^\infty} \| h \|_{L^2}.
$$

We now estimate $L_1$. Indeed, by (7.17) and (6.18), we have

$$
L_1 \leq C \| \Omega \|_{L^1(S^{n-1})} \left( \int_{1}^{2^2} \sum_{j \in \mathbb{Z}} \| \nabla [b, \Delta_{k-j}^2 f] \|_{L^2}^2 \frac{dt}{t} \right)^{1/2}
$$

$$
\leq C \| \Omega \|_{L^1(S^{n-1})} \| \nabla b \|_{L^\infty} \| f \|_{L^2}.
$$

Similarly, by (7.18) and Littlewood-Paley theory, we have

$$
L_2 \leq C \| \Omega \|_{L^1(S^{n-1})} \| \nabla b \|_{L^\infty} \left( \int_{1}^{2^2} \sum_{j \in \mathbb{Z}} \| \Delta_{k-j}^2 f \|_{L^2}^2 \frac{dt}{t} \right)^{1/2}
$$

$$
\leq C \| \Omega \|_{L^1(S^{n-1})} \| \nabla b \|_{L^\infty} \| f \|_{L^2}.
$$

Combining the estimates of $L_1$ and $L_2$, we get

$$
\| I_{2,k} f \|_{L^2} \leq C \| \Omega \|_{L^1(S^{n-1})} \| \nabla b \|_{L^\infty} \| f \|_{L^2}.
$$

Then combined with the estimate of (7.15), we get for $k \in \mathbb{Z}$,

$$
\| S_{2,k}(\mathcal{C}^2 f) \|_{L^2}^2 \leq C \min \{ 2^{-\gamma k} \| \Omega \|_{L^\infty(S^{n-1})}^2, 2^{\nu k} \| \Omega \|_{L^1(S^{n-1})}^2 \} \| \nabla b \|_{L^\infty} \| f \|_{L^2}^2.
$$
This finishes the proof of (7.1) and (7.2).

So to prove Lemma 7.1, it suffices to prove (7.3). Let

\[ B = \left\{ (a_{j,t})_{j \in \mathbb{Z}, t \in [1,2]} : \|a_{j,t}\|_B := \left( \sum_{j \in \mathbb{Z}} a_{j,t}^2 \right)^{1/2} < \infty \right\}. \]

Clearly, \((B, \| \cdot \|_B)\) is a Banach space. Then,

\[ S_{2,k}(\mathcal{E}f)(x) = \left( \sum_{j \in \mathbb{Z}} \sup_{t_1 < \cdots < t_N} \sum_{l=1}^{N-1} \left| [b, T_{j,l} \Delta_{k-l}^2 f(x) - [b, T_{j,l} \Delta_{k-l}^2 f(x)]^2 \right|^{1/2} \right) \]

where

\[ T_{j,l} f(x) = \int_{2^j t_l < |y| \leq 2^{j+1} t_l} f(x - y) \frac{\Omega(y)}{|y|^{n+1}} dy \quad \text{and} \quad [t_l, t_{l+1}] \subset [1, 2). \]

Then by

\[ [b, T_{j,l} \Delta_{k-l}^2 f] = [b, T_{j,l} \Delta_{k-l}^2 f] + T_{j,l} \Delta_{k-l}^2 [b, \Delta_{k-l}^2 f], \]

we get

\[ S_{2,k}(\mathcal{E}f)(x) \leq \left( \sum_{j \in \mathbb{Z}} \sup_{t_1 < \cdots < t_N} \sum_{l=1}^{N-1} \left| [b, T_{j,l} \Delta_{k-l}^2 f(x)]^2 \right|^{1/2} \right) \]

By the mean value zero property of \(\Omega\), we have

\[ T_{j,l} f(x) = \int_{2^j t_l}^{2^{j+1} t_l} \int_{S^{n-1}} f(x - r y') \Omega(y') d\sigma(y') \frac{dr}{r^2} \]

\[ = \int_{2^j t_l}^{2^{j+1} t_l} \int_{S^{n-1}} \Omega(y') \left( f(x - r y') - f(x) \right) d\sigma(y') \frac{dr}{r^2} \]

\[ = \sum_{|\beta|=1} \int_{2^j t_l}^{2^{j+1} t_l} \int_0^1 \int_{S^{n-1}} \Omega(y') D^\beta f(x + s r y') (ry')^\beta d\sigma(y') ds \frac{dr}{r^2}. \]

For \([t_l, t_{l+1}] \subset [1, 2)\), let

\[ \tilde{T}_{j,l} f(x) = \int_{2^j t_l}^{2^{j+1} t_l} \int_0^1 \int_{S^{n-1}} |\Omega(y')| ||\nabla f(x + s r y')|| d\sigma(y') ds \frac{dr}{r} \]

and

\[ T_{j,l} f(x) = \int_{2^j t_l < |y| \leq 2^{j+1} t_l} |f(x - y)| \frac{|\Omega(y')|}{|y|^n} dy. \]
Then,
\[ S_{2,k}(\mathcal{C}f)(x) \leq C\|\nabla b\|_{L^\infty} \left( \sum_{j \in \mathbb{Z}} \sup_{t_1, \ldots, t_N} \sum_{l=1}^{N-1} |T_{j, t_l, t_{l+1}}^* \Delta_{k-j}^2 f(x)|^2 \right)^{1/2} + C \left( \sum_{j \in \mathbb{Z}} \sup_{t_1, t_N} |T_{j, t_l, t_N} \Delta_{k-j}^2 f(x)|^2 \right)^{1/2} \]
\[ = C\|\nabla b\|_{L^\infty} \left( \sum_{j \in \mathbb{Z}} \sup_{t_1, t_N} |T_{j, t_l, t_N} \Delta_{k-j}^2 f(x)|^2 \right)^{1/2} + C \left( \sum_{j \in \mathbb{Z}} \sup_{t_1, t_N} |\tilde{T}_{j, t_l, t_N} \Delta_{k-j}^2 f(x)|^2 \right)^{1/2}. \]

Therefore, we get
\[ S_{2,k}(\mathcal{C}f)(x) \leq C\|\nabla b\|_{L^\infty} \left( \sum_{j \in \mathbb{Z}} |M_\Omega(\Delta_{k-j}^2 f(x)|^2 \right)^{1/2} + C \left( \sum_{j \in \mathbb{Z}} |M_\Omega(\nabla [b, \Delta_{k-j}^2 f(x)|^2 \right)^{1/2}. \]

To proceeding with the estimates, we need the following inequality, for \(1 < p < \infty\)
\[ \left\| \left( \sum_{j \in \mathbb{Z}} |M_\Omega f_j|^2 \right)^{1/2} \right\|_{L^p} \leq C\|\Omega\|_{L^1(\mathbb{S}^{n-1})} \left\| \left( \sum_{j \in \mathbb{Z}} |f_j|^2 \right)^{1/2} \right\|_{L^p}, \]
which was established in [11]. Then by Littlewood-Paley theory and [6,18], we have for \(1 < p < \infty\)
\[
\|S_{2,k}(\mathcal{C}f)\|_{L^p} \leq C\|\nabla b\|_{L^\infty} \|\Omega\|_{L^1(\mathbb{S}^{n-1})} \left\| \left( \sum_{j \in \mathbb{Z}} |\Delta_{k-j}^2 f|^2 \right)^{1/2} \right\|_{L^p} + C\|\Omega\|_{L^1(\mathbb{S}^{n-1})} \left\| \left( \sum_{j \in \mathbb{Z}} |\nabla [b, \Delta_{k-j}^2 f]|^2 \right)^{1/2} \right\|_{L^p} 
\leq C\|\nabla b\|_{L^\infty} \|\Omega\|_{L^1(\mathbb{S}^{n-1})} \|f\|_{L^p}. \]

This gives (7.3). Therefore, we complete the proof Lemma 7.1. \(\square\)

## 8 Proof of Corollary 1.2

For \(\varepsilon > 0\), write
\[ [b, T_\varepsilon] \nabla f(x) = -T_\varepsilon [b, \nabla] f(x) + [b, \nabla T_\varepsilon] f(x). \]

For the first term, since \([b, \nabla] f = -(\nabla b) f\), then by Theorem 1.2 in [18], for \(1 < p < \infty\), we have
\[ \|\Lambda \sqrt{N_\lambda} (\{T_\varepsilon [b, \nabla] f\}_{\varepsilon > 0})\|_{L^p} \leq C\|K\|_{L(\log^+ L)^2(\mathbb{S}^{n-1})}\|f\|_{L^p} \]
\[ \leq C\|K\|_{L(\log^+ L)^2(\mathbb{S}^{n-1})} \|\nabla b\|_{L^\infty} \|f\|_{L^p}. \]
For the second term, it is easy to verify that $\nabla K(x)$ is homogeneous of degree $-n - 1$ and

$$(x_k \nabla K(x))^\wedge(\xi) = i\xi_k \widehat{\nabla K}(\xi) = i \frac{\partial}{\partial \xi_k} (i \xi_1 \widehat{K}(\xi), \ldots, i \xi_n \widehat{K}(\xi)).$$

Moreover, if $j = k$, then $\frac{\partial}{\partial \xi_k} (\xi_j \widehat{K})(\xi) = \widehat{K}(\xi) + \xi_j \frac{\partial}{\partial \xi_k} \widehat{K}(\xi)$. If $j \neq k$, then $\frac{\partial}{\partial \xi_k} (\xi_j \widehat{K})(\xi) = \xi_j \frac{\partial}{\partial \xi_k} \widehat{K}(\xi)$.

So we get for $k = 1, \ldots, n$, $(x_k \nabla K(x))^\wedge(0) = 0$. Additionally, $\widehat{\nabla K}(\xi) = i \xi \widehat{K}(\xi)$, then $\widehat{\nabla K}(0) = 0$.

This says that

$$\int_{S^{n-1}} (x'_k)^N \nabla K(x') \, d\sigma(x') = 0$$

for any $k = 1, \ldots, n$ and $N = 0, 1$. Since $|\nabla K(x')| \in L(\log^+ L)^2(S^{n-1})$, then by Theorem 1.1, the family of the operators

$$\{[b, \nabla T_\varepsilon]f(x)\}_{\varepsilon > 0} = \left\{ \int_{|x-y| > \varepsilon} \nabla K(x-y)(b(x) - b(y)) f(y) \, dy \right\}_{\varepsilon > 0}$$

satisfies

$$(8.2) \quad \|\lambda \sqrt{N\lambda}([b, \nabla T_\varepsilon]f)\|_{L^p} \leq C\|\nabla K\|_{L(\log^+ L)^2(S^{n-1})} \|\nabla b\|_{L^\infty} \|f\|_{L^p}, \ 1 < p < \infty.$$ 

Combining the estimates of (S.1) and (S.2), we get

$$\|\lambda \sqrt{N\lambda}([b, \nabla T_\varepsilon]f)\|_{L^p} \leq C(\|K\|_{L(\log^+ L)^2(S^{n-1})} + \|\nabla K\|_{L(\log^+ L)^2(S^{n-1})}) \|\nabla b\|_{L^\infty} \|f\|_{L^p}, \ 1 < p < \infty.$$ 

For $\{[b, T_\varepsilon]f\}_{\varepsilon > 0}$, we have

$$\nabla [b, T_\varepsilon]f(x) = -[b, \nabla] T_\varepsilon f(x) + [b, \nabla T_\varepsilon]f(x) = -(\nabla b)(x) T_\varepsilon f(x) + [b, \nabla T_\varepsilon]f(x).$$

Similarly to the proof of Theorem 1.2 in [18], we can get for $1 < p < \infty$,

$$\|\lambda \sqrt{N\lambda}([\nabla b, T_\varepsilon]f)\|_{L^p} \leq C\|\nabla K\|_{L(\log^+ L)^2(S^{n-1})} \|\nabla b\|_{L^\infty} \|f\|_{L^p}.$$ 

Then by (S.2), we get for $1 < p < \infty$,

$$\|\lambda \sqrt{N\lambda}(\nabla T_\varepsilon f)\|_{L^p} \leq C(\|K\|_{L(\log^+ L)^2(S^{n-1})} + \|\nabla K\|_{L(\log^+ L)^2(S^{n-1})}) \|\nabla b\|_{L^\infty} \|f\|_{L^p}.$$ 

\[\square\]

References

[1] M. Akcoglu, R. L. Jones and P. Schwartz, Variation in probability, ergodic theory and analysis, Illinois. J. Math. 42 (1998), 154-177.

[2] P. Auscher and M. Taylor, Parabolic differential operators and commutator estimates: Paradifferential operators, Com. PDE. 20 (1995), 1743-1775.
Jump and variational inequalities

[3] J. M. Bony, *Calcul symbolique et propagation des singularites pour les equations aux derivees partielles non lineaires*, Ann. Sci. Ecole Norm. Sup. (4), 14 (1981), 209-246.

[4] J. Bourgain, *Pointwise ergodic theorems for arithmetic sets*, Inst. Hautes Études SCI, Publ. Math. 69 (1989), 5-41.

[5] J. Campbell, R. Jones, K. Reinhold and M. Wierdl, *Oscillation and variation for the Hilbert transform*, Duke Math. J. 105 (2000), 59-83.

[6] J. Campbell, R. Jones, K. Reinhold and M. Wierdl, *Oscillation and variation for singular integrals in higher dimensions*, Trans. Amer. Math. Soc. 355 (2002), 2115-2137.

[7] A. P. Calderón, *Uniqueness in the Cauchy problem for partial differential equations*, Amer. J. Math. 80 (1958), 16-36.

[8] A. P. Calderón, *Commutators of singular integrals*, Proc. Nat. Acad. Sci. USA. 53 (1965), 1092-1099.

[9] A. P. Calderón, *Commutators, singular integrals on Lipschitz curves and application*, Proc. Inter. Con. Math. Helsinki, 1978, 86-95.

[10] J. Chen, D. Fan and Y. M. Ying, *Certain operators with rough singular kernels*, Canadian J. Math., 55 (2003), 504-532.

[11] Y. Chen and Y. Ding, *Rough singular integrals on Triebel-Lizorkin space and Besov space*, J. Math. Anal. Appl. 347 (2008) 493-501.

[12] Y. Chen and Y. Ding, *Necessary and sufficient conditions for the bounds of the Calderón type commutator for the Littlewood-Paley operator*, Nonl. Anal. TMA.130(2016), 279-297.

[13] Y. Chen, Y. Ding and G. Hong, *Commutators with fractional differentiation and new characterizations of BMO-Sobolev spaces*, Analysis and PDE. To appear.

[14] M. Christ and J.-L. Journé, *Polynomial growth estimates for multilinear singular integral operators*, Acta Math. 159 (1987), 51-80.

[15] J. Cohen, *A sharp estimate for a multilinear singular integral in $\mathbb{R}^n$*, Indiana Univ. Math. J. 30 (1981), 693-702.

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

[17] R. Coifman and Y. Meyer, *Au delà des operateurs pseudo-differentiels*, Asterisque, no.57, Soc. Math. de France, 1978.

[18] Y. Ding, G. Hong and H. Liu, *Jump and variational inequalities for rough operators*, J. Four. Anal. Appl. (2016), DOI 10.1007/s00041-016-9484-8.

[19] Y. Do, C. Muscalu and C. Thiele, *Variational estimates for paraproducts*, Rev. Mat. Iberoam. 28 (2012), 857-878.
[20] J. Duoandikoetxea and J. Rubio de Francia, Maximal and singular integral operators via Fourier transform estimates, Invent. Math. 84 (1986), 541-561.

[21] J. Duoandikoetxea, Fourier analysis, Graduate Studies in Math, AMS Providence, Rhode Island. 29 (2001), 106.

[22] C. Fefferman, Recent Progress in Classical Fourier Analysis, Proc. Inter. Con. Math. Vancouver, 1974, 95-118.

[23] C. Fefferman and D. Phong, On positivity of pseudodifferential operators, Proc. Nat. Acad. Sci. USA. 75 (1978), 4673-4674.

[24] C. Fefferman and E. M. Stein, Some maximal inequalities, Amer. J. Math. 93 (1971), 107-115.

[25] T. Gillespie and J. Torrea, Dimension free estimates for the oscillation of Riesz transforms, Israel J. Math. 141 (2004), 125-144.

[26] L. Grafakos, Classic Fourier Analysis, Graduate Texts in Mathematics, Springer, New York. 249 (2008).

[27] L. Grafakos, Modern Fourier Analysis, Graduate Texts in Mathematics, Springer, New York. 250 (2008).

[28] L. Grafakos and P. Honzík, A weak-type estimate for commutators, Int. Math. Res. Not. 2012(2012), 4785-4796.

[29] S. Hofmann, Weighted inequalities for commutators of rough singular integrals, Indian U. Math. J. 39 (1990), 1275-1304.

[30] S. Hofmann, Parabolic singular integrals of Calderón-type, rough operators, and caloric layer potentials, Duke Math. J. 90 (1997), 209-259.

[31] L. Hörmander, Pseudo-differential operators of type (1, 1), Comm. PDE. 13 (1988), 1085-1111.

[32] G. Hong, The behaviour of square functions from ergodic theory in $L^\infty$, Proc. Amer. Math. Soc. 143 (2015), 4797-4802.

[33] T. Hytönen, M. Lacey and I. Parissis, A variation norm Carleson theorem for vector-valued Walsh-Fourier series, Rev. Mat. Iberoam. 30 (2014), 979-1014.

[34] R. Jones, R. Kaufman, J. Rosenblatt and M. Wierdl, Oscillation in ergodic theory, Ergodic Theory Dynam. Systems. 18 (1998), 889-935.

[35] R. Jones, J. Rosenblatt and M. Wierdl, Oscillation inequalities for rectangles, Proc. Amer. Math. Soc. 129 (2000), 1349-1358.

[36] R. Jones, J. Rosenblatt and M. Wierdl, Oscillation in ergodic theory: Higher dimensional results, Israel J. Math. 135 (2003), 1-27.
[37] R. Jones, A. Seeger and J. Wright, Strong variational and jump inequalities in harmonic analysis, Trans. Amer. Math. Soc. **360** (2008), 6711-6742.

[38] B. Krause, Polynomial Ergodic Averages converge rapidly: variations on a theorem of Bourgain, arXiv: 1402.1803.

[39] B. Krause and P. Zorin-Kranich, Weighted and vector-valued variational estimates for ergodic averages, arXiv: 1409.7120.

[40] C. Le Merdy and Q. Xu, Strong q-variation inequalities for analytic semigroups, Ann. Inst. Fourier. **62** (2012), 2069-2097.

[41] D. Lépingle, La variation d’ordre $p$ des semi-martingales, Z. Wahrsch. Verw. Gebiete. **36** (1976), 295-316.

[42] T. Ma, J. Torrea and Q. Xu, Weighted variation inequalities for differential operators and singular integrals, J. Funct. Anal. **268** (2015), 376-416.

[43] A. Mas and X. Tolsa, Variation for the Riesz transform and uniform rectifiability, J. Eur. Math. Soc. **16** (2014), 2267-321.

[44] Y. Meyer, Ondelettes et Opérateurs, Vol. II, Hermann, Paris, 1990.

[45] Y. Meyer and R. Coifman, Ondelettes et Opérateurs, Vol. III, Hermann, Paris, 1991.

[46] M. Mirek, B. Trojan, Discrete maximal functions in higher dimensions and applications to ergodic theory, arXiv: 1405.5566.

[47] M. Mirek, E. M. Stein and B. Trojan, $\ell_p(Z^d)$-estimates for discrete operators of Radon type: Variational estimates, arXiv: 1512.07523.

[48] M. Mirek, B. Trojan and P. Zorin-Kranich, Variational estimates for averages and truncated singular integrals along the prime numbers, arXiv: 1410.3255.

[49] I. Mitrea, M. Mitrea and M. Taylor, Cauchy integrals, Calderón projectors, and Toeplitz operators on uniformly rectifiable domains, Adv. Math. **268** (2015), 666-757.

[50] B. Muckenhoupt, R. L. Wheeden, Weighted norm inequalities for singular and fractional integrals, Tran. Amer. Math. Soc. **161** (1971), 249-258.

[51] T. Murai, Boundedness of singular integral operators of Calderón-type, Adv. Math. **59** (1986), 71-81.

[52] R. Oberlin, A. Seeger, T. Tao, C. Thiele and J. Wright, A variation norm Carleson theorem, J. Eur. Math. Soc. **14** (2012), 421-464.

[53] G. Pisier and Q. Xu, The strong $p$-variation of martingales and orthogonal series, Probab. Theory Related Fields. **77** (1988), 497-514.
[54] J. Qian, *The p-variation of partial sum processes and the empirical process*, Ann. Prob. **77** (1998), 1370-1383.

[55] E. M. Stein, *Harmonic Analysis, Real Variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton Univ. Press, Princeton, NJ, 1993.

[56] E. M. Stein, *Singular Integrals and Pseudo-differential Operators*, Graduate Lecture Notes, Princeton Univ., 1972.

[57] M. Taylor, *Pseudodifferential Operators and Nonlinear PDE*, Birkhauser, Boston, 1991.

[58] M. Taylor, *Tools for PDE: pseudodifferential operators, paradifferential operators, and layer potentials*, Amer. Math. Soc. **81** (2000).

[59] M. Taylor, *Commutator estimates for Hölder continuous and bmo-Sobolev multipliers*, Proc. Amer. Math. Soc. **143** (2015), 5265-5274.

[60] A. Youssfi, *Regularity properties of commutators and BMO-Triebel-Lizorkin spaces*, Ann. I. Fourier **45** (1995), 795-807.

[61] P. Zorin-Kranich, *Variation estimates for averages along primes and polynomials*, arXiv: 1403.4085.