

A CHARACTERIZATION OF TWO WEIGHT NORM INEQUALITY FOR LITTLEWOOD-PALEY g_λ^* -FUNCTION

MINGMING CAO, KANGWEI LI, AND QINGYING XUE

ABSTRACT. Let $n \geq 2$ and g_λ^* be the well-known high dimensional Littlewood-Paley function which was defined and studied by E. M. Stein,

$$g_\lambda^*(f)(x) = \left(\iint_{\mathbb{R}_+^{n+1}} \left(\frac{t}{t + |x - y|} \right)^{n\lambda} |\nabla P_t f(y, t)|^2 \frac{dy dt}{t^{n-1}} \right)^{1/2}, \quad \lambda > 1$$

where $P_t f(y, t) = p_t * f(x)$, $p_t(y) = t^{-n} p(y/t)$ and $p(x) = (1 + |x|^2)^{-(n+1)/2}$, $\nabla = (\frac{\partial}{\partial y_1}, \dots, \frac{\partial}{\partial y_n}, \frac{\partial}{\partial t})$. In this paper, we give a characterization of two weight norm inequality for g_λ^* -function. We show that, $\|g_\lambda^*(f\sigma)\|_{L^2(w)} \lesssim \|f\|_{L^2(\sigma)}$ if and only if the two weight Muckenhoupt A_2 condition holds, and a testing condition holds :

$$\sup_{Q: \text{cubes in } \mathbb{R}^n} \frac{1}{\sigma(Q)} \int_{\mathbb{R}^n} \iint_{\widehat{Q}} \left(\frac{t}{t + |x - y|} \right)^{n\lambda} |\nabla P_t(\mathbf{1}_Q \sigma)(y, t)|^2 \frac{w dx dt}{t^{n-1}} dy < \infty,$$

where \widehat{Q} is the Carleson box over Q and (w, σ) is a pair of weights. We actually prove this characterization for g_λ^* function associated with more general fractional Poisson kernel $p^\alpha(x) = (1 + |x|^2)^{-(n+\alpha)/2}$. Moreover, the corresponding results for intrinsic g_λ^* -function are also presented.

1. INTRODUCTION

The g_λ^* -function originated in the work of Littlewood and Paley [8] in the 1930s. It is a basic tool to prove the L^p -boundedness of various linear operators. Later, the classical g_λ^* function of higher dimension was first introduced and studied by Stein [13] in 1961, a certain sublinear operator arises in Littlewood-Paley theory [1, 14]. It plays important roles in harmonic analysis and other fields. Let $n \geq 2$, we recall its definition as follows:

$$g_\lambda^*(f)(x) = \left(\iint_{\mathbb{R}_+^{n+1}} \left(\frac{t}{t + |x - y|} \right)^{n\lambda} |\nabla P_t f(y, t)|^2 \frac{dy dt}{t^{n-1}} \right)^{1/2}, \quad \lambda > 1$$

where $P_t f(y, t) = p_t * f(x)$, $p_t(y) = t^{-n} p(\frac{y}{t})$ denotes the Poisson kernel and $\nabla = (\frac{\partial}{\partial y_1}, \dots, \frac{\partial}{\partial y_n}, \frac{\partial}{\partial t})$. It is easy to show that g_λ^* is an isometry on $L^2(\mathbb{R}^n)$. With much greater difficulty, it can be proved that for any $1 < p < \infty$, $\|g_\lambda^*(f)\|_{L^p(\mathbb{R}^n)}$ and $\|f\|_{L^p(\mathbb{R}^n)}$

Date: April 20, 2017.

Key words and phrases. Two weight inequality; Littlewood-Paley g_λ^* -function; Pivotal Condition; Random dyadic grids.

M.C. and Q.X were supported partly by NSFC (No. 11471041), the Fundamental Research Funds for the Central Universities (NO. 2012CXQT09 and NO. 2014KJJCA10) and NCET-13-0065. K.L. is supported by the Basque Government through the BERC 2014-2017 program and by Spanish Ministry of Economy and Competitiveness MINECO: BCAM Severo Ochoa excellence accreditation SEV-2013-0323.

are equivalent norms [13]. Moreover, in [13], Stein also proved that if $\lambda > 2$, then g_λ^* is of weak type $(1, 1)$, and is of strong type (p, p) for $1 < p < \infty$. In 1970, as a replacement of weak $(1, 1)$ bounds for $1 < \lambda < 2$, Fefferman [1] considered the end-point weak (p, p) estimates of g_λ^* -function when $p > 1$ and $\lambda = 2/p$.

Recently, Lacey and the second named author [5] gave a characterization of two weight norm inequalities for the classical g -function and the corresponding intrinsic square function. Recall that the classical g -function is defined by

$$g(f)(x) = \left(\int_0^\infty |\nabla P_t f(x, t)|^2 t dt \right)^{1/2}.$$

It was shown that the following two weight norm inequality for the classical Littlewood-Paley g -function for a pair of weights (w, σ) on \mathbb{R}^n :

$$(1.1) \quad \|g(f\sigma)\|_{L^2(w)} \lesssim \mathcal{N} \|f\|_{L^2(\sigma)}$$

holds if and only if (w, σ) satisfies

$$(1.2) \quad \mathcal{A}_2^2 := \sup_Q \frac{\sigma(Q) w(Q)}{|Q| |Q|} < \infty;$$

and the testing condition holds, uniformly over all cubes $Q \subset \mathbb{R}^n$,

$$(1.3) \quad \iint_{\widehat{Q}} |\nabla P_t(\mathbf{1}_Q \sigma)(x, t)|^2 w dx t dt \lesssim \sigma(Q), \quad \widehat{Q} = Q \times [0, \ell(Q)].$$

The condition (1.3) is called the Sawyer testing condition, which can be traced back to [12]. It is known that Littlewood-Paley g -function is point wisely controlled by g_λ^* -function. Thus it is quite natural to ask if one can establish a characterization for the Littlewood-Paley g_λ^* -function. But the g_λ^* -function also involves additional difficulties since more integrals appear in the definition. One also needs to find the new suitable testing condition to replace condition (1.3).

In order to state our results, we first introduce the definition of the Littlewood-Paley g_λ^* -function with fractional Poisson kernels.

Definition 1.1. Let $\lambda > 1$, for any $x \in \mathbb{R}^n$, the Littlewood-Paley g_λ^* -function with fractional Poisson kernels is defined by

$$g_\lambda^{*,\alpha}(f)(x) = \left(\iint_{\mathbb{R}_+^{n+1}} \left(\frac{t}{t + |x - y|} \right)^{n\lambda} |\nabla P_t^\alpha f(y, t)|^2 \frac{dy dt}{t^{n-1}} \right)^{1/2}, \quad 0 < \alpha \leq 1,$$

where $P_t^\alpha f(y, t) = p_t^\alpha * f(x)$, $p_t^\alpha(y) = t^{-n} p^\alpha(y/t)$ and $p^\alpha(x) = (1 + |x|^2)^{-\frac{n+\alpha}{2}}$.

Remark 1.2. If $\alpha = 1$, then $g_\lambda^{*,1}$ coincides with the classical Littlewood-Paley g_λ^* -function of higher dimension defined and studied by E. M. Stein [13] in 1961.

Motivated by the above work, in this paper, we will focus on the characterization of the two weight inequality for the Littlewood-Paley g_λ^* -function.

$$(1.4) \quad \|g_\lambda^{*,\alpha}(f\sigma)\|_{L^2(w)} \lesssim \mathcal{N} \|f\|_{L^2(\sigma)}.$$

In addition, we introduce the corresponding testing condition:

$$(1.5) \quad \mathcal{B}^2 := \sup_{Q: \text{cubes in } \mathbb{R}^n} \frac{1}{\sigma(Q)} \int_{\mathbb{R}^n} \iint_{\widehat{Q}} \left(\frac{t}{t + |x - y|} \right)^{n\lambda} |\nabla P_t^\alpha(\mathbf{1}_Q \sigma)(y, t)|^2 \frac{w dx dt}{t^{n-1}} dy < \infty.$$

Here we formulate the main result of this paper as follows.

Theorem 1.1. *Let $\lambda > 2$, $0 < \alpha \leq \min\{1, n(\lambda - 2)/2\}$ and σ, w be two weights. Then the two weight inequality (1.4) holds if and only if the two weight condition (1.2) and testing condition (1.5) hold. Moreover, $\mathcal{N} \simeq \mathcal{A}_2 + \mathcal{B}$, where \mathcal{N} is the best constant in the inequality (1.4).*

Remark 1.3. The characterization of the two weight inequality for the classical Littlewood-Paley g_λ^* -function is contained in Theorem 1.1 ($\alpha = 1$, $\lambda \geq 2(1 + 1/n)$). Actually, when $\lambda \geq 2(1 + 1/n)$, we have $0 < \alpha \leq 1$. It not only includes the classical case, but also extends to the case for $0 < \alpha < 1$. Another notable fact is that we are able to improve the result of [5] with the fractional Poisson kernel p^α , $0 < \alpha \leq 1$.

To state another main result, we begin with one more definition.

Definition 1.4. For $0 < \alpha \leq 1$, let \mathcal{C}_α be the family of functions φ satisfying $\text{supp } \varphi \subset \{x \in \mathbb{R}^n; |x| \leq 1\}$, $\int_{\mathbb{R}^n} \varphi(x) dx = 0$, and such that $|\varphi(x) - \varphi(x')| \leq |x - x'|^\alpha$, for all $x, x' \in \mathbb{R}^n$. If $f \in L_{loc}^1(\mathbb{R}^n)$ and $(y, t) \in \mathbb{R}_+^{n+1}$, we define $A_\alpha f(y, t) = \sup_{\varphi \in \mathcal{C}_\alpha} |f * \varphi_t(y)|$, where $\varphi_t(x) = t^{-n} \varphi(x/t)$. Then the *intrinsic* g_λ^* -function is defined by setting, for all $x \in \mathbb{R}^n$,

$$g_{\lambda, \alpha}^*(f)(x) = \left(\iint_{\mathbb{R}_+^{n+1}} \left(\frac{t}{t + |x - y|} \right)^{n\lambda} [A_\alpha f(y, t)]^2 \frac{dy dt}{t^{n+1}} \right)^{1/2}.$$

For the intrinsic $g_{\lambda, \alpha}^*$ function, we have the following result.

Theorem 1.2. *Let $\lambda > 2$, $0 < \alpha \leq \min\{1, n(\lambda - 2)/2\}$ and σ, w be two weights. Then the two weight inequality*

$$\|g_{\lambda, \alpha}^*(f\sigma)\|_{L^2(w)} \lesssim \mathcal{N}_\alpha \|f\|_{L^2(\sigma)}$$

holds if and only if

- (i) (w, σ) satisfies the A_2 condition (1.2);
- (ii) the testing condition holds :

$$\mathcal{B}_\alpha^2 := \sup_{Q: \text{cubes in } \mathbb{R}^n} \frac{1}{\sigma(Q)} \int_{\mathbb{R}^n} \iint_{\widehat{Q}} \left(\frac{t}{t + |x - y|} \right)^{n\lambda} [A_\alpha(\mathbf{1}_Q \sigma)(y, t)]^2 \frac{w dx dt}{t^{n+1}} dy < \infty.$$

Moreover, the best constants satisfy $\mathcal{N}_\alpha \simeq \mathcal{A}_2 + \mathcal{B}_\alpha$.

Note that $g_\lambda^{*, \alpha} f(x) \leq g_{\lambda, \alpha}^* f(x)$, for all $x \in \mathbb{R}^n$. Since the main steps in the proof of Theorem 1.2 are the same as the Theorem 1.1, we omit the proof of Theorem 1.2.

The rest of this article is organized as follows. The necessary condition is shown in the Section 2. In Section 3, applying the random dyadic grids and martingale difference decomposition, we give the final reduction of the main theorem. In order to prove the sufficiency, some lemmas and elementary estimates are established in Section 4. Finally, in Section 5, by splitting into four parts, we prove the sufficiency in Theorem 1.1.

Notation We write $A \lesssim B$, if there is a constant $C > 0$ so that $A \leq CB$. We may also write $A \approx B$ if $B \lesssim A \lesssim B$.

We then set some dyadic notation. For cubes Q and R , we denote

- $\ell(Q)$ is the side length of Q ;
- $d(Q, R)$ denotes the distance between the cubes Q and R ;
- $D(Q, R) := \ell(Q) + \ell(R) + d(Q, R)$ is the long distance;
- $\widehat{Q} := Q \times (0, \ell(Q)]$ is the Carleson box over Q ;
- $W_Q := Q \times (\frac{\ell(Q)}{2}, \ell(Q)]$ is the Whitney region associated with Q ;
- $Q^{(k)}$ denotes the unique dyadic cube for which $\ell(Q^{(k)}) = 2^k \ell(Q)$ and $Q \subset Q^{(k)}$ and $Q \subset Q^{(k)}$;
- $\text{ch}(Q)$ denotes the dyadic children of Q .

2. THE NECESSITY AND CONSTANT ESTIMATES

2.1. Proposition. The inequality (1.4) implies the inequality (1.2).

Proof. For some fixed cube Q , we have

$$|\nabla P_t^\alpha(\mathbf{1}_Q \sigma)(y, t)| \geq |\partial_t P_t^\alpha * (\mathbf{1}_Q \sigma)(y)| = \left| \int_Q \frac{nt^{\alpha+1} - \alpha t^{\alpha-1} |y-z|^2}{(t^2 + |y-z|^2)^{\frac{n+\alpha}{2}+1}} \sigma dz \right|.$$

If $x, y, z \in Q$ and $2\ell(Q) \leq t \leq 3\ell(Q)$, then

$$\frac{nt^{\alpha+1} - \alpha t^{\alpha-1} |y-z|^2}{(t^2 + |y-z|^2)^{\frac{n+\alpha}{2}+1}} \gtrsim \frac{1}{t^{n+1}}.$$

Thus,

$$|\nabla P_t^\alpha(\mathbf{1}_Q \sigma)(y, t)| \gtrsim \frac{\sigma(Q)}{t^{n+1}}.$$

Furthermore, for $x \in Q$,

$$\begin{aligned} g_\lambda^{*,\alpha}(\mathbf{1}_Q \sigma)(x)^2 &\geq \int_Q \int_{2\ell(Q)}^{3\ell(Q)} \left(\frac{t}{t + |x-y|} \right)^{n\lambda} \left| \int_Q \frac{nt^{\alpha+1} - \alpha t^{\alpha-1} |y-z|^2}{(t^2 + |y-z|^2)^{\frac{n+\alpha}{2}+1}} \sigma dz \right|^2 \frac{dt}{t^{n-1}} dy \\ &\gtrsim \int_Q \int_{2\ell(Q)}^{3\ell(Q)} \frac{\sigma(Q)^2}{t^{3n+1}} dt dy \gtrsim \frac{\sigma(Q)^2}{|Q|^2}. \end{aligned}$$

Therefore, the boundedness of $g_\lambda^{*,\alpha}$ gives that

$$\frac{\sigma(Q)w(Q)}{|Q|^2} \lesssim \frac{1}{\sigma(Q)} \|g_\lambda^{*,\alpha}(\mathbf{1}_Q \sigma)\|_{L^2(w)}^2 \leq \mathcal{N}^2.$$

That is, $\mathcal{A}_2 \lesssim \mathcal{N}$. □

Moreover, it is trivial that (1.4) implies (1.5). Thus, we have proved the necessity of Theorem 1.1.

2.2. Random Dyadic Grids. Let us recall random dyadic grids defined in [2]. Denote by $\mathcal{D} = \mathcal{D}(\beta)$ the random dyadic grid, where $\beta = \{\beta_j\}_{j=-\infty}^\infty \in (\{0, 1\}^n)^\mathbb{Z}$. That is

$$\mathcal{D} = \{Q + \sum_{j:2^{-j} < \ell(Q)} 2^{-j}\beta_j : Q \in \mathcal{D}_0\},$$

where \mathcal{D}_0 is the standard dyadic grid in \mathbb{R}^n .

A cube $I \in \mathcal{D}$ is said to be bad if there exists a $J \in \mathcal{D}$ with $\ell(J) \geq 2^r \ell(I)$ such that $\text{dist}(I, \partial J) \leq \ell(I)^\gamma \ell(J)^{1-\gamma}$, where $r \in \mathbb{Z}_+$ and $\gamma \in (0, \frac{1}{2})$ are given parameters. Otherwise, I is called good.

Throughout this article, we take $\gamma = \frac{\alpha}{2(n+\alpha)}$ and r will be determined in the following. Moreover, roughly speaking, a dyadic cube I will be bad if it is relatively close to the boundary of a much bigger dyadic cube. Denote $\pi_{\text{good}} = \mathbb{P}_\beta(Q + \beta \text{ is good}) = \mathbb{E}_\beta(1_{\text{good}}(Q + \beta))$. Then π_{good} is independent of $Q \in \mathcal{D}_0$. And we choose r large enough so that $\pi_{\text{good}} > 0$.

2.3. Definition. Given a dyadic cube I , we set \mathcal{W}_I to be the maximal dyadic cubes $K \subset I$ such that $2^r \ell(K) \leq \ell(I)$ and $\text{dist}(K, \partial I) \geq \ell(K)^\gamma \ell(I)^{1-\gamma}$.

2.4. Proposition. The following statements hold.

- (1) For any good $J \Subset I$, there is a cube $K \in \mathcal{W}_I$ which contains J ;
- (2) For any $C > 0$, provided r is sufficiently large, depending upon γ , there holds

$$\sum_{K \in \mathcal{W}_I} \mathbf{1}_{CK} \lesssim \mathbf{1}_I.$$

Here, $J \Subset I$ means that $J \subset I$ and $2^r \ell(J) \leq \ell(I)$; in words, J is strongly contained in I .

2.5. The Pivotal Condition. The *pivotal constant* \mathcal{P} is the smallest constant in the following inequality. For any cube I^0 , and any partition of I^0 into dyadic cubes $\{I_j; j \in \mathbb{N}\}$, there holds

$$(2.1) \quad \sum_{j \in \mathbb{N}} \sum_{K \in \mathcal{W}_{I_j}} \mathcal{P}_\alpha(K, \mathbf{1}_{I^0} \sigma)^2 w(K) \leq \mathcal{P}^2 \sigma(I^0),$$

where Poisson term

$$\mathcal{P}_\alpha(I, f) = \int_{\mathbb{R}^n} \frac{\ell(I)^\alpha}{(\ell(I) + \text{dist}(x, I))^{n+\alpha}} f(x) dx.$$

To estimate the best constants, we give the following Proposition.

2.6. Proposition. The A_2 condition (1.2) and testing condition (1.5) imply the finiteness of the pivotal constant \mathcal{P} . In particular, there holds $\mathcal{P} \lesssim \mathcal{A}_2 + \mathcal{B}$.

Proof. We follow the strategy used in [5]. Taking the large enough constant C in Proposition 2.4 such that $\frac{\alpha}{2} \geq n(\frac{2}{C-1})^2$. The A_2 condition and Proposition 2.4 give that

$$\sum_{j \in \mathbb{N}} \sum_{K \in \mathcal{W}_{I_j}} \mathcal{P}_\alpha(K, \mathbf{1}_{CK} \sigma)^2 w(K) \lesssim \mathcal{A}_2^2 \sum_{j \in \mathbb{N}} \sum_{K \in \mathcal{W}_{I_j}} \sigma(CK) \lesssim \mathcal{A}_2^2 \sigma(I^0).$$

Thus, it is enough to treat the Poisson terms $\mathcal{P}_\alpha(K, \mathbf{1}_{I^0 \setminus CK} \sigma)$.

It is easy to verify

$$\mathcal{P}_\alpha(K, \mathbf{1}_{I^0 \setminus CK} \sigma) \lesssim t \partial_t P_t^\alpha(\mathbf{1}_{I^0 \setminus CK} \sigma)(y, t), \quad \text{for any } y \in K, t \sim \ell(K).$$

Therefore,

$$\mathcal{P}_\alpha(K, \mathbf{1}_{I^0 \setminus CK} \sigma)^2 w(K) \lesssim \int_K \iint_{W_K} \left(\frac{t}{t + |x - y|} \right)^{n\lambda} |\nabla P_t^\alpha(\mathbf{1}_{I^0 \setminus CK} \sigma)(y, t)|^2 \frac{w dx dt}{t^{n-1}} dy.$$

Since we have

$$\begin{aligned} & \sum_{j \in \mathbb{N}} \sum_{K \in \mathcal{W}_{I_j}} \int_K \iint_{W_K} \left(\frac{t}{t + |x - y|} \right)^{n\lambda} |\nabla P_t^\alpha(\mathbf{1}_{I^0} \sigma)(y, t)|^2 \frac{w dx dt}{t^{n-1}} dy \\ & \leq \sum_{j \in \mathbb{N}} \sum_{K \in \mathcal{W}_{I_j}} \int_{\mathbb{R}^n} \iint_{W_K} \left(\frac{t}{t + |x - y|} \right)^{n\lambda} |\nabla P_t^\alpha(\mathbf{1}_{I^0} \sigma)(y, t)|^2 \frac{w dx dt}{t^{n-1}} dy \\ & \leq \int_{\mathbb{R}^n} \int_{\widehat{I^0}} \left(\frac{t}{t + |x - y|} \right)^{n\lambda} |\nabla P_t^\alpha(\mathbf{1}_{I^0} \sigma)(y, t)|^2 \frac{w dx dt}{t^{n-1}} dy \\ & \leq \mathcal{B}^2 \sigma(I^0), \end{aligned}$$

and

$$\begin{aligned} & \sum_{j \in \mathbb{N}} \sum_{K \in \mathcal{W}_{I_j}} \int_K \iint_{W_K} \left(\frac{t}{t + |x - y|} \right)^{n\lambda} |\nabla P_t^\alpha(\mathbf{1}_{CK} \sigma)(y, t)|^2 \frac{w dx dt}{t^{n-1}} dy \\ & \leq \sum_{j \in \mathbb{N}} \sum_{K \in \mathcal{W}_{I_j}} \int_{\mathbb{R}^n} \iint_{\widehat{CK}} \left(\frac{t}{t + |x - y|} \right)^{n\lambda} |\nabla P_t^\alpha(\mathbf{1}_{CK} \sigma)(y, t)|^2 \frac{w dx dt}{t^{n-1}} dy \\ & \leq \mathcal{B}^2 \sum_{j \in \mathbb{N}} \sum_{K \in \mathcal{W}_{I_j}} \sigma(CK) \\ & \lesssim \mathcal{B}^2 \sigma(I^0), \end{aligned}$$

the desired estimate follows immediately. \square

3. THE PROBABILISTIC REDUCTION

Our next task is to simplify the proof of sufficiency. The probabilistic techniques we will use are taken from [2]. We here need some fundamental tools, including the random dyadic grids, the probabilistic good/bad decompositions, and the martingale difference expansions, which can be found in [2, 5, 6], and essentially goes back to [11].

3.1. The Generalized Result. In order to prove the main theorem, it is enough to show the following generalized result.

$$(3.1) \quad \|g_{\psi, \lambda}^*(f \cdot \sigma)\|_{L^2(w)} \lesssim (\mathcal{A}_2 + \mathcal{B}) \|f\|_{L^2(\sigma)},$$

where

$$g_{\psi, \lambda}^*(f)(x) = \left(\iint_{\mathbb{R}_+^{n+1}} \left(\frac{t}{t + |x - y|} \right)^{n\lambda} |\psi_t * f(y)|^2 \frac{dy dt}{t^{n+1}} \right)^{1/2},$$

$\psi_t(x) = \frac{1}{t^n} \psi\left(\frac{x}{t}\right)$ and ψ satisfies the following conditions:

- (1) $|\psi(x)| \lesssim (1 + |x|)^{-n-\alpha}$;
- (2) $|\psi(x) - \psi(y)| \lesssim |x - y|^\alpha (1 + |x|)^{-n-\alpha}$.

A particular case of the above function class was introduced by Wilson [15, p. 114]. However, we do not need the cancelation property of ψ in this paper.

3.2. Averaging over Good Whitney Regions. Let $f \in L^2(\sigma)$. For $R \in \mathcal{D}$, let $W_R = R \times (\ell(R)/2, \ell(R)]$ be the associated Whitney region. Note that the position and goodness of $R \dot{+} \beta$ are independent (see [2]). Therefore, one can write

$$\begin{aligned}
& \|g_\lambda^*(f \cdot \sigma)\|_{L^2(w)}^2 \\
&= \int_{\mathbb{R}^n} \iint_{\mathbb{R}_+^{n+1}} \left(\frac{t}{t+|y|}\right)^{n\lambda} |\psi_t * (f \cdot \sigma)(x-y)|^2 \frac{dy dt}{t^{n+1}} w dx \\
&= \iint_{\mathbb{R}_+^{n+1}} \int_{\mathbb{R}^n} |\psi_t * (f \cdot \sigma)(x-y)|^2 \left(\frac{t}{t+|y|}\right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \\
&= \mathbb{E}_\beta \sum_{R \in \mathcal{D}_0} \iint_{W_{R \dot{+} \beta}} \int_{\mathbb{R}^n} |\psi_t * (f \cdot \sigma)(x-y)|^2 \left(\frac{t}{t+|y|}\right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \\
&= \frac{1}{\pi_{good}} \sum_{R \in \mathcal{D}_0} \mathbb{E}_\beta(\mathbf{1}_{good}(R \dot{+} \beta)) \mathbb{E}_\beta \iint_{W_{R \dot{+} \beta}} \int_{\mathbb{R}^n} |\psi_t * (f \cdot \sigma)(x-y)|^2 \left(\frac{t}{t+|y|}\right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \\
&= \frac{1}{\pi_{good}} \sum_{R \in \mathcal{D}_0} \mathbb{E}_\beta \left(\mathbf{1}_{good}(R \dot{+} \beta) \iint_{W_{R \dot{+} \beta}} \int_{\mathbb{R}^n} |\psi_t * (f \cdot \sigma)(x-y)|^2 \left(\frac{t}{t+|y|}\right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \right) \\
&= \frac{1}{\pi_{good}} \mathbb{E}_\beta \sum_{R \in \mathcal{D}_{good}} \iint_{W_R} \int_{\mathbb{R}^n} |\psi_t * (f \cdot \sigma)(x-y)|^2 \left(\frac{t}{t+|y|}\right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t}.
\end{aligned}$$

With the monotone convergence theorem, it suffices to show that there exists a constant $\mathcal{C} > 0$ such that for any $s \in \mathbb{N}$, we have

$$\sum_{\substack{R \in \mathcal{D}_{good} \\ R \subset [-2^s, 2^s]^n \\ 2^{-s} \leq \ell(R) \leq 2^s}} \iint_{W_R} \int_{\mathbb{R}^n} |\psi_t * (f \cdot \sigma)(x-y)|^2 \left(\frac{t}{t+|y|}\right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \leq \mathcal{C}(\mathcal{A}_2 + \mathcal{B})^2 \|f\|_{L^2(\sigma)}^2.$$

3.3. The Final Reduction. In order to get the further reduction, we introduce the martingale difference decomposition. Define

$$\mathbb{E}_Q^\sigma f := \frac{1}{\sigma(Q)} \int_Q f d\sigma,$$

assuming that $\sigma(Q) > 0$, otherwise set it to be zero. For the martingale differences,

$$\Delta_Q^\sigma f := \sum_{Q' \in ch(Q)} (\mathbb{E}_{Q'}^\sigma f - \mathbb{E}_Q^\sigma f) \mathbf{1}_{Q'}.$$

For fixed $s \in \mathbb{N}$, by Lebesgue differentiation theorem, we can write

$$f = \sum_{\substack{Q \in \mathcal{D} \\ \ell(Q) \leq 2^s}} \Delta_Q^\sigma f + \sum_{\substack{Q \in \mathcal{D} \\ \ell(Q) = 2^s}} (\mathbb{E}_Q^\sigma f) \mathbf{1}_Q.$$

Since $\{\Delta_Q^\sigma f\}_{Q \in \mathcal{D}}$ is a family of orthogonal, we have

$$\|f\|_{L^2(\sigma)}^2 = \sum_{\substack{Q \in \mathcal{D} \\ \ell(Q) \leq 2^s}} \|\Delta_Q^\sigma f\|_{L^2(\sigma)}^2 + \sum_{\substack{Q \in \mathcal{D} \\ \ell(Q) = 2^s}} \|(\mathbb{E}_Q^\sigma f) \mathbf{1}_Q\|_{L^2(\sigma)}^2.$$

Now we claim that we can assume that f is compactly supported, say $\text{supp } f \subset Q^0$. Let \mathcal{F} denote the subspace of $L^2(\sigma)$ which has compact support. We shall show that

$$(3.2) \quad \mathcal{K} := \sup_{\substack{f \in \mathcal{F} \\ \|f\|_{L^2(\sigma)} = 1}} \|g_\lambda^*(f\sigma)\|_{L^2(w)} < \infty.$$

Indeed, if (3.2) is proved, then for any $f \in L^2(\sigma)$ and $\varepsilon > 0$, there exists some cube Q such that

$$\|f - f\chi_Q\|_{L^2(\sigma)} < \varepsilon \|f\|_{L^2(\sigma)},$$

For simplicity, set $g := f - f\chi_Q$. Then we have

$$\begin{aligned} & \sum_{\substack{R \in \mathcal{D}_{good} \\ RC[-2^s, 2^s]^n \\ 2^{-s} \leq \ell(R) \leq 2^s}} \iint_{W_R} \int_{\mathbb{R}^n} |\psi_t * (f \cdot \sigma)(x - y)|^2 \left(\frac{t}{t + |y|}\right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \\ & \leq 2 \sum_{\substack{R \in \mathcal{D}_{good} \\ RC[-2^s, 2^s]^n \\ 2^{-s} \leq \ell(R) \leq 2^s}} \iint_{W_R} \int_{\mathbb{R}^n} |\psi_t * (f\chi_Q \cdot \sigma)(x - y)|^2 \left(\frac{t}{t + |y|}\right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \\ & \quad + 2 \sum_{\substack{R \in \mathcal{D}_{good} \\ RC[-2^s, 2^s]^n \\ 2^{-s} \leq \ell(R) \leq 2^s}} \iint_{W_R} \int_{\mathbb{R}^n} |\psi_t * (g \cdot \sigma)(x - y)|^2 \left(\frac{t}{t + |y|}\right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t}. \end{aligned}$$

Substitute with

$$f(x) = \frac{\ell(Q)^\alpha}{(\ell(Q) + \text{dist}(x, Q))^{n+\alpha}} \chi_{Q \setminus 4\sqrt{n}Q},$$

in (3.2) and using similar arguments as that in [5], we get

$$\int_Q \frac{\ell(Q)^{2\alpha}}{(\ell(Q) + \text{dist}(z, Q))^{2(n+\alpha)}} d\sigma(z) w(Q) \lesssim \mathcal{K}^2 + \mathcal{A}_2^2.$$

Then by letting Q' increase to \mathbb{R}^n , we know that (3.2) and the A_2 condition imply the Poisson type A_2 condition. Therefore,

$$\begin{aligned}
& \iint_{W_R} \int_{\mathbb{R}^n} |\psi_t * (g \cdot \sigma)(x - y)|^2 \left(\frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \\
& \leq \iint_{W_R} \int_{\mathbb{R}^n} \|g\|_{L^2(\sigma)}^2 \|\psi_t(x - y - \cdot)\|_{L^2(\sigma)}^2 \left(\frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \\
& \leq C_n \int_{\mathbb{R}^n} \frac{\ell(R)^{2\alpha}}{(\ell(R) + \text{dist}(z, R))^{2(n+\alpha)}} d\sigma(z) w(R) \|g\|_{L^2(\sigma)}^2 \\
& \leq C_n (\mathcal{K}^2 + \mathcal{A}_2^2) \varepsilon^2 \|f\|_{L^2(\sigma)}^2
\end{aligned}$$

Then by taking sufficiently large cube Q such that $2^{(2s+2)n} C_n (\mathcal{K}^2 + \mathcal{A}_2^2) \varepsilon^2 < \mathcal{K}^2$. We finally get

$$\sum_{\substack{R \in \mathcal{D}_{good} \\ RC[-2^s, 2^s]^n \\ 2^{-s} \leq \ell(R) \leq 2^s}} \iint_{W_R} \int_{\mathbb{R}^n} |\psi_t * (f \cdot \sigma)(x - y)|^2 \left(\frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \leq 4\mathcal{K}^2 \|f\|_{L^2(\sigma)},$$

which means that we reduce the problem to prove (3.2). Then by repeating the previous arguments, we further reduce the problem to estimate

$$\sum_{\substack{R \in \mathcal{D}_{good} \\ RC[-2^s, 2^s]^n \\ 2^{-s} \leq \ell(R) \leq 2^s}} \iint_{W_R} \int_{\mathbb{R}^n} |\psi_t * (f \cdot \sigma)(x - y)|^2 \left(\frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t},$$

where f has compact support. Assume that $\text{supp } f \subset [-2^{s'}, 2^{s'}]^n$. Without loss of generality, we can assume that $s \geq s' + 1$. Then it suffices to estimate

$$\sum_{\substack{R \in \mathcal{D}_{good} \\ RC[-2^s, 2^s]^n \\ 2^{-s} \leq \ell(R) \leq 2^s}} \iint_{W_R} \int_{\mathbb{R}^n} |\psi_t * (f \mathbf{1}_{[-2^{s-1}, 2^{s-1}]^n} \cdot \sigma)(x - y)|^2 \left(\frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t}.$$

Denote \mathcal{F}_s the subspace of \mathcal{F} which supported in $[-2^{s-1}, 2^{s-1}]^n$.

$$\mathcal{H}_s := \sup_{\substack{f \in \mathcal{F}_s \\ \|f\|_{L^2(\sigma)}=1}} \sum_{\substack{R \in \mathcal{D}_{good} \\ RC[-2^s, 2^s]^n \\ 2^{-s} \leq \ell(R) \leq 2^s}} \iint_{W_R} \int_{\mathbb{R}^n} |\psi_t * (f \cdot \sigma)(x - y)|^2 \left(\frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t}$$

Similar arguments as the previous show that

$$\begin{aligned}
& \iint_{W_R} \int_{\mathbb{R}^n} |\psi_t * (f \cdot \sigma)(x - y)|^2 \left(\frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \\
& \leq C_n \int_{[-2^{s-2}, 2^{s-2}]^n} \frac{\ell(R)^{2\alpha}}{(\ell(R) + \text{dist}(z, R))^{2(n+\alpha)}} d\sigma(z) w(R) \|f\|_{L^2(\sigma)}^2 \\
& \leq 2^{2sn} C_n \mathcal{A}_2^2 \|f\|_{L^2(\sigma)}^2,
\end{aligned}$$

which means that $\mathcal{K}_s \leq 2^{(4s+2)n} C_n \mathcal{A}_2^2 < \infty$. Using the martingale decomposition, we can write

$$f = \sum_{\substack{Q \in \mathcal{D} \\ \ell(Q) \leq 2^s}} \Delta_Q^\sigma f,$$

when $\ell(Q) = 2^s$, Δ_Q^σ should be understood as E_Q^σ . Denote

$$f_{\text{good}} = \sum_{\substack{Q \in \mathcal{D}_{\text{good}} \\ \ell(Q) \leq 2^s}} \Delta_Q^\sigma f.$$

Again, we can set $\tilde{g} := f - f_{\text{good}}$. For any $\varepsilon > 0$, choosing r sufficiently large such that $\|\tilde{g}\|_{L^2(\sigma)} < \varepsilon$, see [3]. Then we have

$$\begin{aligned} \mathcal{K}_s \leq & 2 \sup_{\substack{f \in \mathcal{F} \\ \|f\|_{L^p(\sigma)}=1}} \sum_{\substack{R \in \mathcal{D}_{\text{good}} \\ RC[-2^s, 2^s]^n \\ 2^{-s} \leq \ell(R) \leq 2^s}} \iint_{W_R} \int_{\mathbb{R}^n} |\psi_t * (f_{\text{good}} \cdot \sigma)(x-y)|^2 \left(\frac{t}{t+|y|} \right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \\ & + 2\mathcal{K}_s \|\tilde{g}\|_{L^2(\sigma)}^2. \end{aligned}$$

By taking $\varepsilon = 1/2$, (which means that r is independent of s) we reduce the problem to prove

$$(3.3) \quad \sum_{\substack{R \in \mathcal{D}_{\text{good}} \\ \ell(R) \leq 2^s}} \iint_{W_R} \int_{\mathbb{R}^n} \left| \sum_{\substack{Q \in \mathcal{D}_{\text{good}} \\ \ell(Q) \leq 2^s}} \psi_t * (\Delta_Q^\sigma f \cdot \sigma)(x-y) \right|^2 \left(\frac{t}{t+|y|} \right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \lesssim (\mathcal{A}_2 + \mathcal{B})^2 \|f\|_{L^2(\sigma)}^2.$$

4. SOME LEMMAS AND ELEMENTARY ESTIMATES

To prove the boundedness of $g_\lambda^*(\cdot\sigma)$ from $L^2(\sigma)$ to $L^2(w)$, we here present some crucial estimates and lemmas.

4.1. Elementary Estimate 1. Let $0 < \alpha \leq n(\lambda - 2)/2$. For given cubes $Q, R \in \mathcal{D}$ and $(x, t) \in W_R$, we have the following estimate

$$(4.1) \quad \left(\int_{\mathbb{R}^n} |\psi_t * ((\Delta_Q^\sigma f)\sigma)(x-y)|^2 \left(\frac{t}{t+|y|} \right)^{n\lambda} \frac{dy}{t^n} \right)^{1/2} \lesssim \frac{\ell(R)^\alpha \sigma(Q)^{1/2}}{(\ell(R) + d(Q, R))^{n+\alpha}} \|\Delta_Q^\sigma f\|_{L^2(\sigma)}.$$

Proof. By the size condition, we obtain

$$|\psi_t * ((\Delta_Q^\sigma f) \cdot \sigma)(x-y)| \lesssim \int_{\mathbb{R}^n} \frac{t^\alpha}{(t + |x-y-z|)^{n+\alpha}} |\Delta_Q^\sigma f(z)| d\sigma(z).$$

Since $z \in Q$ and $x \in R$, $|x-z| \geq d(Q, R)$.

If $|y| \leq \frac{1}{2}d(Q, R)$, then $|x-y-z| \geq |x-z| - |y| \geq \frac{1}{2}d(Q, R)$. Thus,

$$\begin{aligned} |\psi_t * ((\Delta_Q^\sigma f)\sigma)(x-y)| & \lesssim \frac{\ell(R)^\alpha}{(\ell(R) + d(Q, R))^{n+\alpha}} \|\Delta_Q^\sigma f\|_{L^1(\sigma)} \\ & \leq \frac{\ell(R)^\alpha}{(\ell(R) + d(Q, R))^{n+\alpha}} \sigma(Q)^{1/2} \|\Delta_Q^\sigma f\|_{L^2(\sigma)}, \end{aligned}$$

and

$$\left(\int_{|y| \leq \frac{1}{2}d(Q,R)} |\psi_t * ((\Delta_Q^\sigma f)\sigma)(x-y)|^2 \left(\frac{t}{t+|y|} \right)^{n\lambda} \frac{dy}{t^n} \right)^{1/2} \lesssim \frac{\ell(R)^\alpha \sigma(Q)^{1/2}}{(\ell(R) + d(Q,R))^{n+\alpha}} \|\Delta_Q^\sigma f\|_{L^2(\sigma)}.$$

If $|y| > \frac{1}{2}d(Q,R)$, then

$$\left(\frac{t}{t+|y|} \right)^{n\lambda} \frac{1}{t^n} \lesssim \frac{\ell(R)^{n\lambda-n}}{(\ell(R) + d(Q,R))^{n\lambda}}.$$

Hence, by Young's inequality, it yields that

$$\begin{aligned} & \left(\int_{|y| > \frac{1}{2}d(Q,R)} |\psi_t * ((\Delta_Q^\sigma f)\sigma)(x-y)|^2 \left(\frac{t}{t+|y|} \right)^{n\lambda} \frac{dy}{t^n} \right)^{1/2} \\ & \lesssim \frac{\ell(R)^{\frac{n\lambda}{2} - \frac{n}{2}}}{(\ell(R) + d(Q,R))^{\frac{n\lambda}{2}}} \|\psi_t * ((\Delta_Q^\sigma f)\sigma)(x-\cdot)\|_{L^2(\mathbb{R}^n)} \\ & \leq \frac{\ell(R)^{\frac{n\lambda}{2} - \frac{n}{2}}}{(\ell(R) + d(Q,R))^{\frac{n\lambda}{2}}} \|\psi_t\|_{L^2(\mathbb{R}^n)} \|\Delta_Q^\sigma f\|_{L^1(\sigma)} \\ & \lesssim \frac{\ell(R)^{\frac{n\lambda}{2} - \frac{n}{2}} t^{-\frac{n}{2}}}{(\ell(R) + d(Q,R))^{\frac{n\lambda}{2}}} \sigma(Q)^{1/2} \|\Delta_Q^\sigma f\|_{L^2(\sigma)} \\ & \lesssim \frac{\ell(R)^{\frac{n\lambda}{2} - n}}{(\ell(R) + d(Q,R))^{\frac{n\lambda}{2}}} \sigma(Q)^{1/2} \|\Delta_Q^\sigma f\|_{L^2(\sigma)} \\ & \leq \frac{\ell(R)^\alpha}{(\ell(R) + d(Q,R))^{n+\alpha}} \sigma(Q)^{1/2} \|\Delta_Q^\sigma f\|_{L^2(\sigma)}, \end{aligned}$$

where we have used the condition $0 < \alpha \leq n(\lambda - 2)/2$ in the last step.

This completes the proof of (4.1). \square

4.2. Elementary Estimate 2. Let $0 < \alpha \leq n(\lambda - 2)/2$. Assume that $Q, R \in \mathcal{D}$ are given cubes with $\ell(Q) < \ell(R)$, $\ell(Q) < 2^s$ and $(x, t) \in W_R$. Then we have the following estimate

$$(4.2) \quad \left(\int_{\mathbb{R}^n} |\psi_t * (\Delta_Q^\sigma f \cdot \sigma)(x-y)|^2 \left(\frac{t}{t+|y|} \right)^{n\lambda} \frac{dy}{t^n} \right)^{1/2} \lesssim \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2} \sigma(Q)^{1/2}}{(\ell(R) + d(Q,R))^{n+\alpha}} \|\Delta_Q^\sigma f\|_{L^2(\sigma)}.$$

Proof. Let z_Q be the center of Q . By the cancellation condition $\int_Q \Delta_Q^\sigma f \sigma dx = 0$, we have

$$\psi_t * (\Delta_Q^\sigma f \cdot \sigma)(x-y) = \int_Q (\psi_t(x-y-z) - \psi_t(x-y-z_Q)) \Delta_Q^\sigma f(z) d\sigma(z).$$

Since $|z - z_Q| \leq \frac{\sqrt{n}}{2} \ell(Q) \leq \frac{\sqrt{n}}{4} \ell(R) < \frac{\sqrt{n}}{2} t \leq \frac{\sqrt{n}}{2} \ell(R)$, we have

$$|\psi_t(x-y-z) - \psi_t(x-y-z_Q)| \lesssim \frac{|z - z_Q|^\alpha}{(t + |x-y-z|)^{n+\alpha}} \lesssim \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2}}{(t + |x-y-z|)^{n+\alpha}}.$$

Making use the similar arguments as in the preceding subsection, we will obtain the inequality (4.2).

4.3. Some Lemmas. For the sake of talking convenience, we here present two key lemmas, which will be used later.

Lemma 4.1 ([5]). *Let*

$$A_{QR}^\alpha = \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2}}{D(Q, R)^{n+\alpha}} \sigma(Q)^{1/2} w(R)^{1/2},$$

where the long distance $D(Q, R) = \ell(Q) + \ell(R) + d(Q, R)$, $Q, R \in \mathcal{D}$ and $\alpha > 0$. Then for any $x_Q, y_R \geq 0$, we have the following estimate

$$\left(\sum_{Q, R \in \mathcal{D}} A_{QR}^\alpha x_Q y_R \right)^2 \lesssim \mathcal{A}_2^2 \sum_{Q \in \mathcal{D}} x_Q^2 \times \sum_{R \in \mathcal{D}} y_R^2.$$

Lemma 4.2. *Let $0 < \alpha \leq n(\lambda - 2)/2$. Given three cubes $R \subset K \subset S$, and function f satisfies $\text{supp}(f) \cap S = \emptyset$. If $\text{dist}(R, \partial K) \geq \ell(R)^\gamma \ell(K)^{1-\gamma}$, then there holds*

$$(4.3) \quad \int_{\mathbb{R}^n} \iint_{W_R} |\psi_t * (f \cdot \sigma)(x - y)|^2 w dx \left(\frac{t}{t + |y|} \right)^{n\lambda} \frac{dt dy}{t^{n+1}} \lesssim \left(\frac{\ell(R)}{\ell(K)} \right)^\alpha \mathcal{P}_\alpha(K, |f| \sigma)^2 w(R).$$

Proof. First, we shall prove, for any $z \notin S$,

$$(4.4) \quad \frac{\ell(R)^\alpha}{(\ell(R) + \text{dist}(z, R))^{n+\alpha}} \leq \left[\frac{\ell(R)}{\ell(K)} \right]^{\alpha/2} \frac{\ell(K)^\alpha}{(\ell(K) + \text{dist}(z, K))^{n+\alpha}}.$$

In fact, since $\text{dist}(z, R) \geq \text{dist}(z, K) + \text{dist}(R, \partial K)$, we have

$$\begin{aligned} \frac{\ell(R)^\alpha}{(\ell(R) + \text{dist}(z, R))^{n+\alpha}} &= \left(\frac{\ell(R)}{\ell(K)} \right)^\alpha \frac{\ell(K)^\alpha}{(\ell(R) + \text{dist}(z, R))^{n+\alpha}} \\ &\lesssim \left(\frac{\ell(R)}{\ell(K)} \right)^{\alpha - (n+\alpha)\gamma} \frac{\ell(K)^\alpha}{(\ell(K) + \text{dist}(z, K))^{n+\alpha}}. \end{aligned}$$

Secondly, we turn to the estimate of (4.3). Decompose

$$\begin{aligned} &\int_{\mathbb{R}^n} |\psi_t * (f \cdot \sigma)(x - y)|^2 \left(\frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t^n} \\ &\leq \int_{\mathbb{R}^n} \left(\int_{z: |y| \leq \frac{1}{2} \text{dist}(z, R)} |\psi_t(x - y - z)| |f(z)| d\sigma(z) \right)^2 \left(\frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t^n} \\ &\quad + \int_{\mathbb{R}^n} \left(\int_{z: |y| > \frac{1}{2} \text{dist}(z, R)} |\psi_t(x - y - z)| |f(z)| d\sigma(z) \right)^2 \left(\frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t^n} \\ &:= \mathcal{E}_1 + \mathcal{E}_2. \end{aligned}$$

For $(x, t) \in W_R$, and $z \notin S$, we have

$$|\psi_t(x - y - z)| \lesssim \frac{t^\alpha}{(t + |x - y - z|)^{n+\alpha}} \lesssim \frac{\ell(R)^\alpha}{(\ell(R) + |x - y - z|)^{n+\alpha}}.$$

If $|y| \leq \frac{1}{2} \text{dist}(z, R)$, $|x - y - z| \geq |x - z| - |y| \geq \frac{1}{2} \text{dist}(z, R)$. Then by (4.4)

$$|\psi_t(x - y - z)| \lesssim \frac{\ell(R)^\alpha}{(\ell(R) + \text{dist}(z, R))^{n+\alpha}} \lesssim \left(\frac{\ell(R)}{\ell(K)} \right)^{\alpha/2} \frac{\ell(K)^\alpha}{(\ell(K) + \text{dist}(z, K))^{n+\alpha}}.$$

Hence,

$$\begin{aligned}\mathcal{E}_1 &\lesssim \left(\frac{\ell(R)}{\ell(K)}\right)^\alpha \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} \frac{\ell(K)^\alpha}{(\ell(K) + \text{dist}(z, K))^{n+\alpha}} |f(z)| d\sigma(z) \right)^2 \left(\frac{t}{t+|y|}\right)^{n\lambda} \frac{dy}{t^n} \\ &\lesssim \left(\frac{\ell(R)}{\ell(K)}\right)^\alpha \mathcal{P}_\alpha(K, |f|\sigma)^2.\end{aligned}$$

If $|y| > \frac{1}{2}\text{dist}(z, R)$, the inequality (4.4) and Young's inequality imply that

$$\begin{aligned}\mathcal{E}_2 &\lesssim t^n \int_{\mathbb{R}^n} \left(\int_{z:|y|>\frac{1}{2}\text{dist}(z,R)} |\psi_t(x-y-z)| \frac{\ell(R)^{\frac{n\lambda}{2}-n}}{(\ell(R) + \text{dist}(z, R))^{\frac{n\lambda}{2}}} |f(z)| d\sigma(z) \right)^2 dy \\ &\leq t^n \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |\psi_t(x-y-z)| \frac{\ell(R)^\alpha}{(\ell(R) + \text{dist}(z, R))^{n+\alpha}} |f(z)| d\sigma(z) \right)^2 dy \\ &\lesssim t^n \left(\frac{\ell(R)}{\ell(K)}\right)^\alpha \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |\psi_t(x-y-z)| \frac{\ell(K)^\alpha}{(\ell(K) + \text{dist}(z, K))^{n+\alpha}} |f(z)| d\sigma(z) \right)^2 dy \\ &\leq t^n \left(\frac{\ell(R)}{\ell(K)}\right)^\alpha \|\psi_t\|_{L^2(\mathbb{R}^n)}^2 \left(\int_{\mathbb{R}^n} \frac{\ell(K)^\alpha}{(\ell(K) + \text{dist}(z, K))^{n+\alpha}} |f(z)| d\sigma(z) \right)^2 \\ &\lesssim \left(\frac{\ell(R)}{\ell(K)}\right)^\alpha \mathcal{P}_\alpha(K, |f|\sigma)^2.\end{aligned}$$

Consequently, the inequality (4.3) is concluded from the above estimates. \square

5. THE SUFFICIENCY IN THE MAIN THEOREM

In this section, we undertake to prove the sufficiency. We shall divide the collection $\{Q; Q \in \mathcal{D}_{good}, \ell(Q) \leq 2^s\}$ into the following four parts. The last one is the core and quite complicated.

5.1. The Case $\ell(Q) < \ell(R)$. In this case, we must have $\ell(Q) < 2^s$. It follows from (4.2) and Lemma 4.1 that

$$\begin{aligned}&\sum_{\substack{R \in \mathcal{D}_{good} \\ \ell(R) \leq 2^s}} \iint_{W_R} \int_{\mathbb{R}^n} \left| \sum_{\substack{Q \in \mathcal{D}_{good} \\ \ell(Q) < \ell(R)}} \psi_t * (\Delta_Q^\sigma f \cdot \sigma)(x-y) \right|^2 \left(\frac{t}{t+|y|}\right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \\ &\leq \sum_{\substack{R \in \mathcal{D}_{good} \\ \ell(R) \leq 2^s}} \iint_{W_R} \left[\sum_{\substack{Q \in \mathcal{D}_{good} \\ \ell(Q) < \ell(R)}} \left(\int_{\mathbb{R}^n} |\psi_t * (\Delta_Q^\sigma f \cdot \sigma)(x-y)|^2 \left(\frac{t}{t+|y|}\right)^{n\lambda} \frac{dy}{t^n} \right)^{1/2} \right]^2 w dx \frac{dt}{t} \\ &\lesssim \sum_{\substack{R \in \mathcal{D}_{good} \\ \ell(R) \leq 2^s}} \iint_{W_R} \left[\sum_{\substack{Q \in \mathcal{D}_{good} \\ \ell(Q) < \ell(R)}} \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2}}{(\ell(R) + d(Q, R))^{n+\alpha}} \sigma(Q)^{1/2} \|\Delta_Q^\sigma f\|_{L^2(\sigma)} \right]^2 w dx \frac{dt}{t} \\ &\lesssim \sum_{R \in \mathcal{D}_{good}} \left(\sum_{Q \in \mathcal{D}_{good}} A_{QR}^\alpha \|\Delta_Q^\sigma f\|_{L^2(\sigma)} \right)^2 \lesssim \mathcal{A}_2^2 \|f\|_{L^2(\sigma)}^2.\end{aligned}$$

5.2. The Case $\ell(Q) \geq \ell(R)$ and $d(Q, R) > \ell(R)^\gamma \ell(Q)^{1-\gamma}$. We claim that there holds in this case

$$(5.1) \quad \frac{\ell(R)^\alpha}{(\ell(R) + d(Q, R))^{n+\alpha}} \lesssim \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2}}{D(Q, R)^{n+\alpha}}.$$

Indeed, if $\ell(Q) \leq d(Q, R)$, it is obvious that

$$\frac{\ell(R)^\alpha}{(\ell(R) + d(Q, R))^{n+\alpha}} \lesssim \frac{\ell(R)^\alpha}{D(Q, R)^{n+\alpha}} \leq \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2}}{D(Q, R)^{n+\alpha}}.$$

If $\ell(Q) > d(Q, R)$, then $D(Q, R) \sim \ell(Q)$. Using $d(Q, R) > \ell(R)^\gamma \ell(Q)^{1-\gamma}$ and $\gamma = \frac{\alpha}{2(n+\alpha)}$, we obtain

$$\ell(Q) = \left(\frac{\ell(Q)}{\ell(R)} \right)^\gamma \ell(R)^\gamma \ell(Q)^{1-\gamma} < \left(\frac{\ell(Q)}{\ell(R)} \right)^\gamma d(Q, R),$$

and

$$\frac{\ell(R)^\alpha}{(\ell(R) + d(Q, R))^{n+\alpha}} \leq \frac{\ell(R)^\alpha}{d(Q, R)^{n+\alpha}} \leq \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2}}{\ell(Q)^{n+\alpha}} \sim \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2}}{D(Q, R)^{n+\alpha}}.$$

Then Lemma 4.1 and the inequalities (4.1), (5.1) give that

$$\begin{aligned} & \sum_{\substack{R \in \mathcal{D}_{good} \\ \ell(R) \leq 2^s}} \iint_{W_R} \int_{\mathbb{R}^n} \left| \sum_{\substack{Q \in \mathcal{D}_{good}: \ell(Q) \geq \ell(R) \\ d(Q, R) > \ell(R)^\gamma \ell(Q)^{1-\gamma}}} \psi_t * (\Delta_Q^\sigma f \cdot \sigma)(x - y) \right|^2 \left(\frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \\ & \leq \sum_{\substack{R \in \mathcal{D}_{good} \\ \ell(R) \leq 2^s}} \iint_{W_R} \left[\sum_{\substack{Q \in \mathcal{D}_{good}: \ell(Q) \geq \ell(R) \\ d(Q, R) > \ell(R)^\gamma \ell(Q)^{1-\gamma}}} \left(\int_{\mathbb{R}^n} |\psi_t * (\Delta_Q^\sigma f \cdot \sigma)(x - y)|^2 \left(\frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t^n} \right)^{1/2} \right]^2 w dx \frac{dt}{t} \\ & \lesssim \sum_{\substack{R \in \mathcal{D}_{good} \\ \ell(R) \leq 2^s}} \iint_{W_R} \left[\sum_{\substack{Q \in \mathcal{D}_{good}: \ell(Q) \geq \ell(R) \\ d(Q, R) > \ell(R)^\gamma \ell(Q)^{1-\gamma}}} \frac{\ell(R)^\alpha}{(\ell(R) + d(Q, R))^{n+\alpha}} \sigma(Q)^{1/2} \|\Delta_Q^\sigma f\|_{L^2(\sigma)} \right]^2 w dx \frac{dt}{t} \\ & \lesssim \sum_{R \in \mathcal{D}_{good}} \left(\sum_{Q \in \mathcal{D}_{good}} A_{QR}^\alpha \|\Delta_Q^\sigma f\|_{L^2(\sigma)} \right)^2 \lesssim \mathcal{A}_2^2 \|f\|_{L^2(\sigma)}^2. \end{aligned}$$

5.3. The Case $\ell(R) \leq \ell(Q) \leq 2^r \ell(R)$ and $d(Q, R) \leq \ell(R)^\gamma \ell(Q)^{1-\gamma}$. In this case, it is trivial that $D(Q, R) \sim \ell(Q) \sim \ell(R)$. Thus

$$\frac{\ell(R)^\alpha}{(\ell(R) + d(Q, R))^{n+\alpha}} \leq \ell(R)^{-n} \sim \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2}}{D(Q, R)^{n+\alpha}}.$$

Then proceeding as we did in the previous subsection, we obtain that

$$\begin{aligned}
& \sum_{\substack{R \in \mathcal{D}_{good} \\ \ell(R) \leq 2^s}} \iint_{W_R} \int_{\mathbb{R}^n} \left| \sum_{\substack{\ell(R) \leq \ell(Q) \leq 2^r \ell(R) \\ d(Q,R) \leq \ell(R)^\gamma \ell(Q)^{1-\gamma}}} \psi_t * (\Delta_Q^\sigma f \cdot \sigma)(x-y) \right|^2 \left(\frac{t}{t+|y|} \right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \\
& \lesssim \sum_{\substack{R \in \mathcal{D}_{good} \\ \ell(R) \leq 2^s}} \iint_{W_R} \left[\sum_{\substack{\ell(R) \leq \ell(Q) \leq 2^r \ell(R) \\ d(Q,R) \leq \ell(R)^\gamma \ell(Q)^{1-\gamma}}} \frac{\ell(R)^\alpha}{(\ell(R) + d(Q,R))^{n+\alpha}} \sigma(Q)^{1/2} \|\Delta_Q^\sigma f\|_{L^2(\sigma)} \right]^2 w dx \frac{dt}{t} \\
& \lesssim \mathcal{A}_2^2 \|f\|_{L^2(\sigma)}^2.
\end{aligned}$$

5.4. The Case $\ell(Q) > 2^r \ell(R)$ and $d(Q,R) \leq \ell(R)^\gamma \ell(Q)^{1-\gamma}$. Recall $R^{(k)}$ is the k generations older dyadic ancestor of R . In this case, since R is good, it must actually have $R \subset Q$. That is, Q is the ancestor of R . Then we can write

$$\begin{aligned}
& \sum_{\substack{R \in \mathcal{D}_{good} \\ \ell(R) \leq 2^s}} \iint_{W_R} \int_{\mathbb{R}^n} \left| \sum_{\substack{\ell(Q) > 2^r \ell(R) \\ d(Q,R) \leq \ell(R)^\gamma \ell(Q)^{1-\gamma}}} \psi_t * (\Delta_Q^\sigma f \cdot \sigma)(x-y) \right|^2 \left(\frac{t}{t+|y|} \right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \\
& = \sum_{\substack{R \in \mathcal{D}_{good} \\ \ell(R) \leq 2^{s-r-1}}} \iint_{W_R} \int_{\mathbb{R}^n} \left| \sum_{k=r+1}^{s-\log_2 \ell(R)} \psi_t * (\Delta_{R^{(k)}}^\sigma f \cdot \sigma)(x-y) \right|^2 \left(\frac{t}{t+|y|} \right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \\
& \leq \sum_{\substack{R \in \mathcal{D}_{good} \\ \ell(R) \leq 2^{s-r-1}}} \iint_{W_R} \int_{\mathbb{R}^n} \left| \sum_{k=r+1}^{s-\log_2 \ell(R)} \psi_t * ((\mathbf{1}_{R^{(k)} \setminus R^{(k-1)}}) \Delta_{R^{(k)}}^\sigma f) \sigma)(x-y) \right|^2 \left(\frac{t}{t+|y|} \right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \\
& \quad + \sum_{\substack{R \in \mathcal{D}_{good} \\ \ell(R) \leq 2^{s-r-1}}} \iint_{W_R} \int_{\mathbb{R}^n} \left| \sum_{k=r+1}^{s-\log_2 \ell(R)} \psi_t * ((\mathbf{1}_{R^{(k-1)}}) \Delta_{R^{(k)}}^\sigma f) \sigma)(x-y) \right|^2 \left(\frac{t}{t+|y|} \right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \\
& := \mathcal{J} + \mathcal{K}.
\end{aligned}$$

Fix the summing variable $k \geq r+1$. Then, the inequality (4.3) implies that

$$\begin{aligned}
& \sum_{\substack{R \in \mathcal{D}_{good} \\ \ell(R) \leq 2^{s-r-1}}} \iint_{W_R} \int_{\mathbb{R}^n} \left| \psi_t * ((\mathbf{1}_{R^{(k)} \setminus R^{(k-1)}}) \Delta_{R^{(k)}}^\sigma f) \sigma)(x-y) \right|^2 \left(\frac{t}{t+|y|} \right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \\
& \lesssim 2^{-k\alpha} \sum_{\substack{R \in \mathcal{D}_{good} \\ \ell(R) \leq 2^{s-r-1}}} \mathcal{P}_\alpha(R^{(k)}, |\Delta_{R^{(k)}}^\sigma f| \sigma)^2 w(R) \\
& \lesssim 2^{-k\alpha} \sum_I \|\Delta_I^\sigma f\|_{L^2(\sigma)}^2 \frac{\sigma(I)}{|I|} \frac{w(I)}{|I|} \lesssim 2^{-k\alpha} \mathcal{A}_2^2 \|f\|_{L^2(\sigma)}^2,
\end{aligned}$$

where we reindexed the sum over R above. By the geometric decay in k , we deduce

$$\mathcal{J} \lesssim \mathcal{A}_2^2 \|f\|_{L^2(\sigma)}^2.$$

It remains only to analyze the contribution made to \mathcal{K} by the term $(\Delta_{R^{(k)}}^\sigma f)\mathbf{1}_{R^{(k-1)}}$. Our goal is to prove

$$(5.2) \quad \mathcal{K} \lesssim (\mathcal{A}_2 + \mathcal{B})^2 \|f\|_{L^2(\sigma)}^2.$$

To finish this, we here need an extra concept : *Stopping cubes*. For more applications and consequences associated with stopping cubes, we refer readers to the works [3], [4], [9]. The following argument is essentially taken from [5].

Stopping Cubes. We make the following construction of stopping cubes \mathcal{S} . Set \mathcal{S}_0 to be all the maximal dyadic children of Q_0 , which are in \mathcal{D}_f . Then set $\tau(S) = \mathbb{E}_S^\sigma f$, for $S \in \mathcal{S}_0$. In the recursive step, assuming that \mathcal{S}_k is constructed, for $S \in \mathcal{S}_k$, set $\text{ch}_S(S)$ to be the maximal subcubes $I \subset S$, $I \in \mathcal{D}_f$, such that either

- (a) $\mathbb{E}_I^\sigma f > 2\tau(S)$;
- (b) The first condition fails, and $\sum_{K \in \mathcal{W}_I} \mathcal{P}_\alpha(K, \mathbf{1}_S \sigma)^2 w(K) \geq C_0 \mathcal{P}^2 \sigma(I)$.

Then, define $\mathcal{S}_{k+1} := \bigcup_{S \in \mathcal{S}_k} \text{ch}_S(S)$, and for any $\dot{S} \in \text{ch}_S(S)$

$$\tau(\dot{S}) := \begin{cases} \mathbb{E}_{\dot{S}}^\sigma |f| & \mathbb{E}_{\dot{S}}^\sigma |f| > 2\tau(S), \\ \tau(S) & \text{otherwise .} \end{cases}$$

Finally, $\mathcal{S} := \bigcup_{k=0}^\infty \mathcal{S}_k$. Note that $\ell(\dot{S}) \leq 2^{-r-1}\ell(S)$ for all $\dot{S} \in \text{ch}_S(S)$. In particular, it follows that

$$(5.3) \quad \dot{S}^{(1)} \subset K, \text{ for some } K \in \mathcal{W}_S.$$

This holds since $\dot{S}^{(1)}$ is good, and strongly contained in S , so that Proposition 2.4 gives the implication above.

Some Notations . For any dyadic cube I , $S(I)$ will denote its father in \mathcal{S} , the minimal cube in \mathcal{S} that contains it. Note that there maybe the case $S(I) = I$. For any stopping cube S , $\mathcal{F}(S)$ will denote its father in the stopping tree, inductively, $\mathcal{F}^{k+1}S = \mathcal{F}(\mathcal{F}^k S)$.

The construction enjoys the following properties.

Lemma 5.1 ([9]). *The following statements hold.*

- (i) For all cubes I , $|\mathbb{E}_I^\sigma f| \lesssim \tau(S(I))$.
- (ii) The quasi-orthogonality bound holds :

$$(5.4) \quad \sum_{S \in \mathcal{S}} \tau(S)^2 \sigma(S) \lesssim \|f\|_{L^2(\sigma)}^2.$$

Applying the tool of stopping cubes, we can make the following decomposition.

$$\begin{aligned}
& \sum_{k=r+1}^{s-\log_2 \ell(R)} (\mathbb{E}_{R^{(k-1)}}^\sigma \Delta_{R^{(k)}}^\sigma f) \mathbf{1}_{R^{(k-1)}} \\
(5.5) \quad &= \sum_{m=1}^{\infty} \sum_{k=r+1}^{s-\log_2 \ell(R)} \mathbf{1}_{\mathcal{F}^m S(R^{(r)}) \subset S(R^{(k-1)})} (\mathbb{E}_{R^{(k-1)}}^\sigma \Delta_{R^{(k)}}^\sigma f) \mathbf{1}_{\mathcal{F}^m S(R^{(r)}) \setminus \mathcal{F}^{m-1} S(R^{(r)})} \\
& \quad + \sum_{k=r+1}^{s-\log_2 \ell(R)} (\mathbb{E}_{R^{(k-1)}}^\sigma \Delta_{R^{(k)}}^\sigma f) \mathbf{1}_{S(R^{(r)})} - \sum_{k=r+1}^{s-\log_2 \ell(R)} (\mathbb{E}_{R^{(k-1)}}^\sigma \Delta_{R^{(k)}}^\sigma f) \mathbf{1}_{S(R^{(k-1)}) \setminus R^{(k-1)}}.
\end{aligned}$$

Now, we are in the position to consider the contribution of \mathcal{K} , which is defined in the beginning of Section 5.4. Recall that

$$\mathcal{K} = \sum_{\substack{R \in \mathcal{D}_{good} \\ \ell(R) \leq 2^{s-r-1}}} \iint_{W_R} \int_{\mathbb{R}^n} \left| \sum_{k=r+1}^{s-\log_2 \ell(R)} \psi_t * ((\mathbf{1}_{R^{(k-1)}} \Delta_{R^{(k)}}^\sigma f) \sigma)(x-y) \right|^2 \left(\frac{t}{t+|y|} \right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t}.$$

Thus, \mathcal{K} is bounded by corresponding three parts, which are written as \mathcal{K}_{Glo} , \mathcal{K}_{Par} , and \mathcal{K}_{Loc} , respectively. We next shall estimate each one successively.

• **The Global Part.** First, we analyze the first term on the right of (5.5), It is worth noting that reindexing the corresponding sum is crucial to our estimates. To do this, we here borrow an idea from [5].

Fix a stopping cube S and integer m . Note that for a choice of constant $|c| \lesssim 1$, there holds

$$(5.6) \quad \sum_{k=r+1}^{s-\log_2 \ell(R)} \mathbf{1}_{\mathcal{F}^m S \subset S(R^{(k-1)})} (\mathbb{E}_{R^{(k-1)}}^\sigma \Delta_{R^{(k)}}^\sigma f) \mathbf{1}_{\mathcal{F}^m S \setminus \mathcal{F}^{m-1} S} = c \cdot \tau(\mathcal{F}^m S) \mathbf{1}_{\mathcal{F}^m S \setminus \mathcal{F}^{m-1} S}.$$

Note that the restriction is on $S(R^{(k-1)})$ above. We are going to reindex the sum above. Consider $\dot{S} \in \mathcal{S}$, and split integer $m = p + q$, where $p = \lceil m/2 \rceil$. Consider the subpartition of \dot{S} given by $\mathcal{P}(m, \dot{S}) = \{\dot{S} \in \mathcal{S} : \mathcal{F}^p \dot{S} = \dot{S}\}$. Now, for stopping cube S with $\mathcal{F}^q S = \dot{S}$, and good $R \in S$, we have $R \subset \dot{K}$ for some $\dot{K} \in \mathcal{W}_{\dot{S}}$, where $\dot{S} \in \mathcal{P}(m, \dot{S})$. Note that we have $R \subset \dot{K} \subset \dot{S}$. It follows from the goodness of R that the assumption of Lemma 4.2 holds for these three intervals. The above argument is saying that

$$\bigcup_{\substack{R: \text{good}, R \in S \\ \mathcal{F}^q S = \dot{S}}} R \subset \bigcup_{\dot{S} \in \mathcal{P}(m, \dot{S})} \bigcup_{\dot{K} \in \mathcal{W}_{\dot{S}}} \bigcup_{R: R \subset \dot{K}} R.$$

For each $\dot{S} \in \mathcal{S}$, using (5.6) to bound the sum over martingale differences, (4.3) and pivotal condition, we obtain

$$\begin{aligned}
& \sum_{R: \mathcal{F}^m S(R^{(r)}) = \dot{S}} \iint_{W_R} \int_{\mathbb{R}^n} \left| \sum_{\substack{k=r+1 \\ \dot{S} \subset S(R^{(k-1)})}}^{s-\log_2 \ell(R)} \mathbb{E}_{R^{(k-1)}}^\sigma \Delta_{R^{(k)}}^\sigma f \cdot \psi_t * (\sigma \mathbf{1}_{\dot{S} \setminus \mathcal{F}^{m-1} S})(x-y) \right|^2 \\
& \quad \times \left(\frac{t}{t+|y|} \right)^{n\lambda} dy \frac{w dx dt}{t^{n+1}} \\
& \lesssim \tau(\dot{S})^2 \sum_{\dot{S} \in \mathcal{P}(m, \dot{S})} \sum_{\dot{K} \in \mathcal{W}_{\dot{S}}} \sum_{R: R \in \dot{K}} \int_{\mathbb{R}^n} \iint_{W_R} |\psi_t * (\sigma \mathbf{1}_{\dot{S} \setminus \mathcal{F}^{p-1} \dot{S}})(x-y)|^2 \left(\frac{t}{t+|y|} \right)^{n\lambda} w dx \frac{dt dy}{t^{n+1}} \\
& \lesssim \tau(\dot{S})^2 \sum_{\dot{S} \in \mathcal{P}(m, \dot{S})} \sum_{\dot{K} \in \mathcal{W}_{\dot{S}}} \mathcal{P}_\alpha(\dot{K}, \sigma \mathbf{1}_{\dot{S}})^2 \sum_{R: R \in \dot{K}} \left[\frac{\ell(R)}{\ell(\dot{K})} \right]^\alpha w(R) \\
& \lesssim 2^{-m\alpha/2} \tau(\dot{S})^2 \sum_{\dot{S} \in \mathcal{P}(m, \dot{S})} \sum_{\dot{K} \in \mathcal{W}_{\dot{S}}} \mathcal{P}_\alpha(\dot{K}, \sigma \mathbf{1}_{\dot{S}})^2 w(\dot{K}) \\
& \lesssim 2^{-m\alpha/2} (\mathcal{A}_2 + \mathcal{B})^2 \tau(\dot{S})^2 \sigma(\dot{S}).
\end{aligned}$$

In the last step, we used Lemma 2.6. The sum over $\dot{S} \in \mathcal{S}$ is controlled by the quasi-orthogonality bound (5.4).

Let us next explain how to obtain the geometric factor. We can assume that $q > 2$. Now, $S(R) = S$ and $\mathcal{F}^q S = \dot{S}$. Write the stopping cubes between S and \dot{S} as

$$R \subset S = S_1 \subsetneq S_2 \subsetneq \cdots \subsetneq S_q := \dot{S}, \quad S_t \in \mathcal{S}, \quad 1 \leq t \leq q.$$

Observing (5.3), we have $S_{q-1} \subset \dot{K}$, for $\dot{K} \in \mathcal{W}_{\dot{S}}$ as above. Then, we have $\ell(R) \leq 2^{-q+1} \ell(\dot{K})$. Since $q \simeq m/2$, we obtain the geometric decay in m above.

• **The Paraproduct Estimate.** Next, we bound the second term on the right of (5.5). It is worth noting that the sum over the martingale differences is controlled by the stopping value $\tau(S)$. That is,

$$\left| \sum_{k=r+1}^{s-\log_2 \ell(R)} \mathbb{E}_{R^{(k-1)}}^\sigma \Delta_{R^{(k)}}^\sigma f \right| = |\mathbb{E}_{R^{(r)}}^\sigma f| \lesssim \tau(S).$$

Therefore, for fixed $S \in \mathcal{S}$, an application of testing condition (1.5) gives that

$$\begin{aligned}
& \sum_{\substack{R: \ell(R) \leq 2^{s-r-1} \\ S(R^{(r)}) = S}} \iint_{W_R} \int_{\mathbb{R}^n} \left| \sum_{k=r+1}^{s-\log_2 \ell(R)} \psi_t * (\mathbb{E}_{R^{(k-1)}}^\sigma \Delta_{R^{(k)}}^\sigma f \cdot \sigma \mathbf{1}_S)(x-y) \right|^2 \left(\frac{t}{t+|y|} \right)^{n\lambda} dy \frac{w dx dt}{t} \\
& \lesssim \tau(S)^2 \int_{\mathbb{R}^n} \iint_{\dot{S}} |\psi_t * (\sigma \mathbf{1}_S)(x-y)|^2 \left(\frac{t}{t+|y|} \right)^{n\lambda} w dx \frac{dt dy}{t^{n+1}} \lesssim \mathcal{B}^2 \tau(S)^2 \sigma(S).
\end{aligned}$$

And the quasi-orthogonality bound controls the sum over S .

• **The Local Bound.** Finally, let us estimate the third term on the right in (5.5). We will see that the stopping rule on the pivotal condition is now essential. Fix an $S \in \mathcal{S}$, and fix a $k \geq r$. In addition, fix a (good) cube \dot{R} which intersects S , and child \ddot{R} of \dot{R} .

Recall the construction of the stopping cubes. Since $S(\ddot{R}) = S$, this means that the cube \ddot{R} must fail the conditions of the stopping cube construction, in particular it must fail the pivotal stopping condition. Thus, by (4.3), for any $K \in \mathcal{W}_{\ddot{R}}$,

$$\sum_{\substack{R:S(\ddot{R})=S \\ R^{(k-1)}=\ddot{R}, R \subset K}} \iint_{W_R} \int_{\mathbb{R}^n} |\psi_t * (\mathbf{1}_{S(\ddot{R}) \setminus \ddot{R}})(x-y)|^2 \left(\frac{t}{t+|y|} \right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \lesssim 2^{-k\alpha/2} \mathcal{P}_\alpha(K, \sigma \mathbf{1}_S)^2 w(K).$$

Furthermore, since \ddot{R} is not a stopping cube,

$$\begin{aligned} & |\mathbb{E}_{\ddot{R}}^\sigma \Delta_{\ddot{R}}^\sigma f|^2 \sum_{\substack{R:S(\ddot{R})=S \\ R^{(k-1)}=\ddot{R}}} \iint_{W_R} \int_{\mathbb{R}^n} |\psi_t * (\mathbf{1}_{S(\ddot{R}) \setminus \ddot{R}})(x-y)|^2 \left(\frac{t}{t+|y|} \right)^{n\lambda} \frac{dy}{t^n} w dx \frac{dt}{t} \\ & \lesssim 2^{-k\alpha/2} (\mathcal{A}_2 + \mathcal{B})^2 |\mathbb{E}_{\ddot{R}}^\sigma \Delta_{\ddot{R}}^\sigma f|^2 \sigma(\ddot{R}). \end{aligned}$$

It is clear that we can sum over the various fixed quantities to complete the proof in this case.

So far, we have proved (5.2). Consequently, we complete the proof of sufficiency in Theorem 1.1. \square

REFERENCES

- [1] C. Fefferman, *Inequalities for strongly singular convolution operators*, Acta. Math., 124 (1970), 9-36.
- [2] T. Hytönen, *The sharp weighted bound for general Calderon-Zygmund operators*, Ann. Math., 175 (3) (2012), 1473-1506.
- [3] M. T. Lacey, *The two weight inequality for the Hilbert transform: a primer*, submitted (2013), available at <http://www.arxiv.org/abs/1304.5004>.
- [4] M. T. Lacey, *Two weight inequality for the Hilbert transform: a real variable characterization, II*, Duke Math. J., **163** (2014), no. 15, 2821-2840.
- [5] M. T. Lacey, K. Li, *Two weight norm inequalities for g function*, Math. Res. Lett., 21 (2014), no. 03, 521-536.
- [6] M. T. Lacey and H. Martikainen, *Local Tb theorem with L^2 testing conditions and general measures: square functions*, to appear in J. Anal. Math. available at <http://www.arxiv.org/abs/1308.4571>.
- [7] M. T. Lacey, E. T. Sawyer and I. Uriarte-Tuero, *A two weight inequality for the Hilbert transform assuming an energy hypothesis*, J. Funct. Anal., 263 (2012) 305-363.
- [8] J. Littlewood, R. Paley, *Theorems on Fourier series and power series, II*, Proc. Lond. Math. Soc. 42 (1936), 52-89.
- [9] M. T. Lacey, E. T. Sawyer, I. Uriarte-Tuero and C.-Y. Shen, *Two weight inequality for the Hilbert transform: a real variable characterization, I*, Duke Math. J., 163(2014) 2795-2820.
- [10] H. Martikainen and M. Mouroglou, *Square functions with general measures*, Proc. Amer. Math. Soc. 142 (2014), 3923-3931.
- [11] F. Nazarov, S. Treil and A. Volberg, *The Tb-theorem on non-homogeneous spaces*, Acta Math. 190 (2) (2003), 151-239.
- [12] E. Sawyer, *A characterization of a two weight norm inequality for maximal operators*, Studia Math., 75 (1982), 1-11.

- [13] E. M. Stein, *On some function of Littlewood-Paley and Zygmund*, Bull. Amer. Math. Soc. 67(1961)99-101.
- [14] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton Univ. Press, Princeton, NJ, 1970.
- [15] M. Wilson, *Weighted Littlewood-Paley theory and exponential-square integrability*, Lecture Notes in Mathematics, 1924, Springer, (2008).

MINGMING CAO, SCHOOL OF MATHEMATICAL SCIENCES, BEIJING NORMAL UNIVERSITY, LABORATORY OF MATHEMATICS AND COMPLEX SYSTEMS, MINISTRY OF EDUCATION, BEIJING 100875, PEOPLE'S REPUBLIC OF CHINA

E-mail address: m.cao@mail.bnu.edu.cn

KANGWEI LI, BCAM, BASQUE CENTER FOR APPLIED MATHEMATICS, MAZARREDO, 14. 48009 BILBAO BASQUE COUNTRY, SPAIN

E-mail address: kli@bcamath.org

QINGYING XUE, SCHOOL OF MATHEMATICAL SCIENCES, BEIJING NORMAL UNIVERSITY, LABORATORY OF MATHEMATICS AND COMPLEX SYSTEMS, MINISTRY OF EDUCATION, BEIJING 100875, PEOPLE'S REPUBLIC OF CHINA

E-mail address: qyxue@bnu.edu.cn