Analysis Math., **49 (4)** (2023), 971–1006
DOI: 10.1007/s10476-023-0246-y

BESOV SPACES, SCHATTEN CLASSES AND WEIGHTED VERSIONS OF THE QUANTISED DERIVATIVE

Z. GONG[1,†], J. LI[2,*,‡] and B. D. WICK[3,§]

[1]Department of Applied Mathematics, School of Mathematics and Physics,
University of Science and Technology Beijing, Beijing 100083, P. R. China
e-mail: gongzhenb@126.com

[2]School of Mathematical and Physical Sciences, Macquarie University, Sydney NSW, 2109, Australia
e-mail: ji.li@mq.edu.au

[3]Department of Mathematics, Washington University – St. Louis, St. Louis, MO 63130-4899, USA
e-mail: wick@math.wustl.edu

(Received October 5, 2022; revised October 13, 2023; accepted October 16, 2023)

Dedicated to Professor Oleg Besov on the occasion of his 90th birthday

Abstract. In this paper, we establish the Schatten class and endpoint weak Schatten class estimates for the commutator of Riesz transforms on weighted L^2 spaces. As an application a weighted version for the estimate of the quantised derivative introduced by Alain Connes and studied recently by Lord–McDonald–Sukochev–Zanin and Frank–Sukochev–Zanin is provided.

1. Introduction

The commutator $[b, T]$ of the singular integral operator T with a symbol b, which is defined by

$$[b, T]f(x) = b(x)Tf(x) - T(bf)(x),$$

has played a vital role in harmonic analysis, complex analysis, and partial differential equations. We refer to the fundamental work by Nehari [24],

* Corresponding author.
† ZG's research is supported by the China Scholarship Council, grant number 202006460063.
‡ JL's research is supported by Australian Research Council DP 220100285.
§ BDW's research is supported in part by National Science Foundation Grants DMS #1800057, #2054863, and #2000510 and Australian Research Council DP 220100285.
Key words and phrases: Schatten class, commutator, Riesz transform, Besov space.
Mathematics Subject Classification: 47B10, 42B20, 43A85.

Calderón [2], and Coifman–Rochberg–Weiss [5]. It has been extensively studied by many authors in different aspects with various applications, see for example [3,4,16,25].

Besides the boundedness and compactness, the Schatten class estimates of the commutator have been an important topic, as it connects to non-commutative analysis. For example, the commutator of the Riesz transforms $[b, R_j]$, $j = 1, \dots, n$, links to the quantised derivative

$$\bar{d}b := i\big[\operatorname{sgn}(\mathcal{D}), 1 \otimes M_b\big]$$

of Alain Connes introduced in [7, Chapter IV], where M_b is the multiplication operator defined as $M_b f(x) = b(x)f(x)$. This has been intensively studied in [10,12,13,18,22,30]. We note that in [22] they implemented a new approach to prove that for $b \in L^\infty(\mathbb{R}^n)$, $\bar{d}b$ is in the weak Schatten class if and only if b is in the Sobolev space.

In [20], the authors have considered the Schatten class estimate of the commutator of the Hilbert transform in the two-weight setting, along the line in [1] and [19], and made a fundamental first step.

Theorem A. *Let H be the Hilbert transform on $\mathbb{R}$, $\mu, \lambda \in A_2(\mathbb{R})$ and set $\nu = \mu^{\frac{1}{2}}\lambda^{-\frac{1}{2}}$. Suppose $b \in \mathrm{VMO}(\mathbb{R})$, then the commutator $[b, H]$ belong to $S^2(L^2_\lambda(\mathbb{R}), L^2_\mu(\mathbb{R}))$ if and only if $b \in B^2_\nu(\mathbb{R})$.*

As observed in [20], the full characterization of Schatten class estimates of $[b, H]$ is not known, nor, the full characterization for the commutator of Riesz transforms. In fact, even the one weight setting has not been characterized before leading to the problem considered in this paper of determining the characterization of the Schatten class S^p $(0 < p < \infty)$ of the commutator of Riesz transforms in the one-weight setting. Here, we will consider the Schatten–Lorentz membership of the commutators acting on weighted spaces $L^2(w)$ for w in the Muckenhoupt A_2 class. The main approach used in this paper is based around dyadic harmonic analysis, the decomposition of the cubes via the median of a VMO function, and the use of nearly weakly orthonormal sequences from [30].

To state our result, we first recall the Schatten classes. Let $\mathcal{G}_1$ and $\mathcal{G}_2$ be separable complex Hilbert spaces. Suppose T is a compact operator from $\mathcal{G}_1$ to $\mathcal{G}_2$ and T^* the adjoint operator. It is clear that $|T| = (T^*T)^{\frac{1}{2}}$ is a compact, self-adjoint, and non-negative operator on $\mathcal{G}_1$. Let $(\psi_k)_k$, $k \in \mathbb{Z}_+$, be an orthonormal basis for $\mathcal{G}_1$ consisting of the eigenvectors of $|T|$, and let $s_k(T)$ be the eigenvalue corresponding to the eigenvector ψ_k. The numbers $s_1(T) \geq s_2(T) \geq \cdots \geq s_n(T) \geq \cdots \geq 0$, are called the singular values of T.

If $0 < p < \infty$, $0 < q \leq \infty$ and the sequence of singular values is $\ell^{p,q}$-summable (with respect to a weight), then T is said to belong to the

Schatten–Lorentz class $S^{p,q}(\mathcal{G}_1,\mathcal{G}_2)$. That is, $T\in S^{p,q}(\mathcal{G}_1,\mathcal{G}_2)$ if and only if

$$\|T\|_{S^{p,q}(\mathcal{G}_1,\mathcal{G}_2)}=\Big(\sum_{k\in\mathbb{Z}^+}(s_k(T))^q\,(1+k)^{\frac{q}{p}-1}\Big)^{1/q},\quad q<\infty,$$

and

$$\|T\|_{S^{p,\infty}(\mathcal{G}_1,\mathcal{G}_2)}=\sup_{k\in\mathbb{Z}^+}s_k(T)(1+k)^{1/p},\quad q=\infty.$$

Set, $S^{p,p}(\mathcal{G}_1,\mathcal{G}_2)=S^p(\mathcal{G}_1,\mathcal{G}_2)$. Moreover, see for example [25], we also have

$$S^{p_1,q_1}(\mathcal{G}_1,\mathcal{G}_2)\subset S^{p_2,q_2}(\mathcal{G}_1,\mathcal{G}_2)\ \text{ for all} q_1,q_2\quad\text{if } p_1<p_2, \tag{1.1}$$

$$S^{p,q_1}(\mathcal{G}_1,\mathcal{G}_2)\subset S^{p,q_2}(\mathcal{G}_1,\mathcal{G}_2)\quad\text{if } q_1<q_2. \tag{1.2}$$

If $\mathcal{G}_1=\mathcal{G}_2=\mathcal{G}$, we will simply write $S^{p,q}(\mathcal{G},\mathcal{G})=S^{p,q}(\mathcal{G})$.

Suppose $w\in A_2$, which will be defined in the next section. It is easy to see that $[b,R_j]$ is bounded, respectively compact on $L^2(\mathbb{R}^n,w)$, if and only if b is in BMO, respectively VMO; see for example [3]. We now consider the Besov space $B^{p,p}_{n/p}(\mathbb{R}^n)$, $0<p<\infty$, defined as the set of $b\in L^1_{\rm loc}(\mathbb{R}^n)$ such that

$$\|b\|_{B^{p,p}_{n/p}(\mathbb{R}^n)}:=\Big(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}\frac{|b(x)-b(y)|^p}{|x-y|^{2n}}\,dy\,dx\Big)^{1/p}<\infty. \tag{1.3}$$

Note that $B^{p,p}_{n/p}(\mathbb{R}^n)\subset \mathrm{VMO}(\mathbb{R}^n)$. Thus, for $b\in B^{p,p}_{n/p}(\mathbb{R}^n)$, $[b,R_j]$ is both bounded and compact. Our first result then provides a characterization of when the commutator is in the Schatten class $S_p(L^2(w))$ in terms of membership of the symbol in the Besov space $B^{p,p}_{n/p}(\mathbb{R}^n)$:

THEOREM 1.1. *Suppose $n>1$, $0<p<\infty$, $w\in A_2$ and $b\in \mathrm{VMO}(\mathbb{R}^n)$. Then for any $j=1,2,\dots,n$, the commutator $[b,R_j]\in S^p(L^2(\mathbb{R}^n,w))$ if and only if*

(1) *$b\in B^{p,p}_{n/p}(\mathbb{R}^n)$ for $n<p<\infty$. Moreover, we have $\|b\|_{B^{p,p}_{n/p}(\mathbb{R}^n)}\approx \|[b,R_j]\|_{S^p(L^2(\mathbb{R}^n,w))}$;*

(2) *b is a constant when $0<p\le n$.*

In Theorem 1.1, we note that there is a “cut-off” in the sense that the function space collapses to constants when p is less than the critical index, $p=n$, of the dimension. This suggests that at the endpoint $p=n$ there might be a more interesting phenomenon going on when one replaces membership in the Schatten–Lorentz space by its membership in weak-type versions. This leads to the following result at the critical index.

THEOREM 1.2. *Suppose $n > 1$, $b \in \mathrm{VMO}(\mathbb{R}^n)$, $w \in A_2$. Then for any $j = 1, 2, \dots, n$, the commutator $[b, R_j] \in S^{n,\infty}(L^2(\mathbb{R}^n, w))$ if and only if $b \in \dot{W}^{1,n}(\mathbb{R}^n)$. Moreover,*

$$\|b\|_{\dot{W}^{1,n}(\mathbb{R}^n)} \approx \|[b, R_j]\|_{S^{n,\infty}(L^2(\mathbb{R}^n,w))}.$$

Here $\dot{W}^{1,n}(\mathbb{R}^n)$ is the homogeneous Sobolev space on $\mathbb{R}^n$ defined by $\dot{W}^{1,n}(\mathbb{R}^n) = \{b \in (\mathcal{S}(\mathbb{R}^n))' : \nabla b \in L^n(\mathbb{R}^n)\}$ with the seminorm $\|b\|_{\dot{W}^{1,n}(\mathbb{R}^n)} = \|\nabla b\|_{L^n(\mathbb{R}^n)}$.

Once we have this, we provide a new application to the quantised derivative of Connes, which is the following result:

THEOREM 1.3. *Suppose $n > 1$, $f \in \mathrm{VMO}(\mathbb{R}^n)$, $w \in A_2$. Then $\bar{d}f \in S^{n,\infty}(\mathbb{C}^N \otimes L^2(w))$ if and only if $f \in \dot{W}^{1,n}(\mathbb{R}^n)$. Moreover,*

$$\|\bar{d}f\|_{S^{n,\infty}(\mathbb{C}^N \otimes L^2(w))} \approx \|f\|_{\dot{W}^{1,n}(\mathbb{R}^n)}.$$

Details of the proof of this application and how it follows immediately from Theorem 1.2 are given in Section 7.

The rest of this paper is organized as follows. Section 2 provides preliminary background information and notation. The proof of Theorem 1.1 is started in Section 3 where the proof of (1) in Theorem 1.1 is given. In Section 4, we give the proof of (2) of Theorem 1.1, and Section 5 provides the proof of Theorem 1.2. In Section 6, we discuss the one-dimensional case.

Throughout this paper, using $A \lesssim B$ and $A \gtrsim B$ to denote the statement that $A \le CB$ and $A \ge CB$ for some constant $C > 0$, and $A \approx B$ to denote the statement that $A \lesssim B$ and $A \gtrsim B$. The letter "C" will denote a positive constant whose value can change at each appearance. As usual, for $p \ge 1$, $\frac{1}{p} + \frac{1}{p'} = 1$, and $\ell(Q)$ denotes the sidelength of Q.

2. Preliminaries

2.1. A_2 weights.

We now recall the definition of Muckenhoupt weights.

DEFINITION 2.1. Let $w(x)$ be a nonnegative locally integrable function on $\mathbb{R}^n$. We say w is an A_2 weight, written $w \in A_2(\mathbb{R}^n)$, if

$$[w]_{A_2} := \sup_Q \frac{1}{|Q|} \int_Q w(x)\,dx \cdot \frac{1}{|Q|} \int_Q w(x)^{-1}\,dx < \infty.$$

Here the supremum is taken over all cubes $Q \subset \mathbb{R}^n$. The quantity $[w]_{A_2}$ is called the A_2 constant of w.

It is well known that $A_2(\mathbb{R}^n)$ weights are doubling. Namely, Lemma 2.1 [14]. *Let $w \in A_2(\mathbb{R}^n)$. Then for every $\lambda > 1$ and for every cube $Q \subset \mathbb{R}^n$,*

$$w(\lambda Q) \lesssim \lambda^{2n} w(Q).$$

In this article, we will also use the reverse Hölder inequality for $A_2(\mathbb{R}^n)$ weights.

Lemma 2.2 [14]. *Let $w \in A_2(\mathbb{R}^n)$. There is a reverse Hölder exponent $\sigma_w > 0$, such that for every cube $Q \subset \mathbb{R}^n$*

$$\left[\frac{1}{|Q|}\int_Q w^{1+\sigma_w}(x)\,dx\right]^{\frac{1}{1+\sigma_w}} \lesssim \frac{w(Q)}{|Q|}. \tag{2.1}$$

2.2. Dyadic systems in $\mathbb{R}^n$.

Definition 2.2. Let the collection $\mathscr{D}^0 = \mathscr{D}^0(\mathbb{R}^n)$ denote the standard system of dyadic cubes on $\mathbb{R}^n$, where

$$\mathscr{D}^0(\mathbb{R}^n) = \bigcup_{k\in\mathbb{Z}} \mathscr{D}^0_k(\mathbb{R}^n)$$

with

$$\mathscr{D}^0_k(\mathbb{R}^n) = \left\{2^{-k}([0,1)^n + m) : k \in \mathbb{Z},\ m \in \mathbb{Z}^n\right\}.$$

Next, we recall a shifted system of dyadic cubes on $\mathbb{R}^n$.

Definition 2.3 [17]. For $\omega = (\omega_1, \omega_2, \dots, \omega_n) \in \left\{0, \frac{1}{3}, \frac{2}{3}\right\}^n$, define the shifted dyadic system $\mathscr{D}^\omega = \mathscr{D}^\omega(\mathbb{R}^n)$,

$$\mathscr{D}^\omega(\mathbb{R}^n) = \bigcup_{k\in\mathbb{Z}} \mathscr{D}^\omega_k(\mathbb{R}^n),$$

where

$$\mathscr{D}^\omega_k(\mathbb{R}^n) = \left\{2^{-k}([0,1)^n + m + (-1)^k\omega) : k \in \mathbb{Z},\ m \in \mathbb{Z}^n\right\}.$$

It is straightforward to check that $\mathscr{D}^\omega$ inherits the nestedness property of $\mathscr{D}^0$: if $Q, Q' \in \mathscr{D}^\omega$, then $Q \cap Q' \in \{Q, Q', \varnothing\}$. See [17] for more details. When the particular choice of ω is unimportant, the notation $\mathscr{D}$ is sometimes used for a generic dyadic system.

2.3. An expression of Haar functions. Next we recall the Haar basis on $\mathbb{R}^n$. For any dyadic cube $Q \in \mathscr{D}$, there exist dyadic intervals I_1, I_2, …, I_n on $\mathbb{R}$ with common length $l(Q)$, such that $Q = I_1 \times I_2 \times \cdots \times I_n$. Then Q is associated with 2^n Haar functions:

$$h^\varepsilon_Q(x) := h^{(\varepsilon_1,\varepsilon_2,\dots,\varepsilon_n)}_{I_1\times I_2\times\cdots\times I_n}(x_1, x_2, \dots, x_n) := \prod_{i=1}^n h^{(\varepsilon_i)}_{I_i}(x_i)$$

where $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n) \in \{0,1\}^n$ and

$$h_{I_i}^{(1)} := \frac{1}{\sqrt{I_i}} \chi_{I_i} \quad \text{and} \quad h_{I_i}^{(0)} := \frac{1}{\sqrt{I_i}} (\chi_{I_i -} - \chi_{I_i +}).$$

Writing $\varepsilon \equiv 1$ when $\varepsilon_i \equiv 1$ for all $i = 1, 2, \ldots, n$, $h_Q^1 := \frac{1}{\sqrt{Q}} \chi_Q$ is noncancellative; on the other hand, when $\varepsilon \not\equiv 1$, the rest of the $2^n - 1$ Haar functions h_Q^ε associated with Q satisfy the following properties.

LEMMA 2.3. *For $\varepsilon \not\equiv 1$, we have*

(1) *h_Q^ε is supported on Q and $\int_{\mathbb{R}^n} h_Q^\varepsilon(x)\, dx = 0$;*

(2) *h_Q^ε is constant on each $R \in ch(Q)$, where $ch(Q) = \{R \in \mathscr{D}_{k+1} : R \subseteq Q\}$ denotes the dyadic sub-cubes of the cube $Q \in \mathscr{D}_k$;*

(3) *$\langle h_Q^\varepsilon, h_Q^\eta \rangle = 0$, for $\varepsilon \not\equiv \eta$;*

(4) *if $h_Q^\varepsilon \neq 0$, then*

$$\|h_Q^\varepsilon\|_{L^p(\mathbb{R}^n)} = |Q|^{\frac{1}{p} - \frac{1}{2}} \quad \text{for } 1 \le p \le \infty;$$

(5) *$\|h_Q^\varepsilon\|_{L^1(\mathbb{R}^n)} \cdot \|h_Q^\varepsilon\|_{L^\infty(\mathbb{R}^n)} = 1$;*

(6) *the average of a function b over a dyadic cube Q, $\langle b \rangle_Q := \frac{1}{|Q|} \int_Q b(x)\, dx$ can be expressed as*

$$\langle b \rangle_Q = \sum_{\substack{P \in \mathscr{D}, Q \subsetneq P \\ \varepsilon \not\equiv 1}} \langle b, h_P^\varepsilon \rangle h_P^\varepsilon(Q)$$

where $h_P^\varepsilon(Q)$ is a constant.

(7) *fixing a cube Q, and expanding b in the Haar basis, we have*

$$(b(x) - \langle b \rangle_Q) \chi_Q(x) = \sum_{\substack{R \in \mathscr{D}, R \subset Q \\ \varepsilon \not\equiv 1}} \langle b, h_R^\varepsilon \rangle h_R^\varepsilon;$$

(8) *the conditional expectation of a locally integrable function b on $\mathbb{R}^n$ with respect to the increasing family of σ-algebras $\sigma(\mathscr{D}_k)$ is given by the expression:*

$$E_k(b)(x) = \sum_{Q \in \mathscr{D}_k} \langle b \rangle_Q \chi_Q(x), \quad x \in \mathbb{R}^n,$$

where $\langle b \rangle_Q$ is the average of b over Q as defined in (6) above. Note that we have

$$E_{k+1}(b)(x) - E_k(b)(x) = \sum_{Q \in \mathscr{D}_k} \sum_{\varepsilon \not\equiv 1} \langle b, h_Q^\varepsilon \rangle h_Q^\varepsilon(x).$$

2.4. Characterization of Schatten class. In 1989, Rochberg and Semmes introduced the notion of nearly weakly orthogonal (NWO) sequences of functions contained in the following definition, see [30].

DEFINITION 2.4. Let $\{e_Q\}_{Q\in\mathscr{D}}$ be a collection of functions. We say $\{e_Q\}_{Q\in\mathscr{D}}$ is a NWO sequence, if $\operatorname{supp} e_Q \subset Q$ and the maximal function f^* is bounded on $L^p(\mathbb{R}^n)$, where f^* is defined as

$$f^*(x) = \sup_Q \frac{|\langle f, e_Q\rangle|}{|Q|^{1/2}} \chi_Q(x).$$

In this paper, we work with weighted versions of NWO sequences. We will use the following result proved by Rochberg and Semmes.

LEMMA 2.4 [30]. *If the collection of functions $\{e_Q : Q \in \mathscr{D}\}$ are supported on Q and satisfy for some $2 < r < \infty$, $\|e_Q\|_r \lesssim |Q|^{1/r-1/2}$, then $\{e_Q\}_{Q\in\mathscr{D}}$ is an* NWO *sequence.*

NWO sequences are very useful in studying Schatten class properties of operators. In [30], Rochberg and Semmes developed a substitute for the Schmidt decomposition of the operator T. If an operator T has a representation of the form

$$T = \sum_{Q\in\mathscr{D}} \lambda_Q \langle \cdot, e_Q\rangle f_Q \tag{2.2}$$

with $\{e_Q\}_{Q\in\mathscr{D}}$ and $\{f_Q\}_{Q\in\mathscr{D}}$ being NWO sequences and $\{\lambda_Q\}_{Q\in\mathscr{D}}$ is a sequence of scalars. It is easy to see that

$$\|T\|_{S^{p,q}(\mathcal{G})} \lesssim \|\lambda_Q\|_{\ell^{p,q}}, \quad 0 < p < \infty,\ 0 < q < \infty. \tag{2.3}$$

A sort of converse also holds. When $1 < p = q < \infty$, Rochberg and Semmes also obtained:

LEMMA 2.5 [30]. *For any (bounded) compact operator T on $L^2(\mathbb{R}^n)$ and $\{e_Q\}_{Q\in\mathscr{D}}$ and $\{f_Q\}_{Q\in\mathscr{D}}$* NWO *sequences, then for $1 < p < \infty$,*

$$\Big[\sum_{Q\in\mathscr{D}} |\langle Te_Q, f_Q\rangle|^p\Big]^{1/p} \lesssim \|T\|_{S^p(L^2(\mathbb{R}^n))}.$$

Lacey and the last two authors in [20] provided a relationship between Schatten norms on weighted and unweighted $L^2(\mathbb{R})$ (which also generalizes to $\mathbb{R}^n$).

LEMMA 2.6 [20]. *Suppose $1 \le p < \infty$, and $w \in A_2(\mathbb{R}^n)$. Then T belongs to $S^p(L^2(\mathbb{R}^n, w))$ if and only if $w^{\frac12} T w^{-\frac12}$ belong to $S^p(L^2(\mathbb{R}^n))$. Moreover,*

$$\|T\|_{S^p(L^2(\mathbb{R}^n,w))} \approx \|w^{\frac12} T w^{-\frac12}\|_{S^p(L^2(\mathbb{R}^n))}.$$

Using this Lemma from [20], we can also obtain the following result.

LEMMA 2.7. *Suppose* $1 \le p < \infty$, *and* $w \in A_2(\mathbb{R}^n)$. *Then* T *belongs to* $S^{p,\infty}(L^2(\mathbb{R}^n, w))$ *if and only if* $w^{\frac{1}{2}} T w^{-\frac{1}{2}}$ *belong to* $S^{p,\infty}(L^2(\mathbb{R}^n))$. *Moreover,*

$$\|T\|_{S^{p,\infty}(L^2(\mathbb{R}^n,w))} \approx \left\|w^{\frac{1}{2}} T w^{-\frac{1}{2}}\right\|_{S^{p,\infty}(L^2(\mathbb{R}^n))}.$$

2.5. Description of the Besov space.

DEFINITION 2.5. Suppose $0 < p, q < \infty$ and $0 < \alpha < 1$. Let $b \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$. Then b belongs to the Besov space $B^{p,q}_\alpha(\mathbb{R}^n)$ if and only if

$$\left(\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n} \frac{|b(x)-b(y)|^p}{|x-y|^{\frac{p}{q}n+p\alpha}}\,dy\right)^{q/p} dx\right)^{1/q} < \infty.$$

In particular, note that if $\alpha = \frac{n}{p}$ and $p = q$ then

$$\|b\|_{B^{p,p}_{n/p}(\mathbb{R}^n)} := \left(\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|b(x)-b(y)|^p}{|x-y|^{2n}}\,dy\,dx\right)^{1/p}.$$

Useful in the proof below will be dyadic norms and so we give the norm of the dyadic Besov space next.

DEFINITION 2.6. Suppose $0 < p < \infty$. Let $b \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and $\mathscr{D}$ be an arbitrary dyadic system in $\mathbb{R}^n$. Then b belongs to the dyadic Besov space $B^p_d(\mathbb{R}^n, \mathscr{D})$ if and only if

$$\|b\|_{B^p_d(\mathbb{R}^n,\mathscr{D})} := \left(\sum_{\substack{Q\in\mathscr{D}\\ \varepsilon\not\equiv 1}} \left(|\langle b, h^\varepsilon_Q\rangle||Q|^{-\frac{1}{2}}\right)^p\right)^{1/p} < \infty.$$

Suppose $w \in A_2(\mathbb{R}^n)$. By the definition of $A_2(\mathbb{R}^n)$ weights, we obtain that

$$\frac{w(Q)w^{-1}(Q)}{|Q||Q|} \approx 1.$$

Then we have an equivalent norm on this dyadic space given by

$$\|b\|^p_{B^p_d(\mathbb{R}^n,\mathscr{D})} = \sum_{\substack{Q\in\mathscr{D}\\ \varepsilon\not\equiv 1}} \left(|\langle b, h^\varepsilon_Q\rangle||Q|^{-\frac{1}{2}}\right)^p$$

$$\approx \sum_{\substack{Q\in\mathscr{D}\\ \varepsilon\not\equiv 1}} \left(\frac{w(Q)^{\frac{1}{2}}(w^{-1}(Q))^{\frac{1}{2}}|\langle b, h^\varepsilon_Q\rangle|}{|Q|^{\frac{3}{2}}}\right)^p \approx \sum_{\substack{Q\in\mathscr{D}\\ \varepsilon\not\equiv 1}} \left(\frac{|\langle b, h^\varepsilon_Q\rangle||Q|^{\frac{1}{2}}}{w(Q)^{\frac{1}{2}}(w^{-1}(Q))^{\frac{1}{2}}}\right)^p.$$

Key to our analysis will be the fact that a suitable family of dyadic norms is equivalent to the norm in the continuous setting, the content of the next lemma.

LEMMA 2.8. *Suppose* $w \in A_2(\mathbb{R}^n)$, $n < p < \infty$. *There are dyadic systems* $\mathscr{D}^\omega$, $\omega \in \left\{0, \frac{1}{3}, \frac{2}{3}\right\}^n$, *such that*

$$\bigcap_{\omega \in \{0, \frac{1}{3}, \frac{2}{3}\}^n} B_d^p(\mathbb{R}^n, \mathscr{D}^\omega) = B_{n/p}^{p,p}(\mathbb{R}^n),$$

with

$$\sum_{\omega \in \{0, \frac{1}{3}, \frac{2}{3}\}^n} \|b\|_{B_d^p(\mathbb{R}^n, \mathscr{D}^\omega)} \approx \|b\|_{B_{n/p}^{p,p}(\mathbb{R}^n)}.$$

PROOF. On the one hand, we first prove the dyadic Besov norm is dominated by the continuous Besov norm, that is, for every $\omega \in \left\{0, \frac{1}{3}, \frac{2}{3}\right\}^n$,

$$\|b\|_{B_d^p(\mathbb{R}^n, \mathscr{D}^\omega)} \lesssim \|b\|_{B_{n/p}^{p,p}(\mathbb{R}^n)}.$$

Choose a dyadic cube $Q \in \mathscr{D}^\omega$. Let $\hat{Q} = Q + \{2\ell(Q)\}^n$. By applying Lemma 2.3 and Hölder's inequality, we have

$$\begin{aligned}
|\langle b, h_Q^\varepsilon \rangle| |Q|^{-\frac{1}{2}} &\lesssim \left| \int_Q b(x) h_Q^\varepsilon(x)\, dx \right| \frac{|\hat{Q}|}{|Q|^{\frac{3}{2}}} \\
&\lesssim \int_{\hat{Q}} \int_Q |b(x) - b(y)| |h_Q^\varepsilon(x)|\, dx\, dy |Q|^{-\frac{3}{2}} \\
&\lesssim \left(\int_{\hat{Q}} \int_Q \frac{|b(x) - b(y)|^p}{|x - y|^{2n}}\, dx\, dy \right)^{1/p} \left(\int_{\hat{Q}} \int_Q |h_Q^\varepsilon(x)|^{p'}\, dx\, dy \right)^{1/p'} |Q|^{-\frac{2}{p'} + \frac{1}{2}} \\
&\lesssim \left(\int_{\hat{Q}} \int_Q \frac{|b(x) - b(y)|^p}{|x - y|^{2n}}\, dx\, dy \right)^{1/p}.
\end{aligned}$$

Hence, we can obtain

$$\begin{aligned}
\|b\|_{B_d^p(\mathbb{R}^n, \mathcal{D}^\omega)}^p &= \sum_{\substack{Q \in \mathscr{D}^\omega \\ \varepsilon \neq 1}} \left(|\langle b, h_Q^\varepsilon \rangle| |Q|^{-\frac{1}{2}} \right)^p \lesssim \sum_{Q \in \mathscr{D}^\omega} \int_{\hat{Q}} \int_Q \frac{|b(x) - b(y)|^p}{|x - y|^{2n}}\, dx\, dy \\
&\lesssim \sum_{k \in \mathbb{Z}} \sum_{Q \in \mathscr{D}_k^\omega} \int_Q \int_{\{y \in \mathbb{R}^n : 2^{-k} \le |x - y| \le 3 \cdot 2^{-k}\}} \frac{|b(x) - b(y)|^p}{|x - y|^{2n}}\, dy\, dx \lesssim \|b\|_{B_{n/p}^{p,p}(\mathbb{R}^n)}^p.
\end{aligned}$$

On the other hand, we need to show that

$$\|b\|_{B^{p,p}_{n/p}(\mathbb{R}^n)} \lesssim \sum_{\omega\in\{0,\frac{1}{3},\frac{2}{3}\}^n} \|b\|_{B^p_d(\mathbb{R}^n,\mathscr{D}^\omega)}. \tag{2.4}$$

Observe that

$$\|b\|^p_{B^{p,p}_{n/p}(\mathbb{R}^n)} = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{|b(x)-b(y)|^p}{|x-y|^{2n}}\,dy\,dx$$
$$= \sum_{k\in\mathbb{Z}}\sum_{Q\in\mathscr{D}_k}\int_Q\int_{\{y\in\mathbb{R}^n:2^{-k}<|x-y|\le 2^{-k+1}\}} \frac{|b(x)-b(y)|^p}{|x-y|^{2n}}\,dy\,dx$$
$$\le \sum_{k\in\mathbb{Z}}\sum_{Q\in\mathscr{D}_k}\int_Q\int_{\{y\in\mathbb{R}^n:0<|x-y|\le 3\cdot 2^{-k}\}} \frac{|b(x)-b(y)|^p}{|x-y|^{2n}}\,dy\,dx.$$

It is clear that for $x\in Q$, that $\{y\in\mathbb{R}^n : 0<|x-y|\le 3\cdot 2^{-k}\}\subset 7Q$. Then there is $J^\omega\in\mathscr{D}^\omega$ such that

$$7Q\subset \bigcup_{\omega\in\{0,\frac{1}{3},\frac{2}{3}\}^n} J^\omega,$$

(see for example [17]). Then, we have

$$\|b\|_{B^{p,p}_{n/p}(\mathbb{R}^n)} \lesssim \Big(\sum_{\omega\in\{0,\frac{1}{3},\frac{2}{3}\}^n}\sum_{J^\omega\in\mathscr{D}^\omega}\frac{1}{|J^\omega|^2}\int_{J^\omega}\int_{J^\omega}|b(x)-b(y)|^p\,dy\,dx\Big)^{1/p}$$
$$\le \Big(\sum_{\omega\in\{0,\frac{1}{3},\frac{2}{3}\}^n}\sum_{J^\omega\in\mathscr{D}^\omega}\frac{1}{|J^\omega|^2}\int_{J^\omega}\int_{J^\omega}|b(x)-\langle b\rangle_{J^\omega}|^p\,dx\,dy\Big)^{1/p}$$
$$+\Big(\sum_{\omega\in\{0,\frac{1}{3},\frac{2}{3}\}^n}\sum_{J^\omega\in\mathscr{D}^\omega}\frac{1}{|J^\omega|^2}\int_{J^\omega}\int_{J^\omega}|b(y)-\langle b\rangle_{J^\omega}|^p\,dy\,dx\Big)^{1/p}$$
$$=2\Big(\sum_{\omega\in\{0,\frac{1}{3},\frac{2}{3}\}^n}\sum_{J^\omega\in\mathscr{D}^\omega}\frac{1}{|J^\omega|}\int_{J^\omega}|b(x)-\langle b\rangle_{J^\omega}|^p\,dx\Big)^{1/p} =: 2\mathfrak{I}.$$

To continue, we first denote

$$E^\omega_k(b)(x) := \sum_{J^\omega\in\mathscr{D}^\omega_k}\langle b\rangle_{J^\omega}\chi_{J^\omega}(x).$$

Then we have

$$\mathfrak{I} = \Big(\sum_{\omega\in\{0,\frac13,\frac23\}^n} \sum_{k=-\infty}^{\infty} \sum_{J^\omega\in\mathscr{D}_k^\omega} \frac{1}{|J^\omega|} \int_{J^\omega} \big|b(x) - E_k^\omega(b)(x)\big|^p\,dx\Big)^{1/p} \tag{2.5}$$

$$\leq \Big(\sum_{\omega\in\{0,\frac13,\frac23\}^n} \sum_{k=-\infty}^{\infty} 2^{kn} \sum_{J^\omega\in\mathscr{D}_k^\omega} \int_{J^\omega} \big|b(x) - E_{k+1}^\omega(b)(x)\big|^p\,dx\Big)^{1/p}$$

$$+ \Big(\sum_{\omega\in\{0,\frac13,\frac23\}^n} \sum_{k=-\infty}^{\infty} 2^{kn} \sum_{J^\omega\in\mathscr{D}_k^\omega} \int_{J^\omega} \big|E_{k+1}^\omega(b)(x) - E_k^\omega(b)(x)\big|^p\,dx\Big)^{1/p} =: \mathfrak{I}_1 + \mathfrak{I}_2.$$

We consider $\mathfrak{I}_2$. Note that for $x\in J^\omega$, that

$$E_{k+1}^\omega(b)(x) - E_k^\omega(b)(x) = \Big(\sum_{\varepsilon\not\equiv1} \langle b, h_{J^\omega}^\varepsilon\rangle h_{J^\omega}^\varepsilon(x)\Big)\chi_{J^\omega}(x)$$

via definition of the conditional expectation and the connection with the Haar functions. Using this fact and continuing the estimate we have

$$\mathfrak{I}_2 = \Big(\sum_{\omega\in\{0,\frac13,\frac23\}^n} \sum_{J^\omega\in\mathscr{D}^\omega} \frac{1}{|J^\omega|} \int_{J^\omega} \Big|\sum_{\varepsilon\not\equiv1} \langle b, h_{J^\omega}^\varepsilon\rangle h_{J^\omega}^\varepsilon(x)\Big|^p\,dx\Big)^{1/p} \tag{2.6}$$

$$\lesssim \Big(\sum_{\omega\in\{0,\frac13,\frac23\}^n} \sum_{J^\omega\in\mathscr{D}^\omega} \sum_{\varepsilon\not\equiv1} \frac{1}{|J^\omega|} \int_{J^\omega} \big|\langle b, h_{J^\omega}^\varepsilon\rangle\big|^p \big|h_{J^\omega}^\varepsilon(x)\big|^p\,dx\Big)^{1/p}$$

$$\lesssim \Big(\sum_{\omega\in\{0,\frac13,\frac23\}^n} \sum_{J^\omega\in\mathscr{D}^\omega} \sum_{\varepsilon\not\equiv1} \frac{\big|\langle b, h_{J^\omega}^\varepsilon\rangle\big|^p}{|J^\omega|^{\frac p2}}\Big)^{1/p} \lesssim \sum_{\omega\in\{0,\frac13,\frac23\}^n} \|b\|_{B_d^p(\mathbb{R}^n,\mathscr{D}^\omega)},$$

where the implicit constant depends on p and n.

Moreover, from the estimate of $\mathfrak{I}_2$ we also see that for each k,

$$\Big(\sum_{\omega\in\{0,\frac13,\frac23\}^n} 2^{kn} \sum_{J^\omega\in\mathscr{D}_k^\omega} \int_{J^\omega} \big|E_{k+1}^\omega(b)(x) - E_k^\omega(b)(x)\big|^p\,dx\Big)^{1/p}$$

$$\leq \sum_{\omega\in\{0,\frac13,\frac23\}^n} \|b\|_{B_d^p(\mathbb{R}^n,\mathscr{D}^\omega)},$$

which gives that

$$\Big(\sum_{\omega\in\{0,\frac13,\frac23\}^n} \sum_{J^\omega\in\mathscr{D}_k^\omega} \int_{J^\omega} \big|E_{k+1}^\omega(b)(x) - E_k^\omega(b)(x)\big|^p\,dx\Big)^{1/p}$$

$$\le 2^{-kn/p} \sum_{\omega\in\{0,\frac13,\frac23\}^n} \|b\|_{B_d^p(\mathbb{R}^n,\mathscr{D}^\omega)}.$$

This implies that for any large positive integer M,

$$\Big(\sum_{\omega\in\{0,\frac13,\frac23\}^n}\sum_{J^\omega\in\mathscr{D}_k^\omega}\int_{J^\omega}\big|b(x)-E_{M+1}^\omega(b)(x)\big|^p\,dx\Big)^{1/p} \tag{2.7}$$
$$\le \sum_{k=M+1}^{\infty}\Big(\sum_{\omega\in\{0,\frac13,\frac23\}^n}\sum_{J^\omega\in\mathscr{D}_k^\omega}\int_{J^\omega}\big|E_{k+1}^\omega(b)(x)-E_k^\omega(b)(x)\big|^p\,dx\Big)^{1/p}$$
$$\le C\,2^{-(M+1)n/p}\sum_{\omega\in\{0,\frac13,\frac23\}^n}\|b\|_{B_d^p(\mathbb{R}^n,\mathscr{D}^\omega)}.$$

Then to continue, we take a truncation of $\mathfrak{I}$ at the level $-L$ and M for large positive integers L and M:

$$\mathfrak{I}_{L,M}=\Big(\sum_{\omega\in\{0,\frac13,\frac23\}^n}\sum_{k=-L}^{M}\sum_{J^\omega\in\mathscr{D}_k^\omega}\frac{1}{|J^\omega|}\int_{J^\omega}\big|b(x)-E_k^\omega(b)(x)\big|^p\,dx\Big)^{1/p}$$
$$\le \mathfrak{I}_{1,L,M}+\mathfrak{I}_2,$$

where

$$\mathfrak{I}_{1,L,M}:=\Big(\sum_{\omega\in\{0,\frac13,\frac23\}^n}\sum_{k=-L}^{M}2^{kn}\sum_{J^\omega\in\mathscr{D}_k^\omega}\int_{J^\omega}\big|b(x)-E_{k+1}^\omega(b)(x)\big|^p\,dx\Big)^{1/p}.$$

Via the decomposition $J^\omega\in\mathscr{D}_k^\omega$ into subcubes in $\mathscr{D}_{k+1}^\omega$ we see that

$$\mathfrak{I}_{1,L,M}=\Big(\sum_{\omega\in\{0,\frac13,\frac23\}^n}\sum_{k=-L}^{M}2^{-n}2^{(k+1)n}\sum_{J^\omega\in\mathscr{D}_{k+1}^\omega}\int_{J^\omega}\big|b(x)-E_{k+1}^\omega(b)(x)\big|^p\,dx\Big)^{1/p}$$
$$=\Big(\sum_{\omega\in\{0,\frac13,\frac23\}^n}\sum_{k=-L+1}^{M+1}2^{-n}\cdot 2^{kn}\sum_{J^\omega\in\mathscr{D}_k^\omega}\int_{J^\omega}\big|b(x)-E_k^\omega(b)(x)\big|^p\,dx\Big)^{1/p}$$
$$\le 2^{-\frac{n}{p}}\Big(\sum_{\omega\in\{0,\frac13,\frac23\}^n}\sum_{k=-L+1}^{M}2^{kn}\sum_{J^\omega\in\mathscr{D}_k^\omega}\int_{J^\omega}\big|b(x)-E_k^\omega(b)(x)\big|^p\,dx\Big)^{1/p}$$
$$+\Big(\sum_{\omega\in\{0,\frac13,\frac23\}^n}2^{Mn}\sum_{J^\omega\in\mathscr{D}_k^\omega}\int_{J^\omega}\big|b(x)-E_{M+1}^\omega(b)(x)\big|^p\,dx\Big)^{1/p}$$

$$\leq 2^{-\frac{n}{p}}\mathfrak{I}_{1,L,M} + C2^{Mn/p}2^{-(M+1)n/p} \sum_{\omega\in\{0,\frac{1}{3},\frac{2}{3}\}^n} \|b\|_{B_d^p(\mathbb{R}^n,\mathscr{D}^\omega)},$$

where the last inequality follows from (2.7). Since $2^{-\frac{n}{p}} < 1$, absorbing the first term on the right hand side back into the left hand side and simplifying implies that for every L and M,

$$\mathfrak{I}_{1,L,M} \lesssim \sum_{\omega\in\{0,\frac{1}{3},\frac{2}{3}\}^n} \|b\|_{B_d^p(\mathbb{R}^n,\mathscr{D}^\omega)},$$

where the implicit constant is independent of L and M, but does depend upon n and p. Hence, letting M and L go to ∞ the above argument gives that $\mathfrak{I}_1 \lesssim \sum_{\omega\in\{0,\frac{1}{3},\frac{2}{3}\}^n} \|b\|_{B_d^p(\mathbb{R}^n,\mathscr{D}^\omega)}$, which, together with (2.6), gives

$$\mathfrak{I} \lesssim \sum_{\omega\in\{0,\frac{1}{3},\frac{2}{3}\}^n} \|b\|_{B_d^p(\mathbb{R}^n,\mathscr{D}^\omega)}.$$

Thus, we see that (2.4) holds completing the proof of Lemma 2.8. □

3. Proof of (1) in Theorem 1.1

From Lemma 2.8 we know that the continuous Besov space is the intersection of 3^n dyadic Besov spaces with $n < p < \infty$. Thus, the proof of (1) in Theorem 1.1 can be completed by discussing the following two properties.

PROPOSITION 3.1. *For $n < p < \infty$, let $w \in A_2(\mathbb{R}^n)$, and $b \in \mathrm{VMO}(\mathbb{R}^n)$ with*

$$\|[b, R_j]\|_{S^p(L^2(\mathbb{R}^n,w))} < \infty.$$

Then we have

$$\|b\|_{B_d^p(\mathbb{R}^n)} \lesssim \|[b, R_j]\|_{S^p(L^2(\mathbb{R}^n,w))}.$$

PROPOSITION 3.2. *For $n < p < \infty$, suppose that $w \in A_2(\mathbb{R}^n)$, and $b \in B_{n/p}^{p,p}(\mathbb{R}^n)$, then we have*

$$\|[b, R_j]\|_{S^p(L^2(\mathbb{R}^n,w))} \lesssim \|b\|_{B_{n/p}^{p,p}(\mathbb{R}^n)}.$$

3.1. Proof of Proposition 3.1. Below, we consider cubes $Q \in \mathscr{D}$, a fixed dyadic system. To prove Proposition 3.1, we first need to use a known result.

LEMMA 3.3. *For each dyadic cube Q, there exists another dyadic cube $\hat{Q}$ such that*

(i) $|Q| = |\hat{Q}|$, *and* $\operatorname{dist}(Q, \hat{Q}) \approx |Q|$.

(ii) *The kernel of the Riesz transform* $K_j(x - \hat{x})$ *does not change sign for all* $(x, \hat{x}) \in Q \times \hat{Q}$, *and*

$$|K_j(x - \hat{x})| \gtrsim \frac{1}{|Q|}. \tag{3.1}$$

Let $m_b(\hat{Q})$ be a median value of b over $\hat{Q}$ (see for example the definition in [23]). This means $m_b(\hat{Q})$ is a real number such that

$$\max\Big\{\big|\{y \in \hat{Q} : b(y) < m_b(\hat{Q})\}\big|, \big|\{y \in \hat{Q} : b(y) > m_b(\hat{Q})\}\big|\Big\} \le |\hat{Q}|/2. \tag{3.2}$$

By $\int_Q h_Q^\varepsilon(x)\,dx = 0$ and using (3.2), a simple calculation gives

$$\begin{aligned}
&\Big|\int_Q b(x) h_Q^\varepsilon(x)\ dx\Big| \\
&= \Big|\int_Q (b(x) - m_b(\hat{Q})) h_Q^\varepsilon(x)\,dx\Big| \le \frac{1}{|Q|^{\frac{1}{2}}} \int_Q \big|b(x) - m_b(\hat{Q})\big|\,dx \\
&\le \frac{1}{|Q|^{\frac{1}{2}}} \int_{Q\cap E_1^Q} \big|b(x) - m_b(\hat{Q})\big|\,dx + \frac{1}{|Q|^{\frac{1}{2}}} \int_{Q\cap E_2^Q} \big|b(x) - m_b(\hat{Q})\big|\,dx \\
&=: \mathrm{Term}_1^Q + \mathrm{Term}_2^Q\,.
\end{aligned}$$

Where

$$E_1^Q := \big\{x \in Q : b(x) < m_b(\hat{Q})\big\} \quad \text{and} \quad E_2^Q := \big\{x \in Q : b(x) \ge m_b(\hat{Q})\big\}.$$

Now we denote

$$F_1^{\hat{Q}} := \big\{y \in \hat{Q} : b(y) \ge m_b(\hat{Q})\big\} \quad \text{and} \quad F_2^{\hat{Q}} := \big\{y \in \hat{Q} : b(y) \le m_b(\hat{Q})\big\}.$$

Then by the definition of $m_b(\hat{Q})$, we have $|F_1^{\hat{Q}}| = |F_2^{\hat{Q}}| \approx |\hat{Q}|$ and $F_1^{\hat{Q}} \cup F_2^{\hat{Q}} = \hat{Q}$.

Note that for $s = 1, 2$, if $x \in E_s^Q$ and $y \in F_s^{\hat{Q}}$, then

$$\begin{aligned}
\big|b(x) - m_b(\hat{Q})\big| &\le \big|b(x) - m_b(\hat{Q})\big| + \big|m_b(\hat{Q}) - b(y)\big| \\
&= \big|b(x) - m_b(\hat{Q}) + m_b(\hat{Q}) - b(y)\big| = |b(x) - b(y)|.
\end{aligned}$$

Therefore, for $s = 1, 2$, by using (3.1) and by the fact that $|F_s^{\hat{Q}}| \approx |Q|$, we have

$$\mathrm{Term}_s^Q \lesssim \frac{1}{|Q|^{\frac{1}{2}}} \int_{Q\cap E_s^Q} \big|b(x) - m_b(\hat{Q})\big|\,dx\,\frac{|F_s^{\hat{Q}}|}{|Q|}$$

$$= \frac{1}{|Q|^{\frac{1}{2}}} \int_{Q\cap E_s^Q} \int_{F_s^{\hat{Q}}} \big|b(x) - m_b(\hat{Q})\big| \ \frac{1}{|Q|} \ dy\,dx$$

$$\lesssim \frac{1}{|Q|^{\frac{1}{2}}} \int_{Q\cap E_s^Q} \int_{F_s^{\hat{Q}}} \big|b(x) - m_b(\hat{Q})\big| |K_j(x-y)|\,dy\,dx$$

$$\lesssim \frac{1}{|Q|^{\frac{1}{2}}} \int_{Q\cap E_s^Q} \int_{F_s^{\hat{Q}}} |b(x) - b(y)||K_j(x-y)|\,dy\,dx.$$

To continue, by noting that $K_j(x-y)$ and $b(x)-b(y)$ do not change sign for $(x,y) \in \big(Q \cap E_s^Q\big) \times F_s^{\hat{Q}}$, $s = 1,2$, we have that

$$\mathrm{Term}_s^Q \lesssim \frac{1}{|Q|^{\frac{1}{2}}} \bigg| \int_{Q\cap E_s^Q} \int_{F_s^{\hat{Q}}} \big(b(x)-b(y)\big) K_j(x-y)\,dy\,dx \bigg|$$

$$= \frac{1}{|Q|^{\frac{1}{2}}} \bigg| \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \big(b(x)-b(y)\big) K_j(x-y) \chi_{F_s^{\hat{Q}}}(y)\,dy\, \chi_{Q\cap E_s^Q}(x)\,dx \bigg|.$$

We now insert the weight w to get

$$\mathrm{Term}_s^Q \lesssim \frac{1}{|Q|^{\frac{1}{2}}} \bigg| \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} (b(x)-b(y)) w^{\frac{1}{2}}(x) K_j(x-y) w^{-\frac{1}{2}}(y)$$

$$\times \big(w^{\frac{1}{2}}(y)\chi_{F_s^{\hat{Q}}}(y)\big)\,dy \big(w^{-\frac{1}{2}}(x)\chi_{Q\cap E_s^Q}(x)\big)\,dx \bigg|.$$

Thus, we further have

$$\sum_{Q\in\mathscr{D},\varepsilon\not\equiv 1} \bigg(\frac{|\langle b, h_Q^\varepsilon\rangle||Q|^{\frac{1}{2}}}{w(Q)^{\frac{1}{2}}(w^{-1}(Q))^{\frac{1}{2}}} \bigg)^p$$

$$\lesssim \sum_{Q\in\mathscr{D},\varepsilon\not\equiv 1} \sum_{s=1}^{2} \bigg| \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} (b(x)-b(y)) w^{\frac{1}{2}}(x) K_j(x-y) w^{-\frac{1}{2}}(y)$$

$$\times \big(w^{\frac{1}{2}}(y)\chi_{F_s^{\hat{Q}}}(y)\big)\,dy \frac{\big(w^{-\frac{1}{2}}(x)\chi_{Q\cap E_s^Q}(x)\big)}{w(Q)^{\frac{1}{2}}(w^{-1}(Q))^{\frac{1}{2}}}\,dx \bigg|^p$$

$$\lesssim \sum_{Q\in\mathscr{D},\varepsilon\not\equiv 1} \sum_{s=1}^{2} \bigg| \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} (b(x)-b(y)) w^{\frac{1}{2}}(x) K_j(x-y) w^{-\frac{1}{2}}(y)$$

$$\times \frac{w^{\frac{1}{2}}(y)\chi_{F_s^{\hat{Q}}}(y)}{w(\hat{Q})^{\frac{1}{2}}}\,dy \frac{w^{-\frac{1}{2}}(x)\chi_{Q\cap E_s^Q}(x)}{(w^{-1}(Q))^{\frac{1}{2}}}\,dx \bigg|^p$$

$$=: \sum_{Q\in\mathscr{D},\varepsilon\not\equiv 1}\sum_{s=1}^{2}\left|\left\langle w^{\frac12}[b,R_j]w^{-\frac12}G^s_{\hat Q}, H^s_Q\right\rangle\right|^p,$$

where

$$G^s_{\hat Q}(y) := \frac{w^{\frac12}(y)\chi_{F^{\hat Q}_s}(x)}{w(\hat Q)^{\frac12}} \quad\text{and}\quad H^s_Q(x) := \frac{w^{-\frac12}(x)\chi_{Q\cap E^Q_s}(x)}{(w^{-1}(Q))^{\frac12}}.$$

Applying Lemma 2.2, there is a reverse Hölder exponent $\sigma_w > 0$, so

$$\|G^s_{\hat Q}\|_{L^{2(\sigma_w+1)}} \lesssim \frac{1}{w(\hat Q)^{\frac12}}\left(\int_{\hat Q} w^{(\sigma_w+1)}(x)\,dx\right)^{\frac{1}{2(\sigma_w+1)}} \lesssim |\hat Q|^{\frac{1}{2(\sigma_w+1)}-\frac12}.$$

Similarly, $\|H^s_Q\|_{L^{2(\sigma_w+1)}} \lesssim |Q|^{\frac{1}{2(\sigma_w+1)}-\frac12}$. Then, $\{G^s_{\hat Q}\}_{\hat Q\in\mathscr{D}}$ and $\{H^s_Q\}_{Q\in\mathscr{D}}$ are NWO sequences for $L^2(\mathbb{R}^n)$. It follows from Lemma 2.6 and Lemma 2.5 that

$$\|b\|_{B^p_d(\mathbb{R}^n)} \lesssim \|w^{\frac12}[b,R_j]w^{-\frac12}\|_{S^p(L^2(\mathbb{R}^n))} \approx \|[b,R_j]\|_{S^p(L^2(\mathbb{R}^n,w))}.$$

The proof of Proposition 3.1 is complete.

3.2. Proof of Proposition 3.2. In [3], the authors have obtained that for $w\in A_2(\mathbb{R}^n)$, for every $b\in B^{p,p}_{n/p}(\mathbb{R}^n)\subset \mathrm{VMO}(\mathbb{R}^n)$, $[b,R_j]$ is compact form $L^2(\mathbb{R}^n,w)$ to $L^2(\mathbb{R}^n,w)$. On the other hand, Petermichl, Treil and Volberg have shown that Riesz transforms are averages of dyadic shifts as in [28] (see also [27]). For a choice of dyadic system $\mathscr{D}$ with Haar basis $\{h^\varepsilon_Q\}$, let $\sigma\colon \mathscr{D}\to\mathscr{D}$ with $|\sigma(Q)| = 2^{-n}|Q|$, for all $Q\in\mathscr{D}$. Use the same notation for a map $\sigma\colon \{0,1\}^n-\{1\}^n \to \{\{0,1\}^n-\{1\}^n\}\cup\{0\}$, and so if $\sigma(\varepsilon)=0$ then $h^{\sigma(\varepsilon)} := 0$. In [28], and as utilized in [21], the dyadic shift operator III is given by

$$\mathrm{III} f(x) := \sum_{Q\in\mathscr{D},\varepsilon\not\equiv 1}\langle f, h^\varepsilon_Q\rangle h^{\sigma(\varepsilon)}_{\sigma(Q)}(x). \tag{3.3}$$

It is clear that

$$\mathrm{III} h^\varepsilon_Q = h^{\sigma(\varepsilon)}_{\sigma(Q)}.$$

We further have $\|\mathrm{III}\|_{L^2(\mathbb{R}^n,w)\to L^2(\mathbb{R}^n,w)} \lesssim 1$ and the Riesz transforms are in the convex hull of the operators III. Therefore, we only need to prove that

$$\|[b,\mathrm{III}]\|_{S^p(L^2(\mathbb{R}^n,w))} \lesssim \|b\|_{B^p_d(\mathbb{R}^n)}.$$

As proved in [15], $[b, \mathrm{III}]$ can be decomposed into a composition of the shift III and paraproduct operators as follows:

$$(\Pi_b^{\mathscr{D}} + \Pi_b^{*\mathscr{D}} + \Gamma_b^{\mathscr{D}})(\mathrm{III}f) - \mathrm{III}(\Pi_b^{\mathscr{D}} + \Pi_b^{*\mathscr{D}} + \Gamma_b^{\mathscr{D}})f + \Pi_{\mathrm{III}f}^{\mathscr{D}} b - \mathrm{III}(\Pi_f^{\mathscr{D}} b),$$

where

$$\Pi_b^{\mathscr{D}} f = \sum_{Q\in\mathscr{D},\varepsilon\not\equiv 1} \langle b, h_Q^\varepsilon\rangle \langle f\rangle_Q h_Q^\varepsilon, \quad \Pi_b^{*\mathscr{D}} f = \sum_{Q\in\mathscr{D},\varepsilon\not\equiv 1} \langle b, h_Q^\varepsilon\rangle \langle f, h_Q^\varepsilon\rangle \frac{\chi_Q}{|Q|}$$

and

$$\Gamma_b^{\mathscr{D}} f = \sum_{Q\in\mathscr{D}} \sum_{\substack{\varepsilon,\eta\not\equiv 1\\ \varepsilon\not\equiv\eta}} \langle b, h_Q^\varepsilon\rangle \langle f, h_Q^\eta\rangle h_Q^\varepsilon h_Q^\eta$$

are the paraproduct operators with symbol b. Then we have

$$\begin{aligned}\|[b, \mathrm{III}]\|_{S^p(L^2(\mathbb{R}^n,w))} \le 2\|\Pi_b^{\mathscr{D}}\|_{S^p(L^2(\mathbb{R}^n,w))} \|\mathrm{III}\|_{L^2(\mathbb{R}^n,w)\to L^2(\mathbb{R}^n,w)} \\ + 2\|\Pi_b^{*\mathscr{D}}\|_{S^p(L^2(\mathbb{R}^n,w))} \|\mathrm{III}\|_{L^2(\mathbb{R}^n,w)\to L^2(\mathbb{R}^n,w)} \\ + 2\|\Gamma_b^{\mathscr{D}}\|_{S^p(L^2(\mathbb{R}^n,w))} \|\mathrm{III}\|_{L^2(\mathbb{R}^n,w)\to L^2(\mathbb{R}^n,w)} + \|\Pi_{\mathrm{III}f}^{\mathscr{D}} b - \mathrm{III}(\Pi_f^{\mathscr{D}} b)\|_{S^p(L^2(\mathbb{R}^n,w))}.\end{aligned}$$

Thus, in order to show that Proposition 3.2 holds, using Lemma 2.6, we need to obtain the following two lemmas.

LEMMA 3.4. *Suppose that $w \in A_2(\mathbb{R}^n)$, and $b \in \mathrm{VMO}(\mathbb{R}^n)$ for $n < p < \infty$. Then we have $w^{\frac12}\Pi_b^{\mathscr{D}} w^{-\frac12}$, $w^{\frac12}\Pi_b^{*\mathscr{D}} w^{-\frac12}$, and $w^{\frac12}\Gamma_b^{\mathscr{D}} w^{-\frac12}$ belong to $S^p(L^2(\mathbb{R}^n))$ respectively, if and only if $b \in B_d^p(\mathbb{R}^n, \mathscr{D})$. Moreover,*

$$\|w^{\frac12}\Pi_b^{\mathscr{D}} w^{-\frac12}\|_{S^p(L^2(\mathbb{R}^n))} \approx \|b\|_{B_d^p(\mathbb{R}^n,\mathscr{D})}; \tag{3.4}$$

$$\|w^{\frac12}\Pi_b^{*\mathscr{D}} w^{-\frac12}\|_{S^p(L^2(\mathbb{R}^n))} \approx \|b\|_{B_d^p(\mathbb{R}^n,\mathscr{D})}; \tag{3.5}$$

$$\|w^{\frac12}\Gamma_b^{\mathscr{D}} w^{-\frac12}\|_{S^p(L^2(\mathbb{R}^n))} \approx \|b\|_{B_d^p(\mathbb{R}^n,\mathscr{D})}. \tag{3.6}$$

In particular, when $b \in B_{n/p}^{p,p}(\mathbb{R}^n))$, then we know all these operators are bounded with norm at most $\|b\|_{B_{n/p}^{p,p}(\mathbb{R}^n)}$.

PROOF. The last statement follows from Lemma 2.8 which gives the equivalence between the continuous and dyadic norms, and so in particular that the continuous norm controls the dyadic one.

Lacey and the last two authors in [20] proved (3.4) and (3.5) in one dimension. The proof generalizes to n dimensions by direct modifications we do not include, but are easy exercises left for the reader.

It only remains to prove (3.6). We turn first to sufficiency.

Sufficiency. Suppose $b \in B_d^p(\mathbb{R}^n, \mathscr{D})$. By the definition of $\Gamma_b^{\mathscr{D}} f$, we have

$$\begin{aligned}
\big(w^{\frac12}\Gamma_b^{\mathscr{D}} w^{-\frac12}\big)(f)(x) &= \sum_{Q\in\mathscr{D}} \sum_{\substack{\varepsilon,\eta\not\equiv 1\\ \varepsilon\not\equiv\eta}} \langle b, h_Q^\varepsilon\rangle \langle w^{-\frac12} f, h_Q^\eta\rangle h_Q^\varepsilon(x) h_Q^\eta(x) w^{\frac12}(x)\\
&= \sum_{Q\in\mathscr{D}} \sum_{\substack{\varepsilon,\eta\not\equiv 1\\ \varepsilon\not\equiv\eta}} \frac{\langle b, h_Q^\varepsilon\rangle w(Q)^{\frac12} (w^{-1}(Q))^{\frac12}}{|Q|^{\frac32}} \cdot \frac{h_Q^\varepsilon(x) h_Q^\eta(x) w^{\frac12}(x)|Q|}{w(Q)^{\frac12}}\\
&\qquad\times \int_{\mathbb{R}^n} \frac{w^{-\frac12}(y) h_Q^\eta(y)|Q|^{\frac12}}{(w^{-1}(Q))^{\frac12}} f(y)\, dy\\
&=: \sum_{Q\in\mathscr{D}} \sum_{\substack{\varepsilon,\eta\not\equiv 1\\ \varepsilon\not\equiv\eta}} B(Q)\cdot G_Q(x) \int_{\mathbb{R}^n} f(y) H_Q(y)\, dy.
\end{aligned}$$

As in the proof of Proposition 3.1, applying Lemma 2.2 leads to the fact that $\{G_Q\}_{Q\in\mathscr{D}}$ and $\{H_Q\}_{Q\in\mathscr{D}}$ are NWO sequences for $L^2(\mathbb{R}^n)$, and by Lemma 2.6 and (2.3), that

$$\begin{aligned}
&\|\Gamma_b^{\mathscr{D}}\|^p_{S^p(L^2(\mathbb{R}^n,w))} \approx \|w^{\frac12}\Gamma_b^{\mathscr{D}} w^{-\frac12}\|^p_{S^p(L^2(\mathbb{R}^n))}\\
&\lesssim \sum_{\substack{Q\in\mathscr{D}\\ \varepsilon\not\equiv 1}} |B(Q)|^p = \sum_{\substack{Q\in\mathscr{D}\\ \varepsilon\not\equiv 1}} \frac{|\langle b, h_Q^\varepsilon\rangle|^p w(Q)^{\frac p2} (w^{-1}(Q))^{\frac p2}}{|Q|^{\frac{3p}{2}}} = \|b\|^p_{B_d^p(\mathbb{R}^n,\mathscr{D})}.
\end{aligned}$$

We next turn to the necessity.

Necessity. For any dyadic cube Q, we have $\langle b, h_Q^\varepsilon\rangle = \langle \Gamma_b^{\mathscr{D}}(h_Q^\eta), h_Q^\varepsilon h_Q^\eta |Q|\rangle$. Therefore,

$$\begin{aligned}
\sum_{\substack{Q\in\mathscr{D}\\ \varepsilon\not\equiv 1}} \left(\frac{|\langle b, h_Q^\varepsilon\rangle||Q|^{\frac12}}{w(Q)^{\frac12}(w^{-1}(Q))^{\frac12}}\right)^p &= \sum_{Q\in\mathscr{D}} \sum_{\substack{\varepsilon,\eta\not\equiv 1\\ \varepsilon\not\equiv\eta}} \left(\frac{|\langle \Gamma_b^{\mathscr{D}}(h_Q^\eta), h_Q^\varepsilon h_Q^\eta|Q|\rangle||Q|^{\frac12}}{w(Q)^{\frac12}(w^{-1}(Q))^{\frac12}}\right)^p\\
&= \sum_{Q\in\mathscr{D}} \sum_{\substack{\varepsilon,\eta\not\equiv 1\\ \varepsilon\not\equiv\eta}} \left(\frac{|\langle w^{\frac12}\Gamma_b^{\mathscr{D}} w^{-\frac12}(w^{\frac12} h_Q^\eta), w^{-\frac12} h_Q^\varepsilon h_Q^\eta |Q|\rangle||Q|^{\frac12}}{w(Q)^{\frac12}(w^{-1}(Q))^{\frac12}}\right)^p\\
&= \sum_{Q\in\mathscr{D}} \sum_{\substack{\varepsilon,\eta\not\equiv 1\\ \varepsilon\not\equiv\eta}} \left|\left\langle w^{\frac12}\Gamma_b^{\mathscr{D}} w^{-\frac12}\left(\frac{w^{\frac12}\sqrt{|Q|}h_Q^\varepsilon}{w(Q)^{\frac12}}\right), \frac{w^{-\frac12} h_Q^\varepsilon h_Q^\eta|Q|}{(w^{-1}(Q))^{\frac12}}\right\rangle\right|^p
\end{aligned}$$

$$=: \sum_{Q\in\mathscr{D}} \sum_{\substack{\varepsilon,\eta\not\equiv 1\\ \varepsilon\not\equiv\eta}} \left|\left\langle w^{\frac{1}{2}}\Gamma_b^{\mathscr{D}} w^{-\frac{1}{2}}(G'_Q), H'_Q\right\rangle\right|^p,$$

where

$$G'_Q := \frac{w^{\frac{1}{2}}\sqrt{|Q|}h_Q^{\varepsilon}}{w(Q)^{\frac{1}{2}}} \quad \text{and} \quad H'_Q := \frac{w^{-\frac{1}{2}}h_Q^{\varepsilon}h_Q^{\eta}|Q|}{(w^{-1}(Q))^{\frac{1}{2}}}.$$

By Lemma 2.2, and a computation similar inside the proof of Proposition 3.1, we can obtain that the above two collections of functions are NWO sequences. Thus, we establish by Lemma 2.5 that

$$\|b\|^p_{B_d^p(\mathbb{R}^n,\mathscr{D})} = \sum_{\substack{Q\in\mathscr{D}\\ \varepsilon\not\equiv 1}} \left(\frac{|\langle b, h_Q^{\varepsilon}\rangle||Q|^{\frac{1}{2}}}{w(Q)^{\frac{1}{2}}(w^{-1}(Q))^{\frac{1}{2}}}\right)^p$$

$$\lesssim \|w^{\frac{1}{2}}\Gamma_b^{\mathscr{D}} w^{-\frac{1}{2}}\|^p_{S^p(L^2(\mathbb{R}^n))} \approx \|\Gamma_b^{\mathscr{D}}\|^p_{S^p(L^2(\mathbb{R}^n,w))}.$$

The proof is complete. □

LEMMA 3.5. *For $n<p<\infty$, suppose that $w\in A_2(\mathbb{R}^n)$, and $b\in B_d^p(\mathbb{R}^n,\mathscr{D})$. Let $\mathfrak{R}^{\mathscr{D}} f := \Pi^{\mathscr{D}}_{\mathrm{III} f} b - \mathrm{III}(\Pi^{\mathscr{D}}_f b)$, then we have*

$$\|w^{\frac{1}{2}}\mathfrak{R}^{\mathscr{D}} w^{-\frac{1}{2}}\|_{S^p(L^2(\mathbb{R}^n))} \lesssim \|b\|_{B_d^p(\mathbb{R}^n,\mathscr{D})}. \tag{3.7}$$

PROOF. A direct computation gives

$$\mathfrak{R}^{\mathscr{D}} f(x) := \Pi^{\mathscr{D}}_{\mathrm{III} f} b - \mathrm{III}(\Pi^{\mathscr{D}}_f b)$$

$$= \sum_{P\in\mathscr{D},\eta\not\equiv 1} \langle \mathrm{III} f, h_P^{\eta}\rangle\langle b\rangle_P h_P^{\eta}(x) - \sum_{Q\in\mathscr{D},\varepsilon\not\equiv 1} \langle \Pi_f^{\mathscr{D}} b, h_Q^{\varepsilon}\rangle h_{\sigma(Q)}^{\sigma(\varepsilon)}(x)$$

$$= \sum_{Q\in\mathscr{D},\varepsilon\not\equiv 1} \langle f, h_Q^{\varepsilon}\rangle(\langle b\rangle_{\sigma(Q)} - \langle b\rangle_Q) h_{\sigma(Q)}^{\sigma(\varepsilon)}(x)$$

$$= \sum_{\substack{Q\in\mathscr{D}\\ \varepsilon,\eta\not\equiv 1}} \langle f, h_Q^{\varepsilon}\rangle\langle b, h_Q^{\eta}\rangle h_Q^{\eta}(\sigma(Q)) h_{\sigma(Q)}^{\sigma(\varepsilon)}(x),$$

where the last equality is from [15, (2.2)]. Therefore, we have that

$$w^{\frac{1}{2}}(x)\mathfrak{R}^{\mathscr{D}}(w^{-\frac{1}{2}} f)(x)$$

$$= \sum_{\substack{Q\in\mathscr{D}\\ \varepsilon,\eta\not\equiv 1}} \int_{\mathbb{R}^n} w^{-\frac{1}{2}}(y) f(y) h_Q^{\varepsilon}(y)\, dy w^{\frac{1}{2}}(x)\langle b, h_Q^{\eta}\rangle h_Q^{\eta}(\sigma(Q)) h_{\sigma(Q)}^{\sigma(\varepsilon)}(x)$$

$$= \sum_{\substack{Q\in\mathscr{D}\\ \varepsilon,\eta\not\equiv 1}} \frac{\langle b, h_Q^\eta\rangle h_Q^\eta(\sigma(Q)) w(Q)^{\frac12}(w^{-1}(Q))^{\frac12}}{|Q|} \cdot \frac{h_{\sigma(Q)}^{\sigma(\varepsilon)}(x) w^{\frac12}(x)|Q|^{\frac12}}{w(Q)^{\frac12}}$$

$$\times \int_{\mathbb{R}^n} \frac{w^{-\frac12}(y) h_Q^\varepsilon(y)|Q|^{\frac12}}{(w^{-1}(Q))^{\frac12}} f(y)\,dy =: \sum_{\substack{Q\in\mathscr{D}\\ \varepsilon,\eta\not\equiv 1}} B^1(Q)\cdot G_Q^1(x)\int_{\mathbb{R}^n} f(y) H_Q^1(y)\,dy,$$

where

$$B^1(Q) := \frac{\langle b, h_Q^\eta\rangle h_Q^\eta(\sigma(Q)) w(Q)^{\frac12}(w^{-1}(Q))^{\frac12}}{|Q|},$$

$$G_Q^1(x) := \frac{h_{\sigma(Q)}^{\sigma(\varepsilon)}(x) w^{\frac12}(x)|Q|^{\frac12}}{w(Q)^{\frac12}} \quad \text{and} \quad H_Q^1(y) := \frac{w^{-\frac12}(y) h_Q^\varepsilon(y)|Q|^{\frac12}}{(w^{-1}(Q))^{\frac12}}.$$

By Lemma 2.2, and a repeat of a computation inside the proof of Proposition 3.1, we know that $\{G_Q^1\}_{Q\in\mathscr{D}}$ and $\{H_Q^1\}_{Q\in\mathscr{D}}$ are NWO sequences for $L^2(\mathbb{R}^n)$. Therefore, by (2.3), and $|h_Q^\eta(\sigma(Q))| \approx |Q|^{-\frac12}$ we get

$$\|w^{\frac12}\mathfrak{R}^{\mathscr{D}} w^{-\frac12}\|_{S^p(L^2(\mathbb{R}^n))} \lesssim \|B^1(Q)\|_{\ell^p}$$

$$\approx \Bigg(\sum_{Q\in\mathscr{D},\ \eta\not\equiv 1}\bigg(\frac{|\langle b, h_Q^\eta\rangle||Q|^{\frac12}}{w(Q)^{\frac12}(w^{-1}(Q))^{\frac12}}\bigg)^p\Bigg)^{1/p} = \|b\|_{B_d^p(\mathbb{R}^n,\mathscr{D})}.$$

The proof is complete. □

4. Proof of Theorem 1.1: the case of $0 < p \le n$

In this section, we prove (2) in Theorem 1.1. That is, for $0 < p \le n$, the commutator $[b, R_j] \in S^p(L^2(\mathbb{R}^n, w))$ if and only if b is a constant. Here $n \ge 1$ (and when $n = 1$ we mean of course the Hilbert transform as opposed to the Riesz transforms).

The sufficient condition is obvious, since $[b, R_j] = 0$ when b is a constant. Thus, it suffices to show the necessity. By the inclusion $S^p(L^2(\mathbb{R}^n, w)) \subset S^q(L^2(\mathbb{R}^n, w))$ for $p < q$, then the proof of (2) in Theorem 1.1 can be proved on the basis of the following property.

PROPOSITION 4.1. *Suppose $w \in A_2(\mathbb{R}^n)$, and $b \in \mathrm{VMO}(\mathbb{R}^n)$ with $[b, R_j] \in S^n(L^2(\mathbb{R}^n, w))$, then b is a constant.*

PROOF. In order to obtain Proposition 4.1, we recall the following standard notation on martingale differences and conditional expectation. Let $\mathscr{D}_k$

be the collection of dyadic cubes at level k as in Section 2.2. Next, we choose h_Q among these 2^n-1 Haar functions such that

$$\Big|\int_Q b(x)h_Q(x)\,dx\Big| = \max_{\varepsilon\not\equiv 1}\Big|\int_Q b(x)h_Q^\varepsilon(x)\,dx\Big|.$$

Note that $Q\in\mathscr{D}_k$, the function

$$(E_{k+1}(b)(x)-E_k(b)(x))\chi_Q(x)=\sum_{\varepsilon\not\equiv 1}\langle b,h_Q^\varepsilon\rangle h_Q^\varepsilon(x).$$

So we have

$$(4.1)\quad \Big(\frac{1}{|Q|}\int_Q|E_{k+1}(b)(x)-E_k(b)(x)|^n\,dx\Big)^{1/n}\le C|Q|^{-1/n}\Big\|\sum_{\varepsilon\not\equiv 1}\langle b,h_Q^\varepsilon\rangle h_Q^\varepsilon\Big\|_n$$
$$\le C|Q|^{-1/n}\sum_{\varepsilon\not\equiv 1}|\langle b,h_Q^\varepsilon\rangle|\big\|h_Q^\varepsilon\big\|_n\le C|Q|^{-1/2}\Big|\int_Q b(x)h_Q(x)\,dx\Big|,$$

where C is a constant depending only on n. Then we obtain that

$$\sum_k 2^{nk}\big\|E_{k+1}(b)-E_k(b)\big\|_n^n=\sum_k\sum_{Q\in\mathscr{D}_k}\frac{1}{|Q|}\int_Q\big|E_{k+1}(b)(x)-E_k(b)(x)\big|^n\,dx$$
$$\le C\sum_k\sum_{Q\in\mathscr{D}_k}|Q|^{-n/2}\Big|\int_Q b(g)h_Q(x)\,dx\Big|^n.$$

Following the proof of the estimate of (2.6) in the proof of Lemma 2.8, we have that

$$\sum_k 2^{nk}\big\|E_{k+1}(b)-E_k(b)\big\|_n^n\lesssim\|b\|_{B_d^p(\mathbb{R}^n,\mathscr{D})}\le C\big\|[b,R_j]\big\|^n_{S^n(L^2(\mathbb{R}^n,w))},$$

where the last inequality follows from Proposition 3.1. This, together with Hölder's inequality, further implies that for a fixed positive integer K_0,

$$\Big\|\Big\{\frac{1}{|Q|}\int_Q\big|E_{k+K_0}(b)(x)-E_k(b)(x)\big|\,dx\Big\}_{k\in\mathbb{Z},Q\in\mathscr{D}_k}\Big\|_{l^n}$$
$$\le\Big(\sum_{k\in\mathbb{Z},Q\in\mathscr{D}_k}\Big|\frac{1}{|Q|}\int_Q\big|E_{k+K_0}(b)(x)-E_k(b)(x)\big|\,dx\Big|^n\Big)^{1/n}$$
$$\le\Big(\sum_{k\in\mathbb{Z},Q\in\mathscr{D}_k}\frac{1}{|Q|}\int_Q\big|E_{k+K_0}(b)(x)-E_k(b)(x)\big|^n\,dx\Big)^{1/n}$$

$$\leq \sum_{j=0}^{K_0-1} \Big(\sum_{k\in\mathbb{Z}, Q\in\mathscr{D}_k} \frac{1}{|Q|} \int_Q \big|E_{k+j+1}(b)(x) - E_{k+j}(b)(x)\big|^n \, dx \Big)^{1/n}$$

$$= \sum_{j=0}^{K_0-1} \Big(\sum_{k\in\mathbb{Z}} 2^{nk} \|E_{k+j+1}(b) - E_{k+j}(b)\|_n^n \Big)^{1/n} \lesssim \big\|[b,R_j]\big\|_{S^n(L^2(\mathbb{R}^n,w))},$$

where the implicit constant depends on K_0. This further implies that

$$\Big\| \Big\{ \frac{1}{|Q|}\int_Q \frac{1}{|Q|}\int_Q \big|E_{k+K_0}(b)(x) - E_{k+K_0}(b)(y)\big| \, dy \, dx \Big\}_{k\in\mathbb{Z}, Q\in\mathscr{D}_k} \Big\|_{l^n} \lesssim \big\|[b,R_j]\big\|_{S^n(L^2(\mathbb{R}^n,w))} \tag{4.2}$$

since $E_k(b)(x) = E_k(b)(y)$ for every $Q\in\mathscr{D}_k$ and for every $x,y\in Q$.

Suppose $b\in C^\infty(\mathbb{R}^n)$ with $\|[b,R_j]\|_{S^n(L^2(\mathbb{R}^n,w))} < +\infty$. If b is not constant, then there exists a point $x_0\in\mathbb{R}^n$ such that $\nabla b(x_0)\neq 0$. By applying [9, Lemma 5.3] with $\mathbb{R}^n$, there is $\varepsilon>0$ and $N>0$ such that if $k>N$, then for any dyadic cube $Q\in\mathscr{D}_k$ with $|C_Q - x_0| < \varepsilon$, and for $\tilde{Q},\hat{Q}\in\mathscr{D}_{k+K_0}$ with $\tilde{Q}\subset Q$, $\hat{Q}\subset Q$, and $\operatorname{dist}(\tilde{Q},\hat{Q})\approx \ell(\tilde{Q})$,

$$|\langle b\rangle_{\tilde{Q}} - \langle b\rangle_{\hat{Q}}| \geq C\ell(Q)|\nabla b(x_0)|.$$

Here C_Q represents the center of Q.

Noting that for $k>N$, the number of $Q\in\mathscr{D}_k$ and $|C_Q - x_0|<\varepsilon$ is at least 2^{kn}. Thus, we obtain

$$\Big\| \Big\{ \frac{1}{|Q|}\int_Q \frac{1}{|Q|}\int_Q \big|E_{k+K_0}(b)(x) - E_{k+K_0}(b)(y)\big| \, dy \, dx \Big\}_{k\in\mathbb{Z}, Q\in\mathscr{D}_k} \Big\|_{l^n}$$

$$\geq \Big\| \Big\{ \frac{1}{|Q|}\int_Q \frac{1}{|Q|}\int_Q \big|E_{k+K_0}(b)(x) - E_{k+K_0}(b)(y)\big| \, dy \, dx \Big\}_{\substack{k\in\mathbb{Z},\ k>N \\ Q\in\mathscr{D}_k,\ |C_Q-x_0|<\varepsilon}} \Big\|_{l^n}$$

$$\geq \Big\| \Big\{ \big|\langle b\rangle_{\tilde{Q}} - \langle b\rangle_{\hat{Q}}\big| \Big\}_{\substack{k\in\mathbb{Z},\ k>N \\ Q\in\mathscr{D}_k,\ |C_Q-x_0|<\varepsilon}} \Big\|_{l^n}$$

$$\geq C \Big(\sum_{k>N} 2^{kn} \big(2^{-k}|\nabla b(x_0)|\big)^n \Big)^{1/n} = \infty.$$

This contradicts (4.2).

Suppose $b\in \mathrm{VMO}(\mathbb{R}^n)$ with $\|[b,R_j]\|_{S^n(L^2(\mathbb{R}^n,w))} < +\infty$. Let $\psi\in C_0^\infty(\mathbb{R}^n)$, non-negative, and $\int\psi(x)\,dx = 1$. Define $\psi_\varepsilon(x) = \frac{1}{\varepsilon^n}\psi(\frac{x}{\varepsilon})$. Then $b_\varepsilon(x) =$

$b * \psi_\varepsilon(x)$ is in $C^\infty(\mathbb{R}^n)$. We note that for ε small enough,

$$\Big\|\Big\{\frac{1}{|Q|}\int_Q \frac{1}{|Q|}\int_Q \big|E_{k+K_0}(b_\varepsilon)(x)-E_{k+K_0}(b_\varepsilon)(y)\big|\,dy\,dx\Big\}_{k\in\mathbb{Z},Q\in\mathscr{D}_k}\Big\|_{l^n}$$

$$\lesssim \sup_{h\in B(0,1)}\Big\|\Big\{\frac{1}{|Q|}\int_Q \frac{1}{|Q|}\int_Q \big|E_{k+K_0}(\tau_h b)(x)-E_{k+K_0}(\tau_h b)(y)\big|\,dy\,dx\Big\}_{\substack{k\in\mathbb{Z}\\ Q\in\mathscr{D}_k}}\Big\|_{l^n},$$

where $\tau_h b(x) = b(x-h)$ for h in the unit ball $B(0,1)$ in $\mathbb{R}^n$. For every fixed $h\in B(0,1)$, by repeating the arguments, especially Proposition 3.1, for $\tau_h b(x)$, and by translating the dyadic system according to h and the translation invariance of the kernel of the Riesz transform R_j, we obtain that

$$\Big\|\Big\{\frac{1}{|Q|}\int_Q \frac{1}{|Q|}\int_Q \big|E_{k+K_0}(\tau_h b)(x)-E_{k+K_0}(\tau_h b)(y)\big|\,dy\,dx\Big\}_{k\in\mathbb{Z},Q\in\mathscr{D}_k}\Big\|_{l^n}$$
$$\lesssim \|[b,R_j]\|_{S^n(L^2(\mathbb{R}^n,w))},$$

where the implicit constant is independent of h. This yields that for ε small enough,

$$\Big\|\Big\{\frac{1}{|Q|}\int_Q \frac{1}{|Q|}\int_Q \big|E_{k+K_0}(b_\varepsilon)(x)-E_{k+K_0}(b_\varepsilon)(y)\big|\,dy\,dx\Big\}_{k\in\mathbb{Z},Q\in\mathscr{D}_k}\Big\|_{l^n}$$
$$\lesssim \|[b,R_j]\|_{S^n(L^2(\mathbb{R}^n,w))}.$$

Thus, b_ε is a constant on $\mathbb{R}^n$. Since $b_\varepsilon \to b$ as $\varepsilon \to 0^+$, we obtain that b is a constant. Therefore, the proposition holds. □

5. Proof of Theorem 1.2: $p = n$

5.1. Proof of the sufficient condition. In this subsection, we assume that $b \in \dot{W}^{1,n}(\mathbb{R}^n)$, then prove

$$[b,R_j] \in S^{n,\infty}\big(L^2(\mathbb{R}^n,w)\big).$$

By Lemma 2.7, we just need to show that $\|w^{\frac12}[b,R_j](w^{-\frac12})\|_{S^{n,\infty}(L^2(\mathbb{R}^n))} \lesssim \|b\|_{\dot{W}^{1,n}(\mathbb{R}^n)}$.

Let $\Lambda = \{(x,y)\in\mathbb{R}^n\times\mathbb{R}^n : x=y\}$, and $\Omega = \{(x,y)\in\mathbb{R}^n\times\mathbb{R}^n\backslash\Lambda : x\neq y\}$. Let $\mathscr{P}$ be a dyadic Whitney decomposition of the open set Ω, that is $\bigcup_{P\in\mathscr{P}} P = \Omega$. Therefore, we write $K_j(x-y) = \Sigma_{P\in\mathscr{P}} K_j(x-y)\chi_P(x,y)$, and P can be the cubes $P_1\times P_2$, where $P_1, P_2 \in \mathscr{D}$, have the same side length and that distance between them must be comparable to this sidelength. Thus, for each dyadic cube $P_1 \in \mathscr{D}$, P_2 is related to P_1 and at most M of the cubes P_2 such that $P_1 \times P_2 \in \mathscr{P}$.

Therefore, for $s = 1, 2, \ldots, M$ and $Q = P_1$, there is $R_{Q,s}$ such that $Q \times R_{Q,s} \in \mathscr{P}$ and we can reorganize the sum

$$K_j(x-y) = \sum_{P \in \mathscr{P}} K_j(x-y)\chi_P(x,y) = \sum_{Q \in \mathscr{D}} \sum_{s=1}^{M} K_j(x-y)\chi_{(Q \times R_{Q,s})}(x,y),$$

where $|Q| = |R_{Q,s}|$ and $\operatorname{dist}(Q, R_{Q,s}) \approx |Q|$.

Next, decomposing in a Fourier series on $Q \times R_{Q,s}$ we can write

$$K_j(x-y)\chi_{(Q \times R_{Q,s})}(x,y) = \sum_{\vec{l} \in \mathbb{Z}^{2n}} c^j_{\vec{l},Q} e^{2\pi i \vec{l'} \cdot \tilde{x}} e^{2\pi i \vec{l''} \cdot \tilde{y}} \chi_Q(x) \cdot \chi_{Q_{R,s}}(y),$$

where $x_i = C_Q^{(i)} + \ell(Q)\tilde{x}_i$, $y_i = C_{R_{Q,s}}^{(i)} + \ell(R_{Q,s})\tilde{y}_i$, $i = 1, 2, \ldots, n$, $\vec{l} = (\vec{l'}, \vec{l''})$ where $\vec{l'} = (l_1, l_2, \ldots, l_n)$, $\vec{l''} = (l_{n+1}, l_{n+2}, \ldots, l_{2n})$, and $c^j_{\vec{l},Q}$ is the Fourier coefficient

$$c^j_{\vec{l},Q} := \int_{R_{Q,s}} \int_Q K_j(x-y)\chi_{(Q \times R_{Q,s})}(x,y) e^{-2\pi i \vec{l'} \cdot \tilde{x}} e^{-2\pi i \vec{l''} \cdot \tilde{y}} \, dx \, dy \, \frac{1}{|Q|} \frac{1}{|R_{Q,s}|}.$$

For the multi-index α, $\gamma \in \mathbb{Z}^n_+$, using the relation $\hat{f}(\vec{l}) = \frac{1}{(2\pi i \vec{l})^{(\alpha,\gamma)}} (\widehat{\partial^{(\alpha,\gamma)} f})(\vec{l})$, and the size condition of $K_j(x-y)$,

$$|\partial_x^\alpha \partial_y^\gamma K_j(x-y)| \le C(\alpha,\gamma) \frac{1}{|x-y|^{n+|\alpha|+|\gamma|}},$$

yield that

$$\begin{aligned} |c^j_{\vec{l},Q}| &\lesssim \frac{1}{(1+|\vec{l}|)^{|\alpha|+|\gamma|}} \ell(Q)^{|\alpha|} \ell(R_{Q,s})^{|\gamma|} \\ &\quad \times \int_{R_{Q,s}} \int_Q |\partial_x^\alpha \partial_y^\gamma K_j(x-y)| \, dx \, dy \, \frac{1}{|Q|} \frac{1}{|R_{Q,s}|} \\ &\lesssim \frac{1}{(1+|\vec{l}|)^{|\alpha|+|\gamma|}} \ell(Q)^{|\alpha|} \ell(R_{Q,s})^{|\gamma|} \int_{R_{Q,s}} \int_Q \frac{1}{|x-y|^{n+|\alpha|+|\gamma|}} \, dx \, dy \, \frac{1}{|Q|} \frac{1}{|R_{Q,s}|} \\ &\lesssim \frac{1}{|Q|} \frac{1}{(1+|\vec{l}|)^{|\alpha|+|\gamma|}}. \end{aligned}$$

Let $\lambda^j_{\vec{l},Q} = |Q|^{\frac{1}{2}} |R_{Q,s}|^{\frac{1}{2}} c^j_{\vec{l},Q}$, then

$$|\lambda^j_{\vec{l},Q}| \lesssim \frac{1}{(1+|\vec{l}|)^{|\alpha|+|\gamma|}},$$

and

$$K_j(x-y)\chi_{(Q\times R_{Q,s})}(x,y) = \sum_{\vec{l}\in\mathbb{Z}^{2n}} \lambda^j_{\vec{l},Q} \frac{1}{|Q|^{1/2}} F_{\vec{l}',Q}(x) \frac{1}{|R_{Q,s}|^{1/2}} G_{\vec{l}'',R_{Q,s}}(y),$$

where $F_{\vec{l}',Q}(x) = e^{2\pi i\vec{l}'\cdot\tilde{x}}\chi_Q(x)$ and $G_{\vec{l}'',R_{Q,s}}(y) = e^{2\pi i\vec{l}''\cdot\tilde{y}}\chi_{R_{Q,s}}(y)$. Then, we get

$$K(x-y) = \sum_{Q\in\mathscr{D}}\sum_{s=1}^{M} K_j(x-y)\chi_{(Q\times R_{Q,s})}(x,y)$$

$$= \sum_{Q\in\mathscr{D}}\sum_{s=1}^{M}\sum_{\vec{l}\in\mathbb{Z}^{2n}} \lambda^j_{\vec{l},Q} \frac{1}{|Q|^{1/2}} F_{\vec{l}',Q}(x) \frac{1}{|R_{Q,s}|^{1/2}} G_{\vec{l}'',R_{Q,s}}(y).$$

Thus, the kernel of $w^{\frac{1}{2}}[b,R_j](w^{-\frac{1}{2}})$ can be represented as

$$K^w_b(x,y) = \sum_{Q\in\mathscr{D}}\sum_{s=1}^{M}\sum_{\vec{l}\in\mathbb{Z}^{2n}} (b(x)-b(y))\lambda^j_{\vec{l},Q} \frac{1}{|Q|^{1/2}} w^{\frac{1}{2}}(x)$$

$$\times F_{\vec{l}',Q}(x) \frac{1}{|R_{Q,s}|^{1/2}} G_{\vec{l}'',R_{Q,s}}(y) w^{-\frac{1}{2}}(y).$$

For each Q rewrite $b(x)-b(y)$ as $(b(x)-\langle b\rangle_{KQ}) + (\langle b\rangle_{KQ} - b(y))$ yielding

$$K^w_b(x,y) = C_{[w]_{A_2(\mathbb{R}^n)}} \sum_{Q\in\mathscr{D}}\sum_{s=1}^{M}\sum_{\vec{l}\in\mathbb{Z}^{2n}}\sum_{m=0}^{1} \lambda^j_{\vec{l},Q}$$

$$\times \frac{(b(x)-b_{KQ})^m w^{\frac{1}{2}}(x) F_{\vec{l}',Q}(x)}{[w(Q)]^{\frac{1}{2}}} \cdot \frac{(b_{KQ}-b(y))^{1-m} w^{-\frac{1}{2}}(y) G_{\vec{l}'',R_{Q,s}}(y)}{[w^{-1}(Q)]^{\frac{1}{2}}},$$

where $K>1$ is the constant chosen such that KQ contains $Q\cup R_{Q,s}$. We introduce the notation

$$\operatorname{osc}_r(b,Q) = \left[|KQ|^{-1}\int_{KQ} |b(u)-\langle b\rangle_{KQ}|^r du\right]^{1/r},$$

where $r > \frac{2(1+\sigma_w)}{\sigma_w}$ with σ_w the reverse Hölder exponent. Then

$$F_{\vec{l}',Q,m}(x) = (\operatorname{osc}_r(b,Q))^{-m} \frac{(b(x)-\langle b\rangle_{KQ})^m\, w^{\frac{1}{2}}(x) F_{\vec{l}',Q}(x)}{[w(Q)]^{\frac{1}{2}}}$$

and

$$G_{\vec{l}'',R_{Q,s},m}(y) = (\operatorname{osc}_r(b,Q))^{-(1-m)} \frac{(\langle b\rangle_{KQ} - b(y))^{1-m} w^{-\frac{1}{2}}(y) G_{\vec{l}'',R_{Q,s}}(y)}{[w^{-1}(Q)]^{\frac{1}{2}}}.$$

For $m = 0, 1$, let $\beta = \frac{2(1+\sigma_w)r}{r+2(1+\sigma_w)}$ and $p = \frac{r}{\beta}$. It is clear that $\beta > 2$, $p > 1$, and $\beta p' = 2(1+\sigma_w)$. By Hölder's inequality, Lemma 2.2, $\operatorname{supp}(F_{\vec{l}',Q}) \subset Q$, and $\|F_{\vec{l}',Q}\|_\infty \leqslant 1$ yields that

$$\begin{aligned}
&\big\|F_{\vec{l}',Q,m}\big\|_{L^\beta(\mathbb{R})} \\
&\le \bigg(\int_{\mathbb{R}^n} \bigg|(\operatorname{osc}_r(b,Q))^{-m} \frac{(b(x) - \langle b\rangle_{KQ})^m w^{\frac{1}{2}}(x) F_{\vec{l}',Q}(x)}{[w(Q)]^{\frac{1}{2}}}\bigg|^\beta dx\bigg)^{1/\beta} \\
&\le C\,[w(Q)]^{-\frac{1}{2}} (\operatorname{osc}_r(b,Q))^{-m} \bigg(\frac{1}{|KQ|}\int_{KQ} |b(x) - \langle b\rangle_{KQ}|^{\beta m} w^{\frac{\beta}{2}}(x)\,dx\bigg)^{1/\beta} |Q|^{\frac{1}{\beta}} \\
&\le C\,[w(Q)]^{-\frac{1}{2}} (\operatorname{osc}_r(b,Q))^{-m} \bigg(\frac{1}{|KQ|}\int_{KQ} |b(x) - \langle b\rangle_{KQ}|^{p\beta m}\,dx\bigg)^{\frac{1}{p\beta}} \\
&\quad \times \bigg(\frac{1}{|KQ|}\int_{KQ} w^{\frac{p'\beta}{2}}(x)\,dx\bigg)^{\frac{1}{p'\beta}} |Q|^{\frac{1}{\beta}} \\
&\le C\,[w(Q)]^{-\frac{1}{2}} (\operatorname{osc}_r(b,Q))^{-m} \bigg(\frac{1}{|KQ|}\int_{KQ} |b(x) - \langle b\rangle_{KQ}|^{r}\,dx\bigg)^{m/r} \\
&\quad \times \bigg(\frac{1}{|KQ|}\int_{KQ} w^{1+\sigma_w}(x)\,dx\bigg)^{1/2(1+\sigma_w)} |Q|^{\frac{1}{\beta}} \le C_{[w]_{A_2(\mathbb{R}^n)},K} |Q|^{\frac{1}{\beta}-\frac{1}{2}}.
\end{aligned}$$

We now consider $G_{\vec{l}'',R_{Q,s},m}$ and use a method similar to the proof above.

$$\begin{aligned}
&\big\|G_{\vec{l}'',R_{Q,s},m}\big\|_{L^\beta(\mathbb{R}^n)} \\
&\le \bigg(\int_{\mathbb{R}^n} \bigg|(\operatorname{osc}_r(b,Q))^{-(1-m)} \frac{(\langle b\rangle_{KQ} - b(y))^{1-m} w^{-\frac{1}{2}}(y) G_{\vec{l}'',R_{Q,s}}(y)}{[w^{-1}(Q)]^{\frac{1}{2}}}\bigg|^\beta dy\bigg)^{1/\beta} \\
&\le C\,[w^{-1}(Q)]^{-\frac{1}{2}} (\operatorname{osc}_r(b,Q))^{-(1-m)} \bigg(\int_{KQ} |b(x) - \langle b\rangle_{KQ}|^{\beta(1-m)} w^{-\frac{\beta}{2}}(y)\,dx\bigg)^{1/\beta} \\
&\le C\,[w^{-1}(Q)]^{-\frac{1}{2}} |Q|^{\frac{1}{\beta}} (\operatorname{osc}_r(b,Q))^{-(1-m)} \\
&\times \bigg(\frac{1}{|KQ|}\int_{KQ} |b(x) - \langle b\rangle_{KQ}|^{p(1-m)\beta}\,dx\bigg)^{\frac{1}{p\beta}} \bigg(\frac{1}{|KQ|}\int_{KQ} w^{-\frac{p'\beta}{2}}(y)\,dy\bigg)^{\frac{1}{p'\beta}}
\end{aligned}$$

$$\leq C\,[w^{-1}(Q)]^{-\frac{1}{2}}|Q|^{\frac{1}{\beta}}(\mathrm{osc}_r(b,Q))^{-(1-m)}$$

$$\times\left(\frac{1}{|KQ|}\int_{KQ}|b(x)-\langle b\rangle_{KQ}|^r\,dx\right)^{1-m/r}\left(\frac{1}{|KQ|}\int_{KQ}w^{-(1+\sigma_w)}(y)\,dy\right)^{\frac{1}{2(1+\sigma_w)}}$$

$$\leq C\,[w^{-1}(Q)]^{-\frac{1}{2}}|Q|^{\frac{1}{\beta}}\left(\frac{1}{|KQ|}\int_{KQ}w^{-(1+\sigma_w)}(y)\,dy\right)^{\frac{1}{2(1+\sigma_w)}}$$

$$\leq [w^{-1}(Q)]^{-\frac{1}{2}}|Q|^{\frac{1}{\beta}}\left(\frac{w^{-1}(KQ)}{|KQ|}\right)^{1/2}\leq C_{[w]_{A_2(\mathbb{R}^n)},K}|Q|^{\frac{1}{\beta}-\frac{1}{2}},$$

where the last inequality comes from Lemma 2.1. Therefore,

$$w^{\frac{1}{2}}[b,R_j](w^{-\frac{1}{2}})$$

$$=C_{[w]_{A_2(\mathbb{R}^n)}}\sum_{\vec{l}\in\mathbb{Z}^{2n}}\sum_{m=0}^{1}\sum_{s=1}^{M}\sum_{Q\in\mathscr{D}}\lambda_{Q,\vec{l}}\,\mathrm{osc}_r(b,Q)\big\langle f,G_{\vec{l}'',R_{Q,s},m}\big\rangle F_{\vec{l}',Q,m},$$

where $\{G_{Q,\vec{l}'',m}\}$ and $\{F_{Q,\vec{l}',m}\}$ are NWO sequences and the coefficients $\{\lambda_{Q,\vec{l}}\}$ satisfy $|\lambda_{Q,\vec{l}}|\lesssim\frac{1}{(1+|\vec{l}|)^{|\alpha|+|\gamma|}}$ for all multi-indices $\alpha,\gamma\in\mathbb{Z}^n_+$. Thus, by (2.3), we have

$$\big\|w^{\frac{1}{2}}[b,R_j](w^{-\frac{1}{2}})\big\|_{S^{n,\infty}(L^2(\mathbb{R}^n))}\lesssim\big\|\mathrm{osc}_r(b,Q)\big\|_{\ell^{n,\infty}}.$$

By [11, Theorem 1 and Remark (d)] (see also [29, Theorem 2.2]), we know that $\mathrm{osc}_r(b,Q)\in\ell^{n,\infty}$ follows from $b\in\dot{W}^{1,n}(\mathbb{R}^n)$. Then $w^{\frac{1}{2}}[b,R_j](w^{-\frac{1}{2}})\in S^{n,\infty}\big(L^2(\mathbb{R}^n)\big)$. Hence, we are done with the proof of the sufficient condition in Theorem 1.2.

5.2. Proof of the necessary condition. In this subsection, we assume that $[b,R_j]\in S^{n,\infty}(L^2(\mathbb{R}^n,w))$, then prove that $b\in\dot{W}^{1,n}(\mathbb{R}^n)$.

First, choosing two cubes Q and $\hat{Q}$ in $\mathscr{D}$, as Lemma 3.3. Define

$$J_Q(x,y)=|Q|^{-2}K_j^{-1}(x-y)\chi_Q(x)\chi_{\hat{Q}}(y).$$

For $K_j^{-1}(x-y)$, decomposing in a multiple Fourier series on $Q\times\hat{Q}$, we can write

$$K_j^{-1}(x-y)=\sum_{\vec{l}\in\mathbb{Z}^{2n}}c^j_{\vec{l},Q}e^{2\pi i\vec{l}'\cdot\widetilde{x}}e^{2\pi i\vec{l}''\cdot\widetilde{y}}\chi_Q(x)\cdot\chi_{\hat{Q}}(y),$$

where $x_i = C_Q^{(i)} + \ell(Q)\tilde{x}_i$, $y_i = C_{\hat{Q}}^{(i)} + \ell(\hat{Q})\tilde{y}_i$, $i = 1, 2, \dots, n$, $\vec{l} = (\vec{l'}, \vec{l''})$ where $\vec{l'} = (l_1, l_2, \dots, l_n)$, $\vec{l''} = (l_{n+1}, l_{n+2}, \dots, l_{2n})$, and the Fourier coefficient $c^j_{\vec{l},Q}$ are given by

$$c^j_{\vec{l},Q} := \int_{\hat{Q}} \int_Q K_j^{-1}(x-y)\chi_{(Q\times\hat{Q})} e^{-2\pi i\vec{l'}\cdot\widetilde{x}} e^{-2\pi i\vec{l''}\cdot\widetilde{y}}\, dx\, dy\, \frac{1}{|Q|}\, \frac{1}{|\hat{Q}|}.$$

Similar to the estimate in the previous subsection, using

$$\left|\partial_x^\alpha \partial_y^\beta K_j^{-1}(x-y)\right| \le C(\alpha,\beta)\, |x-y|^{n-|\alpha|-|\beta|},$$

$|Q| = |\hat{Q}|$ and $\mathrm{dist}(Q, \hat{Q}) \approx |Q|$ yields

$$\begin{aligned} |c^j_{\vec{l},Q}| &\lesssim \frac{1}{(1+|\vec{l}|)^{|\alpha|+|\beta|}}\, \ell(Q)^{|\alpha|} \ell(\hat{Q})^{|\beta|} \int_{\hat{Q}} \int_Q |\partial_x^\alpha \partial_y^\beta K_j^{-1}(x-y)|\, dx\, dy\, \frac{1}{|Q|}\, \frac{1}{|\hat{Q}|} \\ &\lesssim \frac{1}{(1+|\vec{l}|)^{|\alpha|+|\beta|}}\, \ell(Q)^{|\alpha|} \ell(\hat{Q})^{|\beta|} \int_{\hat{Q}} \int_Q |x-y|^{n-|\alpha|-|\beta|}\, dx\, dy\, \frac{1}{|Q|}\, \frac{1}{|\hat{Q}|} \\ &\lesssim |Q|\, \frac{1}{(1+|\vec{l}|)^{|\alpha|+|\beta|}} \end{aligned}$$

where $\alpha, \beta \in \mathbb{Z}_+^n$ are multi-indices. Therefore, we can denote $\lambda^j_{\vec{l},Q} = \frac{1}{|Q|} c^j_{\vec{l},Q}$, and then the estimate

$$|\lambda^j_{\vec{l},Q}| \lesssim \frac{1}{(1+|l|)^{|\alpha|+|\beta|}}$$

holds. Obviously,

$$J_Q(x,y) = \sum_{\vec{l}\in\mathbb{Z}^{2n}} \lambda^j_{\vec{l},Q}\, \frac{1}{|Q|^{1/2}}\, F_{\vec{l'},Q}(x)\, \frac{1}{|\hat{Q}|^{1/2}}\, G_{\vec{l''},\hat{Q}}(y),$$

where $F_{\vec{l'},Q}(x) = e^{2\pi i\vec{l'}\cdot\widetilde{x}}\chi_Q(x)$ and $G_{\vec{l''},\hat{Q}}(y) = e^{2\pi i\vec{l''}\cdot\widetilde{y}}\chi_{\hat{Q}}(y)$.

Next, recall that for each $Q \in \mathscr{D}$, there is $\hat{Q}$ in $\mathscr{D}$ as in Lemma 3.3. We set the function $\varepsilon_{Q,\hat{Q}}(x,y) = \mathrm{sgn}(b(x) - b(y))\chi_Q(x)\chi_{\hat{Q}}(y)$. Define the operator L_Q as

$$w^{\frac{1}{2}}(x) L_Q(w^{-\frac{1}{2}} f)(x) = \int_{\mathbb{R}^n} w^{\frac{1}{2}}(x)\varepsilon_{Q,\hat{Q}}(x,y) J_Q(x,y) w^{-\frac{1}{2}}(y) f(y)\, dy.$$

Considering an arbitrary sequence $\{a_Q\}_{Q\in\mathscr{D}} \in \ell^{\frac{n}{n-1},1}$. Here $\ell^{\frac{n}{n-1},1}$ is the Lorentz sequence space defined as the set of all sequences $\{a_Q\}_{Q\in\mathscr{D}}$ such that

$$\big\|\{a_Q\}_{Q\in\mathscr{D}}\big\|_{\ell^{\frac{n}{n-1},1}}=\sum_{k=1}^{\infty}k^{\frac{n}{n-1}-1}a_k^*,$$

where the sequence $\{a_k^*\}$ is the sequence $\{|a_Q|\}$ rearranged in a decreasing order.

Define the operator L as

$$w^{\frac{1}{2}}(x)L(w^{-\frac{1}{2}}f)(x)=\sum_{Q\in\mathscr{D}}a_Q w^{\frac{1}{2}}(x)L_Q(w^{-\frac{1}{2}}f)(x).$$

Therefore, we also write

$$w^{\frac{1}{2}}(x)L(w^{-\frac{1}{2}}f)(x)=C_{[w]_{A_2(\mathbb{R}^n)}}\sum_{Q\in\mathscr{D}}\sum_{\vec{l}\in\mathbb{Z}^{2n}}\lambda^j_{\vec{l},Q}a_Q\langle f,\tilde{G}_{\vec{l}'',\hat{Q}}\rangle\tilde{F}_{\vec{l}',Q}(x),$$

where

$$\tilde{G}_{\vec{l}'',\hat{Q}}(y)=\frac{G_{\vec{l}'',\hat{Q}}(y)w^{-\frac{1}{2}}(y)}{(w^{-1}(Q))^{\frac{1}{2}}}\quad\text{and}\quad\tilde{F}_{\vec{l}',Q}(x)=\frac{F_{\vec{l}',Q}(x)w^{\frac{1}{2}}(x)}{(w(Q))^{\frac{1}{2}}}.$$

By Lemma 2.2, and repeating the argument from inside the proof Proposition 3.1, they are NWO sequences. Thus, applying Lemma 2.7 gives

$$\|L\|_{S^{\frac{n}{n-1},1}(L^2(\mathbb{R}^n,w))}\approx\big\|w^{\frac{1}{2}}Lw^{-\frac{1}{2}}\big\|_{S^{\frac{n}{n-1},1}(L^2(\mathbb{R}^n))}\le\|a_Q\|_{\ell^{\frac{n}{n-1},1}}.$$

Using the idea of [30, p. 262], we also can obtain

$$\begin{aligned}\operatorname{Trace}\big(w^{\frac{1}{2}}[b,R_j]L_Q(w^{-\frac{1}{2}})\big)&=|Q|^{-2}\int_Q\int_{\hat{Q}}(b(x)-b(y))\varepsilon_{Q,\hat{Q}}(x,y)\,dy\,dx\\&=|Q|^{-2}\int_Q\int_{\hat{Q}}|b(x)-b(y)|\,dy\,dx.\end{aligned}$$

Then we have

$$\begin{aligned}\operatorname{Trace}\big(w^{\frac{1}{2}}[b,R_j]L_Q(w^{\frac{1}{2}})\big)&\gtrsim\frac{1}{|Q|}\int_Q|b(x)-\langle b\rangle_{\hat{Q}}|\,dx\\&\gtrsim\frac{1}{|Q|}\int_Q|b(x)-\langle b\rangle_Q|\,dx=:M(b,Q).\end{aligned}$$

Therefore, by duality, there exists a sequence $\{a_Q\}_{Q\in\mathscr{D}}$ with $\|a_Q\|_{\ell^{\frac{n}{n-1},1}}\le 1$ such that

$$\|b\|_{\dot{W}^{1,n}(\mathbb{R}^n)}\lesssim\|M(b,Q)\|_{\ell^{n,\infty}}\lesssim\big\|\operatorname{Trace}(w^{\frac{1}{2}}[b,R_j]L_Q(w^{-\frac{1}{2}}))\big\|_{\ell^{n,\infty}}$$

$$= \sup_{\|a_Q\|_{\ell^{\frac{n}{n-1},1}}\leq 1} \sum_{Q\in\mathscr{D}} \mathrm{Trace}\big(w^{\frac{1}{2}}[b,R_j]L_Q(w^{-\frac{1}{2}})\big)\cdot a_Q$$

$$= \sup_{\|a_Q\|_{\ell^{\frac{n}{n-1},1}}\leq 1} \mathrm{Trace}\big(w^{\frac{1}{2}}[b,R_j]L(w^{-\frac{1}{2}})\big)$$

$$\lesssim \sup_{\|a_Q\|_{\ell^{\frac{n}{n-1},1}}\leq 1} \big\|w^{\frac{1}{2}}[b,R_j](w^{-\frac{1}{2}})\big\|_{S^{n,\infty}(L^2(\mathbb{R}^n))}\big\|w^{\frac{1}{2}}L(w^{-\frac{1}{2}})\big\|_{S^{\frac{n}{n-1},1}(L^2(\mathbb{R}^n))}$$

$$\lesssim \|[b,R_j]\|_{S^{n,\infty}(L^2(\mathbb{R}^n,w))},$$

where the first inequality comes from [11, Theorem 1 and Remark (d)], see also [30]. Hence, the proof of the necessary condition in Theorem 1.2 is complete.

6. Discussion on the one dimensional case

When $n=1$, Peller [25] obtained the following result in the unweighted case:

THEOREM 6.1 [25]. *For $b\in \mathrm{VMO}(\mathbb{R})$, and $0<p<\infty$, we have*

$$\|[b,H]\|_{S^p(L^2(\mathbb{R}))} \approx \|b\|_{B^{p,p}_{1/p}(\mathbb{R})}.$$

For $p=2$, Lacey and the last two authors in [20] considered Schatten classes and the commutator $[b,H]$ in the two weight setting (see Theorem A in Section 1). In [20, Section 7], the authors raised two questions about the one weight question in one dimension:

(i) For $b\in \mathrm{VMO}(\mathbb{R})$, and $1<p<\infty$. Is $\|[b,H]\|_{S^p(L^2(w))} \approx \|b\|_{B^{p,p}_{1/p}(\mathbb{R})}$ true?

(ii) Can the above conclusion be extended to $0<p\leq 1$?

Similar to the proof of (1) in Theorem 1.1, we can give a positive answer to problem (i). The reader can see that all of Section 3 works for $n=1$ and replacing the Riesz transform with the Hilbert transform. The main details are an equivalence with the Besov space and a dyadic counterpart and the ability to study the commutator by the dyadic shift operator. Section 3 does the analysis in the case of the Riesz transforms, but the case of the Hilbert transform is similar, and in fact slightly easier since we can use Petermichl's Haar shift [26] and the resulting paraproducts in the one variable case are easier to work with (there is no paraproduct like Γ_b which required an additional argument). We omit the details. Coupling these results and making the direct modifications in Section 3 exactly answers (i) above.

However, we can not come up with a good way to solve problem (ii) in this paper because of the method used. There are a couple of obstacles in answering this question using the methods from this paper. A first obstruction is that Lemma 2.5 is used to provide a lower bound for the Schatten norm of the commutator $[b, H]$; resulting in the restriction that $n = 1 < p < \infty$. Consequently, we lose a tool to study the case $0 < p \le 1$ for $[b, H]$.

While Section 4 does carry over to the case of $n = 1$, this section is unfortunately not applicable to the situation for (ii). A main obstacle to handling the case $n = 1$ and $0 < p \le 1$ is that the norm of the Besov space requires more derivatives to characterize it. For $n = 1$, the corresponding Besov space $B^{p,p}_{1/p}(\mathbb{R})$, $p \le 1$, is defined by

$$B^{p,p}_{1/p}(\mathbb{R}) = \left\{ b \in \mathrm{BMO}(\mathbb{R}) : \int_0^\infty \int_{\mathbb{R}} |t^k \nabla^k P_t(b)(x)|^p \frac{dx\, dt}{t^2} < \infty \right\}, \tag{6.1}$$

where $P_t(b)(x)$ is the Poisson integral of b on $\mathbb{R}^2$ and $\nabla = (\partial_x, \partial_t)$ and k must satisfy $k > \frac{1}{p}$. In particular, because of the condition $k > \frac{1}{p}$ one will have to utilize a norm involving more derivatives. In [30, Section 5] it is pointed out that $b \in B^{p,p}_{1/p}(\mathbb{R})$ implies that the sequence $\{\mathrm{osc}(b, Q, r, K, L)\} \in \ell^p$ where $L > 1/p$, $1 \le r < \infty$, $K \ge 1$, and

$$\mathrm{osc}(b, Q, r, K, L) = \inf_{\deg(P) \le L} \left\{ \frac{1}{|KQ|} \int_{KQ} |b(x) - P(x)|^r \, dx \right\}^{1/r}, \quad Q \in \mathscr{D}$$

with $P(x)$ the corresponding polynomial of degree less than or equal to L.

Note that the norm (6.1) with $k > 1/p$ does not connect to the norm we introduced in (1.3) which essentially uses only a first derivative. In fact, the norm in (1.3) for $p > 1$ is equivalent to the sequence $\{\mathrm{osc}(b, Q, r, K)\} \in \ell^p$, $1 \le r < \infty$, $K \ge 1$, and

$$\mathrm{osc}(b, Q, r, K) = \left\{ \frac{1}{|KQ|} \int_{KQ} |b(x) - \langle b \rangle_{KQ}|^r \, dx \right\}^{1/r}, \quad Q \in \mathscr{D},$$

here we refer to [30, Section 4, pp. 266–267]. Moreover, the space given via the norm in (1.3) for $p > 1$ coincides with the classical Besov space, we refer to [31, Sections 2.2.2, 2.5.7]. The proof of Proposition 4.1 only handles this simpler oscillation condition and not the one that is more closely connected to the Besov space $B^{p,p}_{1/p}(\mathbb{R})$ and the oscillation condition involving polynomials.

So, this remains an open problem. While Peller in [25] proved Theorem 6.1 by using Hankel operators exploiting the connection with analyticity, a possible alternate approach will be to develop an alternate dyadic norm on the Besov space but using a wavelet with more cancellation.

7. An application: the quantised derivative

Let $n > 1$, and let $x_1, x_2, \ldots, x_n$ be the coordinates of $\mathbb{R}^n$. For $j = 1, \ldots, n$, we define D_j to be the derivative in the direction x_j,

$$D_j = \frac{1}{i}\frac{\partial}{\partial x_j} = -i\partial_j.$$

When $f \in L^\infty(\mathbb{R}^n)$ is not a smooth function then $D_j f$ denotes the distributional derivative of f. We also consider D_j as a self-adjoint operator on $L^2(\mathbb{R}^n)$ with its standard domain of square integrable functions with a square integrable weak derivative in the direction x_j. This is equivalent to the closure of the symmetric operator D_j restricted to Schwartz functions. We use the notation $\nabla f = i\,(D_1 f, D_2 f, \ldots, D_n f)$ for an essentially bounded function $f \in L^\infty(\mathbb{R}^n)$. For a square integrable function f with a square integrable derivative in each direction we consider ∇ as an unbounded operator from $L^2(\mathbb{R}^n)$ to the Bochner space $L^2(\mathbb{R}^n, \mathbb{C}^n)$.

Let $N = 2^{\lfloor n/2 \rfloor}$. We use n-dimensional Euclidean gamma matrices, which are $N \times N$ self-adjoint complex matrices $\gamma_1, \ldots, \gamma_n$ satisfying the anticommutation relation

$$\gamma_j\gamma_k + \gamma_k\gamma_j = 2\delta_{j,k}, \quad 1 \le j, k \le n,$$

where δ is the Kronecker delta. The precise choice of matrices satisfying this relation is unimportant so we assume that a choice is fixed for the rest of the discussion.

Using this choice of gamma matrices, we can define the n-dimensional Dirac operator by

$$\mathcal{D} = \sum_{j=1}^{n} \gamma_j \otimes D_j.$$

This is a linear operator on the Hilbert space $\mathbb{C}^N \otimes L^2(\mathbb{R}^n)$ initially defined with dense domain $\mathbb{C}^N \otimes \mathcal{S}(\mathbb{R}^n)$, where $\mathcal{S}(\mathbb{R}^n)$ is the Schwartz space of functions on $\mathbb{R}^n$. It is easily seen that $\mathcal{D}$ is symmetric on this domain. Taking the closure we obtain a self-adjoint operator which we also denote by $\mathcal{D}$. We then define the sign of $\mathcal{D}$ as the operator $\operatorname{sgn}(\mathcal{D})$ via the Borel functional calculus, i.e., $\operatorname{sgn}(\mathcal{D}) = \frac{\mathcal{D}}{|\mathcal{D}|}$.

Given $f \in L^\infty(\mathbb{R}^n)$, denote by M_f the operator of pointwise multiplication by f on the Hilbert space $L^2(\mathbb{R}^n)$. The operator $1 \otimes M_f$ is a bounded linear operator on $\mathbb{C}^N \otimes L^2(\mathbb{R}^n)$, where 1 denotes the identity operator on $\mathbb{C}^N$. The commutator

$$\bar{d}f := i\left[\operatorname{sgn}(\mathcal{D}), 1 \otimes M_f\right]$$

denotes the quantised derivative of Alain Connes introduced in [7, IV]. It is of particular interest in the quantised calculus to determine conditions on f such that $\bar{d}f \in S^{n,\infty}(\mathbb{C}^N \otimes L^2(\mathbb{R}^n))$. The asymptotic behaviour of the singular values of the quantised derivative denotes the dimension of the infinitesimal in the quantised calculus. That the sequence of singular values belongs to the weak space $\ell^{n,\infty}$ when the dimension of the Euclidean space is n indicates analogous behaviour between quantum derivatives and differential forms. Specifically, a product of n derivatives lies in the space $S^{1,\infty}(\mathbb{C}^N \otimes L^2(\mathbb{R}^n))$, which is the only weak space admitting a non-trivial trace that acts as the integral.

In one dimension, necessary and sufficient conditions on $f \in L^\infty(\mathbb{R})$ such that $\left[\mathrm{sgn}(-i\frac{d}{dx}), M_f\right] \in S^{p,q}(\mathbb{C}^N \otimes L^2(\mathbb{R}))$ where $p, q \in (0,\infty]$ are provided by Peller in [25, Chapter 4, Theorem 4.4]. Janson and Wolff [18], and Connes, Sullivan and Teleman [8] have studied necessary and sufficient conditions for $\bar{d}f \in S^{p,q}(\mathbb{C}^N \otimes L^2(\mathbb{R}^n))$ with $p, q \in (0,\infty]$ in the higher dimensional case $n > 1$. The case of $p = q$ was studied by Janson and Wolff in their paper [18]. They proved that when $p > n$, a necessary and sufficient condition for $\bar{d}f \in S^p(\mathbb{C}^N \otimes L^2(\mathbb{R}^n))$ is that f is in the Besov space $B^{p,p}_{n/p}(\mathbb{R}^n)$. They also show that if $p \le n$, then $\bar{d}f \in S^p$ if and only if f is a constant.

The case of $p \neq q$ with $p \in [1,\infty)$ and $q \in [1,\infty]$ was answered by Rochberg and Semmes in [30, Corollary 2.8, Theorem 3.4]. Necessary and sufficient conditions on $f \in L^\infty(\mathbb{R}^n)$ are given so that

$$\bar{d}f \in S^{p,q}(\mathbb{C}^N \otimes L^2(\mathbb{R}^n)).$$

These conditions are given in terms of the mean oscillation of f, and it is not obvious whether an equivalent condition could be given in terms of more familiar function spaces. In the Appendix of Connes, Sullivan and Teleman's paper [8, p. 679], it is proved that necessary and sufficient conditions for $\bar{d}f \in S^{n,\infty}(\mathbb{C}^N \otimes L^2(\mathbb{R}^n))$ are that $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ and $\nabla f \in L^n(\mathbb{R}^n, \mathbb{C}^n)$.

Recently, Lord–McDonald–Sukochev–Zanin [22] gave a different proof of this result under the assumption that $f \in L^\infty(\mathbb{R}^n)$ using double operator integrals. Their method gave sharp bounds on the quasinorm $\|\bar{d}f\|_{S^{n,\infty}(\mathbb{C}^N \otimes L^2(\mathbb{R}^n))}$. For the norm $\nabla f \in L^n(\mathbb{R}^n, \mathbb{C}^n)$, they implicitly assumed that the essentially bounded function f has weak partial derivatives and that the Bochner norm of ∇f in $L^n(\mathbb{R}^n, \mathbb{C}^n)$,

$$\|\nabla f\|_{L^n(\mathbb{R}^n,\mathbb{C}^n)} = \left(\int_{\mathbb{R}^n} \|(\nabla f)(x)\|_n^n \, dx\right)^{1/n} = \left(\int_{\mathbb{R}^n} \sum_{j=1}^n |D_j f(x)|^n \, dx\right)^{1/n},$$

is finite. The key step that they established is a new trace formula described as follows, which is analogous to Connes [6].

Recall that a trace on $S^{1,\infty}(\mathbb{C}^N \otimes L^2(\mathbb{R}^n))$ is a linear functional

$$\varphi\colon S^{1,\infty}(\mathbb{C}^N \otimes L^2(\mathbb{R}^n)) \to \mathbb{C}$$

such that $\varphi([A,B]) = 0$ for all bounded operators A and for all operators $B \in S^{1,\infty}(\mathbb{C}^N \otimes L^2(\mathbb{R}^n))$. The trace φ is called continuous when it is continuous with respect to the $S^{1,\infty}(\mathbb{C}^N \otimes L^2(\mathbb{R}^n))$ quasinorm. Given an orthonormal basis $\{e_n\}_{n=0}^{\infty}$ of $\mathbb{C}^N \otimes L^2(\mathbb{R}^n)$, define the operator $T := \operatorname{diag}\{\frac{1}{n+1}\}_{n=0}^{\infty}$ by $\langle e_n, Te_m\rangle = \delta_{n,m}\frac{1}{n+1}$. The linear functional φ is called normalised when

$$\varphi\Big(\operatorname{diag}\Big\{\frac{1}{n+1}\Big\}_{n=0}^{\infty}\Big) = 1.$$

The property that φ is normalised is independent of the choice of orthonormal basis, since for all unitary operators U and all bounded operators B we have $\varphi(UBU^*) = \varphi(B)$.

Proposition 7.1 [22]. *Let $f \in L^\infty(\mathbb{R}^n)$ be real valued and such that $\nabla f \in L^n(\mathbb{R}^n, \mathbb{C}^n)$. Then there is a constant $c_n > 0$ such that for any continuous normalised trace φ on $S^{1,\infty}(\mathbb{C}^N \otimes L^2(\mathbb{R}^n))$ we have*

$$\varphi\big(|\bar{d}f|^n\big) = c_n \int_{\mathbb{R}^n} \|\nabla f(x)\|_2^n \, dx.$$

Proposition 7.1 is the analogue of [6, Theorem 3(3)] for functions on the non-compact manifold $\mathbb{R}^n$. It is also stated for a larger class of functions than [6, Theorem 3(3)] which is proved for smooth functions. Based on this trace formula, in [22] they obtained that

Proposition 7.2 [22]. *Let $n > 1$ and $f \in L^\infty(\mathbb{R}^n)$. Then, for $\bar{d}f \in S^{n,\infty}(\mathbb{C}^N \otimes L^2(\mathbb{R}^n))$, it is necessary and sufficient that $\nabla f \in L^n(\mathbb{R}^n, \mathbb{C}^n)$. Further, there exist positive constants c and C depending only on n such that*

$$c\|\nabla f\|_{L^n(\mathbb{R}^n,\mathbb{C}^n)} \le \|\bar{d}f\|_{S^{n,\infty}(\mathbb{C}^N\otimes L^2(\mathbb{R}^n))} \le C\|\nabla f\|_{L^n(\mathbb{R}^n,\mathbb{C}^n)}.$$

From our Theorem 1.2, we have the following result in this direction:

Theorem 7.3. *Suppose $n > 1$, $f \in \mathrm{VMO}(\mathbb{R}^n)$, $w \in A_2$. Then $\bar{d}f \in S^{n,\infty}(\mathbb{C}^N \otimes L^2(w))$ if and only if $f \in \dot{W}^{1,n}(\mathbb{R}^n)$. Moreover,*

$$\|\bar{d}f\|_{S^{n,\infty}(\mathbb{C}^N\otimes L^2(w))} \approx \|f\|_{\dot{W}^{1,n}(\mathbb{R}^n)}.$$

Proof. We provide the details of the link between $\bar{d}f$ and $[f, \nabla\Delta^{-1/2}]$, where Δ is the standard Laplacian on $\mathbb{R}^n$. In fact, from the definition of $\mathcal{D}$ and the property of these self-adjoint complex matrices $\gamma_1, \ldots, \gamma_n$, we see that

$$\mathcal{D}^2 = -1 \otimes \Delta.$$

Moreover, $\operatorname{sgn}(\mathcal{D})$ which can equivalently be expressed as

$$\operatorname{sgn}(\mathcal{D}) = \sum_{j=1}^{n} \gamma_j \otimes D_j \Delta^{-1/2} = \sum_{j=1}^{n} \gamma_j \otimes R_j,$$

where R_j is the jth Riesz transform. Hence,

$$\begin{aligned} \bar{d}f &= i\big[\operatorname{sgn}(\mathcal{D}), 1 \otimes M_f\big] = i\Big[\sum_{j=1}^{n} \gamma_j \otimes R_j, 1 \otimes M_f\Big] = i\sum_{j=1}^{n}\big[\gamma_j \otimes R_j, 1 \otimes M_f\big] \\ &= i\sum_{j=1}^{n}\big(\gamma_j \otimes R_j M_f - \gamma_j \otimes M_f R_j\big) = i\sum_{j=1}^{n} \gamma_j \otimes [R_j, M_f]. \end{aligned}$$

Thus, the result follows from Theorem 1.2. □

Acknowledgements. The authors greatly appreciate the efforts and valuable comments of the referee while reading and checking the paper.

Ji Li would like to thank E. McDonald, F. Sukochev and D. Zanin for helpful discussions on the quantised derivative.

References

[1] S. Bloom, A commutator theorem and weighted BMO, *Trans. Amer. Math. Soc.*, **292** (1985), 103–122.

[2] A.-P. Calderón, Commutators of singular integral operators, *Proc. Natl. Acad. Sci. USA*, **53** (1965), 1092–1099.

[3] P. Chen, X. Duong, J. Li and Q. Wu, Compactness of Riesz transform commutator on stratified Lie groups, *J. Funct. Anal.*, **277** (2019), 1639–1676.

[4] R. Coifman, P. L. Lions, Y. Meyer and S. Semmes, Compensated compactness and Hardy spaces, *J. Math. Pures Appl.*, **72** (1993), 247–286.

[5] R. Coifman, R. Rochberg and G. Weiss, Factorization theorems for Hardy spaces in several variables, *Ann. of Math.*, **103** (1976), 611–635.

[6] A. Connes, The action functional in noncommutative geometry, *Comm. Math. Phys.*, **117** (1988) 673–683.

[7] A. Connes, *Noncommutative Geometry*, Academic Press, Inc. (San Diego, CA, 1994).

[8] A. Connes, D. Sullivan and N. Teleman, Quasiconformal mappings, operators on Hilbert space, and local formulae for characteristic classes, *Topology*, **33** (1994), 663–681.

[9] Z. Fan, M. T. Lacey and J. Li, Schatten classes and commutators of Riesz Transform on Heisenberg group and applications, *J. Fourier Anal. Appl.*, **29** (2023), Paper No. 17, 28 pp.

[10] Z. Fan, J. Li, E. McDonald, F. Sukochev and D. Zanin, Endpoint weak Schatten class estimates and trace formula for commutators of Riesz transforms with multipliers on Heisenberg groups, *J. Funct. Anal.*, **286** (2024), Paper No. 110188.

[11] R. L. Frank, A characterization of $\dot{W}^{1,p}(\mathbb{R}^n)$, arXiv: 2203.01001 (2022)

[12] R. L. Frank, F. Sukochev and D. Zanin, Asymptotics of singular values for quantum derivatives, *Trans. Amer. Math. Soc.*, **376** (2023), 2047–2088.

[13] H. Gimperlein and M. Goffeng. Nonclassical spectral asymptotics and Dixmier traces: from circles to contact manifolds. *Forum Math. Sigma*, **5** (2017), Paper No. e3, 57 pp.

[14] L. Grafakos, *Modern Fourier Analysis*, 3rd ed., Graduate Texts in Math., vol. 250, Springer (New York, 2014).

[15] I. Holmes, M. T. Lacey and B. D. Wick, Commutators in the two-weight setting, *Math. Ann.*, **367** (2017), 51–80.

[16] T. Hytönen, The L^p-to-L^q boundedness of commutators with applications to the Jacobian operator, *J. Math. Pures Appl.*, **156** (2021), 351–391.

[17] T. Hytönen A. Kairema, Systems of dyadic cubes in a doubling metric space, *Colloq. Math.*, **126** (2012), 1–33.

[18] S. Janson and T. H. Wolff, Schatten classes and commutators of singular integral operators, *Ark. Mat.*, **20** (1982), 301–310.

[19] M. Lacey and J. Li, Compactness of commutator of Riesz transforms in the two weight setting, *J. Math. Anal. Appl.*, **508** (2022), Paper No. 125869, 11 pp..

[20] M. T. Lacey, J. Li and B. D. Wick, Schatten classes and commutator in the two weight setting. I. Hilbert transform, *Potential Anal.* (2023), https://doi.org/10.1007/s11118-023-10072-x.

[21] M. T. Lacey, J. Pipher, S. Petermichl and B. D. Wick, Iterated Riesz commutators: A simple proof of boundedness, in:*Proceedings of 8th Intl. Conf. on Harmonic Analysis and PDE*, (El Escorial, Madrid (Spain), 2008), arXiv: 0808.0832 (2008).

[22] S. Lord, E. McDonald, F. Sukochev and D. Zanin, Quantum differentiability of essentially bounded functions on Euclidean space, *J. Funct. Anal.*, **273** (2017) 2353–2387.

[23] A. K. Lerner, A simple proof of the A_2 conjecture, *Int. Math. Res. Not. IMRN*, **14** (2013), 3159–3170.

[24] Z. Nehari, On bounded bilinear forms, *Ann. of Math.*, **65** (1957), 153–162.

[25] V. V. Peller, *Hankel Operators and their Applications*, Springer Monographs in Math., Springer-Verlag (New York, 2003).

[26] S. Petermichl, Dyadic shifts and a logarithmic estimate for Hankel operators with matrix symbol, *C. R. Acad. Sci. Paris Sér. I Math.*, **330** (2000), 455–460.

[27] S. Petermichl, The sharp weighted bound for the Riesz transforms, *Proc. Amer. Math. Soc.*, **136** (2008), 1237–1249.

[28] S. Petermichl, S. Treil and A. Volberg, Why are the Riesz transforms averages of the dyadic shift?, in: *Proc. 6th International Conference on Harmonic Analysis* (El Escorial, 2000), *Publ. Mat.*, Vol. Extra (2002), 209–228.

[29] R. Rochberg and S. Semmes, End point results for estimates of singular values of singular integral operators, in: *Contributions to Operator Theory and its Applications* (Mesa, AZ, 1987), Oper. Theory Adv. Appl., vol. 35, Birkhäuser Verlag (Basel, 1988), 217–231.

[30] R. Rochberg and S. Semmes, Nearly weakly orthonormal sequences, singular value estimates, and Calderon– Zygmund operators, *J. Funct. Anal.*, **86** (1989), 237–306.

[31] H. Triebel, *Theory of Function Spaces*, Monogr. Math., vol. 78, Birkhäuser Verlag (Basel, 1983).