A CHARACTERIZATION OF TWO WEIGHT NORM INEQUALITY FOR LITTLEWOOD-PALEY $g^*_\lambda$-FUNCTION

MINGMING CAO, KANGWEI LI, AND QINGYING XUE

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

$$g^*_\lambda(f)(x) = \left( \int \int \int_{R_n^{n+1}} \left( \frac{t}{t + |x-y|} \right)^n |\nabla P_t f(y,t)|^2 \frac{dydt}{t^{n-1}} \right)^{1/2}, \quad \lambda > 1$$

where $P_t f(y,t) = p_t * f(x)$, $p_t(y) = t^{-n} p(y/t)$ and $p(x) = (1 + |x|^2)^{-(n+1)/2}$, $\nabla = \left( \frac{\partial}{\partial y_1}, \ldots, \frac{\partial}{\partial y_n}, \frac{\partial}{\partial t} \right)$. In this paper, we give a characterization of two weight norm inequality for $g^*_\lambda$-function. We show that, if $\lambda > 1$, then $g^*_\lambda$ is an isometry on $L^2(R^n)$. Moreover, in $L^2(R^n)$ condition holds, and a testing condition holds:

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

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

1. Introduction

It is well known that, $g^*_\lambda$-function was originated in the work of Littlewood and Paley [8] in the 1930’s. It can be used as a basic technical tool to prove the $L^p$-boundedness of various linear operators. Later, the classical $g^*_\lambda$ function of higher dimension was first introduced and studied by Stein [13] in 1961, a certain sublinear operator arises in Littlewood-Paley theory [11], [14]. It plays important roles in Harmonic analysis and other fields. Let $n \geq 2$, we recall its definition as follows:

$$g^*_\lambda(f)(x) = \left( \int \int \int_{R_n^{n+1}} \left( \frac{t}{t + |x-y|} \right)^n |\nabla P_t f(y,t)|^2 \frac{dydt}{t^{n-1}} \right)^{1/2}, \quad \lambda > 1$$

where $P_t f(y,t) = p_t * f(x)$, $p_t(y) = t^{-n} p(y/t)$ denotes the Poisson kernel and $\nabla = \left( \frac{\partial}{\partial y_1}, \ldots, \frac{\partial}{\partial y_n}, \frac{\partial}{\partial t} \right)$. It is easy to show that $g^*_\lambda$ is an isometry on $L^2(R^n)$. With much greater difficulty, it can be proved that for any $1 < p < \infty$, $\|g^*_\lambda(f)\|_{L^p(R^n)}$ and $\|f\|_{L^p(R^n)}$ are equivalent norms [13]. Moreover, in [13], Stein also proved that if $\lambda > 2$, then $g^*_\lambda$ 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 $g_\lambda^*$-function when $p > 1$ and $\lambda = 2/p$.

Recently, Lacey and Li [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 $(w, \sigma)$ on $\mathbb{R}^n$:

\begin{equation}
\|g(f\sigma)\|_{L^2(w)} \lesssim \mathcal{N}\|f\|_{L^2(\sigma)}
\end{equation}

holds if and only if $(w, \sigma)$ satisfies

\begin{equation}
\mathcal{A}_2^2 := \sup_Q \frac{\sigma(Q) w(Q)}{|Q|} < \infty;
\end{equation}

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

\begin{equation}
\int\int_{\tilde{Q}} |\nabla P_t (1_Q \sigma)(x, t)|^2 w dx \ t dt \lesssim \sigma(Q), \quad \tilde{Q} = Q \times [0, \ell(Q)].
\end{equation}

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 $g_\lambda^*$-function. Thus it is quite natural to ask if one can establish a characterization for the Littlewood-Paley $g_\lambda^*$-function. But, of course, $g_\lambda^*$-function is pretty much difficult to be dealt with, since additional integrals appears 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 $g_\lambda^*$-function with fractional Poisson kernels.

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

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

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 $g_\lambda^*$-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 $g_\lambda^*$-function.

\begin{equation}
\|g_\lambda^{*,\alpha}(f\sigma)\|_{L^2(w)} \lesssim \mathcal{N}\|f\|_{L^2(\sigma)}.
\end{equation}
In addition, we introduce the corresponding testing condition:

\[(1.5)\]

\[
\mathcal{B}^2 := \sup_{Q: \text{cubes in } \mathbb{R}^n} \frac{1}{\sigma(Q)} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \left( \frac{t}{t + |x - y|} \right)^{n\lambda} |\nabla P_t^{\alpha}(1_Q \sigma)(y, t)|^2 \frac{wdxdt}{t^{n-1}} \ dy < \infty.
\]

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 \(\sigma, w\) 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 \(g^*_\lambda\)-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 \(p^\alpha, 0 < \alpha \leq 1\).

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

**Definition 1.4.** For \(0 < \alpha \leq 1\), let \(\mathcal{C}_\alpha\) be the family of functions \(\varphi\) 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 \(f \in L^1_{\text{loc}}(\mathbb{R}^n)\) 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 \(g^*_\lambda\)-function is defined by setting, for all \(x \in \mathbb{R}^n\),

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

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 \(\sigma, w\) 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) \((w, \sigma)\) satisfies the \(A_2\) condition \((1.2)\);

(ii) the testing condition holds:

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

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.
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

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

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,

$$|\nabla P^\alpha_t(1_{Q^\sigma})(y, t)| \gtrsim \frac{\sigma(Q)}{t^{n+1}}. $$

Furthermore, for $x \in Q$,

$$g^{*, \alpha}_\lambda(1_{Q^\sigma})(x)^2 \geq \int_Q \int_{2\ell(Q)} \frac{t^2 + |x - y|^2}{(t + |x - y|^2)^{\frac{n\lambda}{2}}} \left| \int_Q \frac{nt^{\alpha+1} - \alpha t^{\alpha-1}|y - z|^2}{(t^2 + |y - z|^2)^{\frac{n+\alpha}{2} + 1}} \sigma dz \right|^2 \frac{dt}{t^{n-1}} dy \geq \int_Q \int_{2\ell(Q)} \frac{\sigma(Q)^2}{t^{3n+1}} dtdy \gtrsim \frac{\sigma(Q)^2}{|Q|^2}. $$

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

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

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.

In order to meet the demands of below, we will introduce some notions which appeared in [5].

2.2. 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 $\text{dist}(K, \partial I) \geq \ell(K)^\gamma \ell(I)^{1-\gamma}$.

2.3. 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 $\gamma$, there holds

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

Here, $J \Subset I$ means that $J \subset I$ and $2^r\ell(J) \leq \ell(I)$; in words, $J$ is strongly contained in $I$. 
2.4. The Pivotal Condition. The pivotal constant $P$ 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

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

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.

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

Proof. We follow the strategy used in [5]. Taking the large enough constant $C$ in Proposition 2.3 such that $\frac{a}{2} \geq n(\frac{2}{c-1})^2$. The $A_2$ condition and Proposition 2.3 give that

$$\sum_{j \in \mathbb{N}} \sum_{K \in W_{I_j}} \mathcal{P}_\alpha(K, 1_{I^0 \setminus CK})^2 w(K) \lesssim A_2^2 \sum_{j \in \mathbb{N}} \sum_{K \in W_{I_j}} \sigma(CK) \lesssim A_2^2 \sigma(I^0).$$

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

It is easy to verify

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

Therefore,

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

Since we have

$$\sum_{j \in \mathbb{N}} \sum_{K \in W_{I_j}} \int_K \int_{W_K} \frac{t}{t + |x - y|} ^{n\lambda} |\nabla \mathcal{P}_\alpha^\alpha(1_{I^0})(y, t)|^2 \frac{wdx dt}{t^{n-1}} dy$$

$$\leq \sum_{j \in \mathbb{N}} \sum_{K \in W_{I_j}} \int_{\mathbb{R}^n} \int_{W_K} \frac{t}{t + |x - y|} ^{n\lambda} |\nabla \mathcal{P}_\alpha^\alpha(1_{I^0})(y, t)|^2 \frac{wdx dt}{t^{n-1}} dy$$

$$\leq \int_{\mathbb{R}^n} \int_{I^0} \frac{t}{t + |x - y|} ^{n\lambda} |\nabla \mathcal{P}_\alpha^\alpha(1_{I^0})(y, t)|^2 \frac{wdx dt}{t^{n-1}} dy$$

$$\leq \mathcal{B}^2 \sigma(I^0),$$
and
\[
\sum_{j \in \mathcal{N}} \sum_{K \in W_{ij}} \int \int_{W_K} \left( \frac{t}{t + |x - y|} \right)^n |\nabla P_t^\alpha (1_{C_K}(y,t))|^2 \frac{w dx dt}{t^{n-1}} dy
\]
\[
\leq \sum_{j \in \mathcal{N}} \sum_{K \in W_{ij}} \int \int_{\mathbb{R}^n} \left( \frac{t}{t + |x - y|} \right)^n |\nabla P_t^\alpha (1_{C_K}(y,t))|^2 \frac{w dx dt}{t^{n-1}} dy
\]
\[
\leq \mathcal{B}^2 \sum_{j \in \mathcal{N}} \sum_{K \in W_{ij}} \sigma(C_K)
\]
\[
\lesssim \mathcal{B}^2 \sigma(I^0),
\]
the desired estimate follows immediately.

\[
\square
\]

3. The Probabilistic Reduction

Our next task is to simplify the proof of sufficiency. Before doing it, we first recall 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.

\[
(3.1) \quad \| g_{\psi,\lambda}^*(f \cdot \sigma) \|_{L^2(w)} \lesssim (A_2 + \mathcal{B}) \| f \|_{L^2(\sigma)},
\]

where
\[
g_{\psi,\lambda}^*(f)(x) = \left( \int \int_{\mathbb{R}^{n+1}} \left( \frac{t}{t + |x - y|} \right)^n |\psi_t \ast f(y)|^2 \frac{dy dt}{t^{n-1}} \right)^{1/2},
\]
\[
\psi_t(x) = \frac{1}{t^n} \psi\left( \frac{x}{t} \right)
\]
and \( \psi \) satisfies the following conditions:
1. \( |\psi(x)| \lesssim (1 + |x|)^{-n - \alpha} \);
2. \( |\psi(x) - \psi(y)| \lesssim |x - y|^\alpha (1 + |x|)^{-n - \alpha} \).

3.2. Random Dyadic Grids. We next will introduce the fundamental technique, random dyadic grids. 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,

\[
\mathcal{D} = \left\{ Q + \beta; Q \in \mathcal{D}_0 \right\} := \left\{ Q + \sum_{j:2^{-j} < \ell(Q)} 2^{-j} \beta_j; Q \in \mathcal{D}_0 \right\},
\]
where \( \mathcal{D}_0 \) is the standard dyadic grid of \( \mathbb{R}^n \).

**Good and Bad Cubes.** 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}) =
\[ \mathbb{E}_\beta(1_{\text{good}}(Q \setminus \beta)). \] Then \( \pi_{\text{good}} \) is independent of \( Q \in \mathcal{D}_0 \). And we can choose \( r \) large enough so that \( \pi_{\text{good}} > 0 \).

3.3. **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 \( R \setminus \beta \) are independent (see [2]). Therefore, one can write

\[
\|g_\lambda(f \cdot \sigma)\|_{L^2(w)}^2 = \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \left( \frac{t}{t + |y|} \right)^{n\lambda} |\psi_t * (f \cdot \sigma)(x - y)|^2 \frac{dydt}{t^{n+1}} wdx
\]

\[
= \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \left( \frac{t}{t + |y|} \right)^{n\lambda} |\psi_t * (f \cdot \sigma)(x - y)|^2 \frac{dy}{t} wdx \frac{dt}{t}
\]

\[
= \mathbb{E}_\beta \sum_{R \in \mathcal{D}_0} \int_{\mathbb{R}^n} \int_{W_{R \setminus \beta}} |\psi_t * (f \cdot \sigma)(x - y)|^2 \left( \frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t} wdx \frac{dt}{t}
\]

\[
= \frac{1}{\pi_{\text{good}}} \sum_{R \in \mathcal{D}_0} \mathbb{E}_\beta(1_{\text{good}}(R \setminus \beta)) \mathbb{E}_\beta \int_{\mathbb{R}^n} \int_{W_{R \setminus \beta}} |\psi_t * (f \cdot \sigma)(x - y)|^2 \left( \frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t} wdx \frac{dt}{t}
\]

\[
= \frac{1}{\pi_{\text{good}}} \sum_{R \in \mathcal{D}_0} \mathbb{E}_\beta \left( 1_{\text{good}}(R \setminus \beta) \int_{\mathbb{R}^n} \int_{W_{R \setminus \beta}} |\psi_t * (f \cdot \sigma)(x - y)|^2 \left( \frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t} wdx \frac{dt}{t} \right)
\]

\[
= \frac{1}{\pi_{\text{good}}} \mathbb{E}_\beta \sum_{R \in \mathcal{D}_0} \int_{\mathbb{R}^n} \int_{W_R} |\psi_t * (f \cdot \sigma)(x - y)|^2 \left( \frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t} wdx \frac{dt}{t}.
\]

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

\[
\sum_{R \in \mathcal{D}_0} \int_{\mathbb{R}^n} \int_{W_R} |\psi_t * (f \cdot \sigma)(x - y)|^2 \left( \frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t} wdx \frac{dt}{t} \leq C(\mathcal{A}_2 + \mathcal{B})^2 \|f\|_{L^2(\sigma)}^2.
\]

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

\[
\mathbb{E}^\sigma_Q f := \frac{1}{\sigma(Q)} \int_Q f d\sigma,
\]

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

\[
\Delta^\sigma_Q f := \sum_{Q' \in h(Q)} (\mathbb{E}^\sigma_{Q'} f - \mathbb{E}^\sigma_Q f) 1_{Q'}.
\]

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

\[
f = \sum_{Q \in \mathcal{D}} \Delta^\sigma_Q f + \sum_{Q \in \mathcal{D}} (\mathbb{E}^\sigma_Q f) 1_Q.
\]
Since \( \{ \Delta_Q^\sigma f \}_{Q \in \mathcal{D}} \) is a family of orthogonal, we have
\[
\| f \|_{L^2(\sigma)}^2 = \sum_{Q \in \mathcal{D}, \ell(Q) \leq 2^s} \| \Delta_Q^\sigma f \|_{L^2(\sigma)}^2 + \sum_{Q \in \mathcal{D}} \| (\mathbb{E}_Q^\sigma f) 1_Q \|_{L^2(\sigma)}^2.
\]

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
\[
(3.2) \quad \mathcal{K} := \sup_{f \in \mathcal{F}} \| g^\ast_f(f) \|_{L^2(w)} < \infty.
\]
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
\[
\sum_{R \in \mathcal{D}_{\text{good}}} \int \int_{W_R} \int_{\mathbb{R}^n} |\psi_t \ast (f \cdot \sigma)(x - y)|^2 \left( \frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t^n} \frac{wdx}{t} \frac{dt}{t}
\]
\[
\leq 2 \sum_{R \in \mathcal{D}_{\text{good}}} \int \int_{W_R} \int_{\mathbb{R}^n} |\psi_t \ast (f \chi_Q \cdot \sigma)(x - y)|^2 \left( \frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t^n} \frac{wdx}{t} \frac{dt}{t}
\]
\[
+ 2 \sum_{R \in \mathcal{D}_{\text{good}}} \int \int_{W_R} \int_{\mathbb{R}^n} |\psi_t \ast (g \cdot \sigma)(x - y)|^2 \left( \frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t^n} \frac{wdx}{t} \frac{dt}{t}.
\]
Taking
\[
f(x) = \frac{\ell(Q)^a}{(\ell(Q) + \text{dist}(x, Q))^{n + \alpha}} \chi_{Q' \setminus \sqrt{3}Q}.
\]
By (3.2) and using similar arguments as that in [4], we get
\[
\int_{Q'} \frac{\ell(Q)^{2a}}{(\ell(Q) + \text{dist}(z, Q))^{2(n + \alpha)}} d\sigma(z) w(Q) \lesssim \mathcal{K} + \mathcal{A}^2.
\]
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,
\[
\int \int_{W_R} \int_{\mathbb{R}^n} |\psi_t \ast (g \cdot \sigma)(x - y)|^2 \left( \frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t^n} \frac{wdx}{t} \frac{dt}{t} \leq C_n \int_{\mathbb{R}^n} \frac{\ell(R)^{2a}}{(\ell(R) + \text{dist}(z, R))^{2(n + \alpha)}} d\sigma(z) w(R) \| g \|_{L^2(\sigma)}^2
\]
\[
\leq C_n (\mathcal{K}^2 + \mathcal{A}^2) \varepsilon^2 \| f \|_{L^2(\sigma)}^2.
\]
Then by taking sufficiently large cube $Q$ such that $2^{(2s+2)n}C_n(\mathcal{H}^2 + \mathcal{A}_2^2)^2 < \mathcal{H}^2$. We finally get

$$
\sum_{R \in D_{\text{good}}} \int_{\mathbb{R}^n} \int_{W_R} |\psi_t * (f \cdot \sigma)(x - y)|^2 \left(\frac{t}{t + |y|}\right)^{n\lambda} \frac{dy}{t^n} \frac{wdx dt}{t} \leq 4 \mathcal{H}^2 \|f\|_{L^2(\sigma)},
$$

which means that we reduce the problem to prove \[3.2\]. Then by repeating the previous arguments, we further reduce the problem to estimate

$$
\sum_{R \in D_{\text{good}}} \int_{\mathbb{R}^n} \int_{W_R} \int_{\mathbb{R}^n} |\psi_t * (f \cdot \sigma)(x - y)|^2 \left(\frac{t}{t + |y|}\right)^{n\lambda} \frac{dy}{t^n} \frac{wdx dt}{t},
$$

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

$$
\sum_{R \in D_{\text{good}}} \int_{\mathbb{R}^n} \int_{W_R} \int_{\mathbb{R}^n} |\psi_t * (f \cdot \sigma)(x - y)|^2 \left(\frac{t}{t + |y|}\right)^{n\lambda} \frac{dy}{t^n} \frac{wdx dt}{t}.
$$

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

$$
\mathcal{H}_s := \sup_{\|f\|_{L^2(\sigma)} = 1} \sum_{R \in D_{\text{good}}} \int_{\mathbb{R}^n} \int_{W_R} \int_{\mathbb{R}^n} |\psi_t * (f \cdot \sigma)(x - y)|^2 \left(\frac{t}{t + |y|}\right)^{n\lambda} \frac{dy}{t^n} \frac{wdx dt}{t}.
$$

Similar arguments as the previous show that

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

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

$$
f = \sum_{Q \in D} \Delta_Q^\sigma f,
$$

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

$$
f_{\text{good}} = \sum_{Q \in D_{\text{good}}} \Delta_Q^\sigma f.
$$
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

\[
\mathcal{K}_s \leq 2 \sup_{f \in \mathcal{F}} \sum_{R \in \mathcal{D}_{\text{good}}^{\varepsilon} : R \subset [-2^s,2^s]^n : 2^{-s} \leq \ell(R) \leq 2^s} \int_{W_R} \int_{\mathbb{R}^n} |\psi_t \ast (f_{\text{good}} \cdot \sigma)(x - y)|^2 \left( \frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t^n} wdx \frac{dt}{t} + 2\mathcal{K}_s \|\tilde{g}\|_{L^2(\sigma)}^2.
\]

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

\[
\sum_{Q \in \mathcal{D}_{\text{good}}^{\varepsilon} : \ell(Q) \leq 2^s} \int_{W_R} \int_{\mathbb{R}^n} |\sum_{Q \in \mathcal{D}_{\text{good}}^{\varepsilon} : \ell(Q) \leq 2^s} \psi_t \ast (\Delta_Q^\sigma f \cdot \sigma)(x - y)|^2 \left( \frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t^n} wdx \frac{dt}{t} \lesssim (\mathcal{C}_2 + \mathcal{B}) \|f\|_{L^2(\sigma)}^2.
\]

4. Some Lemmas and Elementary Estimates

To prove the boundedness of \( g_\lambda^\sigma(\cdot, \sigma) \) 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

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

Proof. By the size condition, we obtain

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

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,

\[
|\psi_t \ast (\Delta_Q^\sigma f \cdot \sigma)(x - y)| \lesssim \frac{(\ell(R)^\alpha)}{(\ell(R) + d(Q, R))^{n+\alpha}} \|\Delta_Q^\sigma f\|_{L^1(\sigma)} \lesssim \frac{(\ell(R)^\alpha)}{(\ell(R) + d(Q, R))^{n+\alpha}} \mathcal{G}(Q)^{1/2} \|\Delta_Q^\sigma f\|_{L^2(\sigma)},
\]

and

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

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,

\[
\left(\int_{|y|>\frac{1}{2}d(Q,R)} |\psi_t \ast ((\Delta_Q^\sigma f)\sigma)(x-y)|^2 \left(\frac{t}{t+|y|}\right)^{n\lambda} \frac{dy}{t^n}\right)^{1/2} \leq \frac{\ell(R)^{\frac{n\lambda}{2} - \frac{n\alpha}{2}}}{(\ell(R) + d(Q,R))^{\frac{n\lambda}{2}}} \|\psi_t \ast ((\Delta_Q^\sigma f)\sigma)(x-\cdot)\|_{L^2(\mathbb{R}^n)}
\]

\[
\leq \frac{\ell(R)^{\frac{n\lambda}{2} - \frac{n\alpha}{2}} - \epsilon}{(\ell(R) + d(Q,R))^{\frac{n\lambda}{2}}} \|\psi_t\|_{L^2(\mathbb{R}^n)} \|\Delta_Q^\sigma f\|_{L^1(\sigma)}
\]

\[
\leq \frac{\ell(R)^{\frac{n\lambda}{2} - \frac{n\alpha}{2}} - \epsilon}{(\ell(R) + d(Q,R))^{\frac{n\lambda}{2}}} \sigma(Q)^{1/2} \|\Delta_Q^\sigma f\|_{L^2(\sigma)}
\]

\[
\leq \frac{\ell(R)^{\alpha/2} - \epsilon}{(\ell(R) + d(Q,R))^{n+\alpha}} \sigma(Q)^{1/2} \|\Delta_Q^\sigma f\|_{L^2(\sigma)}
\]

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

This completes the proof of (4.11).

\[\square\]

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

\[
(4.2)
\]

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

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

\[
\psi_t \ast (\Delta_Q^\sigma f \cdot \sigma)(x-y) = \int_Q (\psi_t(x-y-z) - \psi_t(x-y-z_Q)) \Delta_Q^\sigma f(z) d\sigma(z).
\]

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

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

For \(x \in R\) and \(z \in Q\), \(|x-z| \geq d(Q, R)\). We will consider two subcases.

Consider first the contribution made by those \(|y| \leq \frac{1}{2} d(Q, R)\). 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,

\[
|\psi_t \ast (\Delta_Q^\sigma f \cdot \sigma)(x-y)| \lesssim \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2}}{(\ell(R) + d(Q,R))^{n+\alpha}} \|\Delta_Q^\sigma f\|_{L^1(\sigma)}
\]

\[
\leq \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2}}{(\ell(R) + d(Q,R))^{n+\alpha}} \sigma(Q)^{1/2} \|\Delta_Q^\sigma f\|_{L^2(\sigma)}.
\]
and

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

It remains only to analyze the contribution made by those $|y| > \frac{1}{2}d(Q,R)$. When $|y| > \frac{1}{2}d(Q,R)$,

\[
\left( \int_{|y| > \frac{1}{2}d(Q,R)} |\psi_t * (\Delta_Q^\sigma f \cdot \sigma)(x-y)|^2 \left( \frac{t}{t+|y|} \right)^{n \lambda} \frac{dy}{t^n} \right)^{1/2} \lesssim \frac{\ell(R)^{n+\lambda-n}}{(\ell(R) + d(Q,R))^{n+\lambda}}.
\]

Therefore,

\[
\left( \int_{|y| > \frac{1}{2}d(Q,R)} |\psi_t * (\Delta_Q^\sigma f \cdot \sigma)(x-y)|^2 \left( \frac{t}{t+|y|} \right)^{n \lambda} \frac{dy}{t^n} \right)^{1/2} \lesssim \frac{\ell(R)^{n+\lambda-n}}{(\ell(R) + d(Q,R))^{n+\lambda}} \left( \int_{\mathbb{R}^n} \left( \int_{Q} \left( \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2}}{(\ell(R) + d(Q,R))^{n+\alpha}} |\Delta_Q^\sigma f(z)| d\sigma(z) \right)^2 dy \right)^{1/2}
\]

\[
\leq \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2}}{(\ell(R) + d(Q,R))^{n+\alpha}} \|\varphi\|_{L^2(\mathbb{R}^n)} \|\Delta_Q^\sigma f\|_{L^2(\sigma)}
\]

\[
\leq \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2}}{(\ell(R) + d(Q,R))^{n+\alpha}} \sigma(Q)^{1/2} \|\Delta_Q^\sigma f\|_{L^2(\sigma)}
\]

\[
\leq \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2}}{(\ell(R) + d(Q,R))^{n+\alpha}} \sigma(Q)^{1/2} \|\Delta_Q^\sigma f\|_{L^2(\sigma)}
\]

where $\varphi(z) = \frac{1}{(t+|z|)^{n+\alpha}}$. \hfill \Box

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

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

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

\[
\left( \sum_{Q, R \in \mathcal{D}} A_{Q,R}^\alpha x_Q y_R \right)^2 \lesssim \sigma_2^2 \sum_{Q \in \mathcal{D}} x_Q^2 \times \sum_{R \in \mathcal{D}} y_R^2.
\]

**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

\[
(\text{3.3}) \int_{\mathbb{R}^n} \int_{W_R} |\psi_t * (f \cdot \sigma)(x-y)|^2 wdx \left( \frac{t}{t+|y|} \right)^{n \lambda} \frac{dy}{t^n+1} \lesssim \left( \frac{\ell(R)}{\ell(K)} \right)^{\alpha} \mathcal{P}_\alpha(K, |f| \sigma)^2 w(R).
\]
Proof. First, we shall prove, for any \(z \not\in S\),
\[
\frac{\ell(R)^\alpha}{(\ell(R) + \text{dist}(z, R))^{n+\alpha}} \leq \left[ \frac{\ell(R)^\alpha}{\ell(K)^\alpha} \right]^{\alpha/2} \frac{\ell(K)^\alpha}{(\ell(K) + \text{dist}(z, K))^{n+\alpha}}.
\]
In fact, since \(\text{dist}(z, R) \geq \text{dist}(z, K) + \text{dist}(R, \partial K)\), we have
\[
\frac{\ell(R)^\alpha}{(\ell(R) + \text{dist}(z, R))^{n+\alpha}} = \left( \frac{\ell(R)}{\ell(K)} \right)^\alpha \frac{\ell(K)^\alpha}{(\ell(R) + \text{dist}(z, R))^{n+\alpha}} \lesssim \left( \frac{\ell(R)}{\ell(K)} \right)^{\alpha-\gamma} \frac{\ell(K)^\alpha}{(\ell(K) + \text{dist}(z, K))^{n+\alpha}}.
\]
Secondly, we turn to the estimate of (4.4). Decompose
\[
\int_{\mathbb{R}^n} |\psi_t \ast (f \cdot \sigma)(x - y)|^2 \left( \frac{t}{t + |y|} \right)^{n\lambda} dy \quad \lesssim \quad \int_{\mathbb{R}^n} \left( \int_{z: |y| \leq \frac{1}{2} \text{dist}(z, R)} |\psi_t(x - y - z)||f(z)| d\sigma(z) \right)^2 \left( \frac{t}{t + |y|} \right)^{n\lambda} dy
\]
\[
+ \int_{\mathbb{R}^n} \left( \int_{z: |y| > \frac{1}{2} \text{dist}(z, R)} |\psi_t(x - y - z)||f(z)| d\sigma(z) \right)^2 \left( \frac{t}{t + |y|} \right)^{n\lambda} dy
\]
\[
:= \mathcal{E}_1 + \mathcal{E}_2.
\]
For \((x, t) \in W_R\), and \(z \not\in S\), 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,
\[
\mathcal{E}_1 \lesssim \left( \frac{\ell(R)}{\ell(K)} \right)^\alpha \int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^n} \frac{\ell(K)^\alpha}{(\ell(K) + \text{dist}(z, K))^{n+\alpha}} |f(z)| d\sigma(z) \right)^2 \left( \frac{t}{t + |y|} \right)^{n\lambda} dy
\]
\[
\lesssim \left( \frac{\ell(R)}{\ell(K)} \right)^\alpha \mathcal{P}_\alpha(K, |f|\sigma)^2.
\]
If \(|y| > \frac{1}{2} \text{dist}(z, R)\), the inequality (4.4) and Young’s inequality imply that

\[
\mathcal{E}_2 \lesssim t^n \int_{\mathbb{R}^n} \left( \int_{|z-y| > \frac{1}{2} \text{dist}(z, R)} |\psi_t(x-y-z)| \frac{\ell(R)^{\alpha-n}}{(\ell(R) + \text{dist}(z, R))^{\alpha}} |f(z)|d\sigma(z) \right)^2 dy
\]

\[
\leq t^n \int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^n} |\psi_t(x-y-z)| \frac{\ell(R)^{\alpha}}{(\ell(R) + \text{dist}(z, R))^{n+\alpha}} |f(z)|d\sigma(z) \right)^2 dy
\]

\[
\lesssim t^n \left( \frac{\ell(R)}{\ell(K)} \right)^\alpha \int_{\mathbb{R}^n} \left( \int_{\mathbb{R}^n} |\psi_t(x-y-z)| \frac{\ell(K)^{\alpha}}{(\ell(K) + \text{dist}(z, K))^{n+\alpha}} |f(z)|d\sigma(z) \right)^2 dy
\]

\[
\leq t^n \left( \frac{\ell(R)}{\ell(K)} \right)^\alpha \|\psi_t\|^2_{L^2(\mathbb{R}^n)} \left( \int_{\mathbb{R}^n} \frac{\ell(K)^{\alpha}}{(\ell(K) + \text{dist}(z, K))^{n+\alpha}} |f(z)|d\sigma(z) \right)^2
\]

\[
\lesssim \left( \frac{\ell(R)}{\ell(K)} \right)^\alpha \mathcal{P}_\alpha(K, |f|\sigma)^2.
\]

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 \(\{Q; Q \in \mathcal{D}_{\text{good}}, \ell(Q) \leq 2^s\}\) into the following four parts to discuss respectively. 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

\[
\sum_{R \in \mathcal{D}_\text{good}} \int W_R \int_{\mathbb{R}^n} \left| \sum_{Q \in \mathcal{D}_\text{good}, \ell(Q) < \ell(R)} \psi_t * (\Delta_Q^\sigma f \cdot \sigma)(x-y)^2 \left( \frac{t}{t+|y|} \right)^{n\lambda dy} \frac{d\lambda}{t^n} \right. dx dt
\]

\[
\leq \sum_{R \in \mathcal{D}_\text{good}} \int W_R \left[ \sum_{Q \in \mathcal{D}_\text{good}, \ell(Q) < \ell(R)} \left( \int_{\mathbb{R}^n} |\psi_t * (\Delta_Q^\sigma f \cdot \sigma)(x-y)^2 \left( \frac{t}{t+|y|} \right)^{n\lambda dy} \right)^{1/2} dx \right]^{1/2} \frac{dt}{t}
\]

\[
\lesssim \sum_{R \in \mathcal{D}_\text{good}} \int W_R \left[ \sum_{Q \in \mathcal{D}_\text{good}, \ell(Q) < \ell(R)} \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2}}{(\ell(R) + d(Q, R))^{n+\alpha}} \sigma(Q)^{1/2} \|\Delta_Q^\sigma f\|_{L^2(\sigma)} \right]^2 \frac{dx dt}{t}
\]

\[
\lesssim \sum_{R \in \mathcal{D}_\text{good}} \left( \sum_{Q \in \mathcal{D}_\text{good}} A_Q^\sigma \|\Delta_Q^\sigma f\|_{L^2(\sigma)} \right)^2 \lesssim f^2 \|f\|^2_{L^2(\sigma)}.
\]

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

\[
\frac{\ell(R)^{\alpha}}{(\ell(R) + d(Q, R))^{n+\alpha}} \lesssim \frac{\ell(Q)^{\alpha/2} \ell(R)^{\alpha/2}}{D(Q, R)^{n+\alpha}}.
\]
Indeed, if \( \ell(Q) \leq d(Q, R) \), it is obvious that

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

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

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

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

\[
\sum_{R \in D_{good}, \ell(R) \leq 2^s} \int_{W_R} \int_{\mathbb{R}^n} \left| \sum_{Q \in D_{good}, \ell(Q) \geq \ell(R)} \psi_t * (\Delta_Q^{\sigma} f \cdot \sigma)(x-y) \right|^2 \left( \frac{t}{t + |y|} \right)^{n+1} w_{\sigma} dx dt = \sum_{R \in D_{good}, \ell(R) \leq 2^s} \int_{W_R} \left[ \sum_{Q \in D_{good}, \ell(Q) \geq \ell(R)} \left( \int_{\mathbb{R}^n} \psi_t * (\Delta_Q^{\sigma} f \cdot \sigma)(x-y) \left| \left( \frac{t}{t + |y|} \right)^{n+1} w_{\sigma} \right|^2 dx dt \right)^{1/2} \right]^2 w_{\sigma} dx dt \leq \sum_{R \in D_{good}} \left( \sum_{Q \in D_{good}} A_Q^{\sigma} \left\| \Delta_Q^{\sigma} f \right\|_{L^2(\sigma)} \right)^2 \lesssim \sigma^2 \left\| f \right\|_{L^2(\sigma)}^2.
\]

5.3. **The Case** \( \ell(R) \leq \ell(Q) \leq 2^s \ell(R) \) and \( d(Q, R) \leq \ell(R)^{1-\gamma} \). 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 by (4.1) and Lemma 4.1 again, we have

\[
\sum_{R \in D_{good}, \ell(R) \leq 2^s} \int_{W_R} \int_{\mathbb{R}^n} \left| \sum_{Q \in D_{good}, \ell(Q) \leq 2^s \ell(R)} \psi_t * (\Delta_Q^{\sigma} f \cdot \sigma)(x-y) \right|^2 \left( \frac{t}{t + |y|} \right)^{n+1} w_{\sigma} dx dt = \sum_{R \in D_{good}, \ell(R) \leq 2^s} \int_{W_R} \left[ \sum_{Q \in D_{good}, \ell(Q) \leq 2^s \ell(R)} \left( \int_{\mathbb{R}^n} \psi_t * (\Delta_Q^{\sigma} f \cdot \sigma)(x-y) \left| \left( \frac{t}{t + |y|} \right)^{n+1} w_{\sigma} \right|^2 dx dt \right)^{1/2} \right]^2 w_{\sigma} dx dt \leq \sum_{R \in D_{good}} \left( \sum_{Q \in D_{good}} A_Q^{\sigma} \left\| \Delta_Q^{\sigma} f \right\|_{L^2(\sigma)} \right)^2 \lesssim \sigma^2 \left\| f \right\|_{L^2(\sigma)}^2.
\]
where we reindexed the sum over \( R \) above. 

By the geometric decay in \( k \), we deduce 

\[
\mathcal{J} \lesssim C_2^2 \| f \|_{L^2(\sigma)}^2.
\]
It remains only to analyze the contribution made to $\mathcal{K}$ by the term $(\Delta_{R}^{g}f)\mathbf{1}_{R^{k-1}}$. Our goal is to prove

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

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 $S$. Set $S_0$ to be all the maximal dyadic children of $Q_0$, which are in $D_f$. Then set $\tau(S) = \mathbb{E}_{g}^{f}$, for $S \in S_0$. In the recursive step, assuming that $S_k$ is constructed, for $S \in S_k$, set $\text{ch}_{S}(S)$ to be the maximal subcubes $I \subset S$, $I \in D_f$, such that either

1. $E_{I}^{g}|f| > 2\tau(S)$;
2. The first condition fails, and $\sum_{K \in W_{i}}P_{a}(K, 1_{S\sigma})^{2}w(K) \geq C_{0}\mathcal{P}^{2}\sigma(I)$.

Then, define $S_{k+1} := \bigcup_{S \in S_k} \text{ch}_{S}(S)$, and for any $\hat{S} \in \text{ch}_{S}(S)$

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

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

$$
\hat{S}^{(1)} \subset K, \text{ for some } K \in W_{S}.
$$

This holds since $\hat{S}^{(1)}$ is good, and strongly contained in $S$, so that Proposition 2.3 gives the implication above.

**Some Notations.** For any dyadic cube $I$, $S(I)$ will denote its father in $S$, the minimal cube in $S$ 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 $I$, $|\mathbb{E}_{I}^{g}f| \lesssim \tau(S(I))$.

(ii) The quasi-orthogonality bound holds :

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

Applying the tool of stopping cubes, we can make the following decomposition.
of Lemma 4.

Note that we have

\[
S \mid F q K \text{ partition of } \text{Global Part.}
\]

The assumptions enter in this way. Fix a stopping cube \( S \) and integer \( m \). Note that for a choice of constant \( \tau \), there holds

\[
\sum_{k=r+1}^{s-\log_2 \ell(R)} (E^\sigma_{R(k-1)} \Delta^\sigma_{R(k)} f) 1_{S(R(k-1))} = c \cdot \tau^{m} S R m S \setminus m S - 1 S R m S \setminus m S.
\]

Thus, \( K \) is split into three parts. We next shall estimate each one successively.

- **The Global Part.** First, we analyze the first term on the right of (5.5), which concerns the case of \( S(R^{(r)}) \) and \( S(R^{(k)}) \) being separated in the \( S \) tree. The stopping values enter in this way. Fix a stopping cube \( S \) and integer \( m \). Note that for a choice of constant \( \tau \), there holds

\[
\sum_{k=r+1}^{s-\log_2 \ell(R)} (E^\sigma_{R(k-1)} \Delta^\sigma_{R(k)} f) 1_{S(R(k-1))} = c \cdot \tau^{m} S R m S \setminus m S - 1 S R m S \setminus m S.
\]

Note that the restriction is on \( S(R^{(k-1)}) \) above. We are going to reindex the sum above. Consider \( S \in S \), and split integer \( m = p + q \), where \( p = \lceil m/2 \rceil \). Consider the sub-partition of \( S \) given by \( \mathcal{P}(m, S) = \{ \hat{S} \in S : \mathcal{P}^2 \hat{S} = 0 \} \). Now, for stopping cube \( S \) with \( \mathcal{P}^2 S = \hat{S} \), and good \( R \in S \), we have \( R \subset \hat{K} \) for some \( \hat{K} \in \mathcal{W}_{\hat{S}} \), where \( \hat{S} \in P(m, S) \). Note that we have \( R \subset \hat{K} \subset \hat{S} \). It follows from the goodness of \( \hat{R} \) that assumption of of Lemma 3.2 holds for these three intervals. The above argument is saying that

\[
\bigcup_{R \in S \setminus S = \hat{S}} R \subset \bigcup_{R \in S \setminus S = \hat{S}} \bigcup_{R \in R \subset K} R.
\]

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

\[
\sum_{R \in S \setminus S} \left( \int_{W_R} \int_{R^n} \sum_{k=r+1}^{s-\log_2 \ell(R)} (E^\sigma_{R(k-1)} \Delta^\sigma_{R(k)} f) \cdot \psi_t \ast (\sigma 1_{S \setminus S m S - 1 S m S - 1 S}) (x - y) \right)^2
\]

\[
\times \left( \frac{t}{t + |y|} \right)^{n\lambda} \frac{wdx dt}{t^{n+1}}
\]

\[
\lesssim \tau (\hat{S})^2 \sum_{\hat{S} \in P(m, S)} \sum_{K \in W_{\hat{S}}} \sum_{R \in R \subset K} \int_{W_R} \int_{R^n} \left| \psi_t \ast (\sigma 1_{S \setminus S m S - 1 S}) (x - y) \right|^2 \left( \frac{t}{t + |y|} \right)^{n\lambda} \frac{wdx dt}{t^{n+1}}
\]

\[
\lesssim \tau (\hat{S})^2 \sum_{\hat{S} \in P(m, S)} \sum_{K \in W_{\hat{S}}} \mathcal{P}_\alpha (K, \sigma 1_{S})^2 \sum_{R \in R \subset K} \left[ \frac{\ell(R)}{\ell(K)} \right] w(R)
\]
orthogonality bound (5.4).

The sum over \(\hat{S} \in \mathcal{S}\) 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 \(\mathcal{F}^q S = \hat{S}\). Write the stopping cubes between \(S\) and \(\hat{S}\) as

\[
R \subset S = S_1 \subset S_2 \subset \cdots \subset S_q := \hat{S}, \quad S_t \in \mathcal{S}, \quad 1 \leq t \leq q.
\]

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

**The Paraproduct Estimate.** Next, we bound the second term on the right of (5.3). It is worth noting that the sum over the martingale differences is controlled by the stopping value \(\tau(S)\). That is,

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

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

\[
\sum_{R : \ell(R) \leq 2^{s-r-1}} \int_{W_R} \int_{\mathbb{R}^n} \left| \sum_{k=r+1}^{s-\log_2 \ell(R)} \psi_t \ast (\mathbb{E}_{R(k-1)}^{\sigma} \Delta_{R(k)}^{\sigma} f \cdot \sigma 1_S)(x - y) \right|^2 \left( \frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t^n} \frac{w dx dt}{t}
\]

\[
\lesssim \tau(S)^2 \int_{\mathbb{R}^n} \int_{\hat{S}} |\psi_t \ast (\sigma 1_S)(x - y)|^2 \left( \frac{t}{t + |y|} \right)^{n\lambda} \frac{w dx dt dy}{t^{n+1}} \lesssim \mathcal{B}^2 \tau(S)^2 \sigma(S).
\]

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 he 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 \(\hat{R}\) which intersects \(S\), and child \(\hat{R}\) of \(\hat{R}\).

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

\[
\sum_{R : S(\hat{R}) = S \quad \mathbb{R}^{k-1} = \hat{R}, R \subseteq K} \int_{W_R} \int_{\mathbb{R}^n} |\psi_t \ast (1_S(\hat{R}) \setminus \hat{R})(x - y)|^2 \left( \frac{t}{t + |y|} \right)^{n\lambda} \frac{dy}{t^{n+1}} \frac{w dx dt}{t} \lesssim 2^{-\kappa/2} \mathcal{P}_\alpha(K, \sigma 1_S)^2 w(K).
\]
Furthermore, since $\bar{R}$ is not a stopping cube,

$$\left| \mathbb{E}_R^\sigma \Delta_R^\sigma f \right|^2 \leq 2^{-k\alpha/2} (\mathcal{A}_2 + \mathcal{B})^2 \left| \mathbb{E}_R^\sigma \Delta_R^\sigma f \right|^2 \sigma(\bar{R}).$$

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

So far, we have proved \(\Box\).

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*, 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*, submitted, Duke Math. J., to appear, available at http://www.arxiv.org/abs/1201.4319.
[10] H. Martikainen and M. Mourgoglou, *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, School of Mathematical Sciences and LPMC, Nankai University, Tianjin 300071, People’s Republic of China
E-mail address: likangwei9@mail.nankai.edu.cn

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