Maximal function, Littlewood–Paley theory, Riesz transform and atomic decomposition in the multi-parameter flag setting

Yongsheng Han, Ming-Yi Lee, Ji Li and Brett D. Wick

Abstract

In this paper, we develop via real variable methods various characterizations of the Hardy spaces in the multi-parameter flag setting. These characterizations include those via the maximal function, the Littlewood–Paley square function and area integral, Riesz transforms and the atomic decomposition in the multi-parameter flag setting. The novel ingredients in this paper include (1) establishing appropriate discrete Calderón reproducing formulae in the flag setting and a version of the Plancherel–Pólya inequalities for flag quadratic forms; (2) introducing the maximal function and area function via flag Poisson kernels and flag version of harmonic functions; (3) developing an atomic decomposition via the finite speed propagation and area function in terms of flag heat semigroups. As a consequence of these real variable methods, we obtain the full characterizations of the multi-parameter Hardy space with the flag structure.

Contents

1 Introduction 3
1.1 Background and main result . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 3
1.2 Statement of results . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 4
1.3 Strategy of proofs of the main results . . . . . . . . . . . . . . . . . . . . . . . . 9

2 The equivalence of $\|g_F(f)\|_1$ and $\|S_F(f)\|_1$ 11
2.1 Discrete Calderón reproducing formula . . . . . . . . . . . . . . . . . . . . . . . . 11
2.2 Plancherel–Pólya type inequalities . . . . . . . . . . . . . . . . . . . . . . . . . . 21
2.3 The equivalence of $\|g_F(f)\|_1$ and $\|S_F(f)\|_1$ . . . . . . . . . . . . . . . . . . . . 23
   2.3.1 The proof that $\|S_F(f)\|_1 \lesssim \|g_F(f)\|_1$ . . . . . . . . . . . . . . . . . . . 28
   2.3.2 The proof that $\|g_F(f)\|_1 \lesssim \|S_F(f)\|_1$ . . . . . . . . . . . . . . . . . . . 24
3 Estimates of flag Poisson integrals and flag Riesz transforms

| 3.1 The estimate $\|S_F(u)\|_1 \lesssim \|u^*\|_1$ | 25 |
| 3.2 The estimate $\|u^*\|_1 \lesssim \|u^+\|_1$ | 30 |
| 3.3 The estimate $\|u^+\|_1 \lesssim \sum_{j=1}^{n+m} \sum_{k=1}^{m} \|R_{j,k}(f)\|_1 + \|f\|_1$ | 36 |

4 Characterizations of the flag Hardy spaces

| 4.1 The estimate $\|S_F(f)\|_1 \lesssim \|S_F(u)\|_1$ | 40 |
| 4.2 The equivalence $\|u^*\|_1 \approx \|M_F^*(f)\|_1$ | 43 |
| 4.3 The equivalence $\|u^+\|_1 \approx \|M_F^+(f)\|_1$ | 46 |

5 Atomic decompositions of flag Hardy spaces

| 5.1 Assumptions and Notations | 47 |
| 5.2 Atomic decomposition for $H^1_F(\mathbb{R}^n \times \mathbb{R}^m)$ | 48 |
| 5.3 Proof of the atomic decomposition | 49 |

6 The estimate $\sum_{j=1}^{n+m} \sum_{k=1}^{m} \|R_{j,k}(f)\|_1 + \|f\|_1 \lesssim \|S_F(f)\|_1$ via atomic decomposition 65

Keywords: maximal function, Littlewood–Paley square function, Lusin area integral, flag Riesz transforms, atomic decomposition, flag Hardy space

Mathematics Subject Classification 2010: 42B35, 42B20
1 Introduction

1.1 Background and main result

It was well-known that techniques from Fourier series and methods of complex analysis played a seminal role in the classical harmonic analysis. After many improvements, mostly achieved by the Calderón–Zygmund school, the real variable methods, such as, maximal function, Littlewood–Paley square function, Lusin area integral, singular integrals and atomic decomposition have come to more prominence.

For the classical one parameter case, the Hardy–Littlewood maximal function and Calderón–Zygmund singular integrals commute with the usual dilations on $\mathbb{R}^n$, $\delta \cdot x = (\delta x_1, \ldots, \delta x_n)$ for $\delta > 0$. This theory has been extensively studied and is by now well understood, see for example the monograph [32]. On the other hand, the product theory began with the strong maximal function and continued with the Marcienkiewicz multiplier. They commute with the multi-parameter dilations on $\mathbb{R}^n$, $\delta \cdot x = (\delta_1 x_1, \ldots, \delta_n x_n)$ for $\delta = (\delta_1, \ldots, \delta_n) \in \mathbb{R}^n_+$. Product theory has been studied, for example, in Gundy–Stein [15], R. Fefferman and Stein [6], R. Fefferman [7, 8, 9], Chang and R. Fefferman [1, 2, 3], Journé [22], and Pipher [30]. More precisely, R. Fefferman and Stein [6] studied the $L^p$ boundedness ($1 < p < \infty$) for the product convolution singular integral operators. Journé in [22] introduced non-convolution product singular integral operators, established the product $T1$ theorem and proved the $L^\infty \rightarrow \text{BMO}$ boundedness for such operators. The product Hardy space $H^p(\mathbb{R}^n \times \mathbb{R}^m)$ was first introduced by Gundy and Stein [15]. Chang and R. Fefferman [1, 2, 3] developed the atomic decomposition and established the dual space of the Hardy space $H^1(\mathbb{R}^n \times \mathbb{R}^m)$, namely the product $\text{BMO}(\mathbb{R}^n \times \mathbb{R}^m)$ space.

Note that the product theory has an explicit underlying multi-parameter product structure. However, when the underlying multi-parameter structure is not explicit, but only implicit, an appropriate $L^p$ theory, with $1 < p < \infty$, has only recently been developed. To be precise, in [24, 25], Muller, Ricci and Stein studied Marcinkiewicz multipliers on the Heisenberg group and obtained the $L^p$ boundedness for $1 < p < \infty$. This is surprising since these multipliers are invariant under a two parameter group of dilations on $\mathbb{C}^n \times \mathbb{R}$, while there is no two parameter group of automorphic dilations on $\mathbb{H}^n$. Moreover, they showed that Marcinkiewicz multipliers can be characterized by a convolution operator of the form $f \ast K$ where, $K$ is a flag convolution kernel. See Nagel, Ricci, and Stein [27] for flag singular integrals on Euclidean space and applications on certain quadratic CR submanifolds of $\mathbb{C}^n$. Nagel, Ricci, Stein, and Wainger [28, 29] further generalized the theory of singular integrals with flag kernels to a more general setting, namely, that of homogeneous groups. They proved that on a homogeneous group singular integral operators with flag kernels are bounded on $L^p, 1 < p < \infty$, and form an algebra. See also [12, 13, 14, 5] for related work.
At the endpoint estimates, it is natural to expect that certain Hardy space and BMO bounds are available. However, the lack of automorphic dilations underlies the failure of such multipliers to be in general bounded on the classical Hardy space and also precludes a pure product Hardy space theory on the Heisenberg group. This was the original motivation in [20] to develop a theory of flag Hardy spaces $H^p_{\text{flag}}$, $0 < p \leq 1$ on the Heisenberg group $\mathbb{H}^n$, that is, in a sense ‘intermediate’ between the classical Hardy spaces $H^p(\mathbb{H}^n)$ and the product Hardy spaces $H^p_{\text{product}}(\mathbb{C}^n \times \mathbb{R})$. The flag $H^p$ theory on the Heisenberg group developed in [20] includes the discrete version of the Calderón reproducing formula associated with the given multi-parameter structure and the Plancherel–Pólya type inequality in this setting. They established the flag Hardy spaces $H^p_{\text{flag}}(\mathbb{H}^n)$ via the discrete Littlewood–Paley square function, and then studied the dual space $CMO^p_{\text{flag}}(\mathbb{H}^n)$ using the corresponding Carleson measures. Calderón–Zygmund decomposition in terms of functions in $H^p_{\text{flag}}(\mathbb{H}^n)$ and interpolation has also been developed.

In [19] they showed that singular integrals with flag kernels, which include the aforementioned Marcinkiewicz multipliers, are bounded on $H^p_{\text{flag}}(\mathbb{H}^n)$, as well as from $H^p_{\text{flag}}(\mathbb{H}^n)$ to $L^p(\mathbb{H}^n)$, for $0 < p \leq 1$. Moreover, in [20] they constructed a singular integral with a flag kernel on the Heisenberg group, which is not bounded on the classical Hardy spaces $H^1(\mathbb{H}^n)$. Since, as pointed out in [20], the flag Hardy space $H^p_{\text{flag}}(\mathbb{H}^n)$ is contained in the classical Hardy space $H^p(\mathbb{H}^n)$, this counterexample implies that $H^1_{\text{flag}}(\mathbb{H}^n) \subsetneq H^1(\mathbb{H}^n)$.

It was well-known that both of the classical and product multi-parameter Hardy spaces can be characterized by the real variable methods, such as, Riesz transforms, maximal functions, the Littlewood–Paley square function and Lusin area integrals, as well as atomic decompositions. Thus, a natural question arises:

**Can one develop all these real variable methods in the multi-parameter flag structure setting?**

The main purpose of this paper is to address this question. To be precise, the main results of this paper develop the real variable methods, maximal functions, the Littlewood–Paley square function and the Lusin area integrals, Riesz transforms, as well as atomic decompositions, to the more complicated multi-parameter flag structure setting. As a consequence, using these real variable methods, we obtain the full characterizations of flag Hardy spaces.

1.2 Statement of results

To state the main results of this paper, one requires several definitions. To begin with, we first introduce the Littlewood–Paley square function and Lusin area integrals associated with the flag structure on $\mathbb{R}^n \times \mathbb{R}^m$. For this purpose, let $\psi^{(1)} \in \mathcal{S}(\mathbb{R}^{n+m})$ with supp $\psi^{(1)} \subset \left\{ \xi : \frac{1}{2} \leq |\xi| \leq 2 \right\}$
and
\[ \int_0^\infty |\psi^{(1)}(t\xi)|^2 \frac{dt}{t} = 1 \text{ for all } \xi \in \mathbb{R}^n \times \mathbb{R}^m \setminus \{(0, 0)\}. \]

Let \( \psi^{(2)} \in \mathcal{S}(\mathbb{R}^m) \) with \( \text{supp } \hat{\psi}^{(2)} \subset \{ \eta : \frac{1}{2} \leq |\eta| \leq 2 \} \) and
\[ \int_0^\infty |\hat{\psi}^{(2)}(s\eta)|^2 \frac{ds}{s} = 1 \text{ for all } \eta \in \mathbb{R}^m \setminus \{0\}. \]

We set
\[ \psi_{t,s}(x, y) = \psi^{(1)}_t *_{\mathbb{R}^m} \psi^{(2)}_s(x, y) := \int_{\mathbb{R}^m} \psi^{(1)}_t(x, y - z) \psi^{(2)}_s(z) dz, \quad (1.1) \]
where \( \psi^{(1)}_t(x, y) = t^{-(n+m)} \psi^{(1)}(\frac{x}{t}, \frac{y}{t}) \) and \( \psi^{(2)}_s(z) = s^{-m} \psi^{(2)}(\frac{z}{s}) \).

**Definition 1.1.** For \( f \in L^1(\mathbb{R}^{n+m}) \), the Littlewood–Paley square function \( g_F(f) \) is defined by
\[ g_F(f)(x, y) = \left\{ \int_0^\infty \int_0^\infty \left| \psi_{t,s} * f(x, y) \right|^2 \frac{dt}{t} \frac{ds}{s} \right\}^{1/2}, \]
where \( \psi_{t,s}(x, y) \) is the same as in (1.1).

We now introduce the Lusin area integral associated with the flag structure.

**Definition 1.2.** For \( f \in L^1(\mathbb{R}^{n+m}) \), the Lusin area integral of \( f \) is defined by
\[ S_F(f)(x, y) = \left\{ \int_{\mathbb{R}^{n+1}} \int_{\mathbb{R}^{n+1}} \chi_{t,s}(x - x_1, y - y_1) |\psi_{t,s} * f(x_1, y_1)|^2 \frac{dx_1 dt}{t^{n+m+1}} \frac{dy_1 ds}{s^{m+1}} \right\}^{1/2}, \]
where \( \chi_{t,s}(x, y) = \chi^{(1)}_t *_{\mathbb{R}^n} \chi^{(2)}_s(x, y), \chi^{(1)}_t(x, y) = t^{-(n+m)} \chi^{(1)}(\frac{x}{t}, \frac{y}{t}), \chi^{(2)}_s(z) = s^{-m} \chi^{(2)}(\frac{z}{s}), \chi^{(1)}(x, y) \) and \( \chi^{(2)}(z) \) are the indicator functions of the unit balls of \( \mathbb{R}^{n+m} \) and \( \mathbb{R}^m \), respectively.

To define maximal functions associated with the flag structure, we first introduce the following collection of functions that will be used to build the maximal functions.

**Definition 1.3.** Let \( \phi(x, y) = \phi^{(1)} *_{\mathbb{R}^n} \phi^{(2)}(x, y) \), where \( \phi^{(1)} \in \mathcal{S}(\mathbb{R}^{n+m}) \) and \( \phi^{(2)} \in \mathcal{S}(\mathbb{R}^m) \) satisfying
\[ \int_{\mathbb{R}^n \times \mathbb{R}^m} \phi^{(1)}(x, y) dx dy = \int_{\mathbb{R}^m} \phi^{(2)}(z) dz = 1. \]
We denote \( \mathcal{D}_F(\mathbb{R}^n \times \mathbb{R}^m) \) by the collection of all functions \( \phi \) that satisfy the above conditions.

The non-tangential maximal function is defined by
Definition 1.4. Let $\phi \in \mathcal{D}_F(\mathbb{R}^n \times \mathbb{R}^m)$. For each $f \in L^1(\mathbb{R}^{n+m})$, the non-tangential maximal function of $f$ is defined by

$$M_\phi^*(f)(x,y) = \sup_{(x_1,y_1,t,s) \in \Gamma(x,y)} |\phi_{t,s} \ast f(x_1,y_1)|,$$

where $\phi_{t,s}(x,y) = \phi_1(t \cdot R^m \phi_2(s \cdot z))$, $\phi_1(t \cdot x_1) = t^{-m} \phi_1(x_{1,t})$, $\phi_2(s \cdot y_1) = s^{-m} \phi_2(y_{1,s})$ and $\Gamma(x,y) = \{(x_1,y_1,t,s) : |x - x_1| \leq t, |y - y_1| \leq t + s\}$.

Similarly, we define the radial maximal function as follows.

Definition 1.5. Let $\phi \in \mathcal{D}_F(\mathbb{R}^n \times \mathbb{R}^m)$. For any $f \in L^1(\mathbb{R}^{n+m})$, the radial maximal function of $f$ is defined by

$$M_\phi^+(f)(x,y) = \sup_{t,s > 0} |\phi_{t,s} \ast f(x,y)|,$$

where $\phi_{t,s}(x,y)$ is defined as in Definition 1.4.

The main result of this paper is the following theorem.

Theorem 1.6. All the following norms

$$\|g_F(f)\|_1, \|S_F(f)\|_1, \|M_\phi^*(f)\|_1, \|M_\phi^+(f)\|_1$$

are equivalent for $f \in L^1(\mathbb{R}^{n+m})$.

As a consequence of Theorem 1.6, it is natural to introduce the flag Hardy space as follows.

Definition 1.7. The flag Hardy spaces $H^1_F(\mathbb{R}^n \times \mathbb{R}^m)$ is defined to be the collection of $f \in L^1(\mathbb{R}^{n+m})$ such that $g_F(f) \in L^1(\mathbb{R}^{n+m})$. The norm of $H^1_F(\mathbb{R}^n \times \mathbb{R}^m)$ is defined by

$$\|f\|_{H^1_F(\mathbb{R}^n \times \mathbb{R}^m)} = \|g_F(f)\|_1.$$

Remark 1.8. Note that the multi-parameter flag structure is involved in the Littlewood–Paley square function $g_F(f)$, the Lusin area integral of $S_F(f)$, the non-tangential and radial maximal function $M_\phi^*(f)$ and $M_\phi^+(f)$. Therefore, the multi-parameter flag structure is involved in the flag Hardy space. Moreover, the flag Hardy space $H^1_F(\mathbb{R}^n \times \mathbb{R}^m)$ can also be characterized by the maximal functions, the Littlewood–Paley square function and the Lusin area integrals. We would like to point out that the main results in this paper still hold for all $0 < p \leq 1$. The reason this paper only deals with the case $p = 1$ is that we would like to keep the length of this paper more reasonable and present the main ideas necessary for the case $0 < p \leq 1$. The extension to the case for $0 < p < 1$ is a lengthy technical exercise best left to the interested reader.
Characterizations of flag Hardy spaces

It was well-known that the atomic decomposition is a very important tool to study the boundedness of singular integrals for the classical one parameter and product multi-parameter Hardy spaces. However, the lack of the cancellation was a major difficulty in providing the atomic decomposition for the flag Hardy space. In this paper, we develop a new approach to provide an atomic decomposition for the flag Hardy space. To do this, we introduce the atom as follows.

Definition 1.9. Let \( \triangle^{(1)} \) be the Laplace on \( \mathbb{R}^{n+m} \) and \( \triangle^{(2)} \) be the Laplace on \( \mathbb{R}^m \) and let \( M \) be a positive integer. A function \( a(x_1, x_2) \in L^2(\mathbb{R}^n \times \mathbb{R}^m) \) is called a \((1, 2, M)\)-atom if it satisfies

1) supp \( a \subset \Omega \), where \( \Omega \) is an open set of \( \mathbb{R}^n \times \mathbb{R}^m \) with finite measure;

2) \( a \) can be further decomposed into

\[
a = \sum_{R = I_R \times J_R \in m(\Omega)} a_R
\]

where \( m(\Omega) \) is the set of all maximal dyadic subrectangles of \( \Omega \), and there exists a series of function \( b_R \) belonging to the range of \( \triangle^{(1)}k_1 \otimes_2 \triangle^{(2)}k_2 \) in \( L^2(\mathbb{R}^n \times \mathbb{R}^m) \), for each \( k_1, k_2 = 1, \ldots, M \), such that

(i) \( a_R = (\triangle^{(1)}M \otimes_2 \triangle^{(2)}M)b_R \);

(ii) supp \( (\triangle^{(1)}k_1 \otimes_2 \triangle^{(2)}k_2)b_R \subset 10R \), \( k_1, k_2 = 0, 1, \ldots, M \);

(iii) \( ||a||_{L^2(\mathbb{R}^n \times \mathbb{R}^m)} \leq |\Omega|^{-\frac{1}{2}} \) and \( k_1, k_2 = 0, 1, \ldots, M \),

\[
\sum_{R = I_R \times J_R \in m(\Omega)} \ell(I_R)^{-4M} \ell(J_R)^{-4M} \left\| \ell(I_R)^2 \Delta^{(1)}k_1 \otimes_2 \ell(J_R)^2 \Delta^{(2)}k_2 \right\|^2_{L^2(\mathbb{R}^n \times \mathbb{R}^m)} \leq |\Omega|^{-1}.
\]

The atomic decomposition for the flag Hardy space is given by the following definition.

Definition 1.10. Let \( M > \max\{n, m\}/4 \). The Hardy spaces \( H^1_{F, at, M}(\mathbb{R}^n \times \mathbb{R}^m) \) is defined as follows. For \( f \in L^2(\mathbb{R}^{n+m}) \), we say that \( f = \sum_j \lambda_j a_j \) is an atomic \((1, 2, M)\)-representation of \( f \) if \( \{\lambda_j\}_{j=0}^{\infty} \in \ell^1 \), each \( a_j \) is a \((1, 2, M)\)-atom, and the sum converges in \( L^2(\mathbb{R}^n \times \mathbb{R}^m) \). The space \( H^1_{F, at, M}(\Omega) \) is defined to be

\[
H^1_{F, at, M}(\mathbb{R}^n \times \mathbb{R}^m) = \{ f \in L^2(\mathbb{R}^{n+m}) : f \text{ has an atomic } (1, 2, M)\text{-representation} \}.
\]
with the norm
\[ \|f\|_{H^1_{F,at,M}} := \inf \left\{ \sum_{j=0}^{\infty} |\lambda_j| : f = \sum_{j=0}^{\infty} \lambda_j a_j \text{ is an atomic} \ (1,2,M)\text{-representation} \right\}. \]

The atomic Hardy space \( H^1_{F,at,M}(\mathbb{R}^n \times \mathbb{R}^m) \) is defined as the completion of \( \mathbb{H}^1_{F,at,M}(\mathbb{R}^n \times \mathbb{R}^m) \) with respect to this norm.

**Theorem 1.11.** Suppose that \( M > \max\{n,m\}/4 \). Then
\[ H^1_F(\mathbb{R}^n \times \mathbb{R}^m) = H^1_{F,at,M}(\mathbb{R}^n \times \mathbb{R}^m). \]
Moreover,
\[ \|f\|_{H^1_F(\mathbb{R}^n \times \mathbb{R}^m)} \approx \|f\|_{H^1_{F,at,M}(\mathbb{R}^n \times \mathbb{R}^m)}, \]
where the implicit constants depend only on \( M, n \) and \( m \).

As a consequence of Theorem 1.11, we obtain the Riesz transform characterization of the flag Hardy space. For this purpose, we first introduce the flag Riesz transforms. To do this, let \( R_j^{(1)} \) be the \( j \)-th Riesz transform on \( \mathbb{R}^{n+m} \), \( j = 1, 2, \ldots, n+m \), and \( R_k^{(2)} \) be the \( k \)-th Riesz transform on \( \mathbb{R}^m \), \( k = 1, 2, \ldots, m \), respectively. Namely, for each \( f \in L^1(\mathbb{R}^{n+m}) \)
\[ R_j^{(1)} f(x) = \text{p.v.} \ c_{n+m} \int_{\mathbb{R}^{n+m}} \frac{x_j - y_j}{|x - y|^{n+m+1}} f(y) dy, \quad x \in \mathbb{R}^{n+m} \]
and for each \( f \in L^1(\mathbb{R}^m) \)
\[ R_k^{(2)} f(z) = \text{p.v.} \ c_m \int_{\mathbb{R}^m} \frac{w_j - z_j}{|w - z|^{m+1}} f(w) dw, \quad z \in \mathbb{R}^m. \]

We set \( R_{j,k} = R_j^{(1)} \ast_{\mathbb{R}^m} R_k^{(2)} \), that is, \( R_{j,k} \) is the composition of \( R_j^{(1)} \) and \( R_k^{(2)} \) on \( \mathbb{R}^m \). Notice that the flag structure is involved in the Riesz transforms \( R_{j,k} \) for \( j = 1, 2, \ldots, n+m \) and \( k = 1, 2, \ldots, m \).

**Theorem 1.12.** \( f \in H^1_F(\mathbb{R}^n \times \mathbb{R}^m) \) if and only if \( \sum_{j=0}^{n+m} \sum_{k=1}^{m} \|R_j^{(1)} R_k^{(2)}(f)\|_1 + \|f\|_1 < \infty \). Moreover,
\[ \sum_{j=1}^{n+m} \sum_{k=1}^{m} \|R_j^{(1)} R_k^{(2)}(f)\|_1 + \|f\|_1 \approx \|f\|_{H^1_F(\mathbb{R}^n \times \mathbb{R}^m)}. \]

As a corollary to the above theorems, we conclude following:

**Corollary 1.13.** The following norms
\[ \|g_F(f)\|_1, \|S_F(f)\|_1, \|M^*_\phi(f)\|_1, \|M^+_\phi(f)\|_1, \sum_{j=1}^{n+m} \sum_{k=1}^{m} \|R_j^{(1)} R_k^{(2)}(f)\|_1 + \|f\|_1 \]
are equivalent for \( f \in L^1(\mathbb{R}^{n+m}) \).
1.3 Strategy of proofs of the main results

In Section 2 we prove the equivalence between \( \| g_F(f) \|_1 \) and \( \| S_F(f) \|_1 \). We recall that in the classical case to show that the \( L^p \) norms, with \( p \leq 1 \), of the Littlewood–Paley square function and Lusin area integral are equivalent, the crucial tool is the sup-inf inequality, namely the Plancherel–Pólya type inequality. In order to establish such an inequality, one needs to develop the discrete Calderón reproducing formula. See [17] for more details in the setting of spaces of homogeneous type in the sense of Coifman and Weiss. In the present flag setting, to obtain the equivalence between the square function and Lusin area integral, we will first establish a discrete Calderón reproducing formula and then prove the Plancherel–Pólya type inequality associated with the flag structure. As a consequence, the equivalence between \( \| g_F(f) \|_1 \) and \( \| S_F(f) \|_1 \) will follow.

As the second step, we provide the equivalence between \( \| S_F(f) \|_1 \) and \( \| M^*\phi(f) \|_1 \).

We will introduce the Lusin area integral, the non-tangential maximal function, and the radial maximal function via flag Poisson integrals. To do this, the flag Poisson kernel is defined by

\[
P(x, y) = P^{(1)} *_{\mathbb{R}^m} P^{(2)}(x, y) = \int_{\mathbb{R}^m} P^{(1)}(x, y - z)P^{(2)}(z)dz,
\]

where

\[
P^{(1)}(x, y) = \frac{c_{n+m}}{(1 + |x|^2 + |y|^2)^{(n+m+1)/2}} \quad \text{and} \quad P^{(2)}(z) = \frac{c_m}{(1 + |z|^2)^{(m+1)/2}}
\]

are the classical Poisson kernels on \( \mathbb{R}^{n+m} \) and \( \mathbb{R}^m \), respectively.

For any \( f \in L^1(\mathbb{R}^{n+m}) \), we define the flag Poisson integral of \( f \) by

\[
u(x, y, t, s) := P_{t,s} * f(x, y),
\]

where \( P_{t,s}(x, y) = P^{(1)} *_{\mathbb{R}^m} P^{(2)}_s(x, y) \).

Since \( P_{t,s}(x, y) \in L^1(\mathbb{R}^{n+m}) \), it is easy to see that \( u(x, y, t, s) \) is well-defined. Moreover, for any fixed \( t \) and \( s \), \( P_{t,s} * f \) is a bounded \( C^\infty \) function and the function \( u(x, y, t, s) \) is harmonic in \((x, y, t)\) and \((y, s)\), respectively.

We now define the flag Lusin area integral of \( u \) as follows.

**Definition 1.14.** For \( f \in L^1(\mathbb{R}^{n+m}) \) and \( u(x, y, t, s) = P_{t,s} * f(x, y) \), \( S_F(u) \), the flag Lusin area integral of \( u(x, y, t, s) \) is defined by

\[
S_F(u)(x, y) = \left\{ \int_{\mathbb{R}^{n+1}} \int_{\mathbb{R}^m+1} \chi_{t,s}(x - x_1, y - y_1)|t\nabla^{(1)}s\nabla^{(2)}u(x, y, t, s)|^2 \frac{dx_1 dt \, dy_1 ds}{t^{n+m+1} s^{m+1}} \right\}^{1/2},
\]

where \( \chi_{t,s}(x, y) \) is the same as in Definition 1.12, \( \nabla^{(1)} = (\partial_t, \partial_{x_1}, \ldots, \partial_{x_n}, \partial_{y_1}, \ldots, \partial_{y_m}) \) and \( \nabla^{(2)} = (\partial_s, \partial_{y_1}, \ldots, \partial_{y_m}) \).
Next, we define the non-tangential maximal function of $u$.

**Definition 1.15.** Let $f \in L^1(\mathbb{R}^{n+m})$, the non-tangential maximal function of $u$ is defined by

$$u^*(x,y) = \sup_{(x_1,y_1,t,s) \in \Gamma(x,y)} |P_{t,s} * f(x_1,y_1)|,$$

where $\Gamma(x,y) = \{(x_1,y_1,t,s) : |x-x_1| \leq t, |y-y_1| \leq t+s\}$.

Similarly, the radial maximal function of $u$ is given by the following

**Definition 1.16.** Let $f \in L^1(\mathbb{R}^{n+m})$, the radial maximal function of $u$ is defined by

$$u^+(x,y) = \sup_{t>0,s>0} |P_{t,s} * f(x,y)|.$$

In Section 3, we will show the following inequalities:

$$\|SF(u)\|_1 \lesssim \|u^*\|_1 \lesssim \|u^+\|_1 \lesssim \sum_{j=1}^{n+m} \sum_{k=1}^{m} \|R_{j,k}(f)\|_1 + \|f\|_1.$$

In Section 4, the following estimates will be concluded:

(I) $\|SF(f)\|_1 \lesssim \|SF(u)\|_1$,

(II) $\|u^*\|_1 \approx \|M^*_\Phi(f)\|_1$,

(III) $\|u^+\|_1 \approx \|M^+\Phi(f)\|_1$,

In Section 5, applying the atomic decomposition implies the following estimate

(IV) $\sum_{j=1}^{n+m} \sum_{k=1}^{m} \|R_{j,k}(f)\|_1 + \|f\|_1 \lesssim \|g_F(f)\|_1$.

Indeed, for each $f \in L^1(\mathbb{R}^{n+m})$, by $\|g_F(f)\|_1 \approx \|SF(f)\|_1$ together with the above estimates, we have the following chain of inequalities: for $f \in L^1(\mathbb{R}^{n+m})$,

$$\|SF(f)\|_1 \lesssim \|SF(u)\|_1 \lesssim \|u^*\|_1 \lesssim \|M^*_\Phi(f)\|_1 \lesssim \|u^+\|_1 \lesssim \|M^+\Phi(f)\|_1 \lesssim \|u^+\|_1 \lesssim \sum_{j=1}^{n+m} \sum_{k=1}^{m} \|R_{j,k}(f)\|_1 + \|f\|_1 \lesssim \|g_F(f)\|_1 \lesssim \|SF(f)\|_1.$$

This implies the main result Theorem 1.6; it also gives Theorem 1.12 and Corollary 1.13.
2 The equivalence of \( \|g_F(f)\|_1 \) and \( \|S_F(f)\|_1 \)

2.1 Discrete Calderón reproducing formula

We first recall the following test function space \( \widetilde{M}_d \) with the size and smoothness conditions on \( \mathbb{R}^d \) for arbitrary positive integer \( d \), which was introduced in [16].

**Definition 2.1.** Fix two exponents \( 0 < \beta < 1 \) and \( \gamma > 0 \). We say that \( f \) defined on \( \mathbb{R}^d \), belongs to \( \widetilde{M}_d(\beta, \gamma, r, x_0), r > 0 \) and \( x_0 \in \mathbb{R}^d \), if

\[
|f(x)| \leq C \frac{r^\gamma}{(r + |x - x_0|)^{d+\gamma}}, \quad (2.1)
\]

\[
|f(x) - f(x')| \leq C \left( \frac{|x - x'|}{r + |x - x_0|} \right)^\beta \frac{r^\gamma}{(r + |x - x_0|)^{d+\gamma}} \quad (2.2)
\]

for \( |x - x'| \leq \frac{r + |x - x_0|}{2} \). If \( f \in \widetilde{M}_d(\beta, \gamma, r, x_0) \), then the norm of \( f \) is defined by

\[
\|f\|_{\widetilde{M}_d(\beta, \gamma, r, x_0)} = \inf \{ C : (2.1) \text{ and } (2.2) \text{ hold} \}.
\]

Then we recall the test function space \( M_d(\beta, \gamma, r, x_0) \subset \widetilde{M}_d(\beta, \gamma, r, x_0) \) on \( \mathbb{R}^d \) with a cancellation condition.

**Definition 2.2.** Fix two exponents \( 0 < \beta < 1 \) and \( \gamma > 0 \). We say that \( f \) defined on \( \mathbb{R}^d \), belongs to \( M_d(\beta, \gamma, r, x_0) \), \( r > 0 \) and \( x_0 \in \mathbb{R}^d \), if \( f \in \widetilde{M}_d(\beta, \gamma, r, x_0) \) and

\[
\int_{\mathbb{R}^d} f(x) \, dx = 0.
\]

If \( f \in M_d(\beta, \gamma, r, x_0) \), then the norm of \( f \) is defined by

\[
\|f\|_{M_d(\beta, \gamma, r, x_0)} = \|f\|_{\widetilde{M}_d(\beta, \gamma, r, x_0)}.
\]

We now define the test function space on \( \mathbb{R}^{n+m} \times \mathbb{R}^m \) as follows.

**Definition 2.3.** Fix two exponents \( 0 < \beta < 1 \) and \( \gamma > 0 \). We say that \( f \) defined on \( \mathbb{R}^{n+m} \times \mathbb{R}^m \) belongs to \( \widetilde{M}_{(n+m)\times m}(\beta, \gamma, r_1, r_2, x_0, y_0, z_0), r_1, r_2 > 0 \) and \( (x_0, y_0, z_0) \in \mathbb{R}^{n+m} \times \mathbb{R}^m \), if for each fixed \( z \in \mathbb{R}^m, f(\cdot, \cdot, z) \in \widetilde{M}_{n+m}(\beta, \gamma, r_1, x_0, y_0) \) and for each \( (x, y) \in \mathbb{R}^{n+m}, f(x, y, \cdot) \in \widetilde{M}_{m}(\beta, \gamma, r_2, z_0) \) and satisfies the following conditions:

1. \( \|f(\cdot, \cdot, z)\|_{\widetilde{M}_{n+m}(\beta, \gamma, r_1, x_0, y_0)} \leq C \left( \frac{r_2^\gamma}{(r_2 + |z - z_0|)^{m+\gamma}} \right) \)

2. \( \|f(x, y, \cdot)\|_{\widetilde{M}_{m}(\beta, \gamma, r_2, z_0)} \leq C \left( \frac{r_1^\gamma}{(r_1 + |x - x_0| + |y - y_0|)^{n+m+\gamma}} \right) \)

3. \( \|f(\cdot, \cdot, z) - f(\cdot, \cdot, z')\|_{\widetilde{M}_{n+m}(\beta, \gamma, r_1, x_0, y_0)} \leq C \left( \frac{|z - z'|}{r_2 + |z - z_0|} \right)^\beta \frac{r_2^\gamma}{(r_2 + |z - z_0|)^{m+\gamma}} \)
We now prove the following discrete Calderón reproducing formula.

\[\text{for } |z - z'| \leq \frac{r_2 + |z - z_0|}{2},\]

(4) \[\|f(x, y, \cdot) - f(x', y', \cdot)\|_{\tilde{\mathcal{M}}_m(\beta, \gamma, r_2, z_0)} \leq C \left(\frac{|x - x'| + |y - y'|}{r_1 + |x - x_0| + |y - y_0|}\right)^{\beta} \frac{r_1^{\gamma}}{(r_1 + |x - x_0| + |y - y_0|)^{n + m + \gamma}}\]

for \(|x - x'| + |y - y'| \leq \frac{r_1 + |x - x_0| + |y - y_0|}{2}.

If \(f \in \widetilde{\mathcal{M}}(n+m) \times (\beta, \gamma, r_1, r_2, x_0, y_0, z_0)\), the norm of \(f\) is defined by

\[\|f\|_{\widetilde{\mathcal{M}}(n+m) \times (\beta, \gamma, r_1, r_2, x_0, y_0, z_0)} = \inf\{C : (1) - (4) \text{ hold}\}.

Similarly we have the definition for the test function space \(\mathcal{M}(n+m) \times (\beta, \gamma, r_1, r_2, x_0, y_0, z_0)\) as a subset in \(\widetilde{\mathcal{M}}(n+m) \times (\beta, \gamma, r_1, r_2, x_0, y_0, z_0)\) and satisfies the corresponding cancellation conditions for the variables \((x, y)\) and for \(z\), respectively.

We would like to point out that if \(f_1 \in \mathcal{M}_{n+m}(\beta, \gamma, r_1, x_0, y_0)\) and \(f_2 \in \mathcal{M}_m(\beta, \gamma, r_2, z_0)\) then \(f(x, y, z) = f_1(x, y) f_2(z) \in \mathcal{M}_{n+m}(\beta, \gamma, r_1, r_2, x_0, y_0, z_0)\).

The flag test function is defined by:

**Definition 2.4.** Let \(0 < \beta, \gamma < 1\), \(r_1, r_2 > 0\) and \(x_0 \in \mathbb{R}^n\), \(y_0 \in \mathbb{R}^m\). We say that a function \(f\) defined on \(\mathbb{R}^n \times \mathbb{R}^m\) belongs to the flag test function space \(\widetilde{\mathcal{M}}_{\text{flag}}(\beta, \gamma, r_1, r_2, x_0, y_0)\) if there exists a function \(f^\sharp(x, y, z) \in \widetilde{\mathcal{M}}_{n+m}(\beta, \gamma, r_1, r_2, x_0, y_0, z_0)\) such that

\[f(x, y) = \int_{\mathbb{R}^m} f^\sharp(x, y - z, z) dz.

If \(f \in \widetilde{\mathcal{M}}_{\text{flag}}(\beta, \gamma, r_1, r_2, x_0, y_0)\), the norm of \(f\) is defined by

\[\|f\|_{\widetilde{\mathcal{M}}_{\text{flag}}(\beta, \gamma, r_1, r_2, x_0, y_0)} = \inf \left\{\|f^\sharp\|_{\widetilde{\mathcal{M}}_{n+m}(\beta, \gamma, r_1, r_2, x_0, y_0, z_0) : f(x, y) = \int_{\mathbb{R}^m} f^\sharp(x, y - z, z) dz} \right\}.

Similarly we can define the test function space \(\mathcal{M}_{\text{flag}}(\beta, \gamma, r_1, r_2, x_0, y_0)\) with the flag cancellation condition as a subset in \(\widetilde{\mathcal{M}}_{\text{flag}}(\beta, \gamma, r_1, r_2, x_0, y_0)\), which is projected from the product test function space \(\mathcal{M}_{n+m}(\beta, \gamma, r_1, r_2, x_0, y_0, z_0)\).

Observe that the flag structure is involved in the flag test function space \(\mathcal{M}_{\text{flag}}(\beta, \gamma, r_1, r_2, x_0, y_0)\).

We now prove the following discrete Calderón reproducing formula.

**Theorem 2.5.** Let \(\psi_{t,s}\) be the same as in (1.1). Then there exist functions \(\phi_{j,k}(x, y, x_I, y_J) \in \mathcal{M}_{\text{flag}}(\beta, \gamma, 2^{-j}, 2^{-k}, x_I, y_I)\) and a fixed large integer \(N\) such that for \(f(x, y) = \int_{\mathbb{R}^m} f_1(x, y - z) f_2(z) dz\) with \(f_1 \in \mathcal{M}_{n+m}(\beta, \gamma, r_1, r_2, x_0, y_0)\) and \(f_2 \in \mathcal{M}_m(\beta, \gamma, r_3, z_0)\),

\[f(x, y) = \sum_j \sum_k \sum_l \sum_j |I||J| \phi_{j,k}(x, y, x_I, y_J) \int_{2^{-j-N}}^{2^{-j-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} \psi_{t,s} \ast f(x_I, y_J) dt ds, \quad (2.3)\]
where the series converges in $L^2(\mathbb{R}^{n+m})$ and in the flag test function space, $I \subset \mathbb{R}^n$ and $J \subset \mathbb{R}^m$ are dyadic cubes with side-lengths $\ell(I) = 2^{-j-N}$ and $\ell(J) = 2^{-(j+k-N)}$, and $x_I$ and $y_J$ are any fixed points in $I$ and $J$, respectively.

Note that for each $f \in L^1(\mathbb{R}^{n+m})$, $f \in (\mathcal{M}_{flag}(\beta, \gamma, r_1, r_2, x_0, y_0))'$. As a consequence of Theorem 2.5 by duality, if $\psi$ is the same as in (1.1) and $f \in L^1(\mathbb{R}^{n+m})$,

$$\langle f, \psi \rangle = \left\langle \sum_{j} \sum_{k} \sum_{I} \sum_{J} |I| |J| \phi_{j,k}(\cdot, x_I, y_J) \int_{\mathbb{R}^n} \psi_{t,s} * f(x_I, y_J) \frac{dt}{t} \frac{ds}{s}, \psi \right\rangle.$$  \hspace{1cm} (2.4)

Remark 2.6. Indeed, the series in the right-hand side of (2.4) converges in the test function space $\mathcal{M}_{flag}(\beta, \gamma, r_1, r_2, x_0, y_0)$ and in the distribution space $(\mathcal{M}_{flag}(\beta, \gamma, r_1, r_2, x_0, y_0))'$, the dual of $\mathcal{M}_{flag}(\beta, \gamma, r_1, r_2, x_0, y_0)$. However, the proofs of such results are a little bit complicated. In this paper, we focus only on the Hardy space with $p = 1$. Thus, for our purpose, we only need the convergence in the distribution sense as given in (2.4).

Proof of Theorem 2.5. To show Theorem 2.5, observe that if $\psi_{t,s}$ are as in (1.1), by taking the Fourier transform, we have the following Calderón reproducing formula, namely for all $f \in L^2(\mathbb{R}^{n+m})$,

$$f(x, y) = \int_0^\infty \int_0^\infty \psi_{t,s} * \psi_{t,s} * f(x, y) \frac{dt}{t} \frac{ds}{s},$$

where the series converges in $L^2(\mathbb{R}^{n+m})$.

Suppose that $f \in \mathcal{M}_{flag}(\beta, \gamma, r_1, r_2, x_0, y_0)$ with $f(x, y) = \int_{\mathbb{R}^n} f_1(x, y - z) f_2(z) dz$ where $f_1 \in \mathcal{M}_{n+m}(\beta, \gamma, r_1, r_2, x_0, y_0)$ and $f_2 \in \mathcal{M}(\beta, \gamma, r_3, z_0)$. Applying Coifman’s decomposition of the identity yields

$$f(x, y) = \int_0^\infty \int_0^\infty \psi_{t,s} * \psi_{t,s} * f(x, y) \frac{dt}{t} \frac{ds}{s} \hspace{1cm} (2.5)$$

$$= \sum_{j,k} \sum_{I,J} \int_{2^{-k-N}}^{2^{-k-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} \psi_{t,s}(x - u, y - v) \psi_{t,s} * f(u, v) du dv \frac{dt}{t} \frac{ds}{s}$$

$$= \sum_{j,k} \sum_{I,J} \int_{2^{-k-N}}^{2^{-k-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} \psi_{t,s}(x - u, y - v) du dv \psi_{t,s} * f(x_I, y_J) \frac{dt}{t} \frac{ds}{s} + \mathcal{R}(f)(x, y),$$

where $I \subset \mathbb{R}^n$ and $J \subset \mathbb{R}^m$ are dyadic cubes with side-lengths $\ell(I) = 2^{-j-N}$ and $\ell(J) = 2^{-(j+k-N)}$, $x_I$ and $y_J$ are any fixed points in $I$ and $J$, respectively, and

$$\mathcal{R}(f)(x, y)$$
\[
\sum_{j,k} \sum_{I,J} \int_{2^{-k-N}}^{2^{-j-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} \psi_{t,s}(x-u,y-v)[\psi_{t,s} \ast f(u,v) - \psi_{t,s} \ast f(x_I,y_J)] \, dv \, dt \int_{I \times J} ds.
\]

Observing that
\[
\psi_{t,s} \ast f(u,v) - \psi_{t,s} \ast f(x_I,y_J)
\]
\[
= \int_{\mathbb{R}^{n+m}} [\psi_{t,s}(u-u',v-v') - \psi_{t,s}(x_I-u',y_J-v')] f(u',v') \, du' \, dv',
\]
we can write
\[
\mathcal{R}(f)(x,y) = \int_{\mathbb{R}^{n+m}} \sum_{j,k} \sum_{I,J} \int_{2^{-k-N}}^{2^{-j-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} \psi_{t,s}(x-u,y-v) \times [\psi_{t,s}(u-u',v-v') - \psi_{t,s}(x_I-u',y_J-v')] \, dv \, dt \int_{I \times J} ds.
\]

Note that \(\psi_{t,s}(x,y) = \int_{\mathbb{R}^m} \psi_t^{(1)}(x,y-z) \psi_s^{(2)}(z) \, dz\) and \(f(x,y) = \int_{\mathbb{R}^m} f^z(x,y-w,w) \, dw\) with \(f^z \in \mathcal{M}_{(n+m) \times m}(\beta, \gamma, r_1, r_2, x_0, y_0, z_0)\). Thus,
\[
\mathcal{R}(f)(x,y) = \int_{\mathbb{R}^{n+m}} \int_{\mathbb{R}^m} \sum_{j,k} \sum_{I,J} \int_{2^{-k-N}}^{2^{-j-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} \psi_{t,s}^{(1)}(x-u,y-v-z) \psi_{s}^{(2)}(z) \times [\psi_{t,s}^{(1)}(u-u',v-v') - \psi_{t,s}^{(1)}(x_I-u',y_J-w-v')] \psi_{s}^{(2)}(w) \, dw \, dv \, dt \, ds \int_{I \times J}.
\]

We now define
\[
\mathcal{R}^z(f^z)(x,y,z) := \int_{\mathbb{R}^{n+m}} \int_{\mathbb{R}^m} \sum_{j,k} \sum_{I,J} \int_{2^{-k-N}}^{2^{-j-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} \psi_{t,s}^{(1)}(x-u,y-v-z) \times [\psi_{t,s}^{(1)}(u-u',v-v') - \psi_{t,s}^{(1)}(x_I-u',y_J-w-v')] \psi_{s}^{(2)}(w) \, dw \, dv \, dt \, ds \int_{I \times J}.
\]

and then we can rewrite
\[
\mathcal{R}(f)(x,y) = \int_{\mathbb{R}^m} \mathcal{R}^z(f^z)(x,y-z,z) \, dz.
\]

We need an estimate on \(\mathcal{R}^z(f^z)\) that is contained in the following lemma:

**Lemma 2.7.** If \(f^z(x,y,z) = f_1(x,y) f_2(z)\) with \(f_1 \in \mathcal{M}_{(n+m)}(\beta, \gamma, r_1, r_2, x_0, y_0)\) and \(f_2 \in \mathcal{M}_m(\beta, \gamma, r_3, z_0)\) then \(\mathcal{R}^z(f^z)(x,y,z) \in \mathcal{M}_{(n+m) \times m}(\beta, \gamma, r_1, r_2, x_0, y_0, z_0)\) and
\[
||\mathcal{R}^z(f^z)||_{\mathcal{M}_{(n+m) \times m}(\beta, \gamma, r_1, r_2, x_0, y_0, z_0)} \leq C 2^{-N} ||f||_{\mathcal{M}_{(n+m) \times m}(\beta, \gamma, r_1, r_2, x_0, y_0, z_0)}.
\]
Assuming Lemma 2.7 for the moment, it implies that if \( f(x, y) = \int_{\mathbb{R}^m} f_1(x, y - z) f_2(z) dz \) with \( f_1 \in \mathcal{M}_{n+m}(\beta, \gamma, r_1, x_0, y_0) \) and \( f_2 \in \mathcal{M}_m(\beta, \gamma, r_2, z_0) \) then \( \mathcal{R}(f) \in \mathcal{M}_{\text{flag}}(\beta, \gamma, r_1, r_2, x_0, y_0) \) and

\[
||\mathcal{R}(f)||_{\mathcal{M}_{\text{flag}}} \leq C 2^{-N} ||f||_{\mathcal{M}_{\text{flag}}},
\]

Note that for \( 0 < \beta, \gamma < 1, t \sim 2^{-j}, s \sim 2^{-k}, x_I \in I \) and \( y_J \in J \),

\[
\frac{1}{|I||J|} \int_{I \times J} \psi_{t,s}(x - u, y - w) du dw = \int_{\mathbb{R}^m} \frac{1}{|I||J|} \int_{I \times J} \psi_{t}^{(1)}(x - u, y - w) du dw \psi_{s}^{(2)}(z) dz
\]

is a flag test function in \( \mathcal{M}_{\text{flag}}(\beta, \gamma, 2^{-j}, 2^{-k}, x_I, y_J) \) with

\[
\psi_{s}^{(2)}(z) \in \mathcal{M}_m(\beta, \gamma, 2^{-k}, 0).
\]

Therefore, if \( N \) is chosen large enough and by Lemma 2.7 we have

\[
\left( \sum_{i=0}^{\infty} \mathcal{R}^i \left[ \int_{I \times J} \psi_{t,s}(\cdot - u, \cdot - w) du dw \right] \right)(x, y) = |I||J| \phi_{j,k}(x, y, x_I, y_J),
\]

where \( \phi_{j,k}(x, y, x_I, y_J) \in \mathcal{M}_{\text{flag}}(\beta, \gamma, 2^{-j}, 2^{-k}, x_I, y_J) \) and hence,

\[
f(x, y) = \sum_{j} \sum_{k} \sum_{I} \sum_{J} |I||J| \phi_{j,k}(x, y, x_I, y_J) \int_{2^{-j} - N}^{2^{-j} + N} \int_{2^{-k} - N}^{2^{-k} + N} (\psi_{t,s} * f)(x_I, y_J) \frac{dt}{t} \frac{ds}{s}.
\]

This then gives the proof of Theorem 2.5 (assuming Lemma 2.7). \( \square \)

We now turn to demonstrating Lemma 2.7. To do this, we introduce the following definition and key estimates.

**Definition 2.8.** Let \( T \) be a bounded linear operator on \( L^2(\mathbb{R}^d) \) associated with a kernel \( K(x, y) \) given by

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

where \( K(x, y) \) satisfies the following conditions: There exists a constant \( C > 0 \) such that

(i) \( |K(x, y)| \leq C|x - y|^{-d} \),

(ii) \( |K(x, y) - K(x', y)| \leq C|x - x'||x - y|^{-d-1} \quad \text{if } |x - x'| \leq |x - y|/2, \)

(iii) \( |K(x, y) - K(x, y')| \leq C|y - y'||x - y|^{-d-1} \quad \text{if } |y - y'| \leq |x - y|/2, \)

(iv) \( |K(x, y) - K(x', y) - K(x, y')| \leq C|x - x'||y - y'||x - y|^{-d-2} \)
if \(|x - x'| \leq |x - y|/2\) and \(|y - y'| \leq |x - y|/2\).

We denote by \(\|K\|_{\mathbb{R}^d}\) the smallest constant \(C\) that satisfies (i)–(iv) above. The operator norm of \(T\) is defined by \(\|T\| = \|T\|_{L^2 \to L^2} + \|K\|_{\mathbb{R}^d}\). Here we use \(d\) to denote arbitrary positive integer.

We would like to point out that the classical Calderón–Zygmund kernel \(K(x, y)\) only needs to satisfy the conditions (i), (ii) and (iii). For our purpose, namely the boundedness of operators on test function space, condition (iv) is required, see [16] for the classical case. More precisely, we have the following:

**Lemma 2.9.** Suppose that \(T\) is an operator as in Definition 2.8 and \(T(1) = T^*(1) = 0\). Then \(T\) is bounded on the test function space \(M_d(\alpha, \beta, r, x_0)\) for \(0 < \alpha, \beta < 1, r > 0\) and \(x_0 \in \mathbb{R}^d\). Moreover, there exists a constant \(C\) such that

\[
\|T(f)\|_{M_d(\alpha, \beta, r, x_0)} \leq C\|T\| \|f\|_{M_d(\alpha, \beta, r, x_0)}.
\]

See [16] for the definition of \(T(1) = T^*(1) = 0\) and the proof of Lemma 2.9. We now define the product operator as follows.

**Definition 2.10.** The operator \(T\) is said to be a product operator on \(\mathbb{R}^{n+m} \times \mathbb{R}^m\) if \(T\) is bounded on \(L^2(\mathbb{R}^{n+m} \times \mathbb{R}^m)\), and

\[
Tf(x, y, z) = \int_{\mathbb{R}^{n+m}} K(x, y, z, u, v, w) f(u, v, w) du dv dw,
\]

where \(K(x, y, z, u, v, w)\), the kernel of \(T\), satisfies the following conditions:

1. \(\|K(\cdot, \cdot, z', \cdot, \cdot, w)\|_{\mathbb{R}^{n+m}} \leq C|z - w|^{-m},\)
2. \(\|K(x, y, \cdot, u, v, \cdot)\|_{\mathbb{R}^m} \leq C(|x - u| + |y - v|)^{-(n+m)},\)
3. \(\|K(\cdot, z, \cdot, u, v, \cdot)\|_{\mathbb{R}^{n+m}} \leq C \frac{|z - z'|}{|z - w|^{n+1}}\) for \(|z - z'| \leq |z - w|/2,\)
4. \(\|K(\cdot, \cdot, z', \cdot, \cdot, w)\|_{\mathbb{R}^{n+m}} \leq C \frac{|w - w'|}{|z - w|^{n+1}}\) for \(|w - w'| \leq |z - w|/2,\)
5. \(\|K(\cdot, z', \cdot, \cdot, w) - K(\cdot, z', \cdot, \cdot, w) - K(\cdot, \cdot, \cdot, \cdot, w) + K(\cdot, \cdot, \cdot, \cdot, w)\|_{\mathbb{R}^{n+m}} \leq C \frac{|z - z'| |w - w'|}{|z - w|^{m+2}}\) for \(|z - z'| \leq |z - w|/2\) and \(|w - w'| \leq |z - w|/2,\)
6. \(\|K(x, y, \cdot, u, v, \cdot) - K(x', y', \cdot, u, v, \cdot)\|_{\mathbb{R}^m} \leq C \frac{|x - x' + |y - y'|}{(|x - u| + |y - v|)^{n+m+1}}\) for \(|x - x'| + |y - y'| \leq (|x - u| + |y - v|)/2,\)
7. \(\|K(x, y, \cdot, u, v, \cdot) - K(x, y, \cdot, u', v, \cdot)\|_{\mathbb{R}^m} \leq C \frac{|u - u' + |v - v'|}{(|x - u| + |y - v|)^{n+m+1}}\) for \(|u - u'| + |v - v'| \leq (|x - u| + |y - v|)/2,\)
We write a proof for Lemma 2.7, which was one of the main ingredients in Theorem 2.5. Indeed, Proposition 2.11 is enough to provide a proof for such a result, which is a little bit complicated. However, Proposition 2.11 holds for all $f \in M_{n+m,m}(\beta,\gamma,r_1,r_2,x_0,y_0)$ satisfying Lemma 2.9. Moreover, Proposition 2.11 is enough to provide a proof for Lemma 2.7, which was one of the main ingredients in Theorem 2.5.

**Proposition 2.11.** If $T$ is a product operator as in Definition 2.10 and $T_1(1) = T_2(1) = T^*_1(1) = T^*_2(1) = 0$, then

$$
\|Tf\|_{M_{n+m,m}(\beta,\gamma,r_1,r_2,x_0,y_0,z_0)} \leq C\|f\|_{M_{n+m,m}(\beta,\gamma,r_1,r_2,x_0,y_0,z_0)}
$$

for all $f(x,y,z) = f_1(x,y)f_2(z)$ with $f_1 \in M_{n+m,m}(\beta,\gamma,r_1,x_0,y_0)$ and $f_2 \in M_{m}(\beta,\gamma,r_2,z_0)$. See [22] for definitions of $T_1(1) = T_2(1) = T^*_1(1) = T^*_2(1) = 0$.

**Remark 2.12.** Indeed, Proposition 2.11 holds for all $f \in M_{n+m,m}(\beta,\gamma,r_1,r_2,x_0,y_0,z_0)$. The proof for such a result is a little bit complicated. However, Proposition 2.11 is enough to provide a proof for Lemma 2.7, which was one of the main ingredients in Theorem 2.5.

**Proof of Proposition 2.11.** Suppose that $f(x,y,z) = f_1(x,y)f_2(z)$ with

$$
\|f_1\|_{M_{n+m,m}(\beta,\gamma,r_1,x_0,y_0)} = \|f_2\|_{M_{m}(\beta,\gamma,r_2,z_0)} = 1.
$$

We write

$$
Tf(x,y,z) = \int_{\mathbb{R}^{n+m}} K(x,y,z,u,v,w)f(u,v,w)du dv dw
= \int_{\mathbb{R}^m} \int_{\mathbb{R}^{n+m}} K(x,y,z,u,v,w)f_1(u,v)du dv f_2(w)dw
= \int_{\mathbb{R}^m} S(z,w)f_2(w)dw,
$$

where $x$, $y$ and $f_1$ are fixed, and $S(z,w) = \int_{\mathbb{R}^{n+m}} K(x,y,z,u,v,w)f_1(u,v)du dv$.

We claim that for fixed $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, $S(g)(z) = \int S(z,w)g(w)dw$ is an operator bounded on $M_{m}(\beta,\gamma,r_2,z_0)$ with the kernel $S(z,w)$ satisfying Lemma 2.9. Moreover,

1. $|S(z,w)| \leq C|x-w|^{-m-\gamma}|T|||z-w|^{-\gamma/(r_1+|x-x_0|+|y-y_0|)}$, for $|z-w|/2$.

2. $|S(z,w) - S(z',w)| \leq C\frac{|z-z'|}{|z-w|^{-\gamma/(r_1+|x-x_0|+|y-y_0|)}}$, for $|z-z'| \leq |z-w|/2$.

3. $|S(z,w) - S(z,w')| \leq C\frac{|w-w'|}{|z-w|^{-\gamma/(r_1+|x-x_0|+|y-y_0|)}}$,
for \(|w - w'| \leq |z - w|/2\),
\( \quad (4) \quad |S(z, w) - S(z', w') - S(z, w') + S(z', w')| \)
\[ \leq C\frac{|z - z'||w - w'|}{|z - w|^{m+2}} |||T||| \frac{r^\gamma_1}{(r_1 + |x - x_0| + |y - y_0|)^{n+m+\gamma}} \]
for \(|z - z'|, |w - w'| \leq |z - w|/2\),
\( \quad (5) \quad S(1) = S^* (1) = 0. \)

The proof of the claim follows from Lemma 2.9. Indeed, for fixed \(z, w \in \mathbb{R}^m\), the operator \(L\) with the kernel \(K(x, y, z, u, v, w)\) is given by
\[
L(f_1)(x, y, z, w) = \int_{\mathbb{R}^{n+m}} K(x, y, z, u, v, w)f_1(u, v)dudv.
\]

By the condition (1) in Definition 2.10 together with Lemma 2.9, the operator \(L\) is bounded on \(\mathcal{M}_{\mathbb{R}^{n+m}}(\beta, \gamma, r_1, x_0, y_0)\). Thus,
\[
|L(f_1)(x, y, z, w)| \leq C|||T||| \frac{r^\gamma_1}{(r_1 + |x - x_0| + |y - y_0|)^{n+m+\gamma}},
\]
which implies that \(S(z, w)\) satisfies estimate (1) in the above claim, that is,
\[
|S(z, w)| \leq C|z - w|^{-m}|||T||| \frac{r^\gamma_1}{(r_1 + |x - x_0| + |y - y_0|)^{n+m+\gamma}}.
\]

Similarly, applying conditions (3) and (4) in Definition 2.10 together with Lemma 2.9 respectively, we conclude that \(S(z, w)\) satisfies the estimates in (2) and (3) in the above claim, respectively. The condition (5) in Definition 2.10 together with Lemma 2.9 yields the estimate (5) in the above claim for \(S(z, w)\).

Based on the estimates on \(S(z, w)\), the kernel of \(S\), applying Lemma 2.9 gives that the operator \(S\) is bounded on \(\mathcal{M}_{\mathbb{R}^m}(\beta, \gamma, r_2, z_0)\) and hence
\[
|Tf(x, y, z)| = |S(f_2)(z)| \leq C|||T||| \frac{r^\gamma_1}{(r_1 + |x - x_0| + |y - y_0|)^{n+m+\gamma}} \frac{r^\gamma_2}{(r_2 + |z - z_0|)^{m+\gamma}}
\]
and
\[
|Tf(x, y, z) - T(x, y, z')| = |S(f_2)(z) - S(f_2)(z')| \leq C|||T||| \frac{r^\gamma_1}{(r_1 + |x - x_0| + |y - y_0|)^{n+m+\gamma}} \frac{r^\gamma_2}{(r_2 + |z - z_0|)^{m+\gamma}}.
\]
for \(|z - z'| \leq \frac{r_2 + |z - z_0|}{2}\).

Similarly, if write
\[
Tf(x, y, z) = \int_{\mathbb{R}^{n+m+m}} K(x, y, z, u, v, w)f(u, v, w)dudvdw
\]
\[
\begin{aligned}
&= \int_{\mathbb{R}^{n+m}} \int_{\mathbb{R}^m} K(x, y, z, u, v, w) f_2(w) dw f_1(u, v) dudv \\
&= \int_{\mathbb{R}^{n+m}} R(x, y, z, u, v) f_1(u, v) dudv,
\end{aligned}
\]

where \( z \) and \( f_2 \) are fixed, and \( R(x, y, z, u, v) = \int_{\mathbb{R}^m} K(x, y, z, u, v, w) f_2(w) dw \), then applying the same proof implies that the operator \( R \) is bounded on \( \mathcal{M}_{\mathbb{R}^{n+m}}(\beta, \gamma, r_1, x_0, y_0) \) and moreover,

\[
|Tf(x, y, z) - Tf(x', y', z)| \\
\leq |R(f_1)(x, y) - R(f_1)(x', y')| \\
\leq C||T||\left(\frac{|x - x'| + |y - y'|}{r_1 + |x - x_0| + |y - y_0|}\right)^\beta r_1^\gamma \left(\frac{|z - z'|}{r_2 + |z - z_0|}\right)^\gamma r_2^\gamma
\]

for \( |x - x'| + |y - y'| \leq (r_1 + |x - x_0| + |y - y_0|)/2 \) and \( |z - z'| \leq (r_2 + |z - z_0|)/2 \). To do this, write

\[
\begin{aligned}
Tf(x, y, z) - Tf(x, y, z') \\
&= \int_{\mathbb{R}^{n+m}} [K(x, y, z, u, v, w) - K(x, y, z', u, v, w)] f(u, v, w) dudv dw \\
&= \int_{\mathbb{R}^{n+m}} \int_{\mathbb{R}^m} [K(x, y, z, u, v, w) - K(x, y, z', u, v, w)] f_2(w) dw f_1(u, v) dudv \\
&= \int_{\mathbb{R}^{n+m}} H(x, y, z, z', u, v) f_1(u, v) dudv \\
&= H(f_1)(x, y, z, z'),
\end{aligned}
\]

where \( z, z' \) and \( f_2 \) are fixed, and

\[
H(x, y, z, z', u, v) := \int_{\mathbb{R}^m} [K(x, y, z, u, v, w) - K(x, y, z', u, v, w)] f_2(w) dw.
\]

We claim that the operator \( H \) with the kernel \( H(x, y, z, z', u, v) \) defined above is bounded on \( \mathcal{M}_{\mathbb{R}^{n+m}}(\beta, \gamma, r_1, x_0, y_0) \) and moreover,

\[
\begin{aligned}
|Tf(x, y, z) - Tf(x', y', z) - Tf(x, y, z') + Tf(x', y', z')| \\
&= |H(f_1)(x, y, z, z') - H(f_1)(x', y', z, z')|
\end{aligned}
\]
is bounded on $M_{\beta, \gamma, r_2, z_0}$ and hence, for $|z - z'| \leq (r_2 + |z - z_0|)/2$,

$$\left| \int_{\mathbb{R}^n} [K(x, y, z, u, v, w) - K(x, y, z', u, v, w)]f_2(w)dw \right| \leq C(|x - u| + |y - v|) - (n+m) \left( \frac{|z - z'|}{r_2 + |z - z_0|} \right)^{\beta} \frac{r_2^\gamma}{(r_2 + |z - z_0|)^{m+\gamma}},$$

which implies that for $|z - z'| \leq (r_2 + |z - z_0|)/2$,

$$|H(x, y, z, z', u, v)| \leq C(|x - u| + |y - v|) - (n+m) \left( \frac{|z - z'|}{r_2 + |z - z_0|} \right)^{\beta} \frac{r_2^\gamma}{(r_2 + |z - z_0|)^{m+\gamma}}.$$

Write

$$H(x, y, z, u, v) - H(x', y', z', u, v) = \int_{\mathbb{R}^n} [K(x, y, z, u, v, w) - K(x', y', z, u, v, w)]f_2(w)dw - \int_{\mathbb{R}^n} [K(x, y, z', u, v, w) - K(x', y', z', u, v, w)]f_2(w)dw.$$

By condition (6) in Definition 2.10 together with Lemma 2.9 for fixed $x, y, x', y', u$ and $v$ with $|x - x'| + |y - y'| \leq (|x - u| + |y - v|)/2$, the operator

$$\int_{\mathbb{R}^n} [K(x, y, z, u, v, w) - K(x', y', z, u, v, w)]f_2(w)dw$$

is bounded on $M_{\beta, \gamma, r_2, z_0}$ and hence, for $|z - z'| \leq (r_2 + |z - z_0|)/2$ and $|x - x'| + |y - y'| \leq (|x - u| + |y - v|)/2$,

$$|H(x, y, z, u, v) - H(x', y', z', u, v)| \leq C \frac{|x - x'| + |y - y'|}{(|x - u| + |y - v|)^{n+m+1}} \left( \frac{|z - z'|}{r_2 + |z - z_0|} \right)^{\beta} \frac{r_2^\gamma}{(r_2 + |z - z_0|)^{m+\gamma}}.$$

Similarly, for $|z - z'| \leq (r_2 + |z - z_0|)/2$ and $|u - u'| + |v - v'| \leq (|x - u| + |y - v|)/2$,

$$|H(x, y, z, u, v) - H(x, y, z', u', v')| \leq C \frac{|x - x'| + |y - y'|}{(|x - u| + |y - v|)^{n+m+1}} \left( \frac{|z - z'|}{r_2 + |z - z_0|} \right)^{\beta} \frac{r_2^\gamma}{(r_2 + |z - z_0|)^{m+\gamma}}.$$
Finally, we write
\[ H(x, y, z, z', u, v) - H(x', y', z', u, v) = \int_{\mathbb{R}^m} \left[ K(x, y, z, u, v, w) - K(x', y', z, u, v, w) \right] f_2(w)dw \]

Applying condition (7) in Definition 2.10 together with Lemma 2.9 for fixed \( x, y, z, x', u, v \) and \( v' \) with \(|x - x'| + |y - y'| \leq (|x - u| + |y - v|)/2 \) and \(|u - u'| + |v - v'| \leq (|x - u| + |y - v|)/2 \), the operator
\[ \int_{\mathbb{R}^m} \left[ K(x, y, z, u, v, w) - K(x', y', z, u', v, w) + K(x', y', z, u', v', w) \right] f_2(w)dw \]
is bounded on \( \mathcal{M}_{\mathbb{R}^m}(\beta, r_2, z_0) \) and hence, for \(|z - z'| \leq (r_2 + |z - z_0|)/2 \), \(|x - x'| + |y - y'| \leq (|x - u| + |y - v|)/2 \) and \(|u - u'| + |v - v'| \leq (|x - u| + |y - v|)/2 \),
\[ |H(x, y, z, u, v) - H(x', y', z', u, v)| \leq C \frac{|u - u'| + |v - v'|}{(|x - u| + |y - v|)^{n+1}} \left( \frac{|z - z'|}{r_2 + |z - z_0|} \right)^{\beta} \frac{r_2^\gamma}{(r_2 + |z - z_0|)^{m+\gamma}}. \]

Therefore, the operator
\[ \int_{\mathbb{R}^n} \left[ K(x, y, z, u, v, w) - K(x', y', z, u', v, w) + K(x', y', z, u', v', w) \right] f_2(w)dw \]
is bounded on \( \mathcal{M}_{\mathbb{R}^n+m}(\beta, \gamma, r_1, x_0, y_0) \) and this yields the claim. The proof of Proposition 2.11 is concluded.

### 2.2 Plancherel–Pólya type inequalities

Applying the discrete Calderón reproducing formula in (2.5) provides the following Plancherel-Pólya-type inequalities.

**Theorem 2.13.** Suppose \( \psi \) is as in (1.1). Then for \( f \in L^1(\mathbb{R}^{n+m}) \),
\[
\left\| \left\{ \sum_j \sum_k \int_{2^{-k-N}}^{2^{-k-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} \sum_I \sup_{u \in I} |\psi_{t,s} \ast f(u,v)|^2 \chi_I(x)\chi_J(y) \frac{dt}{t} \frac{ds}{s} \right\}^\frac{1}{2} \right\|_1
\]
where $I \subset \mathbb{R}^n, J \subset \mathbb{R}^m$, are the same as in Theorem 2.2 and $\chi_I$ and $\chi_J$ are indicator functions of $I$ and $J$, respectively.

Proof. For $f \in L^1(\mathbb{R}^{n+m})$, by Theorem 2.5

$$\psi_{t,s} * f(u,v) = \sum_{j} \sum_{k'} \sum_{j'} \sum_{J'} |I' ||J'| \psi_{t,s} * \phi_{j',k'}(u,v) \int_{2^{-k'-N}}^{2^{-j'-N+1}} \int_{2^{-2j'-N}}^{2^{-j'-N+1}} \psi_{t',s'} * f(x_{l'}, y_{l'}) \frac{dt'}{t'} \frac{ds'}{s'}. $$

For $2^{-j-N} < t < 2^{-k-N+1}$ and $2^{-k-N} < s < 2^{-k-N+1}$, applying the classical almost orthogonality estimate yields that for $0 < \beta, \gamma < 1$,

$$|\psi_{t,s} * \phi_{j',k'}(u,v)| \leq C_N 2^{-|j-j'||\beta|} 2^{-|k-k'|\beta} \frac{2^{-[(k\wedge k')n(1-\frac{1}{p})]}}{(2^{-j-j'})^{n+\gamma} (2^{-[(k\wedge k')n(1-\frac{1}{p})]})^{n+\gamma}}.$$ 

Observe that

$$\sum_{J'} \sum_{I'} \frac{2^{-|(j\wedge j')\gamma}}{(2^{-j-j'})^{n+\gamma} (2^{-[(k\wedge k')n(1-\frac{1}{p})]})^{m+\gamma}} \psi_{t',s'} * f(x_{l'}, y_{l'}) \frac{dt'}{t'} \frac{ds'}{s'}$$

$$\leq C \left( \int_{2^{-k'-N}}^{2^{-j'-N+1}} \int_{2^{-j-N}}^{2^{-j'-N+1}} \sum_{J'} \sum_{I'} \psi_{t',s'} * f(x_{l'}, y_{l'}) \chi_I \chi_J \frac{dt'}{t'} \frac{ds'}{s'} \right)^{1/r} (u,v),$$

where $M_s$ is the strong maximal function on $\mathbb{R}^n \times \mathbb{R}^m$ and $\frac{n+m}{n+\beta} < r < 1$. See [11] for the proof of the classical case. Note that $x_{l'}$ and $y_{l'}$ are arbitrary points in $I'$ and $J'$, respectively, we have

$$\sup_{u \in I, v \in J} |\psi_{t,s} * f(u,v)| \leq C_N \sum_{j'} \sum_{k'} 2^{-|j-j'||\beta|} 2^{-|k-k'|\beta} 2^{-[(j\wedge j')n(1-\frac{1}{p})]2[(j\wedge j')n(1-\frac{1}{p})2^{-j'-N+1}]} \frac{dt'}{t'} \frac{ds'}{s'}$$

$$\times \left\{ M_s \left( \sum_{j} \sum_{k} \sum_{j} \inf_{u \in I, v \in J} \psi_{t,s} * f(u', v') \chi_I \chi_J \frac{dt'}{t'} \frac{ds'}{s'} \right)^{1/r} (u,v) \right\}. $$

Applying Hölder’s inequality together with the facts that

$$\sum_{j} \sum_{k} 2^{-|j-j'|\beta} 2^{-[j-j'0(1-\frac{1}{p})2^{-j'-N+1}]} 2^{-k'-n(1-\frac{1}{p})2^{-j'-N+1}} \leq C,$$
Plancherel-Pólya-type inequalities: 

\[ \sum_j \sum_I \chi_I(x) \chi_J(y) \leq C \]

and 

\[ \int_{2^{-k-N}}^{2^{-j-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} \frac{dt}{s} ds \leq C \]

gives

\[ \left\| \left\{ \sum_j \sum_k \int_{2^{-k-N}}^{2^{-j-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} \sup_{u \in I} \sup_{v \in J} |\psi_t,s \ast f(u,v)|^2 \chi_I(x) \chi_J(y) \frac{dt}{s} ds \right\}^{1/2} \right\|_1 \]

\[ \lesssim \left\| \left\{ \sum_j \sum_k \left( M_s \left( \int_{2^{-k-N}}^{2^{-j-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} \sum_{u \in I} \sup_{v \in J} \psi_t,s \ast f(u,v) \right) \chi_I(x) \chi_J(y) \right\}^{1/2} \right\|_1. \]

By the Fefferman-Stein vector-valued maximal function inequality with \( r < 1 \), we get

\[ \left\| \left\{ \sum_j \sum_k \int_{2^{-k-N}}^{2^{-j-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} \sum_{u \in I} \sup_{v \in J} |\psi_t,s \ast f(u,v)|^2 \chi_I(x) \chi_J(y) \frac{dt}{s} ds \right\}^{1/2} \right\|_1 \]

\[ \approx \left\| \left\{ \sum_j \sum_k \int_{2^{-k-N}}^{2^{-j-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} \sum_{u \in I} \inf_{v \in J} |\psi_t,s \ast f(u,v)|^2 \chi_I(x) \chi_J(y) \frac{dt}{s} ds \right\}^{1/2} \right\|_1. \]

The proof is completed. \( \square \)

We remark that applying a similar proof, for any fixed constant \( C_0 \) one can get the following Plancherel-Pólya-type inequalities:

\[ \left\| \left\{ \sum_j \sum_k \int_{2^{-k-N}}^{2^{-j-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} \sum_{u \in C_0 I} \sup_{v \in C_0 J} |\psi_t,s \ast f(u,v)|^2 \chi_I(x) \chi_J(y) \frac{dt}{s} ds \right\}^{1/2} \right\|_1 \]

\[ \approx \left\| \left\{ \sum_j \sum_k \int_{2^{-k-N}}^{2^{-j-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} \sum_{u \in C_0 I} \inf_{v \in C_0 J} |\psi_t,s \ast f(u,v)|^2 \chi_I(x) \chi_J(y) \frac{dt}{s} ds \right\}^{1/2} \right\|_1, \quad (2.6) \]

where \( C_0 I \subset R^n \) and \( C_0 J \subset R^m \), are cubes with side-length \( \ell(C_0 I) = C_0 2^{-j-N} \) and \( \ell(C_0 J) = C_0 \ell(J) = 2^{-(j-N \land k-N)} \), respectively.

### 2.3 The equivalence of \( \|g_F(f)\|_1 \) and \( \|S_F(f)\|_1 \)

#### 2.3.1 The proof that \( \|S_F(f)\|_1 \leq \|g_F(f)\|_1 \)

We write

\[ \|S_F(f)(x,y)\|_1 = \left\| \left\{ \sum_{j,k} \int_{2^{-j-N}}^{2^{-j-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} \chi_I(x) \chi_J(y) \frac{dx_I dt}{\ell_I^{n+1} \ell_J^{m+1}} \chi_I(x) \chi_J(y) \frac{dy_I ds}{\ell_J^{m+1} \ell_I^{n+1}} \right\}^{1/2} \right\|_1 \]

\[ \times |\psi_t,s \ast f(x_1, y_1)|^2 \chi_I(x) \chi_J(y) \frac{dx_1 dt}{\ell_I^{n+1} \ell_J^{m+1}} \frac{dy_1 ds}{\ell_J^{m+1} \ell_I^{n+1}} \]
where $N$ is a fixed large integer as in the Plancherel-Pólya-type inequalities and $I \subset \mathbb{R}^n$ and $J \subset \mathbb{R}^m$ are dyadic cubes with side-length $\ell(I) = 2^{-j-N}, \ell(J) = 2^{-(j-N^\chi-I)},$ and $\chi_I$ and $\chi_J$ are indicator functions of $I$ and $J$, respectively.

Observe that there exists a fixed constant $C_0$ such that for $2^{-j-N} \leq t \leq 2^{-j-N+1}, 2^{-k-N} \leq s \leq 2^{-k-N+1}$ and $x_1 \in \mathbb{R}^n$ and $y_1 \in \mathbb{R}^m$,

$$\chi_{t,s}(x-x_1, y-y_1) |\psi_{t,s} * f(x_1,y_1)|^2 \chi_I(x) \chi_J(y) \leq \chi_{t,s}(x-x_1, y-y_1) \sup_{u \in C_0 I \atop v \in C_0 J} |\psi_{t,s} * f(u,v)|^2 \chi_I(x) \chi_J(y).$$

Therefore,

$$\|S_F(f)(x,y)\|_1 \leq \left\{ \sum_{j,k} \sum_{I,J} \int_{2^{-k-N}}^{2^{-k-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} \chi_{t,s}(x-x_1, y-y_1) \right.$$ 
\begin{equation*}
\times \sup_{u \in C_0 I \atop v \in C_0 J} |\psi_{t,s} * f(u,v)|^2 \chi_I(x) \chi_J(y) \frac{dx \, dt \, dy \, ds}{t^{n+1} \, s^{m+1}} \right\}^{1/2} \|.
\end{equation*}

Applying the estimate $\int_{\mathbb{R}^n} \int_{\mathbb{R}^m} \chi_{t,s}(x-x_1, y-y_1) \, dx \, dy_1 \leq C t^{n+m} s^m$ together with the Plancherel-Pólya-type inequalities in (2.6) yields

$$\|S_F(f)(x,y)\|_1 \leq \left\{ \sum_{j,k} \sum_{I,J} \int_{2^{-k-N}}^{2^{-k-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} \sup_{u \in C_0 I \atop v \in C_0 J} |\psi_{t,s} * f(u,v)|^2 \chi_I(x) \chi_J(y) \frac{dt \, ds}{t \, s} \right\}^{1/2} \|_1 \leq \|g_F(f)\|_1.$$

2.3.2 The proof that $\|g_F(f)\|_1 \lesssim \|S_F(f)\|_1$

The proof of this part is similar. To see this, write

$$\|g_F(f)\|_1 = \left\{ \int_0^\infty \int_0^\infty |\psi_{t,s} * f(x,y)|^2 \frac{dt \, ds}{t \, s} \right\}^{1/2} \|_1 \leq \left\{ \sum_{j,k} \sum_{I,J} \int_{2^{-k-N}}^{2^{-k-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} |\psi_{t,s} * f(x,y)|^2 \chi_I(x) \chi_J(y) \frac{dt \, ds}{t \, s} \right\}^{1/2} \|. \quad (2.10)$$

By the Plancherel-Pólya-type inequalities in (2.10), the last term above is dominated by

$$C \left\{ \sum_{j,k} \sum_{I,J} \int_{2^{-k-N}}^{2^{-k-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} \inf_{u \in C_0 I \atop v \in C_0 J} |\psi_{t,s} * f(u,v)|^2 \chi_I(x) \chi_J(y) \frac{dt \, ds}{t \, s} \right\}^{1/2} \|. \quad (2.11)$$
Characterizations of flag Hardy spaces

\[ \leq C \left\{ \left( \sum_{j,k} \sum_{I,J} \int_{2^{-k-N}}^{2^{-j-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} \chi_{t,s}(x-x_1, y-y_1) \times \inf_{u \in C_0 I \atop v \in C_0 J} |\psi_{t,s} \ast (u(v))|^2 \chi_I(x) \chi_J(y) \frac{dx_1 dt}{t^{n+m+1}} \frac{dy_1 ds}{s^{m+1}} \right) \right\}^{1/2} \]

\[ \leq C \|S_F(f)\|_1. \]

3 Estimates of flag Poisson integrals and flag Riesz transforms

In this section, we will show the following estimates:

\[ \|S_F(u)\|_1 \lesssim \|u^\ast\|_1 \lesssim \|u^+\|_1 \lesssim \sum_{j=1}^{n+m} \sum_{k=1}^{m} \|R_j^{(1)} R_k^{(2)}(f)\|_1 + \|f\|_1. \]

3.1 The estimate \( \|S_F(u)\|_1 \lesssim \|u^\ast\|_1 \)

We first introduce the following maximal function associated with the flag structure.

**Definition 3.1.** For \( f \in L^1_{loc}(\mathbb{R}^{n+m}) \), we define the maximal function by

\[ M_F(f)(x,y) = \sup_{t,s>0, (x,y) \in R} \frac{1}{|R|} \int_R |f(u,v)| dudv, \]

where \( R = I \times J \) run over all rectangles with sides parallel to the axes and \( \ell(I) = t, \ell(J) = t+s \).

We now recall the lemma of K. Merryfield.

**Lemma 3.2** ([23]). Let \( \varphi \in C_0^\infty(\mathbb{R}^n) \) satisfy

1. \( \varphi(-x) = \varphi(x) \);
2. \( \text{supp } \varphi \subset B_n(0,1), \) where \( B_n(0,1) \) is the unit ball in \( \mathbb{R}^n \);
3. \( \int_{\mathbb{R}^n} \varphi(x) dx = 1. \)

Then there exists a function \( \psi \in C_0^\infty(\mathbb{R}^n) \) that satisfies \( \text{supp } \psi \subset B_n(0,1) \) and \( \int_{\mathbb{R}^n} \psi(x) dx = 0, \) such that for \( u(x,t) = P_t \ast f(x) \) we have

\[ \int_{\mathbb{R}^{n+1}} |\nabla u(x,t)|^2 |g \ast \varphi_t(x)|^2 t dx dt \leq C \int_{\mathbb{R}^n} f^2(x) g^2(x) dx + \int_{\mathbb{R}^{n+1}} u^2(x,t) |g \ast \psi_t(x)|^2 \frac{dx dt}{t}, \]

where \( C \) is independent of \( f \) and \( g \).
Lemma 3.3. Let all the notation be the same as above. Then we define \( \psi \) and \( \psi \) that satisfies \( \text{supp } \varphi \subseteq B_{n+m}(0,1) \), where \( B_{n+m}(0,1) \) is the unit ball in \( \mathbb{R}^{n+m} \);

1. \( \varphi(1)(-x, -y) = \varphi(1)(x, y) \);
2. \( \text{supp } \varphi(1) \subseteq B_{n+m}(0,1) \);
3. \( \int_{\mathbb{R}^{n+m}} \varphi(x,y)dx\,dy = 1 \).

Let \( \varphi(2)(z) \in C_0^\infty(\mathbb{R}^{n+m}) \) satisfies the same conditions as in Lemma 3.2 and \( \varphi(x,y) = \varphi(1) *_{\mathbb{R}^m} \varphi(2)(x,y) \).

Similarly, we can obtain two functions \( \psi(1)(x, y) \) and \( \psi(2)(z) \) such that \( \psi(1) \in C_0^\infty(\mathbb{R}^{n+m}) \) that satisfies \( \text{supp } \psi(1) \subseteq B_{n+m}(0,1) \) and

\[
\int_{\mathbb{R}^{n+m}} \psi(1)(x, y)dx = 0,
\]

and \( \psi(2) \in C_0^\infty(\mathbb{R}^m) \) that satisfies \( \text{supp } \psi(2) \subseteq B_{m}(0,1) \) and

\[
\int_{\mathbb{R}^m} \psi(2)(z)dz = 0.
\]

Then we define \( \psi(x, y) := \psi(1) *_{\mathbb{R}^m} \psi(2)(x,y) \). We arrive at the following technical lemma.

**Lemma 3.3.** Let all the notation be the same as above.

\[
\begin{align*}
&\int_{\mathbb{R}^{n+1}}\int_{\mathbb{R}^{m+1}} \left| t\nabla(1) s\nabla(2) P_t*s \ast f(x,y) \right|^2 g \ast \varphi_t,s(x,y) \frac{dyds}{s} \frac{dxdt}{t} \\
&\quad \leq C \left\{ \int_{\mathbb{R}^n}\int_{\mathbb{R}^m} f^2(x,y)g^2(x,y)dx\,dy \\
&\quad + \int_{\mathbb{R}^n}\int_{\mathbb{R}^{n+1}} \left| P_s *_{\mathbb{R}^m} f(x,y) \right|^2 \left| \psi_t,s *_{\mathbb{R}^m} g(x,y) \right|^2 \frac{dyds}{s} \frac{dx}{t} \\
&\quad + \int_{\mathbb{R}^m}\int_{\mathbb{R}^{n+1}} \left| P_t * f(x,y) \right|^2 \left| \psi_t,s \ast g(x,y) \right|^2 \frac{dxdt}{t} \frac{dy}{s} \\
&\quad + \int_{\mathbb{R}^n}\int_{\mathbb{R}^{m+1}} \left| P_t \ast f(x,y) \right|^2 \left| \psi_t,s \ast g(x,y) \right|^2 \frac{dyds}{s} \frac{dxdt}{t} \right\}.
\end{align*}
\]

**Proof.** Applying Lemma 3.2 with replacing \( n \) by \( n + m \) gives

\[
\begin{align*}
&\int_{\mathbb{R}^{n+1}}\int_{\mathbb{R}^{m+1}} \left| t\nabla(1) s\nabla(2) u(x,y,t,s) \right|^2 g \ast \varphi_t,s(x,y) \frac{dyds}{s} \frac{dxdt}{t} \\
&\quad = \int_0^\infty \int_0^\infty \int_{\mathbb{R}^n}\int_{\mathbb{R}^m} \left| t\nabla(1) P_t \ast ((s\nabla(2) P_s * f)(x,y)) \right|^2 \left| \varphi_t \ast (\varphi_s * g)(x,y) \right|^2 \frac{dxdy}{t} \frac{dt\,ds}{s} \\
&\quad \leq C \int_0^\infty \int_{\mathbb{R}^n\times\mathbb{R}^m} \left| F_s(x,y) \right|^2 \left| G_s(x,y) \right|^2 \frac{dx\,dy\,ds}{s}.
\end{align*}
\]
\[ + \int_0^\infty \int_0^\infty \int_{\mathbb{R}^n \times \mathbb{R}^m} |P_t^{(1)} \ast F_s(x,y)|^2 |\tilde{Q}_t^{(1)} \ast G_s(x,y)|^2 \, dx \, dy \, dt \, ds \]
\[ \triangleq I_1 + I_2, \]

where \( F_s(x,y) = (s \nabla^{(2)} P_s^{(2)}) \ast_{\mathbb{R}^m} f(x,y) \) and \( G_s(x,y) = \varphi_s^{(2)} \ast_{\mathbb{R}^m} g(x,y) \).

To estimate \( I_1 \), we have

\[ \int_0^\infty \int_{\mathbb{R}^n \times \mathbb{R}^m} |F_s(x,y)|^2 |G_s(x,y)|^2 \, dx \, dy \, ds \]
\[ = \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \left| (s \nabla^{(2)} P_s^{(2)} \ast_{\mathbb{R}^m} f(x,\cdot))(y) \right|^2 \left| (\varphi_s^{(2)} \ast_{\mathbb{R}^m} g(x,\cdot))(y) \right|^2 \frac{dy \, ds}{s} \]
\[ \leq C \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} f^2(x,y) g^2(x,y) \, dx \, dy \]
\[ + C \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} |P_s^{(2)} \ast_{\mathbb{R}^m} f(x,y)|^2 |\tilde{Q}_s^{(2)} \ast_{\mathbb{R}^m} g(x,y)|^2 \frac{dy \, ds}{s} \, dx, \]

where the last inequality follows again from Lemma 32.

Similarly,

\[ I_2 = \int_{\mathbb{R}^{n+1} \times \mathbb{R}^{n+1}} |s \nabla^{(2)} P_s^{(2)} \ast_{\mathbb{R}^m} (P_t^{(1)} \ast f(x,\cdot))(y)|^2 |\varphi_s^{(2)} \ast_{\mathbb{R}^m} (\tilde{Q}_t^{(1)} \ast g(x,\cdot))(y)|^2 \frac{dy \, ds \, dx \, dt}{s} \]
\[ \leq C \int_{\mathbb{R}^m} \int_{\mathbb{R}^{n+1} \times \mathbb{R}^{n+1}} |P_t^{(1)} \ast f(x,y)|^2 |\tilde{Q}_t^{(1)} \ast g(x,y)|^2 \frac{dx \, dt}{t} \, dy \]
\[ + C \int_{\mathbb{R}^{n+1} \times \mathbb{R}^{n+1}} |P_s^{(2)} \ast_{\mathbb{R}^m} P_t^{(1)} \ast f(x,y)|^2 |\tilde{Q}_s^{(2)} \ast_{\mathbb{R}^m} \tilde{Q}_t^{(1)} \ast g(x,y)|^2 \frac{dy \, ds \, dx \, dt}{s} \, t. \]

The estimates of term \( I_1 \) and term \( I_2 \) yield \( \boxed{31} \).

We now begin to prove \( \| S_F(u) \|_1 \lesssim \| u^* \|_1 \). For any \( \alpha > 0 \) and each \( f \in L^1(\mathbb{R}^{n+m}) \) satisfying \( \| u^* \|_1 < \infty \), define

\[ A(\alpha) = \left\{ (x,y) \in \mathbb{R}^n \times \mathbb{R}^m : M_F(\chi_{(u^* > \alpha)})(x,y) < \frac{1}{200} \right\}. \]

Then we have

\[ \int_{A(\alpha)} S_F^2(u)(x,y) \, dx \, dy \leq \int_{\mathbb{R}^{n+1} \times \mathbb{R}^{n+1}} \int_{A(\alpha)} \chi_{t,s}(x-x_1,y-y_1) \, dx \, dy \frac{t}{s} \, |\nabla^{(1)} s \nabla^{(2)} u(x_1,y_1,t,s)|^2 \, dx_1 \, dy_1 \, dt \, ds. \]

By the definition of \( \chi_{t,s}(x-x_1,y-y_1) \), for any fixed \((x_1,y_1,t,s)\), if \( \chi_{t,s}(x-x_1,y-y_1) \neq 0 \), then \( (x,y) \) belongs to \( R \), where \( R = R(x_1,y_1,t,s) \) is a rectangle centered at \((x_1,y_1)\) and with side-length 2t and 2t + 2s. This means that to estimate \( \int_{A(\alpha)} \chi_{t,s}(x-x_1,y-y_1) \, dx \, dy \), we only need to consider those \((x,y) \in A(\alpha) \cap R(x_1,y_1,t,s)\). As a consequence,

\[ M_F(\chi_{(u^* > \alpha)})(x,y) < \frac{1}{200}. \]
Hence for such fixed \((x_1, y_1, t, s)\) mentioned above, we have
\[
\frac{1}{|R(x_1, y_1, t, s)|} |A(\alpha) \cap R(x_1, y_1, t, s)| < \frac{1}{200}.
\]

Let \(R^* = \{(x_1, y_1, t, s) : \frac{1}{|R(x_1, y_1, t, s)|} |A(\alpha) \cap R(x_1, y_1, t, s)| < \frac{1}{200}\}\), then we have
\[
\int_{A(\alpha)} S^2_F(u(x,y)) \, dx \, dy \leq \int_{R^*} \left| t \nabla^{(1)} s \nabla^{(2)} u(x_1, y_1, t, s) \right|^2 |\varphi_{t,s} * g(x_1, y_1)| \, dx_1 \, dy_1 \, ds \, dt.
\]

Let \(g(x,y) = \chi_{\{u^* \leq \alpha\}}(x,y)\) and \(\varphi^{(1)}(x,y) \in C^\infty_0(\mathbb{R}^{n+m})\) satisfy
1. \(\varphi^{(1)}(-x,-y) = \varphi^{(1)}(x,y)\);
2. \(\text{supp } \varphi^{(1)} \subset B_{n+m}(0,1)\), where \(B_{n+m}(0,1)\) is the unit ball in \(\mathbb{R}^{n+m}\);
3. \(\int_{\mathbb{R}^{n+m}} \varphi(x,y) \, dx \, dy = 1\);
4. \(\varphi^{(1)}(x,y) = 1\) when \(|(x,y)| \leq \frac{1}{3}\).

Similarly, \(\varphi^{(2)}(x,y) \in C^\infty_0(\mathbb{R}^m)\) satisfies
1. \(\varphi^{(2)}(-z) = \varphi^{(2)}(z)\);
2. \(\text{supp } \varphi^{(2)} \subset B_m(0,1)\), where \(B_m(0,1)\) is the unit ball in \(\mathbb{R}^m\);
3. \(\int_{\mathbb{R}^m} \varphi^{(2)}(z) \, dz = 1\);
4. \(\varphi^{(2)}(z) = 1\) when \(|z| \leq \frac{1}{3}\).

Set \(\varphi(x,y) = \varphi^{(1)} *_{\mathbb{R}^m} \varphi^{(2)}(x,y)\). Then for \((x_1, y_1) \in R^*\), we have
\[
\varphi_{t,s} * g(x_1, y_1) = \int_{\{u^* \leq \alpha\}} \varphi_{t,s}(x_1 - x_1, y - y_1) \, dx_1 \, dy_1 \geq \int_{\{u^* \leq \alpha\} \cap R(x_1, y_1, t, s)} \varphi_{t,s}(x - x_1, y - y_1) \, dx_1 \, dy_1 \geq C,
\]
where the last inequality follows from the definition of \(R^*\). Combining (3.2) and (3.3), we have
\[
\int_{A(\alpha)} S^2_F(u(x,y)) \, dx \, dy \leq C \int_{R^*} \left| t \nabla^{(1)} s \nabla^{(2)} u(x_1, y_1, t, s) \right|^2 |\varphi_{t,s} * g(x_1, y_1)| \, dx_1 \, dy_1 \, ds \, dt.
\]
\[
\leq C \int_{\mathbb{R}^n+1} \int_{\mathbb{R}^m+1} \left| t \nabla^{(1)} s \nabla^{(2)} u(x, y, t, s) \right|^2 |g * \varphi_{t,s}(x, y)| \, dy \, ds \, tx \, dt.
\]
\[
\leq C \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} f^2(x,y) g^2(x,y) \, dx \, dy.
\]
The proof of the estimate for \( II \) where the last inequality follows from Lemma 3.3.

Combining all estimates above implies that \( II_1 + II_2 + II_3 + II_4 \),

where the last inequality follows from Lemma 3.3.

For the term \( II_1 \), we have

\[
|II_1| \leq C \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} |P_t^{(1)} * f(x,y)|^2 |\psi_t^{(1)} * g(x,y)|^2 \frac{dxdt}{t} ddy
\]

\[
\leq C \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} f^2(x,y)g^2(x,y)dxdy
\]

\[
\leq C \int_{\{u^* \leq \alpha \}} f^2(x,y)dxdy
\]

\[
\leq C \int_{\{u^* \leq \alpha \}} |u^*(x,y)|^2 dxdy.
\]

If \( \psi_s^{(2)} * g(x,y) = \int \psi_s^{(2)}(y-w)g(x,w)dw \neq 0 \), then there exists some \( w \) such that \( |y-w| < s \) and \( (x,w) \in \{u^* \leq \alpha \} \). Hence \( |P_s^{(2)} * f(x,y)| \leq \alpha \). As a consequence,

\[
|II_2| \leq C \alpha^2 \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} |\psi_s^{(2)} * g(x,y)|^2 \frac{dyds}{s} dx
\]

\[
= C \alpha^2 \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} |\psi_s^{(2)} * (1 - g(x,y))|^2 \frac{dyds}{s} dx
\]

\[
\leq \alpha^2 \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} |1 - g(x,y)|^2 dydx
\]

\[
\leq \alpha^2 |\{u^* > \alpha \}|.
\]

The proof of the estimate for \( II_3 \) is similar to \( II_2 \).

For the last term \( II_4 \), if \( \psi_{t,s} * g(x,y) = \int \psi_{t,s}(x-v,y-w)g(v,w)dv dw \neq 0 \), similarly as term \( II_2 \), there exists \( (v,w) \) such that \( (v,w) \in \{u^* \leq \alpha \} \) and \( |x-v| < t, |y-w| < t + s \). Hence \( |P_{t,s} * f(x,y)| \leq \alpha \). Following the same strategy of (3.4), we have

\[
|II_4| \leq C \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |P_{t,s} * f(x,y)|^2 |\psi_{t,s} * g(x,y)|^2 \frac{dyds dt}{s t} \leq C \alpha^2 |\{u^* > \alpha \}|.
\]

Combining all estimates above implies that

\[
\int_{\{M_F(\chi_{u^* > \alpha}) \leq \frac{1}{300} \}} S_F^2(u)(x,y)dxdy \leq C \left( \alpha^2 |\{u^* > \alpha \}| + \int_{\{u^* \leq \alpha \}} |u^*(x,y)|^2 dxdy \right).
\]
By the definition of the maximal function $M_F$, we have
\[
\left| \left\{ (x, y) : M_F(\chi_{\{u^* > \alpha\}}) > \frac{1}{200} \right\} \right| \leq C \left| \left\{ (x, y) : M_F(\chi_{\{u^* > \alpha\}})(x, y) > \frac{1}{200} \right\} \right|
\leq C \int_{\mathbb{R}^{n+m}} M_S(\chi_{\{u^* > \alpha\}})^2(x, y) dxdy
\leq C \int_{\mathbb{R}^{n+m}} \chi_{\{u^* > \alpha\}}(x, y) dxdy
\leq C \left| \left\{ u^* > \alpha \right\} \right|.
\] (3.6)

The estimates in (3.5) and (3.6) yield
\[
\left| \left\{ (x, y) : S_F(u)(x, y) > \alpha \right\} \right| \leq \left| \left\{ (x, y) : M_F(\chi_{\{u^* > \alpha\}}) \leq \frac{1}{200} \text{ and } S_F(u)(x, y) > \alpha \right\} \right|
+ \left| \left\{ (x, y) : M_F(\chi_{\{u^* > \alpha\}}) > \frac{1}{200} \text{ and } S_F(u)(x, y) > \alpha \right\} \right|
\leq C \left( \left| \left\{ u^* > \alpha \right\} \right| + \alpha^{-2} \int_{\left\{ u^* \leq \alpha \right\}} u^*(x, y)^2 \right),
\]
which implies that $\|S_F(u)\|_1 \leq C\|u^*\|_1$.

### 3.2 The estimate $\|u^*\|_1 \lesssim \|u^+\|_1$

As mentioned in the introduction, the flag Hardy space is in some sense intermediate between the classical one parameter and the product Hardy spaces. To deal with the flag non-tangential maximal function, we decompose it by the classical one parameter and the product cases. More precisely, we write, for any $(\bar{x}, \bar{y}) \in \mathbb{R}^{n+m}$, that
\[
u^*(\bar{x}, \bar{y}) = \sup_{(x, y, t, s) \in \Gamma(\bar{x}, \bar{y})} |u(x, y, t, s)|
\leq \sup_{(x, y, t, s) \in \Gamma(\bar{x}, \bar{y})} |u(x, y, t, s)| + \sup_{(x, y, t, s) \in \Gamma(\bar{x}, \bar{y})} |u(x, y, t, s)|
=: u^+_1(\bar{x}, \bar{y}) + u^+_2(\bar{x}, \bar{y}),
\]
where $\Gamma(\bar{x}, \bar{y}) = \{(x, y, t, s) : |x - \bar{x}| \leq t, |y - \bar{y}| \leq t + s\}$.

The main idea to show $\|u^+_1\|_1 \lesssim \|u^+\|_1$ is the following lemma which was proved by Fefferman and Stein in [6] for the classical one parameter Hardy space.

**Lemma 3.4.** Suppose $B$ is a ball in $\mathbb{R}^{d+1}$, with center $(x_0, t_0)$. Let $u$ be harmonic in $B$ and continuous on the closure of $B$. For any $p > 0$,
\[
|u(x_0, t_0)|^p \leq C_p \frac{1}{|B|} \int_B |u(x, t)|^p dxdt.
\]

Suppose $f \in L^1(\mathbb{R}^{n+m})$ and $u(x, y, t, s) = P_{t,s} * f(x, y)$. Note that $u(x, y, t, s)$, as a function of $(x, y, t)$ with a fixed $s$, is harmonic on $\mathbb{R}^{n+m+1}_+$. Lemma 3.4 implies that for any $r > 0$ and
$s \leq t,$
\[ |u(x, y, t, s)|^r \leq C_r \frac{1}{|B_1|} \int_{B_1} |u(x_1, y_1, t_1, s)|^r \, dx_1 \, dy_1 \, dt, \]

where $B_1$ is any ball in $\mathbb{R}^{n+m+1}$ with the radius $t$ and the center $(x, y, t) \in \Gamma_1(\bar{x}, \bar{y})$, where
\[ \Gamma_1(\bar{x}, \bar{y}) = \{(x_1, y_1, t) : |\bar{x} - x_1| \leq 2t, |\bar{y} - y_1| \leq 2t\}. \]

Note that the projection of $B$ which, together with the $L^\infty$ norm of $\mathcal{M}_f$ where
\[ \mathcal{M}_f = \sup_{t,s > t} \int_{\mathbb{R}^n \times \mathbb{R}^m} P_t^{(1)}(x - x_1, z) P_s^{(2)}(y - y_1 - z) \, dz \, f(x, y_1) \, dx_1 \, dy_1, \]

where $M_1$ is the standard Hardy–Littlewood maximal function on $\mathbb{R}^{n+m}$.

As a consequence, this implies that
\[ u_1^*(\bar{x}, \bar{y}) \leq C \left( M_1(|u^+|^r)(\bar{x}, \bar{y}) \right)^{\frac{1}{r}}, \]

which, together with the $L^\frac{1}{r}, 0 < r < 1$, boundedness of the Hardy–Littlewood maximal function $M_1(f)$, implies that
\[ \|u_1^*\|_1 \leq C \|u^+\|_1. \]

Now we estimate $u_2^*(\bar{x}, \bar{y})$. Observe that when $s > t$ the cone $\Gamma(\bar{x}, \bar{y}) = \{(x_1, y_1, t) : \bar{x} - x_1| \leq t, |\bar{y} - y_1| \leq t + s\}$ essentially is the cone in the product setting. Therefore, we write that
\[ u_2^*(\bar{x}, \bar{y}) = \sup_{(x,y,t,s) \in \Gamma(\bar{x}, \bar{y}), s > t} |P_{t,s} * f(x, y)| \]
\[ \leq \sup_{(x,y,t,s) \in \Gamma_2(\bar{x}, \bar{y})} \int_{\mathbb{R}^n \times \mathbb{R}^m} \int_{\mathbb{R}^m} P_t^{(1)}(x - x_1, z) P_s^{(2)}(y - y_1 - z) \, dz \, f(x, y_1) \, dx_1 \, dy_1, \]

where
\[ \Gamma_2(\bar{x}, \bar{y}) = \{(x, y, t, s) : |\bar{x} - x| \leq 2t, |\bar{y} - y| \leq 2s\}. \]

The main idea to estimate the last term above is to introduce the following flag grand maximal function $\mathcal{G}_{\beta, \gamma}(f)(x_0, y_0) :$ for $f \in L^1(\mathbb{R}^{n+m})$ and $(x_0, y_0) \in \mathbb{R}^n \times \mathbb{R}^m$,
\[ \mathcal{G}_{\beta, \gamma}(f)(x_0, y_0) := \sup\{ \langle f, \varphi \rangle : \|\varphi\|_{\mathcal{M}_{\text{flag}}(\beta, \gamma, r_1, r_2, x_0, y_0)} \leq 1, r_1, r_2 > 0\}. \]
By Definition 2.4 it is easy to see that as a function of \((x_1, y_1)\),
\[
\int_{\mathbb{R}^m} P^{(1)}_t(x - x_1, z)P^{(2)}_s(y - y_1 - z)dz
\]
is in \(\tilde{\mathcal{M}}_{flag}(1, 1, t, s, \bar{x}, \bar{y})\) with \((x, y, t, s) \in \Gamma_2(\bar{x}, \bar{y})\) since \(P^{(1)}_t(x - x_1, z) \in \tilde{\mathcal{M}}_{n+m}(1, 1, t, \bar{x}, 0)\) and \(P^{(2)}_s(y - y_1) \in \tilde{\mathcal{M}}_{m}(1, 1, s, \bar{y})\). Moreover, it is also easy to check that
\[
\sup_{(x, y, t, s) \in \Gamma_2(\bar{x}, \bar{y})} \left\| \int_{\mathbb{R}^m} P^{(1)}_t(x - x_1, z)P^{(2)}_s(y - y_1 - z)dz \right\|_{\tilde{\mathcal{M}}_{flag}(1, 1, t, s, \bar{x}, \bar{y})} \leq C,
\]
where \(C\) is an absolute constant independent of \((\bar{x}, \bar{y})\).

As a consequence, we obtain that
\[
u^*_2(\bar{x}, \bar{y}) = \sup_{(x, y, t, s) \in \Gamma_2(\bar{x}, \bar{y})} \left| \left\langle \int_{\mathbb{R}^m} P^{(1)}_t(x - \cdot, z)P^{(2)}_s(y - \cdot - z)dz, f(\cdot, \cdot) \right\rangle \right|
\]
\[
\leq C\mathcal{G}_{1,1}(f)(\bar{x}, \bar{y}).
\]

It suffices to prove that for \(f \in L^1(\mathbb{R}^{n+m})\) and \(r > 0\),
\[
\mathcal{G}_{1,1}(f)(\bar{x}, \bar{y}) \leq C\left(M_1\left(M_2(|u^+|^r)\right)(\bar{x}, \bar{y})\right)^{\frac{1}{r}} + C\left(M_2\left(M_1(|u^+|^r)\right)(\bar{x}, \bar{y})\right)^{\frac{1}{r}},
\]
where \(M_1\) and \(M_2\) are the Hardy-Littlewood maximal functions on \(\mathbb{R}^{n+m}\) and \(\mathbb{R}^m\), respectively.

We first claim that
\[
|\langle f, \psi \rangle| \leq C\left(M_1\left(M_2(|u^+|^r)\right)(\bar{x}, \bar{y})\right)^{\frac{1}{r}}
\]
for \(r < 1\) and close to 1, \(f \in L^1(\mathbb{R}^{n+m})\), and for every \(\psi \in \mathcal{M}_{flag}(1, 1, 2^{-j_1}, 2^{-k_1}, \bar{x}, \bar{y})\) with the norm \(\|\psi\|_{\mathcal{M}_{flag}(1, 1, 2^{-j_1}, 2^{-k_1}, \bar{x}, \bar{y})} \leq 1\).

The key idea to show the above claim is to apply the discrete Calderón reproducing formula. To see this, consider the following approximations to the identity on \(\mathbb{R}^{n+m}\): For each \(j \in \mathbb{Z}\), define the operator
\[
P^{(1)}_{2^{-j}} := P^{(1)}_{2^{-j}}
\]
with the kernel \(P^{(1)}_{2^{-j}}(x, y) := P^{(1)}_1(x, y)\).

It is easy to see that
\[
\lim_{j \to \infty} P^{(1)}_{2^{-j}} = \lim_{j \to \infty} P^{(1)}_{2^{-j}} = Id \quad \text{and} \quad \lim_{j \to -\infty} P^{(1)}_{2^{-j}} = \lim_{j \to -\infty} P^{(1)}_{2^{-j}} = 0
\]
in the sense of \(L^2(\mathbb{R}^{n+m})\). And we further have
\[
\int_{\mathbb{R}^{n+m}} P^{(1)}_{2^{-j}}(x, y)dxdy = 1.
\]
Set $Q^{(1)}_j := P^{(1)}_j - P^{(1)}_{j-1}$. Then $Q^{(1)}_j(x, y)$, the kernel of $Q^{(1)}_j$ satisfies the same size and smoothness conditions as $P^{(1)}_j(x, y)$ does, and
\[
\int_{\mathbb{R}^{n+m}} Q^{(1)}_j(x, y) \, dx \, dy = 0.
\]
The operators $P^{(2)}_k$ and $Q^{(2)}_k$ on $\mathbb{R}^m$ are defined similarly.

Repeating the same proof as in Theorem 2.5, we have the following reproducing formula: there exist functions $\phi_{j,k}(x, y, x_I, y_J) \in \mathcal{M}_{flag}(\beta, \gamma, 2^{-j}, 2^{-k}, x_I, y_J)$ and a fixed large integer $N$ such that for $f(x, y) = \int_{\mathbb{R}^m} f_1(x, y - z) f_2(z) \, dz$ with $f_1 \in \mathcal{M}_{n+m}(\beta, \gamma, r_1, x_0, y_0)$ and $f_2 \in \mathcal{M}_{m}(\beta, \gamma, r_2, z_0)$,
\[
f(x, y) = \sum_j \sum_k \sum_I \sum_J |I||J| \phi_{j,k}(x, y, x_I, y_J) Q_{j,k}(f)(x_I, y_J),
\]
where the series converges in $L^2(\mathbb{R}^{n+m})$ and in the flag test function space, and $I \subset \mathbb{R}^n$, $J \subset \mathbb{R}^m$ are dyadic cubes with side-lengths $\ell(I) = 2^{-j-N}$, $\ell(J) = 2^{-(j\wedge k-N)}$, $x_I$ and $y_J$ are any fixed points in $I$ and $J$, respectively, and
\[
Q_{j,k}(f)(x_I, y_J) = \int_{\mathbb{R}^{n+m}} Q_{j,k}(x_I - x, y_J - y) f(x, y) \, dx \, dy
\]
with the kernel
\[
Q_{j,k}(x, y) = \int_{\mathbb{R}^m} Q^{(1)}_j(x, y - z) Q^{(2)}_k(z) \, dz.
\]

Now applying (3.9) to the left-hand side of (3.8), we have
\[
|\langle f, \psi \rangle| = \left| \sum_j \sum_k \sum_I \sum_J |I||J| \langle \psi, \phi_{j,k}(\cdot, \cdot, x_I, y_J) \rangle Q_{j,k}(f)(x_I, y_J) \right|
\leq C \sum_j \sum_k \sum_I \sum_J |I||J| 2^{-|j-j_I|\beta} 2^{-|k-k_I|\beta} 2^{-(j\wedge j_I)\gamma} 2^{-(j\wedge j_I)\gamma} \inf_{z_1 \in \bar{I}, z_2 \in J} \inf_{\bar{z}} |u^+(z_1, z_2)|.
\]

Here in the last inequality we use the following estimates:

(1) the almost orthogonality estimate:
\[
|\langle \psi, \phi_{j,k}(\cdot, \cdot, x_I, y_J) \rangle| \leq C 2^{-|j-j_I|\beta} 2^{-|k-k_I|\beta} 2^{-(j\wedge j_I)\gamma} 2^{-(j\wedge j_I)\gamma} \inf_{z_1 \in \bar{I}, z_2 \in J} \inf_{\bar{z}} |u^+(z_1, z_2)|
\]
for $\beta, \gamma < 1$. For the proof see [18, page 2840] for the one-parameter case and [28] for similar estimates on homogeneous groups.
(2) the fact that \(x_I \) and \(y_J \) are any fixed points in \(I \) and \(J \), implies that we can choose \(x_I \in I \) and \(y_J \in J \) such that
\[
|Q_{j,k}(f)(x_I, y_J) - \inf_{z_1, z_2 \in J} |Q_{j,k}(f)(z_1, z_2)|\]
\[
\leq 2 \inf_{z_1, z_2 \in J} \left| \int_{\mathbb{R}^m} Q_j^{(1)}(z_1 - x, z_2 - y - z) \, dz \right| f(x, y) \, dx \, dy
\]
\[
= 2 \inf_{z_1, z_2 \in J} \left| \int_{\mathbb{R}^m} Q_j^{(1)}(z_1 - x, z_2 - y - z) - P_j^{(1)}(z_1 - x, z_2 - y - z) \right|
\times \left( P_j^{(2)}(z) - P_j^{(2)}(z) \right) \, dz \, dx \, dy
\]
\[
\leq 8 \inf_{z_1, z_2 \in J} |u^+(z_1, z_2)|.
\]
To estimate the last term in (3.10), observe that for \(0 < r < 1 \),
\[
\sum_j \sum_{k} \sum_l |I| |J| \left| 2^{-|j-j_1|/\beta} 2^{-|k-k_1|/\beta} \frac{2^{-(j \wedge j_1)\gamma r}}{(2^{-(j \wedge j_1)\gamma r} + |x_1 - x|)^{(m+\gamma)r}} \right| \left| 2^{-[(k \wedge k_1) \wedge (j \wedge j_1)]\gamma r} \right| \inf_{z_1, z_2 \in J} |u^+(z_1, z_2)|^r \right\}^{1/r}.
\]
Note that \(\ell(I) = 2^{-j-I} \) and \(\ell(J) = 2^{-j-k-J} \). Write
\[
\sum_j \sum_{k} |I| |J| \left| 2^{-|j-j_1|/\beta} 2^{-|k-k_1|/\beta} \frac{2^{-(j \wedge j_1)\gamma r}}{(2^{-(j \wedge j_1)\gamma r} + |x_1 - x|)^{(m+\gamma)r}} \right| \left| 2^{-[(k \wedge k_1) \wedge (j \wedge j_1)]\gamma r} \right| \inf_{z_1, z_2 \in J} |u^+(z_1, z_2)|^r \right\}^{1/r}.
\]
\[
\leq C \frac{2^{-jn(r-1)/2} \gamma (j \wedge k) m(r-1)}{2^{-(j \wedge j_1)\gamma r} \left(2^{-(j \wedge j_1)\gamma r} + |x_1 - x|)^{(m+\gamma)r} \right) \left(2^{-(k \wedge k_1) \wedge (j \wedge j_1)\gamma r} + |y_j - \bar{y}|)^{(m+\gamma)r} \right) \inf_{z_1, z_2 \in J} |u^+(z_1, z_2)|^r \right\}^{1/r}.
\]
\[
\leq C \frac{2^{-jn(r-1)/2} \gamma (j \wedge k) m(r-1)}{2^{-(j \wedge j_1)\gamma r} \left(2^{-(j \wedge j_1)\gamma r} + |x_1 - x|)^{(m+\gamma)r} \right) \left(2^{-(k \wedge k_1) \wedge (j \wedge j_1)\gamma r} + |y_j - \bar{y}|)^{(m+\gamma)r} \right) \times |u^+(x, y)|^r \, dx \, dy
\]
\[
\leq C \frac{2^{-jn(r-1)/2} \gamma (j \wedge k) m(r-1)}{2^{-(j \wedge j_1)\gamma r} \left(2^{-(j \wedge j_1)\gamma r} \right) \left(2^{-(k \wedge k_1) \wedge (j \wedge j_1)\gamma r} + |y_j - \bar{y}|)^{(m+\gamma)r} \right) \times 2^{-[(k \wedge k_1) \wedge (j \wedge j_1)](m+\gamma)r} \left(M_1 \left(M_2 \left(\left|u^+\right|^r\right)\right) \right)(x, y).
\]
A simple computation shows that if \(\frac{m+n}{m+n+\beta} < r < 1 \), then
\[
\sum_j \sum_k 2^{-|j-j_1|/\beta} 2^{-|k-k_1|/\beta} 2^{-jn(r-1)/2} \gamma (j \wedge k) m(r-1) 2^{-[(k \wedge k_1) \wedge (j \wedge j_1)](m+\gamma)r} \leq C.
\]
Thus, we obtain that the right-hand side of (3.10) is bounded by

$$\left( M_1\left( M_2\left( |u^+|''\right) (\bar{x}, \bar{y}) \right) \right)^{1/2},$$

which implies (3.8).

We now prove (3.7). For every $\varphi$ with

$$\varphi(x, y) = \int_{\mathbb{R}^m} \varphi^{(1)}(x, y - z) \varphi^{(2)}(z) dz,$$

where $\varphi^{(1)}(x, y) \in \tilde{\mathcal{M}}_{n+m}(1, 1, t, \bar{x}, 0)$ with $\|\varphi^{(1)}\|_{\tilde{\mathcal{M}}_{n+m}(1, 1, t, \bar{x}, 0)} \leq 1$, and $\varphi^{(2)}(z) \in \tilde{\mathcal{M}}_m(1, 1, s, \bar{y})$ with $\|\varphi^{(2)}\|_{\tilde{\mathcal{M}}_m(1, 1, s, \bar{y})} \leq 1$.

Let

$$\sigma_1 := \int_{\mathbb{R}^{n+m}} \varphi^{(1)}(x, y) dx dy, \quad \sigma_2 := \int_{\mathbb{R}^m} \varphi^{(2)}(z) dz.$$

It is obvious that $|\sigma_1|, |\sigma_2| \leq C$. We set

$$\psi^{(1)}(x, y) := \frac{1}{1 + \sigma_1 C} \left[ \varphi^{(1)}(x, y) - \sigma_1 P_{j_1}^{(1)}(\bar{x} - x, y) \right],$$

$$\psi^{(2)}(z) := \frac{1}{1 + \sigma_2 C} \left[ \varphi^{(2)}(z) - \sigma_2 P_{k_1}^{(2)}(z - \bar{y}) \right],$$

where $j_1 := |\log_2 t| + 1$ and $k_1 := |\log_2 s| + 1$.

Then for an appropriate constant $C$, the function $\psi(x, y) = \int_{\mathbb{R}^m} \psi^{(1)}(x, y - z) \psi^{(2)}(z) dz$ is in $\mathcal{M}_{\text{flag}}(1, 1, t, s, \bar{x}, \bar{y})$ with $\|\psi\|_{\mathcal{M}_{\text{flag}}(1, 1, t, s, \bar{x}, \bar{y})} \leq 1$.

Based on the definition of $\psi$, we have

$$|\langle f, \varphi \rangle| = \left| \int_{\mathbb{R}^{n+m}} f(x, y) \varphi(x, y) dx dy \right|$$

$$= \left| \int_{\mathbb{R}^{n+m}} f(x, y) \left( \int_{\mathbb{R}^m} \varphi^{(1)}(x, y - z) \varphi^{(2)}(z) dz \right) dx dy \right|$$

$$= \left| \int_{\mathbb{R}^{n+m}} f(x, y) \int_{\mathbb{R}^m} \left[ (1 + \sigma_1 C) \psi^{(1)}(x, y - z) + \sigma_1 P_{j_1}^{(1)}(\bar{x} - x, y - z) \right] \right. \times \left. \left[ (1 + \sigma_2 C) \psi^{(2)}(z) + \sigma_2 P_{k_1}^{(2)}(z - \bar{y}) \right] dz dx dy \right|$$

$$\leq \left| \int_{\mathbb{R}^{n+m}} f(x, y) \int_{\mathbb{R}^m} (1 + \sigma_1 C) \psi^{(1)}(x, y - z)(1 + \sigma_2 C) \psi^{(2)}(z) dz dx dy \right|$$

$$+ \left| \int_{\mathbb{R}^{n+m}} f(x, y) \int_{\mathbb{R}^m} \sigma_1 P_{j_1}^{(1)}(\bar{x} - x, y - z)(1 + \sigma_2 C) \psi^{(2)}(z) dz dx dy \right|$$

$$+ \left| \int_{\mathbb{R}^{n+m}} f(x, y) \int_{\mathbb{R}^m} (1 + \sigma_1 C) \psi^{(1)}(x, y - z) \sigma_2 P_{k_1}^{(2)}(z - \bar{y}) dz dx dy \right|$$

$$+ \left| \int_{\mathbb{R}^{n+m}} f(x, y) \int_{\mathbb{R}^m} \sigma_1 P_{j_1}^{(1)}(\bar{x} - x, y - z) \sigma_2 P_{k_1}^{(2)}(z - \bar{y}) dz dx dy \right|$$

$$=: A_1 + A_2 + A_3 + A_4.$$
For the term $A_1$, from [3.8] we obtain that

$$A_1 \leq C \left( M_1 \left( M_2 \left( |u^+| \right) \right)(\bar{x}, \bar{y}) \right)^{\frac{1}{2}}.$$

For the term $A_4$, by definition we have

$$A_4 \leq C u^+(\bar{x}, \bar{y}) = C \left( \left| u^+(\bar{x}, \bar{y}) \right| \right)^{\frac{1}{2}} \leq C \left( M_1 \left( M_2 \left( |u^+| \right) \right)(\bar{x}, \bar{y}) \right)^{\frac{1}{2}}.$$

As for $A_2$, we write

$$A_2 = \left| \frac{\int_{\mathbb{R}^m} \int_{\mathbb{R}^{n+m}} f(x, y) \sigma_1 P_{11}^{(1)}(\bar{x} - x, y - z) dx dy (1 + \sigma_2 C) \psi_2(z) dz}{\int_{\mathbb{R}^m} F_{x,j}^{(1)}(z)(1 + \sigma_2 C) \psi_2(z) dz} \right|,$$

where

$$F_{x,j}^{(1)}(z) := \int_{\mathbb{R}^{n+m}} f(x, y) \sigma_1 P_{11}^{(1)}(\bar{x} - x, y - z) dx dy.$$

Then following the same approach as above, by using the reproducing formula in terms of $Q_{k}^{(2)}$, the almost orthogonality estimates, we obtain that

$$A_2 \leq C \left( M_2 \left( \sup_{s > 0} \left| \int_{\mathbb{R}^m} F_{x,j}^{(1)}(z) P_{s}^{(2)}(z) dz \right|^r \right)(\bar{y}) \right)^{\frac{1}{r}} \left( M_2 \left( \sup_{s > 0} \left| \int_{\mathbb{R}^m} f(x, y) \sigma_1 P_{11}^{(1)}(\bar{x} - x, y - z) dx dy P_{s}^{(2)}(z) dz \right|^r \right)(\bar{y}) \right)^{\frac{1}{r}} \leq C \left( M_2 \left( \sup_{s > 0} \left| \int_{\mathbb{R}^m} f(x, y) \sigma_1 P_{11}^{(1)}(\bar{x} - x, y - z) dx dy P_{s}^{(2)}(z) dz \right|^r \right)(\bar{y}) \right)^{\frac{1}{r}} \left( M_2 \left( \sup_{s > 0} \left| \int_{\mathbb{R}^m} f(x, y) \sigma_1 P_{11}^{(1)}(\bar{x} - x, y - z) dx dy P_{s}^{(2)}(z) dz \right|^r \right)(\bar{y}) \right)^{\frac{1}{r}}.$$

Symmetrically, we obtain that

$$A_3 \leq C \left( M_1 \left( M_2 \left( |u^+| \right) \right)(\bar{x}, \bar{y}) \right)^{\frac{1}{2}}.$$

Combining the estimates of $A_1, A_2, A_3$ and $A_4$, we obtain that [3.7] holds.

### 3.3 The estimate $\|u^+\|_1 \lesssim \sum_{j=1}^{n+m} \sum_{k=1}^{n} \|R_{j,k}(f)\|_1 + \|f\|_1$

Let $P_{t}^{(1)}$ be the Poisson kernel on $\mathbb{R}^{n+m}$ and $Q_{j,t}^{(1)}$ to be the $j$-th conjugate Poisson kernel on $\mathbb{R}^{n+m}$. Then following [6, Section 8], it is easy to verify that $u := u_{0}^{(1)} = P_{t}^{(1)} * f$, $u_{j}^{(1)} = Q_{j,t}^{(1)} * f = P_{t}^{(1)} * (R_{j}^{(1)} * f)$, $j = 1, 2, \ldots, n+m$ is a $(n+m+1)$-tuple of harmonic functions that satisfy the following system of equations:

$$\begin{aligned}
\frac{\partial u_{i}^{(1)}}{\partial x_{j}} &= \frac{\partial u_{j}^{(1)}}{\partial x_{j}}, \quad 0 \leq i, j \leq n + m; \\
\sum_{j=0}^{n+m} \frac{\partial u_{i}^{(1)}}{\partial x_{j}} &= 0.
\end{aligned}$$

(3.11)
Here we use $R_j^{(1)}$ to denote the $j$th Riesz transform on $\mathbb{R}^{n+m}$, $j = 1, 2, \ldots, n + m$. Similarly, we use $P_s^{(2)}$ to denote the Poisson kernel on $\mathbb{R}^m$ and $Q_{k,s}^{(2)}$ to denote the $k$-th conjugate Poisson kernel on $\mathbb{R}^m$.

Again, following [8, Section 8], we can verify that $u := u_0^{(2)} = P_s^{(2)} *_{\mathbb{R}^m} f$, $u_k^{(2)} = Q_{k,s}^{(2)} *_{\mathbb{R}^m} f = P_s^{(2)} *_{\mathbb{R}^m} (R_k^{(2)} *_{\mathbb{R}^m} f)$, $k = 1, 2, \ldots, m$ is a $(m+1)$-tuple of harmonic functions that satisfy the following system of equations:

\[
\begin{align*}
\frac{\partial u_{i,j}}{\partial x_j} &= \frac{\partial u_{i,k}}{\partial x_j}, \quad 0 \leq i, j \leq m; \\
\sum_{j=0}^m \frac{\partial u_{i,j}}{\partial x_j} &= 0.
\end{align*}
\] (3.12)

Here we use $R_k^{(2)}$ to denote the $k$th Riesz transform on $\mathbb{R}^m$, $k = 1, 2, \ldots, m$.

We now set $u(x, y, t, s) = u_0(x, y, t, s) = P_t^{(1)} *_{\mathbb{R}^m} P_s^{(2)} * f(x, y)$. Then we define

\[u_{1,0}(x, y, t, s) = Q_{1,t}^{(1)} *_{\mathbb{R}^m} P_s^{(2)} * f(x, y)\quad \text{and} \quad u_{0,1}(x, y, t, s) = P_t^{(1)} *_{\mathbb{R}^m} Q_{1,s}^{(2)} * f(x, y),\]

and similarly,

\[u_{j,k}(x, y, t, s) = Q_{j,t}^{(1)} *_{\mathbb{R}^m} Q_{k,s}^{(2)} * f(x, y),\]

for $j = 1, \ldots, n + m$ and $k = 1, \ldots, m$.

We first point out that for $k = 1, \ldots, m$, the tuple $(u_{0,k}, u_{1,k}, \ldots, u_{n+m,k})$ satisfies the Cauchy–Riemann equation in (3.11), and that for $j = 1, \ldots, n + m$ the tuple $(u_{j,0}, u_{j,1}, \ldots, u_{j,m})$ satisfies the Cauchy–Riemann equation in (3.12).

Following the idea in [8, Section 8], we consider the matrix-valued function

\[
F = \begin{bmatrix}
    u_{0,0} & \cdots & u_{0,m} \\
    \cdots & \cdots & \cdots \\
    u_{n+m,0} & \cdots & u_{n+m,m}
\end{bmatrix} = P_t^{(1)} *_{\mathbb{R}^m} P_s^{(2)} * \widetilde{F},
\]

where we denote

\[
\widetilde{F} = \begin{bmatrix}
f & \cdots & R_m^{(2)} *_{\mathbb{R}^m} f \\
\cdots & \cdots & \cdots \\
R_{m+n}^{(1)} * f & \cdots & R_{m+n}^{(1)} *_{\mathbb{R}^m} P_m^{(2)} * f
\end{bmatrix}.
\]
We obtain
\[
\sup_{t > 0} \sup_{s > 0} \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} |F(x, y, t, s)| \, dx \, dy \\
\leq \sup_{t > 0} \sup_{s > 0} \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} \left( \sum_{j=0}^{n+m} \sum_{k=0}^{m} |u_{j,k}(x, y, t, s)|^2 \right)^{\frac{1}{2}} \, dx \, dy \\
\leq C \sum_{j=0}^{n+m} \sum_{k=0}^{m} \sup_{t > 0} \sup_{s > 0} \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} |Q_j^{(1)} *_{\mathbb{R}^n} Q_k^{(2)} * f(x, y)| \, dx \, dy \\
\leq C \sum_{j=0}^{n+m} \sum_{k=0}^{m} \sup_{t > 0} \sup_{s > 0} \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} |P_t^{(1)} *_{\mathbb{R}^n} P_s^{(2)} * (R_j^{(1)} *_{\mathbb{R}^m} R_k^{(2)} * f)(x, y)| \, dx \, dy \\
\leq C \sum_{j=0}^{n+m} \sum_{k=0}^{m} \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} |(R_j^{(1)} *_{\mathbb{R}^n} R_k^{(2)} * f)(x, y)| \, dx \, dy,
\]
where the last inequality follows from the fact that
\[
\int_{\mathbb{R}^{n+m}} P_t^{(1)}(x - x_1, y - y_1) \, dx \, dy = C_{n+m} \quad \text{and} \quad \int_{\mathbb{R}^m} P_s^{(2)}(y - y_1) \, dy = C_m
\]
for all \( t, s > 0, x_1 \in \mathbb{R}^n \) and \( y_1 \in \mathbb{R}^m \).

Next it suffices to show
\[
\|u^+\|_{L^1(\mathbb{R}^n \times \mathbb{R}^m)} \leq C \sup_{t > 0} \sup_{s > 0} \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} |F(x, y, t, s)| \, dx \, dy. \tag{3.13}
\]

To see this, we have that for \( q < 1, \)
\[
|F(x, y, t + \epsilon, s + \epsilon_2)|^q = \left| P_t^{(1)} * P_s^{(2)} * F(x, y) \right|^q \\
= \left| P_t^{(1)} * P_{t_1}^{(1)} * P_{s+\epsilon_2}^{(2)} * F(x, y) \right|^q \\
= \left| P_t^{(1)} * P_t^{(1)} * F(x, y, \epsilon_1, s + \epsilon_2) \right|^q \\
\leq C_{q,m} \sum_{k=0}^{m} \left| P_t^{(1)} * F_k(x, y, \epsilon_1, s + \epsilon_2) \right|^q,
\]
where for each \( k, F_k \) is the \( k \)th column in the matrix \( F \). Since \( P_t^{(1)} * F_k \) satisfies the generalised Cauchy–Riemann equations in (3.11) for the variable \( (x, y, t) \), we get that \( |P_t^{(1)} * F_k|^q \) is subharmonic for \( q \geq \frac{n+m-1}{n+m} \). Then from the subharmonic inequality [32, Equation (59), Section 4.2, Chapter 3] we have that for \( q \geq \frac{n+m-1}{n+m}, x \in \mathbb{R}^n, y \in \mathbb{R}^m, t > 0 \) and \( \epsilon_1 > 0, \)
\[
\left| P_t^{(1)} * F_k(x, y, \epsilon_1, s + \epsilon_2) \right|^q \leq P_t^{(1)} * |F_k(x, y, \epsilon_1, s + \epsilon_2)|^q,
\]
which implies that
\[
|F(x, y, t + \epsilon_1, s + \epsilon_2)|^q \leq C_{q,m} \sum_{k=0}^{m} P_t^{(1)} * |F_k(x, y, \epsilon_1, s + \epsilon_2)|^q \tag{3.14}
\]
\[ \leq C_{q,m} P_t^{(1)} * |F(x, y, \epsilon_1, s + \epsilon_2)|^q. \]

And we use the basic fact that \( |F|^q = (\sum_{k=0}^{m} |F_k|^2)^{\frac{q}{2}} \approx \sum_{k=0}^{m} |F_k|^q \).

Again, for \( F(x, y, \epsilon_1, s + \epsilon_2) \), we have
\[
|F(x, y, \epsilon_1, s + \epsilon_2)|^q = |P_s^{(2)} * F(x, y, \epsilon_1, \epsilon_2)|^q \leq C_{q,n+m} \sum_{j=0}^{n+m} |P_s^{(2)} * F_j(x, y, \epsilon_1, \epsilon_2)|^q,
\]

where for each \( j \), \( \tilde{F}_j \) is the \( j \)th row in the matrix \( F \). Since \( P_s^{(2)} * F_j \) satisfies the generalised Cauchy–Riemann equations in (3.11) for the variable \((y, s)\), we get that \( |P_s^{(2)} * F_j|^q \) is subharmonic for \( q \geq \frac{m-1}{m} \). Then again, from the subharmonic inequality [32, Equation (59), Section 4.2, Chapter 3] we have that for \( q \geq \frac{m-1}{m} \), \( y \in \mathbb{R}^m \), \( s > 0 \) and \( \epsilon_2 > 0 \),
\[
|F(x, y, \epsilon_1, s + \epsilon_2)|^q \leq C_{q,n+m} \sum_{j=0}^{n+m} |P_s^{(2)} * F_j(x, y, \epsilon_1, \epsilon_2)|^q.
\]

And we use the basic fact that \( |F|^q = (\sum_{j=0}^{n+m} |F_j|^2)^{\frac{q}{2}} \approx \sum_{j=0}^{n+m} |F_j|^q \).

Combining the estimates of (3.14) and (3.15), we obtain that
\[
|F(x, y, t + \epsilon_1, s + \epsilon_2)|^q \leq C_{q,n,m} P_t^{(1)} * P_s^{(2)} * F(x, y, \epsilon_1, \epsilon_2)|^q.
\]

Then, following the convergence argument in [6, Section 8], also in [32, Section 4.2], we obtain that
\[
\|u^+\| \leq C_{m,n} \sup_{t>0} \sup_{s>0} \int_{\mathbb{R}^n} \int_{\mathbb{R}^m} |F(x, y, t, s)| dx dy.
\]

which implies that the claim (3.13) holds.

4 Characterizations of the flag Hardy spaces

In this section, the following estimates will be established:

(I) \( \|S_F(f)\|_1 \lesssim \|S_F(u)\|_1 \),

(II) \( \|u^+\| \approx \|M^+_\Phi(f)\|_1 \),

(III) \( \|u^+\| \approx \|M^+_\Phi(f)\|_1 \),
4.1 The estimate $\|S_F(f)\|_1 \lesssim \|S_F(u)\|_1$

The estimate $\|S_F(f)\|_1 \lesssim \|S_F(u)\|_1$, follows from the same ideas in Section 2.2 and 2.3. More precisely, we first need to establish the following discrete Calderón reproducing formula. For this purpose, let $\phi^{(1)}(x,y) \in \mathcal{S}(\mathbb{R}^{n+m})$, be radial and satisfy the following conditions:

(i) $\text{supp } \phi^{(1)} \subset B(0,1)$, where $B(0,1)$ is the unit ball in $\mathbb{R}^{n+m}$;

(ii) $\int_{\mathbb{R}^{n+m}} x^\alpha y^\beta \phi^{(1)}(x,y) \, dx dy = 0$, where $|\alpha| + |\beta| \leq 2(n \lor m)$;

(iii) $\int_0^\infty e^{-u} \hat{\phi}^{(1)}(u) \, du = -1$.

In fact, $\phi^{(1)}(x,y)$ can be constructed as follows. Choose $h^{(1)} \in \mathcal{S}(\mathbb{R}^{n+m})$, radial and supported in $B(0,1)$. Let $k = 4(n \lor m)$ and $\phi^{(1)}(x,y) = \Delta^k h^{(1)}(x,y)$. Multiplying by an appropriate constant, we can see that such $\phi^{(1)}(x,y)$ satisfies all the conditions above.

Similarly, choosing $h^{(2)} \in \mathcal{S}(\mathbb{R}^m)$, radial and supported in $B(0,1)$ and $\phi^{(2)}(z) = \Delta^k h^{(2)}(z)$. Multiplying by an appropriate constant, we obtain that $\phi^{(2)}(z) \in \mathcal{S}(\mathbb{R}^m)$, is radial and satisfies the following conditions:

(i) $\text{supp } \phi^{(2)} \subset B(0,1)$, where $B(0,1)$ is the unit ball in $\mathbb{R}^m$;

(ii) $\int_{\mathbb{R}^m} z^\gamma \phi^{(2)}(z) \, dz = 0$, where $|\gamma| \leq 2(n \lor m)$;

(iii) $\int_0^\infty e^{-u} \hat{\phi}^{(2)}(u) \, du = -1$.

Let $\phi(x,y) = \phi^{(1)} \ast_{\mathbb{R}^m} \phi^{(2)}(x,y)$ and $\phi_{t,s}(x,y) = \phi^{(1)} \ast_{\mathbb{R}^m} \phi^{(2)}(x,y)$. Repeating the same proof as in Theorem 2.3 leads to the following statement.

**Theorem 4.1.** There exist $\phi_{j,k,I,J}(x,y) \in \mathcal{M}_{\text{flag}}(\beta, \gamma, 2^{-j}, 2^{-k}, x_I, y_J)$ and a fixed large integer $N$ such that

$$f(x,y) = \sum_j \sum_k \sum_I \sum_J |I| |J| \phi_{j,k,I,J}(x,y) \int_{2^{-k-N}}^{2^{-k-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} \phi_{t,s} \ast \left( ts \frac{\partial}{\partial t} \frac{\partial}{\partial s} P_{t,s} \right) \ast f(x_I, y_J) \, dt \, ds,$$

where $I \subset \mathbb{R}^n$ and $J \subset \mathbb{R}^m$ are dyadic cubes with side-lengths $\ell(I) = 2^{-j-N}$ and $\ell(J) = 2^{-(j-N) \land (k-N)}$, $x_I$ and $y_J$ are any fixed points in $I$ and $J$, respectively. Moreover, for $f \in L^1(\mathbb{R}^{n+m})$ and $\psi$ is the same as in (1.1),

$$\langle f, \psi \rangle = \left\langle \sum_j \sum_k \sum_I \sum_J |I| |J| \phi_{j,k,I,J}(\cdot, \cdot) \int_{2^{-k-N}}^{2^{-k-N+1}} \int_{2^{-j-N}}^{2^{-j-N+1}} \left( ts \frac{\partial}{\partial t} \frac{\partial}{\partial s} P_{t,s} \right) \ast f(x_I, y_J) \, dt \, ds, \psi \right\rangle.$$
Applying the same proof as in Section 2.2 gives the following.

**Theorem 4.2.** Let \( f \in L^1(\mathbb{R}^{n+m}) \), we have

\[
\left\| \sum_j \sum_k \sum_I \sum_J \int_{2^{-j}}^{2^{-j}+1} \int_{2^{-k}}^{2^{-k}+1} \sup_{u \in I, v \in J} \left| \psi_{t,s} * f(u,v) \right| \frac{2dt}{t} \frac{ds}{s} \chi_I(x) \chi_J(y) \right\|_1 \\
\approx \left\| \sum_j \sum_k \sum_I \sum_J \int_{2^{-j}}^{2^{-j}+1} \int_{2^{-k}}^{2^{-k}+1} \inf_{u \in I, v \in J} \left| (ls \frac{\partial}{\partial t} \frac{\partial}{\partial s} P_{t,s}) * f(u,v) \right| \frac{2dt}{t} \frac{ds}{s} \chi_I(x) \chi_J(y) \right\|_1,
\]

where \( I \subset \mathbb{R}^n \) and \( J \subset \mathbb{R}^m \) are dyadic cubes with side-lengths \( \ell(I) = 2^{-j} \) and \( \ell(J) = 2^{-(j \wedge k)} \) and the lower left-corners \( l_1 2^{-j} \) and \( l_2 2^{-(j \wedge k)} \), respectively.

The estimate \( \| S_F(f) \|_1 \lesssim \| S_F(u) \|_1 \), then follows from Theorem 4.2 as in Section 2.3. We leave the details to the reader.

### 4.2 The equivalence \( \| u^* \|_1 \approx \| M^*_\phi(f) \|_1 \)

We first show

\[
\| u^* \|_1 \leq C \| M^*_\phi(f) \|_1.
\]

To do this, we introduce the “tangential” maximal function \( M^{**}_N \) (depending on a parameter \( N \)) by

\[
M^{**}_N(f)(x,y) = \sup_{u \in \mathbb{R}^n, v \in \mathbb{R}^m, t,s > 0} \left| f * \phi_{t,s}(x - u, y - v) \right| \frac{1}{\left( 1 + \frac{|u|}{t} \right)^N \left( 1 + \frac{|v|}{t + s} \right)^N}.
\]

Observe that

\[
M^+_\phi(f)(x,y) \leq M^*_\phi(f)(x,y) \leq 2^{2N} M^{**}_N(f)(x,y).
\]

Next, we introduce the grand maximal functions. For this purpose, we first note that on \( S(\mathbb{R}^{n+m}) \) one has a denumerable collection of seminorms \( \| \cdot \|_{\alpha_1,\alpha_2,\beta_1,\beta_2} \) given by

\[
\| \phi \|_{\alpha_1,\alpha_2,\beta_1,\beta_2} = \sup_{(x,y) \in \mathbb{R}^{n+m}} \left| x^{\alpha_1} y^{\alpha_2} \partial_x^{\beta_1} \partial_y^{\beta_2} \phi(x,y) \right|.
\]

Similarly, on \( S(\mathbb{R}^m) \), seminorms \( \| \cdot \|_{\alpha,\beta} \) are given by

\[
\| \phi \|_{\alpha,\beta} = \sup_{z \in \mathbb{R}^m} \left| z^\alpha \partial_2^\beta \phi(z) \right|.
\]

Let \( F^{(1)} = \{ \| \cdot \|_{\alpha_1,\alpha_2,\beta_1,\beta_2} \} \) be any finite collections of seminorms on \( S(\mathbb{R}^{n+m}) \) and \( F^{(2)} = \{ \| \cdot \|_{\alpha_1,\beta_2} \} \) be any finite collections of seminorms on \( S(\mathbb{R}^m) \). Set

\[
\mathcal{F} = \{ \phi \in \mathcal{S}F(\mathbb{R}^n \times \mathbb{R}^m) : \text{for all } \phi^z \in \mathcal{S}(\mathbb{R}^{n+m} \times \mathbb{R}^m) \text{ satisfying } \phi(x,y) = \int_{\mathbb{R}^m} \phi^z(x,y - z,z)dz, \}
\]
We then define

\[ M_{\mathcal{F}}(f)(x, y) = \sup_{\phi \in \mathcal{F}} M^+_\phi(f)(x, y). \]

We need the following results.

**Lemma 4.3.** If \( M^*_\phi(f) \in L^1(\mathbb{R}^{n+m}) \) and \( N > 2(n \vee m) \), then \( M^*_N(f) \in L^1(\mathbb{R}^{n+m}) \) with

\[ \| M^*_N(f) \|_1 \leq C_{N,p} \| M^*_\phi(f) \|_1. \]  

**Proof.** We point out that if

\[ M^*_{\phi,a,b}(f)(x, y) = \sup_{(x_1, y_1, t, s) \in \Gamma_{a,b}(x, y)} |\phi_{t,s} \ast f(x_1, y_1)|, \]

where \( \Gamma_{a,b}(x, y) = \{(x_1, y_1, t, s) : |x - x_1| \leq at, |y - y_1| \leq b(t + s)\} \), then

\[ \int_{\mathbb{R}^n \times \mathbb{R}^m} |M^*_{\phi,a,b}(f)(x, y)|^p \, dx \, dy \leq C_{n,m} (1 + a)^n (1 + b)^m \int_{\mathbb{R}^n \times \mathbb{R}^m} |M^*_\phi(f)(x, y)|^p \, dx \, dy. \]  

This can be obtained by mimicking the proof in \cite{32} §2.5, Chapter 2. Observing that

\[ \frac{|f \ast \phi_{t,s}(x - u, y - v)|}{(1 + |u|)^N (1 + |v|)^N} \leq \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} 2^{(1-k)N} 2^{(1-\ell)N} |M^*_{\phi,2^k,2^{\ell+1}}(f)(x, y)| \]

for all \( u \in \mathbb{R}^n, v \in \mathbb{R}^m, t, s > 0 \) and \( N > 0 \), and using (4.2), we then get (4.1) with

\[ C_N^p = C_{n,m} \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} (1 + 2^k)^n \cdot (1 + 2^\ell)^m \cdot 2^{(1-k)N} \cdot 2^{(1-\ell)N}, \]

which is finite if \( N > 2(n \vee m) \). The proof of the Lemma 4.3 is concluded.

Next we recall the following lemma from \cite{32} which will be used to pass from one approximation of the identity to another.

**Lemma 4.4** \((\mathbb{R}^2 \text{ Lemma 2, } \S1.3)\). Suppose we are given \( \phi \) and \( \psi \in \mathcal{S}(\mathbb{R}^d) \) with \( \int_{\mathbb{R}^d} \phi = 1 \). Then there is a sequence \( \{\eta^{(k)}\} \subset \mathcal{S}(\mathbb{R}^d) \) so that

\[ \psi = \sum_{k=0}^{\infty} \eta^{(k)} \ast \phi_{2^{-k}} \]  

(4.3)

with \( \eta^{(k)} \to 0 \) rapidly, in the sense that whenever \( \| \cdot \|_{\alpha,\beta} \) is a seminorm and \( M \geq 0 \) is fixed, then

\[ \| \eta^{(k)} \|_{\alpha,\beta} = O(2^{-kM}) \quad \text{as } k \to \infty. \]
From Lemma 4.4 we obtain the following estimate

$$\|M_\Phi(f)\|_1 \leq C\|M^+_\Phi(f)\|_1. \quad (4.4)$$

Indeed, for any \( \phi = \phi^{(1)} *_{\mathbb{R}^m} \phi^{(2)} \in \mathcal{S}_F(\mathbb{R}^n \times \mathbb{R}^m) \), by (4.3) on \( \phi^{(1)} \) and \( \phi^{(2)} \) we have

$$M_\phi(f)(x,y) \leq \sup_{t,s>0} \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} |f * (\phi^{(1)}_{2^{-k}t} *_{\mathbb{R}^m} \phi^{(2)}_{2^{-\ell}s}) * (\eta^{(1),k}_{t} *_{\mathbb{R}^m} \eta^{(2),\ell}_{s})|$$

$$\leq M_N^*(f)(x,y) \sup_{t,s>0} \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \int_{\mathbb{R}^n \times \mathbb{R}^m} \left(1 + \frac{|u|}{2^{-k}t}\right)^N \left(1 + \frac{|v|}{2^{-\ell}(t+s)}\right)^N \times |\eta^{(1),k}_{\ell} *_{\mathbb{R}^m} \eta^{(2),\ell}_{s}(u,v)|dudv$$

$$\leq CM_N^*(f)(x,y),$$

where the last inequality holds if \( \phi \) belongs to an appropriate chosen \( \mathcal{F} \). Thus

$$M_\Phi(f)(x,y) = \sup_{\phi \in \mathcal{F}} M^+_\phi(f)(x,y) \leq CM_N^*(f)(x,y)$$

for all \( x \in \mathbb{R}^n, y \in \mathbb{R}^m \); taking \( N > 2(n \vee m) \) as in (4.1) yields (4.4).

Next, we will show that

$$\|M_\Phi^*(f)\|_1 \leq C\|M_\Phi^+(f)\|_1. \quad (4.5)$$

Let \( \mathcal{F} \) be the same as in (4.4) and for any fixed \( \lambda > 0 \), let

$$F = F_\lambda = \{(x,y) : M_\Phi(f)(x,y) \leq \lambda M_\Phi^*(f)(x,y)\}.$$

We prove (4.5) by showing that, for any \( q > 0 \),

$$M_\Phi^*(f)(x,y) \leq C[M_\Phi(M_\Phi^*(f))^q]^{\frac{1}{q}} \quad \text{for } (x,y) \in F,$$

where \( M_\Phi \) is the strong maximal function. Now for any \( (x,y) \), there exists \( (x_1,y_1,t,s) \) with \( |x - x_1| < t, |y - y_1| < t + s \) and \( f * \phi_{t,s}(x_1,y_1) \geq \frac{1}{2} M_\Phi^*(f)(x,y) \). Choose \( r_1 \) small and consider the ball centered at \( x_1 \) of radius \( r_1 t \), i.e. the points \( u \) so that \( |x_1 - u| < r_1 t \). We have that

$$|f * \phi_{t,s}(x_1,y_1) - f * \phi_{t,s}(u,y_1)| \leq r_1 t \sup_{|u-x_1|<r_1t} |\nabla_u f * \phi_{t,s}(u,y_1)|.$$

Similarly, choose \( r_2 \) small and consider the ball centered at \( y_1 \) of radius \( r_2(t+s) \), i.e. the points \( v \) so that \( |y_1 - v| < r_2(t+s) \). We have that

$$|f * \phi_{t,s}(x_1,y_1) - f * \phi_{t,s}(x_1,v)| \leq r_2(t+s) \sup_{|v-y_1|<r_2(t+s)} |\nabla_v f * \phi_{t,s}(x_1,v)|.$$

Combining the above two case, we have

$$|f * \phi_{t,s}(x_1,y_1) - f * \phi_{t,s}(u,y_1) - f * \phi_{t,s}(x_1,v) + f * \phi_{t,s}(u,v)|$$
Note that the set of functions of the form \( \phi \) and that
\[
\phi \partial_i f \ast \phi \partial_j f = \phi \ast \phi_i \ast \phi_j f(u, v),
\]
where
\[
\phi_i f(u, v) = \int_{\mathbb{R}^m} \frac{\partial \phi_i}{\partial u_i} f(u, v - w) \, dw = \frac{1}{t} \int_{\mathbb{R}^m} \left( \frac{\partial \phi_i}{\partial u_i} \right)_t(u, v - w) \phi_s(w) \, dw.
\]
And \( \frac{\partial}{\partial v_j} f \ast \phi \partial_i f = f \ast \phi_i \ast \phi_j f(u, v) \), where
\[
\phi_i f(u, v) = \frac{1}{t} \int_{\mathbb{R}^m} \frac{\partial \phi_i}{\partial u_i} f(u, v - w) \, dw \quad \text{if } t > s;
\]
\[
\phi_i f(u, v) = \frac{1}{s} \int_{\mathbb{R}^m} \phi_i(u, w) \frac{\partial \phi_i}{\partial u_i} f(u, v - w) \, dw \quad \text{if } t < s.
\]

Note that the set of functions of the form \( \tilde{\phi}^i(x + h_1, y + h_2) \) and \( \tilde{\phi}^j(x + h_1, y + h_2), |h_1| \leq 1 + r_1, |h_2| \leq 1 + r_2 \), \( i = 1, \ldots, n, j = 1, \ldots, m \), is a compact set in \( \mathcal{S}_f(\mathbb{R}^n \times \mathbb{R}^m) \), hence we have \( c\tilde{\phi}^i(x + h_1, y + h_2) \) and \( c\tilde{\phi}^j(x + h_1, y + h_2) \) in \( \mathcal{F} \), where \( c \) is a constant independent of \( \phi, h_1 \) and \( h_2 \). Thus \( |f \ast \phi, s(x, y) - f \ast \phi, s(u, v)| \leq cr_1 M_\mathcal{F}(f)(x, y) \leq cr_1 \lambda M_\phi^*(f)(x, y) \), if \( (x, y) \in F \).

By considering the case \( t > s \), if \( (x, y) \in F \) then we can obtain that
\[
|f \ast \phi, t(s(x, y) - f \ast \phi, t(s(u, v))| \leq cr_2 \lambda M_\phi^*(f)(x, y),
\]
and that
\[
|f \ast \phi, t(s(x, y) - f \ast \phi, t(s(u, v)) + f \ast \phi, t(s(u, v)| \leq Cr_1 \cdot r_2 \lambda M_\phi^*(f)(x, y).
\]

So if we take \( r_1 \) and \( r_2 \) so small that \( cr_1 \lambda, cr_2 \lambda, cr_1 r_2 \lambda < 1/16 \), then we have
\[
|f \ast \phi, t(s(u, v)| > \frac{1}{4} M_\phi^*(f)(x, y) \quad \text{for all } u \in B(x_1, r_1 t) \text{ and } v \in B(y_1, r_2 t).
\]

Thus we get that
\[
\frac{1}{4^q} |M_\phi^*(f)(x, y)|^q \leq \frac{1}{|B(x_1, r_1 t) \times B(y_1, r_2 t)|} \int_{B(x_1, (1 + r_1) t) \times B(y_1, (1 + r_2) t)} |f \ast \phi, t(s(u, v)|^q du dv
\]
\[
\leq \left( \frac{1 + r_1}{r_1} \right)^n \left( \frac{1 + r_2}{r_2} \right)^m M_s[(M_\phi^+(f))^q](x, y),
\]
which is (4.5). Similarly, we can obtain this result when considering the case \( t \leq s \).

Then using the maximal theorem (for \( M_s \)) with \( q < 1 \) leads to
\[
\int_F M_\phi^*(f)(x, y) dx dy \leq C \int_{\mathbb{R}^n \times \mathbb{R}^m} (M_s[(M_\phi^+(f))^q](x, y))^{1/2} dx dy \leq C \int_{\mathbb{R}^n \times \mathbb{R}^m} M_\phi(f)(x, y) dx dy. \quad (4.6)
\]

Now we claim that
\[
\int_{\mathbb{R}^n \times \mathbb{R}^m} M_\phi^*(f)(x, y) dx dy \leq 2 \int_F M_\phi^*(f)(x, y) dx dy. \quad (4.7)
\]
To see this, observe that
\[
\int_{F^c} M_\phi^*(f)(x, y) dxdy \leq \lambda^{-1} \int_{F^c} M_\phi(f)(x, y) dxdy \leq c \lambda^{-1} \int_{\mathbb{R}^n \times \mathbb{R}^m} M_\phi^*(f)(x, y) dxdy,
\]
where the last inequality follows from (4.4). Thus, if we take \( \lambda \geq 2c \), we verify the claim (4.7), which, together with (4.6), yields (4.5).

We recall the result that if \( P^{(1)}(x, y) \) is the Poisson kernel on \( \mathbb{R}^{n+m} \), then
\[
P^{(1)}(x, y) = \frac{c_{n+m}}{(1 + |x|^2 + |y|^2)^{(n+m+1)/2}} = \sum_{k=0}^{\infty} 2^{-k} \phi_{2^k}^{(1), (k)}(x, y),
\]
where \( \{\phi_{2^k}^{(1), (k)}\} \) is a bounded collection of functions in \( \mathcal{S}(\mathbb{R}^{n+m}) \). Similarly, if \( P^{(1)}(z) \) is the Poisson kernel on \( \mathbb{R}^{n+m} \), then
\[
P^{(2)}(z) = \frac{c_m}{(1 + |z|^2)^{(m+1)/2}} = \sum_{\ell=0}^{\infty} 2^{-\ell} \phi_{2^{\ell}}^{(2), (\ell)}(x, y),
\]
where \( \{\phi_{2^{\ell}}^{(2), (\ell)}\} \) is a bounded collection of functions in \( \mathcal{S}(\mathbb{R}^m) \). Then for the Poisson kernel \( P_{t,s}(x, y) \), we have that
\[
P_{t,s}(x, y) = P^{(1)}_t *_{\mathbb{R}^m} P^{(2)}_s(x, y) = \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} 2^{-k} 2^{-\ell} \phi_{2^k t}^{(1), (k)} *_{\mathbb{R}^m} \phi_{2^{\ell} s}^{(2), (\ell)}(x, y),
\]
where obviously, \( \{\phi_{2^k t}^{(1), (k)} *_{\mathbb{R}^m} \phi_{2^{\ell} s}^{(2), (\ell)}\} \) is a bounded collection of functions in \( \mathcal{D}_F(\mathbb{R}^n \times \mathbb{R}^m) \). Thus, we have
\[
\|u^*\|_1 \leq \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} 2^{-k} 2^{-\ell} \| M_{\phi_{2^k t, 2^{\ell} s}}^*(f) \|_1 \leq C \| M_\phi(f) \|_1 \leq C \| M_{\Phi}^*(f) \|_1.
\]

We now prove
\[
\| M_{\Phi}^*(f) \|_1 \leq C \|u^*\|_1.
\]
Following [32] Chapter III, § 1.7, for the Poisson kernel \( P^{(1)}_t(x, y) \), there exists a functions \( \eta^{(1)} \) defined on \((1, \infty)\) such that
\[
\int_1^{\infty} \eta^{(1)}(s) ds = 1, \quad \text{and} \quad \int_1^{\infty} s^k \eta^{(1)}(s) ds = 0, \quad k = 1, 2, \ldots.
\]
We now set
\[
\Phi^{(1)}(x, y) := \int_1^{\infty} \eta^{(1)}(t) P^{(1)}_t(x, y) dt.
\]
Similarly, for the Poisson kernel $P_t^{(2)}(z)$, there exists a functions $\eta^{(2)}$ defined on $(1, \infty)$ such that
\[ \int_1^\infty \eta^{(2)}(s) ds = 1, \quad \text{and} \quad \int_1^\infty s^k \eta^{(2)}(s) ds = 0, \quad k = 1, 2, \ldots. \]

We now set
\[ \Phi^{(2)}(z) := \int_1^\infty \eta^{(1)}(s) P_s^{(2)}(z) ds. \]

Then we have $\Phi^{(1)}(x, y) \in \mathcal{S}(\mathbb{R}^{n+m})$ and $\Phi^{(2)}(z) \in \mathcal{S}(\mathbb{R}^m)$. Moreover, we have
\[ \int_{\mathbb{R}^{n+m}} \Phi^{(1)}(x, y) dxdy = \int_1^\infty \eta^{(1)}(t) dt = 1 \]
and
\[ \int_{\mathbb{R}^m} \Phi^{(2)}(z) dz = \int_1^\infty \eta^{(2)}(s) ds = 1. \]

Hence, define
\[ \tilde{\Phi}(x, y) = \Phi^{(1)} \ast_{\mathbb{R}^m} \Phi^{(2)}(x, y), \]
then we obtain that
\[ M_\Phi^x(f)(x, y) \leq u^*(x, y) \int_1^\infty \eta^{(1)}(t) dt \int_1^\infty \eta^{(2)}(s) ds = u^*(x, y). \]

As a consequence, we obtain that for arbitrary $\Phi \in \mathcal{D}_F(\mathbb{R}^n \times \mathbb{R}^m)$,
\[ \|M_\Phi^x(f)\|_1 \leq C\|M_\Phi^x(f)\|_1 \leq \|u^*\|_1. \]

**4.3 The equivalence $\|u^+\|_1 \approx \|M_\Phi^x(f)\|_1$**

It is clear that $u^+(x) \leq u^*(x)$ for $x \in \mathbb{R}^n$. By Section 4.2, $\|u^*\|_1 \lesssim \|M_\Phi^x(f)\|_1$ and (4.5), we have
\[ \|u^+\|_1 \lesssim \|M_\Phi^x(f)\|_1. \]

On the other hand, by the estimates $\|u^*\|_1 \lesssim \|u^+\|_1$ and $\|M_\Phi^x(f)\|_1 \lesssim \|u^*\|_1$, we get
\[ \|M_\Phi^x(f)\|_1 \lesssim \|M_\Phi^x(f)\|_1 \lesssim \|u^+\|_1. \]
5 Atomic decompositions of flag Hardy spaces

5.1 Assumptions and Notations

Assume that $L$ is a non-negative self-adjoint second order differential operator on $L^2(\mathbb{R}^n)$, whose heat kernel $h_t(x,y)$ of $e^{-tL}$ satisfies the Gaussian upper bound:

$$|h_t(x,y)| \leq \frac{C}{t^n}e^{|x-y|^2/ct}, \quad t > 0,$$

where $c$ and $C$ are two positive constants independent of $x, y$ and $t$.

Let $E_L(\lambda)$ denote its spectral decomposition. Then, for every bounded Borel function $F : [0, \infty) \to \mathbb{C}$, one defines the bounded operator $F(L) : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ by the formula

$$F(L) = \int_0^\infty F(\lambda) \, dE_L(\lambda).$$

In particular, the operator $\cos(t\sqrt{L})$ is then well-defined and bounded on $L^2(\mathbb{R}^n)$. Moreover, it follows from [4, Theorem 3] that if the corresponding heat kernels $p_t(x,y)$ of $e^{-tL}$ satisfy Gaussian bounds (GE), then there exists a finite positive constant $c_0$ such that the Schwartz kernel $K_{\cos(t\sqrt{L})}$ of $\cos(t\sqrt{L})$ satisfies

$$\text{supp} K_{\cos(t\sqrt{L})} \subset \{(x,y) \in \Omega \times \Omega : |x-y| \leq c_0t\}.$$  

(5.2)

See also [31]. By the Fourier inversion formula, whenever $F$ is an even, bounded, Borel function with its Fourier transform $\hat{F} \in L^1(\mathbb{R})$, we can write $F(\sqrt{L})$ in terms of $\cos(t\sqrt{L})$. More specifically, we have

$$F(\sqrt{L}) = (2\pi)^{-1} \int_{-\infty}^{\infty} \hat{F}(t) \cos(t\sqrt{L}) \, dt,$$

which, combined with (5.2), gives

$$K_{F(\sqrt{L})}(x,y) = (2\pi)^{-1} \int_{|t| \geq c_0^{-1}|x-y|} \hat{F}(t) K_{\cos(t\sqrt{L})}(x,y) \, dt, \quad \forall x,y \in \Omega. \quad (5.3)$$

The following result (see [21, Lemma 3.5]) is useful for certain estimates later.

**Lemma 5.1.** Let $\varphi \in C^\infty_0(\mathbb{R})$ be even and satisfy $\text{supp} \varphi \subset (-c_0^{-1}, c_0^{-1})$, where $c_0$ is the constant in (5.2). Let $\Phi$ denote the Fourier transform of $\varphi$. Then for every $\kappa = 0, 1, 2, \ldots$, and for every $t > 0$, the kernel $K_{(t^2L)\kappa \Phi(t\sqrt{L})}(x,y)$ of the operator $(t^2L)^\kappa \Phi(t\sqrt{L})$, defined by spectral theory, satisfies

$$\text{supp} K_{(t^2L)\kappa \Phi(t\sqrt{L})}(x,y) \subset \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n : |x-y| \leq t\}.$$

For $s > 0$, we define

$$\mathcal{F}(s) = \left\{ \psi : \mathbb{C} \to \mathbb{C} \text{ measurable} : |\psi(z)| \leq C \frac{|z|^s}{(1 + |z|^{2s})} \right\}.$$
Then for any non-zero function $\psi \in \mathcal{F}(s)$, we have $\int_0^\infty |\psi(t)|^2 \frac{dt}{t} < \infty$. Denote by $\psi_t(z) = \psi(tz)$.

It follows from the spectral theory in [34] that, for any $f \in L^2(\mathbb{R}^n)$,

$$
\left\{ \int_0^\infty \| \psi(t\sqrt{L}) f \|_{L^2(\mathbb{R}^n)}^2 \frac{dt}{t} \right\}^{1/2}
= \left\{ \int_0^\infty \left( \frac{1}{\psi(t\sqrt{L})} \psi(t\sqrt{L}) f, f \right)_{L^2(\mathbb{R}^n)} \frac{dt}{t} \right\}^{1/2}
= \left\{ \left( \int_0^\infty |\psi(t\sqrt{L}) dt / t, f, f \right)_{L^2(\mathbb{R}^n)} \right\}^{1/2}
\leq \kappa \| f \|_{L^2(\mathbb{R}^n)},
$$

where $\kappa = C_L \{ \int_0^\infty |\psi(t)|^2 dt / t \}^{1/2}$.

### 5.2 Atomic decomposition for $H^1_{F}(\mathbb{R}^n \times \mathbb{R}^m)$.

**Definition 5.2.** Let $\triangle^{(1)}$ be the Laplacian on $\mathbb{R}^{n+m}$ and $\triangle^{(2)}$ be the Laplacian on $\mathbb{R}^m$. For $f \in L^1(\mathbb{R}^{n+m})$, the Lusin area integral of $f$ associated with these Laplacians is defined by

$$
S_{f,\triangle^{(1)},\triangle^{(2)}}(x_1, x_2) = \left( \int_{\mathbb{R}^{n+m+1}}^{\mathbb{R}^{m+1}} \chi_{t_1, t_2}(x_1 - y_1, x_2 - y_2)
\times \left| \left( t_1^2 \triangle^{(1)} e^{-t_1^2 \triangle^{(1)}} \otimes t_2^2 \triangle^{(2)} e^{-t_2^2 \triangle^{(2)}} \right) f(y_1, y_2) \right|^2 \frac{dy_1 dt_1 dy_2 dt_2}{t_1^2 t_2^2} \right)^{1/2},
$$

(5.4)

where $\chi_{t_1, t_2}(x_1, x_2) = \chi^{(1)}(t_1, t_2, x_1, x_2)$, $\chi^{(1)}(x_1, x_2) = t_1^{-(n+m)} \chi^{(1)}(\frac{x_1}{t_1}, \frac{x_2}{t_1})$, $\chi^{(2)}(z) = t_2^{-m} \chi^{(2)}(\frac{z}{t_2})$, $\chi^{(1)}(x_1, x_2)$ and $\chi^{(2)}(z)$ are the indicator function of the unit balls of $\mathbb{R}^{n+m}$ and $\mathbb{R}^m$ respectively.

Based on the discrete reproducing formula as in Theorem 2.5 and the Plancherel–Pólya type inequalities as in Theorem 2.13, we can obtain the estimate

$$
\|S_{f,\triangle^{(1)},\triangle^{(2)}}(f)\|_1 \lesssim \|S_f(f)\|_1.
$$

(5.5)

Since this argument is similar to the estimates as in Section 2.3, we omit it here.

We now define the flag Hardy space $H^1_{F,\triangle^{(1)},\triangle^{(2)}}(\mathbb{R}^n \times \mathbb{R}^m)$ associated with $\triangle^{(1)}$ and $\triangle^{(2)}$ as follows:

$$
H^1_{F,\triangle^{(1)},\triangle^{(2)}}(\mathbb{R}^n \times \mathbb{R}^m) := \{ f \in L^1(\mathbb{R}^{n+m}) : \| S_{f,\triangle^{(1)},\triangle^{(2)}} f \|_{L^1(\mathbb{R}^{n+m})} < \infty \}
$$

with the norm

$$
\| f \|_{H^1_{F,\triangle^{(1)},\triangle^{(2)}}(\mathbb{R}^{n+m})} := \| S_{f,\triangle^{(1)},\triangle^{(2)}} f \|_{L^1(\mathbb{R}^{n+m})}.
$$

We are now recalling the atomic Hardy space $H^1_{F,at,M}(\mathbb{R}^n \times \mathbb{R}^m)$ as in Definition 1.10 and we will later prove that $H^1_{F,at,M}(\mathbb{R}^n \times \mathbb{R}^m)$ is equivalent to $H^1_{F,\triangle^{(1)},\triangle^{(2)}}(\mathbb{R}^n \times \mathbb{R}^m)$ above i.e., Theorem 5.3 below. Then eventually we show that they are both equivalent to the space $H^1_{L}(\mathbb{R}^n \times \mathbb{R}^m)$ via square functions, i.e., Theorem 1.11.

For the convenience of the readers, we repeat the definition of $H^1_{F,at,M}(\mathbb{R}^n \times \mathbb{R}^m)$ here.
Definition 5.3. Let $M > \max\{n,m\}/4$. The Hardy spaces $H^1_{F,\text{at},M}(\mathbb{R}^n \times \mathbb{R}^m)$ is defined as follows. For $f \in L^2(\mathbb{R}^n \times \mathbb{R}^m)$, we say that $f = \sum_j \lambda_j a_j$ is an atomic $(1,2,M)$-representation of $f$ if $\{\lambda_j\}_{j=0}^\infty \in \ell^1$, each $a_j$ is a $(1,2,M)$-atom, and the sum converges in $L^2(\mathbb{R}^n \times \mathbb{R}^m)$. The space $H^1_{F,\text{at},M}(\mathbb{R}^n \times \mathbb{R}^m)$ is defined to be

$$H^1_{F,\text{at},M}(\mathbb{R}^n \times \mathbb{R}^m) = \{f \in L^2(\mathbb{R}^n \times \mathbb{R}^m) : f \text{ has an atomic $(1,2,M)$-representation}\}$$

with the norm

$$\|f\|_{H^1_{F,\text{at},M}(\mathbb{R}^n \times \mathbb{R}^m)} := \inf \left\{ \sum_{j=0}^\infty |\lambda_j| : f = \sum_{j=0}^\infty \lambda_j a_j \text{ is an atomic $(1,2,M)$-representation} \right\}.$$

The atomic Hardy space $H^1_{F,\text{at},M}(\mathbb{R}^n \times \mathbb{R}^m)$ is defined as the completion of $H^1_{F,\text{at},M}(\mathbb{R}^n \times \mathbb{R}^m)$ with respect to this norm.

Theorem 5.4. Suppose that $M > \max\{n,m\}/4$. Then

$$H^1_{F,\Delta^{(1)},\Delta^{(2)}}(\mathbb{R}^n \times \mathbb{R}^m) = H^1_{F,\text{at},M}(\mathbb{R}^n \times \mathbb{R}^m).$$

Moreover,

$$\|f\|_{H^1_{F,\Delta^{(1)},\Delta^{(2)}}(\mathbb{R}^n \times \mathbb{R}^m)} \approx \|f\|_{H^1_{F,\text{at},M}(\mathbb{R}^n \times \mathbb{R}^m)},$$

where the implicit constants depend only on $M,n$ and $m$.

5.3 Proof of the atomic decomposition

We now proceed to the proof of Theorem 5.4. The basic strategy is as follows: by density, it is enough to show that

$$H^1_{F,\text{at},M}(\mathbb{R}^n \times \mathbb{R}^m) = H^1_{F,\Delta^{(1)},\Delta^{(2)}}(\mathbb{R}^n \times \mathbb{R}^m) \cap L^2(\mathbb{R}^n \times \mathbb{R}^m) \quad \text{for } M > \max\{n,m\},$$

with equivalent of norms. The proof of this proceeds in two steps.

Step 1. $H^1_{F,\text{at},M}(\mathbb{R}^n \times \mathbb{R}^m) \subset H^1_{F,\Delta^{(1)},\Delta^{(2)}}(\mathbb{R}^n \times \mathbb{R}^m) \cap L^2(\mathbb{R}^n \times \mathbb{R}^m)$ for $M > \max\{n,m\}/4$.

Step 2. $H^1_{F,\Delta^{(1)},\Delta^{(2)}}(\mathbb{R}^n \times \mathbb{R}^m) \cap L^2(\mathbb{R}^n \times \mathbb{R}^m) \subset H^1_{F,\text{at},M}(\mathbb{R}^n \times \mathbb{R}^m)$ for every $M \in \mathbb{N}$.

The conclusion of Step 1 is an immediate consequence of the following two lemmas.

Lemma 5.5. Fix $M \in \mathbb{N}$. Assume that $T$ is a linear operator or a nonnegative sublinear operator, satisfying the weak-type $(2,2)$ bound

$$|\{x \in \mathbb{R}^n \times \mathbb{R}^m : |Tf(x)| > \eta\}| \leq C_T \eta^{-2} \|f\|^2_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}, \quad \forall \eta > 0.$$

If there is an absolute constant $C > 0$ such that

$$\|Ta\|_{L^1(\mathbb{R}^n \times \mathbb{R}^m)} \leq C \quad \text{for every $(1,2,M)$-atom } a,$$

(5.6)
then $T$ is bounded from $H^1_{F,at,M}(\mathbb{R}^n \times \mathbb{R}^m)$ to $L^1(\mathbb{R}^n \times \mathbb{R}^m)$ and
\[
\|Tf\|_{L^1(\mathbb{R}^n \times \mathbb{R}^m)} \leq C \|f\|_{H^1_{F,at,M}(\mathbb{R}^n \times \mathbb{R}^m)}.
\]
Consequently, by density, $T$ extends to a bounded operator from $H^1_{F,at,M}(\mathbb{R}^n \times \mathbb{R}^m)$ to $L^1(\mathbb{R}^n \times \mathbb{R}^m)$ and $L^1(\mathbb{R}^n \times \mathbb{R}^m)$.

**Proof.** Given $f \in H^1_{F,at,M}(\mathbb{R}^n \times \mathbb{R}^m)$. Then $f = \sum_j \lambda_j a_j$ is an atomic $(1,2,M)$-representation such that
\[
\|f\|_{H^1_{F,at,M}(\mathbb{R}^n \times \mathbb{R}^m)} \approx \sum_{j=0}^{\infty} |\lambda_j|.
\]
Since the sum converges in $L^2$ (by definition), and since $T$ is of weak-type $(2,2)$, we have that at almost every point,
\[
|T(f)| \leq \sum_{j=0}^{\infty} |\lambda_j| |T(a_j)|.
\] (5.7)
Indeed, for every $\eta > 0$, we have that, if $f^N := \sum_{j>N} \lambda_j a_j$, then,
\[
\left| \left\{ x : |Tf(x)| - \sum_{j=0}^{\infty} |\lambda_j| |T(a_j)| > \eta \right\} \right| \leq \limsup_{N \to \infty} \left| \left\{ x : |Tf^N(x)| > \eta \right\} \right|
\leq C_T \eta^{-2} \limsup_{N \to \infty} \|f^N\|_2^2 = 0,
\]
from which (5.7) follows. In turn, (5.7) and (5.6) imply the desired $L^1$ bound for $Tf$. \[ \Box \]

**Lemma 5.6.** Let $S_{F,\Delta(1),\Delta(2)}$ be the square function defined by (5.4) and $M > \max\{n,m\}/4$. Then
\[
\|S_{F,\Delta(1),\Delta(2)} a\|_{L^1(\mathbb{R}^n \times \mathbb{R}^m)} \leq C \quad \text{for every (1,2,$M$)-atom } a,
\]
where $C$ is a positive constant independent of $a$.

By Lemma 5.6 we may apply Lemma 5.5 with $T = S_{F,\Delta(1),\Delta(2)}$ to obtain
\[
\|f\|_{H^1_{F,\Delta(1),\Delta(2)}(\mathbb{R}^n \times \mathbb{R}^m)} = \|S_{F,\Delta(1),\Delta(2)} f\|_{L^1(\mathbb{R}^n \times \mathbb{R}^m)} \leq C \|f\|_{H^1_{F,at,M}(\mathbb{R}^n \times \mathbb{R}^m)}
\]
and Step 1 follows.

Suppose $\Omega \subset \mathbb{R}^n \times \mathbb{R}^m$ is open of finite measure. Denote by $m(\Omega)$ the maximal dyadic subrectangles of $\Omega$. Let $m_1(\Omega)$ denote those dyadic subrectangles $R \subseteq \Omega$, $R = I \times J$ that are maximal in the $x_1$ direction. In other words if $S = I' \times J \supseteq R$ is a dyadic subrectangle of $\Omega$, then $I = I'$. Define $m_2(\Omega)$ similarly. Let
\[
\tilde{\Omega} = \{ x \in \mathbb{R}^n \times \mathbb{R}^m : M_s(\chi_{\Omega})(x) > \frac{1}{2} \},
\]
where \( M_s \) is the strong maximal operator defined as

\[
M_s(f)(x) = \sup_{R: \text{rectangles in } \mathbb{R}^n \times \mathbb{R}^m} \frac{1}{|R|} \int_R |f(y)|dy.
\]

For any \( R = I \times J \in m_1(\Omega) \), we set \( \gamma_1(R) = \gamma_1(R, \Omega) = \sup \frac{|J|}{|R|} \), where the supremum is taken over all dyadic intervals \( l : I \subset l \) so that \( l \times J \subset \tilde{\Omega} \). Define \( \gamma_2 \) similarly. Then Journé’s lemma, (in one of its forms) says, for any \( \delta > 0 \),

\[
\sum_{R \in m_2(\Omega)} |R| \gamma_1^{-\delta}(R) \leq c_\delta |\Omega| \quad \text{and} \quad \sum_{R \in m_1(\Omega)} |R| \gamma_2^{-\delta}(R) \leq c_\delta |\Omega|
\]

for some \( c_\delta \) depending only on \( \delta \), not on \( \Omega \).

**Proof of Lemma 5.6.** Given any \((1, 2, M)\)-atom \( a \), suppose that \( a = \sum_{R \in m(\Omega)} a_R \) is supported in an open set \( \Omega \) with finite measure. For any \( R = I \times J \in m(\Omega) \), let \( \tilde{I} \) be the biggest dyadic cube containing \( I \), so that \( \tilde{I} \times J \subset \tilde{\Omega} \), where \( \tilde{\Omega} = \{ x \in \mathbb{R}^n \times \mathbb{R}^m : M_s(\chi_\Omega)(x) > 1/2 \} \). Next, let \( \tilde{J} \) be the biggest dyadic cube containing \( J \), so that \( \tilde{I} \times \tilde{J} \subset \tilde{\Omega} \), where \( \tilde{\Omega} = \{ x \in \mathbb{R}^n \times \mathbb{R}^m : M_s(\chi_{\tilde{\Omega}})(x) > 1/2 \} \). Now let \( \tilde{R} \) be the 100-fold dilate of \( \tilde{I} \times \tilde{J} \) concentric with \( \tilde{I} \times \tilde{J} \). Clearly, an application of the strong maximal function theorem shows that \( \left| \cup_{R \subset \Omega} \tilde{R} \right| \leq C|\tilde{\Omega}| \leq C|\tilde{\Omega}| \leq C|\Omega| \). From property (iii) of the \((1, 2, M)\)-atom,

\[
\int_{\cup \tilde{R}} |S_{F, \Delta(1), \Delta(2)}(a)(x_1, x_2)|dx_1dx_2 \leq |\cup \tilde{R}|^{1/2} \|S_{F, \Delta(1), \Delta(2)}(a)\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)} \\
\leq C|\Omega|^{1/2} \|a\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)} \\
\leq C|\Omega|^{1/2} |\Omega|^{-1/2} \leq C.
\]

We now prove

\[
\int_{(\cup \tilde{R})^c} |S_{F, \Delta(1), \Delta(2)}(a)(x_1, x_2)|dx_1dx_2 \leq C.
\]

From the definition of \( a \), we write

\[
\int_{(\cup \tilde{R})^c} |S_{F, \Delta(1), \Delta(2)}(a)(x_1, x_2)|dx_1dx_2 \leq \sum_{R \in m(\Omega)} \int_{R^c} |S_{F, \Delta(1), \Delta(2)}(a_R)(x_1, x_2)|dx_1dx_2 \\
\leq \sum_{R \in m(\Omega)} \int_{(100\tilde{I})^c \times \mathbb{R}^m} |S_{F, \Delta(1), \Delta(2)}(a_R)(x_1, x_2)|dx_1dx_2 \\
+ \sum_{R \in m(\Omega)} \int_{\mathbb{R}^n \times (100\tilde{J})^c} |S_{F, \Delta(1), \Delta(2)}(a_R)(x_1, x_2)|dx_1dx_2 \\
= I + II.
\]

(5.10)
For the term $I$, we have
\[
\int_{(100\ell)^c \times \mathbb{R}^m} |S_{F,\triangle(1),\triangle(2)}(a_R)(x_1, x_2)| dx_1 dx_2 = \int_{(100\ell)^c \times J} |S_{F,\triangle(1),\triangle(2)}(a_R)(x_1, x_2)| dx_1 dx_2 \\
+ \int_{(100\ell)^c \times (100J)^c} |S_{F,\triangle(1),\triangle(2)}(a_R)(x_1, x_2)| dx_1 dx_2 \\
= I_1 + I_2.
\]

Let us first estimate the term $I_1$. Set $a_{R,2} = (\mathbb{I}_1 \otimes_2 \triangle(2)^M) b_R$, that is, $a_R = (\triangle(1)^M \otimes_2 \mathbb{I}_2) a_{R,2}$. Using Hölder’s inequality,
\[
I_1 \leq C |J|^{1/2} \int_{(100\ell)^c} \left( \int_{100J} |S_{F,\triangle(1),\triangle(2)}(a_R)(x_1, x_2)|^2 dx_2 \right)^{1/2} dx_1. \tag{5.11}
\]

Let $\chi^{(1)}(x_1, x_2)$ and $\chi^{(2)}(y)$ be the indicator function of the unit balls of $\mathbb{R}^{n+m}$ and $\mathbb{R}^m$, respectively. Set $\chi^{(1)}_t(x_1, x_2) = t^{-(m+m)} \chi^{(1)}(x_1/t, x_2)$ and $\chi^{(2)}_s(y) = s^{-m} \chi^{(2)}(y/s)$. We rewrite $S_{F,\triangle(1),\triangle(2)}$ by
\[
S_{F,\triangle(1),\triangle(2)} f(x_1, x_2) = \left( \int_{\mathbb{R}^{n+1}} \int_{\mathbb{R}^{m+1}} \chi^{(1)}_{t_1} \chi^{(2)}_{t_2} (x_1 - y_1, x_2 - y_2) \\
\times \left| (t_1^2 \Delta(1) e^{-t_1^2 \Delta(2)} \otimes_2 (t_2^2 \Delta(2) e^{-t_2^2 \Delta(2)}) f(y_1, y_2) \right| \frac{dy_1 dt_1}{t_1^{m+1}} \frac{dy_2 dt_2}{t_2^{m+1}} \right)^{1/2}.
\]

Hence, the $L^2$-boundedness of the square function gives
\[
\int_{100J} |S_{F,\triangle(1),\triangle(2)}(a_R)(x_1, x_2)|^2 dx_2 \\
\leq \int_{\mathbb{R}^{n+1}} \int_{\mathbb{R}^{m+1}} \int_{\mathbb{R}^m} \left[ \int_{\mathbb{R}^{m+1}} \chi^{(1)}_{t_1} (x_1 - y_1, x_2 - z_2) \chi^{(2)}_{t_2} (z_2 - y_2) \\
\times \left| (t_1^2 \Delta(2) e^{-t_1^2 \Delta(2)} \otimes_2 (t_2^2 \Delta(1) e^{-t_2^2 \Delta(1)}) a_R(y_1, \cdot) \right| (y_2) \right| \frac{2 dy_2 dt_2}{t_2^{m+1}} \frac{dy_1 dt_1}{t_1^{m+1}} dx_2 \\
\leq \int_{\mathbb{R}^{n+1}} \int_{\mathbb{R}^m} \chi^{(1)}_{t_1} (x_1 - y_1, x_2 - z_2) \left( (t_1^2 \Delta(1) e^{-t_1^2 \Delta(1)} \otimes_2 \mathbb{I}_2) a_R(y_1, x_2) \right) \frac{2 dy_2 dt_2}{t_2^{m+1}} \frac{dy_1 dt_1}{t_1^{m+1}} dx_2 \\
\leq \int_{\mathbb{R}^m} \int_{0}^{\infty} \int_{|y_1 - x_1| \leq t_1} \left( (t_1^2 \Delta(1))^{M+1} e^{-t_1^2 \Delta(1)} \otimes_2 \mathbb{I}_2 \right) a_{R,2}(y_1, x_2) \frac{2 dy_1 dt_1}{t_1^{m+1+4M}} dx_2,
\]

where the last inequality follows from the equality $a_R = (\triangle(1)^M \otimes_2 \mathbb{I}_2) a_{R,2}$. Note that $\text{supp } a_{R,2} \subset \text{supp } (\mathbb{I}_1 \otimes_2 \triangle(2)^M) b_R \subset 10R = 10(I \times J)$.

We then apply the time derivatives of the kernel of $\triangle(1)$ to obtain
\[
\int_{100J} |S_{F,\triangle(1),\triangle(2)}(a_R)(x_1, x_2)|^2 dx_2
\]
\[
\leq C \int_{\mathbb{R}^m} \int_{0}^{\ell(I)} \int_{|x-y|<t_1} \left[ \int_{10I} t_1^{-n} \exp \left( - \frac{|y-z|^2}{ct_1^2} \right) |a_{R,2}(z_1, x_2)| dz_1 \right]^2 \frac{dy_1 dt_1}{t_1^{n+1+4M}} dx_2 \\
+ C \int_{\mathbb{R}^m} \int_{0}^{\ell(I)} \int_{|x-y|<t_1} \left[ \int_{10I} t_1^{-n} \exp \left( - \frac{|y-z|^2}{ct_1^2} \right) |a_{R,2}(z_1, x_2)| dz_1 \right]^2 \frac{dy_1 dt_1}{t_1^{n+1+4M}} dx_2 \\
=: D_1(a_R)(x_1) + D_2(a_R)(x_1).
\]

Let us estimate the term \(D_1(a_R)(x_1)\). Note that if \(x_1 \notin 100I\), \(0 < t_1 < \ell(I)\), \(|x-y| < t_1\) and \(z_1 \in 10I\), then \(|y-z|^2 / |x-y|^2 \leq 1/2\). Hence
\[
D_1(a_R)(x_1) \leq C |I| \int_{0}^{\ell(I)} t_1^{-2n-4M} \int_{|x-y|<t_1} \left[ \int_{10I} t_1^{-n} \exp \left( - \frac{|y-z|^2}{ct_1^2} \right) |a_{R,2}(z_1, x_2)| dz_1 \right]^2 \frac{dt_1}{t_1^{n+1+4M}} dx_2.
\]

We use the fact that \(e^{-s} \leq C s^{-k}\) for any \(k > 0\) to obtain
\[
D_1(a_R)(x_1) \leq C |I| \int_{0}^{\ell(I)} t_1^{-2n-4M-1} \int_{|x-y|<t_1} \left[ \int_{10I} t_1^{-n} \exp \left( - \frac{|y-z|^2}{ct_1^2} \right) |a_{R,2}(z_1, x_2)| dz_1 \right]^2 \frac{dt_1}{t_1^{n+1+4M}} dx_2.
\]

where \((x_I, x_J)\) denotes the center of \(R = I \times J\). In order to estimate the second term \(D_2(a_R)\), observe that if \(x_1 \notin 100I\), \(\ell(I) \leq t_1 < |x-y|/4\), \(|x-y| < t_1\) and \(z_1 \in 10I\), then \(|y-z|^2 / |x-y|^2 \geq 1/2\). Hence,
\[
D_2(a_R)(x_1) \leq C |I| \int_{|x-y|<t_1} \left[ \int_{10I} t_1^{-n} \exp \left( - \frac{|y-z|^2}{ct_1^2} \right) |a_{R,2}(z_1, x_2)| dz_1 \right]^2 \frac{dt_1}{t_1^{n+1+4M}} dx_2.
\]
Combining the estimates of \(D_1(a_R)(x_1)\) and \(D_2(a_R)(x_1)\), we obtain
\[
\int_{10J} |S_{F,\Delta(1),\Delta(2)}(a_R)(x_1, x_2)|^2 \, dx_2 
\lesssim \frac{|I|^{1/n+1}}{|x_1-x_I|^{2n+1}} \ell(I)^{-4M} \ell(J)^{-4M} \|\mathbb{I}_1 \otimes_2 (\ell(J)^2 \Delta(2))^M b_R\|_{L^2(R^n \times R^m)}^2. \tag{5.12}
\]
Putting (5.12) into the term \(I_1\) in (5.11), we have
\[
I_1 \lesssim |R|^{1/2} \int_{100\overline{J}} \frac{|I|^{1/2}}{|x_1-x_I|^{n+1/2}} \, dx_1 \ell(I)^{-2M} \ell(J)^{-2M} \|\mathbb{I}_1 \otimes_2 (\ell(J)^2 \Delta(2))^M b_R\|_{L^2(R^n \times R^m)}^2
\lesssim |R|^{1/2} \gamma_1(R)^{-1/2} \ell(I)^{-2M} \ell(J)^{-2M} \|\mathbb{I}_1 \otimes_2 (\ell(J)^2 \Delta(2))^M b_R\|_{L^2(R^n \times R^m)}^2.
\]
Now we turn to estimate the term \(I_2\). Note that \(a_R = (\Delta(1)^M \otimes_2 \Delta(2)^M) b_R\) and \(\text{supp } b_R \subset 10R = 10(I \times J)\). One can write
\[
(S_{F,\Delta(1),\Delta(2)}(a_R)(x_1, x_2))^2
= \iint_{\Gamma(x_1, x_2)} \left| \left( (t_1^2 \Delta(1))^M e^{-t_2^2 \Delta(2)} \otimes_2 (t_2^2 \Delta(2))^M e^{-t_2^2 \Delta(2)} \right) b_R(y_1, y_2) \right|^2 \frac{dy_1 dy_2}{t_1^{n+M+1} t_2^{n+M+1}} \frac{dy_1 dy_2}{t_1^{n+M+1} t_2^{n+M+1}}
\leq C \left( \int_0^{\ell(I)} \int_0^{\ell(J)} \int_0^{\ell(I)} \int_0^{\ell(J)} \int_0^{\ell(I)} \int_0^{\ell(J)} \int_{|x_1-y_1|<t_1} \int_{|x_2-y_2|<t_1+t_2} \right)
\left[ \int_{10J} t_1^{-n} \exp \left(-\frac{|y_1-z_1|^2}{ct_1^2} \right) t_2^{-m} \exp \left(-\frac{|y_2-z_2|^2}{ct_2^2} \right) \|b_R(z_1, z_2)|dz_1 dz_2 \right]^2 \frac{dy_1 dy_2}{t_1^{n+M+1} t_2^{n+M+1}} \frac{dy_1 dy_2}{t_1^{n+M+1} t_2^{n+M+1}}
= \sum_{i=1}^4 I_{2i}(b_R)(x_1, x_2).
\]
We first estimate the term \(I_{21}(b_R)(x_1, x_2)\). Note that if \(x_1 \notin 100\overline{J}, \, 0 < t_1 < \ell(I), \, |x_1-y_1| < t_1\) and \(z_1 \in 10I, \, \text{then } |y_1-z_1| \geq |x_1-x_1|/2\). If \(x_2 \notin 100\overline{J}, \, 0 < t_2 < \ell(J), \, |x_2-y_2| < t_1 + t_2\) and \(z_2 \in 10J, \, \text{then } |y_2-z_2| \geq |x_2-x_J|/2\) since \(\ell(J) \leq \ell(J)\). Hence
\[
I_{21}(b_R)(x_1, x_2)
\leq C \int_0^{\ell(I)} \int_0^{\ell(J)} \int_{|x_1-y_1|<t_1} \int_{|x_2-y_2|<t_1+t_2} \int_{10I} \int_{10J} \frac{dy_1 dy_2}{t_1^{n+M+1} t_2^{n+M+1}} \frac{dt_1}{t_1^{n+M+1}} \frac{dt_2}{t_2^{n+M+1}} \left[ \int_{10J} \left| b_R(z_1, z_2) \right|dz_1 dz_2 \right]^2
\leq C |R| \int_0^{\ell(I)} \int_0^{\ell(J)} \int_{|x_1-y_1|<t_1} \int_{|x_2-y_2|<t_1+t_2} \int_{10I} \int_{10J} \frac{dt_1}{t_1^{n+M+1}} \frac{dt_2}{t_2^{n+M+1}} \|b_R\|^2_{L^2(R^n \times R^m)}.
\]
We use the fact that \(e^{-s} \leq Cs^{-k}\) for any \(k > 0\) to obtain
\[
I_{21}(b_R)(x_1, x_2)
\]
\[ \leq C|R| \int_0^{\ell(I)} t_1^{2n-4M-1} \left( \frac{t_1}{|x_1 - x_I|} \right)^{2(n+2M+\frac{1}{2})} dt_1 \\
\times \int_0^{\ell(J)} t_2^{-2m-4M-1} \left( \frac{t_2}{|x_2 - x_J|} \right)^{2(m+2M+\frac{1}{2})} dt_2 \|b_R\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}^2 \\
\leq C|R| \int_0^{\ell(I)} t_1^{2n+\frac{1}{2}M} \left( \frac{t_1}{|x_1 - x_I|} \right)^{2(n+2M+\frac{1}{2})} dt_1 \ell(I) \|b_R\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}^2 \\
\leq C|R| \int_0^{\ell(J)} t_2^{-2m+\frac{1}{2}M} \left( \frac{t_2}{|x_2 - x_J|} \right)^{2(m+2M+\frac{1}{2})} dt_2 \ell(J) \|b_R\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}^2. \\
\leq C|R| \frac{|I|^{1/n}}{|x_1 - x_I|^{2n+1} |x_2 - x_J|^{2m+1}} \|b_R\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}^2. \]

We then estimate the term \( I_{22}(b_R)(x_1, x_2) \). Note that if \( x_1 \notin 10I \), \( 0 < t_1 < \ell(I) \), \( |x_1 - y_1| < t_1 \) and \( z_1 \in 10I \), then \( |y_1 - z_1| \geq |x_1 - x_I|/2 \). If \( x_2 \notin 10J \), \( \ell(J) \leq t_2 < |x_2 - x_J|/4 \), \( |x_2 - y_2| < t_1 + t_2 \) and \( z_2 \in 10J \), then \( |y_2 - z_2| \geq |x_2 - x_J|/4 \). Hence,

\[ I_{22}(b_R)(x_1, x_2) \]

\[ \leq C \int_0^{\ell(I)} \int_{\ell(J)} \int_{|x_1 - y_1| < t_1} \int_{|x_2 - y_2| < t_1 + t_2} dy_1 dy_2 t_1^{-2n} \exp \left( - \frac{2|y_1 - z_1|^2}{ct_1^2} \right) \\
\times t_2^{-2m} \exp \left( - \frac{2|y_2 - z_2|^2}{ct_2^2} \right) \frac{dt_1}{t_1^{n+m+4M+1}} \frac{dt_2}{t_2^{m+4M+1}} \left[ \int_{10I} \int_{10J} |b_R(z_1, z_2)| dz_1 dz_2 \right]^2 \\
\leq C|R| \int_0^{\ell(I)} \left( \int_{\ell(J)}^{\infty} + \int_{\ell(J)}^{10J} \right) t_1^{-2n} \exp \left( - \frac{2|y_1 - z_1|^2}{ct_1^2} \right) \\
\times t_2^{-2m} \exp \left( - \frac{2|y_2 - z_2|^2}{ct_2^2} \right) \frac{dt_1}{t_1^{n+m+4M+1}} \frac{dt_2}{t_2^{m+4M+1}} \|b_R\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}^2. \]

We use the fact that \( e^{-s} \leq Cs^{-k} \) for any \( k > 0 \) to obtain

\[ I_{22}(b_R)(x_1, x_2) \]

\[ \leq C|R| \int_0^{\ell(I)} t_1^{2n-4M-1} \left( \frac{t_1}{|x_1 - x_I|} \right)^{2(n+2M+\frac{1}{2})} dt_1 \\
\times \left( \int_{\ell(J)}^{\infty} t_2^{-2m-4M-1} \left( \frac{t_2}{|x_2 - x_J|} \right)^{2(m+2M+\frac{1}{2})} dt_2 + \int_{\ell(J)}^{10J} t_2^{-2m-4M} dt_2 \right) \|b_R\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}^2 \\
\leq C|R| \frac{|I|^{1/n}}{|x_1 - x_I|^{2n+1} |x_2 - x_J|^{2m+1}} \ell(I)^{-4M} \ell(J)^{-4M} \|b_R\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}^2. \]

We now estimate the term \( I_{23}(b_R)(x_1, x_2) \). Note that if \( x_1 \notin 10I \), \( \ell(I) \leq t_1 < |x_1 - x_I|/4 \), \( |x_1 - y_1| < t_1 \) and \( z_1 \in 10I \), then \( |y_1 - z_1| \geq |x_1 - x_I|/4 \). If \( x_2 \notin 10J \), \( 0 < t_2 < \ell(J) \), \( |x_2 - y_2| < t_1 + t_2 \) and \( z_2 \in 10J \), then \( |y_2 - z_2| \geq |x_2 - x_J|/2 \) since \( \ell(I) \leq \ell(J) \). Hence

\[ I_{23}(b_R)(x_1, x_2) \]

\[ \leq C \int_0^{\ell(I)} \int_{\ell(J)} \int_{|x_1 - y_1| < t_1} \int_{|x_2 - y_2| < t_1 + t_2} dy_1 dy_2 t_1^{-2n} \exp \left( - \frac{2|y_1 - z_1|^2}{ct_1^2} \right) \\
\times t_2^{-2m} \exp \left( - \frac{2|y_2 - z_2|^2}{ct_2^2} \right) \frac{dt_1}{t_1^{n+m+4M+1}} \frac{dt_2}{t_2^{m+4M+1}} \left[ \int_{10I} \int_{10J} |b_R(z_1, z_2)| dz_1 dz_2 \right]^2 \\
\leq C|R| \frac{|I|^{1/n}}{|x_1 - x_I|^{2n+1} |x_2 - x_J|^{2m+1}} \ell(I)^{-4M} \ell(J)^{-4M} \|b_R\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}^2. \]
\[
\begin{align*}
&\leq C|R| \left( \int_{\ell(I)}^{x_1 - x_I} + \int_{x_1 - x_I}^{x_2 - x_I} \right) \int_0^{t_1} t_1^{-2n} \exp \left( -\frac{2|y_1 - z_1|^2}{ct_1^2} \right) \\
&\quad \times t_2^{-2m} \exp \left( -\frac{2|y_2 - z_2|^2}{ct_2^2} \right) \frac{dt_1}{t_1^{4M+1}} \frac{dt_2}{t_2^{4M+1}} \|b_R\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}^2.
\end{align*}
\]

Since \( e^{-s} \leq C s^{-k} \) for any \( k > 0 \), we have

\[
I_{23}(b_R)(x_1, x_2) \leq C|R| \left( \int_{\ell(I)}^{x_1 - x_I} + \int_{x_1 - x_I}^{x_2 - x_I} \right) \int_0^{t_1} t_1^{-2n-1-4M} \left( \frac{t_1}{|x_1 - x_I|} \right)^{2(n+2M-\frac{1}{2})} dt_1 + \int_0^{x_1 - x_I} t_1^{-2n-1-4M} dt_1
\]

\[
\times \int_0^{t_2} t_2^{-2m-4M-1} \left( \frac{t_2}{|x_2 - x_J|} \right)^{2(m+2M+\frac{1}{2})} dt_2 \|b_R\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}^2
\]

\[
\leq C|R| \frac{|I|^{1/n}}{|x_1 - x_I|^{2n+1}} \frac{|J|^{1/m}}{|x_2 - x_J|^{2m+1}} \ell(I)^{-4M} \ell(J)^{-4M} \|b_R\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}^2.
\]

Finally, we estimate the term \( I_{24}(b_R)(x_1, x_2) \). If \( x_1 \notin 100\tilde{I}, \ell(I) \leq t_1 < |x_1 - x_I|/4, |x_1 - y_I| < t_1 \) and \( z_1 \in 10I \), then \( |y_1 - z_1| \geq |x_1 - x_I|/4. \) If \( x_2 \notin 100\tilde{J}, \ell(J) \leq t_2 < |x_2 - x_J|/4, |x_2 - y_2| < t_1 + t_2 \) and \( z_2 \in 10J \), then \( |y_2 - z_2| \geq |x_2 - x_J|/4. \) Hence

\[
I_{24}(b_R)(x_1, x_2) \leq C \int_{\ell(I)}^{x_1 - x_I} \int_{\ell(J)}^{x_2 - x_J} \int_{t_1}^{x_1 - y_I} \int_{t_1}^{x_2 - y_2} dy_1 dy_2 t_1^{-2n} \exp \left( -\frac{2|y_1 - z_1|^2}{ct_1^2} \right)
\]

\[
\times t_2^{-2m} \exp \left( -\frac{2|y_2 - z_2|^2}{ct_2^2} \right) \frac{dt_1}{t_1^{n+m+4M+1}} \frac{dt_2}{t_2^{n+m+4M+1}} \left[ \int_{10I}^{10J} \int |b_R(z_1, z_2)|dz_1 dz_2 \right]^2
\]

\[
\leq C|R| \left( \int_{\ell(I)}^{x_1 - x_I} + \int_{x_1 - x_I}^{x_2 - x_I} \right) \left( \int_{\ell(J)}^{x_2 - x_J} + \int_{x_2 - x_J}^{x_2 - x_I} \right) t_1^{-2n} \exp \left( -\frac{2|y_1 - z_1|^2}{ct_1^2} \right)
\]

\[
\times t_2^{-2m} \exp \left( -\frac{2|y_2 - z_2|^2}{ct_2^2} \right) \frac{dt_1}{t_1^{4M+1}} \frac{dt_2}{t_2^{4M+1}} \|b_R\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}^2.
\]

By the fact that \( e^{-s} \leq C s^{-k} \) for any \( k > 0 \),

\[
I_{24}(b_R)(x_1, x_2) \leq C|R| \left( \int_{\ell(I)}^{x_1 - x_I} + \int_{x_1 - x_I}^{x_2 - x_I} \right) \int_0^{t_1} t_1^{-2n-1-4M} \left( \frac{t_1}{|x_1 - x_I|} \right)^{2(n+2M-\frac{1}{2})} dt_1 + \int_0^{x_1 - x_I} t_1^{-2n-1-4M} dt_1
\]

\[
\times \left( \int_{\ell(J)}^{x_2 - x_J} + \int_{x_2 - x_J}^{x_2 - x_I} \right) t_2^{-2m-4M-1} \left( \frac{t_2}{|x_2 - x_J|} \right)^{2(m+2M+\frac{1}{2})} dt_2 + \int_0^{x_2 - x_I} t_2^{-2m-4M-1} dt_2 \|b_R\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}^2
\]

\[
\leq C|R| \frac{|I|^{1/n}}{|x_1 - x_I|^{2n+1}} \frac{|J|^{1/m}}{|x_2 - x_J|^{2m+1}} \ell(I)^{-4M} \ell(J)^{-4M} \|b_R\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}^2.
\]

Combing the estimates of \( I_{2i}(b_R)(x_1, x_2) \), together with Hölder’s inequality and elementary integration, we can show that for every \( i = 1, 2, 3, 4 \),

\[
I_{2i}(b_R)(x_1, x_2) \leq C|R| \frac{|I|^{1/n}}{|x_1 - x_I|^{2n+1}} \times \frac{|J|^{1/m}}{|x_2 - x_J|^{2m+1}} \ell(I)^{-4M} \ell(J)^{-4M} \|b_R\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}^2.
\]
which gives
\[ I_2 \leq C |R|^{1/2} \gamma_1(R)^{-1/2} \ell(I)^{-2M} \ell(J)^{-2M} b_R^2 \ell_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}. \]

Estimates of \( I_1 \) and \( I_2 \), together with Hölder’s inequality and Journé’s covering lemma, shows that

\[
I \leq \sum_{R \in m(\Omega)} \int_{(100\overline{I})^c \times \mathbb{R}^m} |S_{F,\Delta(1),\Delta(2)}(a_R)(x_1, x_2)| \, dx_1 \, dx_2 \\
\leq \sum_{R \in m(\Omega)} |R|^{1/2} \gamma_1(R)^{-1/2} \ell(I)^{-2M} \ell(J)^{-2M} \\
\times \left( \sum_{R \in m(\Omega)} \ell(I)^{-4M} \ell(J)^{-4M} \left( \left( \sum_{(100\overline{I})^c \times \mathbb{R}^m} |S_{F,\Delta(1),\Delta(2)}(a_R)(x_1, x_2)| \, dx_1 \, dx_2 \right)^2 \right) \right)^{1/2} \\
\leq C \left( \sum_{R \in m(\Omega)} |R| \gamma_1(R)^{-1} \right)^{1/2} \\
\times \left( \sum_{R \in m(\Omega)} \ell(I)^{-4M} \ell(J)^{-4M} \left( \left( \sum_{(100\overline{I})^c \times \mathbb{R}^m} |S_{F,\Delta(1),\Delta(2)}(a_R)(x_1, x_2)| \, dx_1 \, dx_2 \right)^2 \right) \right)^{1/2} \\
\leq C |\Omega|^{\frac{1}{2}} |\Omega|^{-\frac{1}{2}} \leq C.
\]

For the term \( I_2 \), we have

\[
\int_{\mathbb{R}^n \times (100\overline{I})^c} |S_{F,\Delta(1),\Delta(2)}(a_R)(x_1, x_2)| \, dx_1 \, dx_2 = \int_{(100\overline{I})^c} \int_{100I} |S_{F,\Delta(1),\Delta(2)}(a_R)(x_1, x_2)| \, dx_1 \, dx_2 \\
+ \int_{(100\overline{I})^c \times (100\overline{I})^c} |S_{F,\Delta(1),\Delta(2)}(a_R)(x_1, x_2)| \, dx_1 \, dx_2 = I_1 + I_2.
\]

The estimate of \( I_2 \) is the same with the estimate of \( I_2 \),

\[ I_2 \leq C |R|^{1/2} \gamma_1(R)^{-1/2} \ell(I)^{-2M} \ell(J)^{-2M} b_R^2 \ell_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}. \]

So we just estimate the term \( I_1 \). Using Hölder’s inequality,

\[ I_1 \leq C |I|^{1/2} \int_{(100\overline{I})^c} \left( \int_{100I} |S_{F,\Delta(1),\Delta(2)}(a_R)(x_1, x_2)|^2 \, dx_1 \right)^{1/2} \, dx_2. \quad (5.13) \]

By \( L^2(\mathbb{R}^{n+m}) \)-boundedness of the square function gives

\[
\int_{100I} |S_{F,\Delta(1),\Delta(2)}(a_R)(x_1, x_2)|^2 \, dx_1 \\
\leq \int_{\mathbb{R}^{n+1}} \int_{\mathbb{R}^n} \left[ \int_{\mathbb{R}^n} \int_{\mathbb{R}^{n+1}} \chi^{(1)}_{t_1}(x_1 - y_1, x_2 - z_2) \\
\times |(t_2^2 \Delta(1) e^{-t_2^2 \Delta(1)}) (\mathbb{I}_1 \otimes (t_2^2 \Delta(2) e^{-t_2^2 \Delta(2)}) a_R(\cdot, y_2)) (y_1) \chi^{(2)}_{t_2}(y_2 - z_2)|^2 \right] \, dy_1 \, dx_1 \, dz_2 \\
\leq C |\Omega|^{\frac{1}{2}} |\Omega|^{-\frac{1}{2}} \leq C.
\]
We then apply the time derivatives of the kernel of $\Delta^{(2)}$ to obtain

$$\int_{10J} |S_{F,\Delta^{(1)}}(a_R(x_1, x_2))|^2 dx_1 \leq C \int_{10J} \int_{0}^{t_0} \int_{|x_2-y_2|<t_2} \int_{|y_2-z_2|<t_2} \left( t_2^{-m} \exp \left( -\frac{|y_2-z_2|^2}{\epsilon t_2^2} \right) |a_{R,1}(x_1, z_2)| d\bar{z}_2 \right)^2 \frac{dy_2 dt_2}{t_2^{m+1+4M}} dx_1$$

where $a_{R,1} = (\Delta^{(1)} [\otimes_2 \mathbb{I}] b_R$. Note that $supp a_{R,1} \subset supp (\Delta^{(1)} [\otimes_2 \mathbb{I}] b_R \subset 10R = 10(I \times J)$. We then apply the time derivatives of the kernel of $\Delta^{(2)}$ to obtain

$$E_1(a_R)(x_2) \leq C \int_{10J} \int_{0}^{t_0} \int_{|x_2-y_2|<t_2} \int_{|y_2-z_2|<t_2} \left( t_2^{-m} \exp \left( -\frac{|y_2-z_2|^2}{\epsilon t_2^2} \right) |a_{R,1}(x_1, z_2)| d\bar{z}_2 \right)^2 \frac{dy_2 dt_2}{t_2^{m+1+4M}} dx_1$$

Let us estimate the term $E_1(a_R)(x_2)$. Note that if $x_2 \notin 10\bar{J}$, $0 < t_2 < \ell(J)$, $|x_2-y_2| < t_2$ and $z_2 \in 10J$, then $|y_2-z_2| \geq |x_2-x_J|/2$. We use the fact that $e^{-s} \leq C s^{-k}$ for any $k > 0$ to obtain

$$E_1(a_R)(x_2) \leq C \int_{10J} \int_{0}^{t_0} \int_{|x_2-y_2|<t_2} \int_{|y_2-z_2|<t_2} \left( t_2^{-m} \exp \left( -\frac{2|x_2-x_J|^2}{2|x_2-x_J|^2} \right) \right) \frac{dt_2}{t_2^{m+1+4M}} \times \left( \int_{10J} \int_{0}^{t_0} \left| a_{R,1}(x_1, z_2) \right|^2 d\bar{z}_2 \right)^2 dx_1$$

In order to estimate the second term $E_2(a_R)$, observe that if $x_2 \notin 10\bar{J}$, $\ell(J) \leq t_2 < |x_2-x_J|/4$, $|x_2-y_2| < t_2$ and $z_2 \in 10J$, then $|y_2-z_2| \geq |x_2-x_J|/4$. Hence,

$$E_2(a_R)(x_2) \leq C \int_{10J} \int_{0}^{t_0} \int_{|x_2-y_2|<t_2} \int_{|y_2-z_2|<t_2} \left( t_2^{-m} \exp \left( -\frac{2|x_2-x_J|^2}{2|x_2-x_J|^2} \right) \right) \frac{dt_2}{t_2^{m+1+4M}} \times \left( \int_{10J} \int_{0}^{t_0} \left| a_{R,1}(x_1, z_2) \right|^2 d\bar{z}_2 \right)^2 dx_1$$
Combining the estimates of $E_1(a_R)(x_2)$ and $E_2(a_R)(x_2)$, we obtain

$$
\int_{100I} |J|^{1/m+1} |S_{F,\triangle(1),\triangle(2)}(a_R)(x_1, x_2)|^2 \, dx_1 \leq C \frac{|J|^{1/m+1} \ell(I)^{-4M} \ell(J)^{-4M} \|((\triangle(1)^M \otimes_2 1))b_R\|_L^2(\mathbb{R}^n \times \mathbb{R}^m)}{|x_2 - x_J|^{2m+1}}.
$$

Putting (5.14) into the term $\Pi_1$ in (5.13), we have

$$
\Pi_1 \leq |R|^{1/2} \int_{100I} \frac{|J|^{1/2m+1} \ell(I)^{-2M} \ell(J)^{-2M} \|((\triangle(1)^M \otimes_2 1))b_R\|_L^2(\mathbb{R}^n \times \mathbb{R}^m)}{|x_2 - x_J|^{2m+1/2}} \, dx_2 \ell(I)^{-2M} \ell(J)^{-2M} \|((\triangle(1)^M \otimes_2 1))b_R\|_L^2(\mathbb{R}^n \times \mathbb{R}^m).
$$

By Hölder’s inequality and Journé’s covering lemma,

$$
\Pi \leq \sum_{R \in m(\Omega)} \int_{\mathbb{R}^n \times (100I)^c} |S_{F,\triangle(1),\triangle(2)}(a_R)(x_1, x_2)| \, dx_1 \, dx_2
\leq \sum_{R \in m(\Omega)} |R|^{1/2} \gamma_2(R)^{-1/2} \ell(I)^{-2M} \ell(J)^{-2M} \|((\triangle(1)^M \otimes_2 1))b_R\|_L^2(\mathbb{R}^n \times \mathbb{R}^m)
\leq C \left( \sum_{R \in m(\Omega)} \frac{|R| \gamma_2(R)^{-1}}{|R|} \right)^{1/2}
\times \left( \sum_{R \in m(\Omega)} \ell(I)^{-4M} \ell(J)^{-4M} \|((\triangle(1)^M \otimes_2 1))b_R\|_L^2(\mathbb{R}^n \times \mathbb{R}^m) \right)^{1/2}
\leq C |\Omega|^\frac{1}{2} |\Omega|^{-\frac{1}{2}} \leq C.
$$

Hence the proof is completed by (5.9) and (5.8). 

We now turn to Step 2. Our goal is to show that every $f \in H^1_{F,\triangle(1),\triangle(2)}(\mathbb{R}^n \times \mathbb{R}^m) \cap L^2(\mathbb{R}^n \times \mathbb{R}^m)$ has a $(1, 2, M)$-atom representation, with appropriate quantitative control of the coefficients. To be more specific,
Proposition 5.7. Suppose $M \geq 1$. If $f \in H^1_{F,\triangle_1,\triangle_2}(\mathbb{R}^n \times \mathbb{R}^m) \cap L^2(\mathbb{R}^n \times \mathbb{R}^m)$, then there exist a family of $(1,2,M)$-atoms $\{a_j\}_{j=0}^{\infty}$ and a sequence of numbers $\{\lambda_j\}_{j=0}^{\infty} \in \ell^1$ such that $f$ can be represented in the form $f = \sum_j \lambda_j a_j$, with the sum converging in $L^2(\mathbb{R}^n \times \mathbb{R}^m)$, and

$$\|f\|_{H^1_{F,\triangle_1,\triangle_2}(\mathbb{R}^n \times \mathbb{R}^m)} \leq C \sum_{j=0}^{\infty} |\lambda_j| \leq C \|f\|_{H^1_{F,\triangle_1,\triangle_2}(\mathbb{R}^n \times \mathbb{R}^m)},$$

where $C$ is independent of $f$. In particular,

$$H^1_{F,\triangle_1,\triangle_2}(\mathbb{R}^n \times \mathbb{R}^m) \cap L^2(\mathbb{R}^n \times \mathbb{R}^m) \subseteq H^1_{F,\triangle_1,\triangle_2}(\mathbb{R}^n \times \mathbb{R}^m).$$

Proof. Let $f \in H^1_{F,\triangle_1,\triangle_2}(\mathbb{R}^n \times \mathbb{R}^m) \cap L^2(\mathbb{R}^n \times \mathbb{R}^m)$. For each $\ell \in \mathbb{Z}$, we define

$$\Omega_{\ell} := \{(x_1,x_2) \in \mathbb{R}^n \times \mathbb{R}^m : S_{F,\triangle_1,\triangle_2}(f)(x_1,x_2) > 2^\ell\},$$

$$B_{\ell} := \{R = I \times J : \ell(J) \geq \ell(I), |R \cap \Omega_{\ell}| > \frac{1}{2}|R|, |R \cap \Omega_{\ell+1}| \leq \frac{1}{2}|R|\},$$

and

$$\bar{\Omega}_{\ell} := \{(x_1,x_2) \in \mathbb{R}^n \times \mathbb{R}^m : M_\omega(\chi_{\Omega_{\ell}}) > \frac{1}{10}\}.$$ For each rectangle $R = I \times J$ in $\mathbb{R}^n \times \mathbb{R}^m$, the tent $T(R)$ is defined as

$$T(R) := \{(y_1,y_2,t_1,t_2) : (y_1,y_2) \in R, t_1 \in (2^{-k_1},2^{-k_1+1}], t_2 \in (2^{-k_2},2^{-k_2+1}]\}.$$ For brevity, in what follows we will write $\chi_{T(R)}$ for $\chi_{T(R)}(y_1,y_2,t_1,t_2)$.

Using the reproducing formula, we can write

$$f(x_1,x_2) = \int_0^\infty \int_0^\infty \psi^{(1)}(t_1 \sqrt{\triangle_1}) \psi^{(2)}(t_2 \sqrt{\triangle_2}) (t_1^2 \triangle_1 e^{-t_1^2 \triangle_1} \otimes_2 t_2^2 \triangle_2 e^{-t_2^2 \triangle_2})(f)(x_1,x_2) \frac{dt_1 dt_2}{t_1 t_2} \cdot \int_\mathbb{R} \int_\mathbb{R} K_{\psi(t_1 \sqrt{\triangle_1})}(x_1,y_1,z_2) K_{\psi(t_2 \sqrt{\triangle_2})}(z_2,y_2) dz_2 \cdot \sum_{\ell \in \mathbb{Z}} \sum_{R \in B_{\ell}} \int_{T(R)} \int_\mathbb{R} K_{\psi(t_1 \sqrt{\triangle_1})}(x_1,y_1,z_2) K_{\psi(t_2 \sqrt{\triangle_2})}(z_2,y_2) dz_2 \cdot \sum_{\ell \in \mathbb{Z}} \lambda_{\ell} \left( \frac{1}{\lambda_{\ell}} \sum_{R \in B_1} \sum_{R \in B_2} \int_{T(R)} \int_\mathbb{R} K_{\psi(t_1 \sqrt{\triangle_1})}(x_1,y_1,z_2) K_{\psi(t_2 \sqrt{\triangle_2})}(z_2,y_2) dz_2 \right) f(y_1,y_2) dy_1 dy_2 \frac{dt_1 dt_2}{t_1 t_2} \cdot \sum_{\ell \in \mathbb{Z}} \lambda_{\ell} a_{\ell}(x_1,x_2).$$
where
\[
\lambda_\ell := \left( \sum_{R \in B_\ell} \int_0^\infty \int_0^\infty \left| (t_1^2 \Delta^{(1)} e^{-t_1^2 \Delta^{(1)}} t_2^2 \Delta^{(2)} e^{-t_2^2 \Delta^{(2)}}) (f)(y_1, y_2) \right|^2 \chi_{T(R)} \frac{dt_1 dt_2}{t_1 t_2} \right)^{1/2} \leq ||\Omega_\ell||^1/2.
\]

We now first claim that each \( a_\ell \) is a flag atom. First, it is direct to see that for each \( \ell \),
\[
a_\ell(x_1, x_2) = \sum_{\bar{R} \in B_\ell, \bar{R} \text{ max}} a_{\ell, \bar{R}}(x_1, x_2).
\]
Next, for each \( \ell \) and \( \bar{R} \in B_\ell \) with \( \bar{R} \) max, we further have
\[
a_{\ell, \bar{R}}(x_1, x_2) := \left( (\Delta^{(1)})^M (\Delta^{(2)})^M \right) (b_{\ell, \bar{R}})(x_1, x_2),
\]
where
\[
(5.15) \quad b_{\ell, \bar{R}}(x_1, x_2) := \frac{1}{\lambda_\ell} \sum_{R \in B_\ell, R \subset \bar{R}} \int_{T(R)} t_1^{2M} t_2^{2M} dt_1 dt_2
\]
\[
\int_{\mathbb{R}^{2m}} K_{\varphi^{(1)}(t_1 \sqrt{\Delta^{(1)}})} (x_1, y_1, x_2, z_2) K_{\varphi^{(2)}(t_2 \sqrt{\Delta^{(2)}})} (z_2, y_2) d\z_2
\]
\[
(t_1^2 \Delta^{(1)} e^{-t_1^2 \Delta^{(1)}} t_2^2 \Delta^{(2)} e^{-t_2^2 \Delta^{(2)}}) (f)(y_1, y_2) dy_1 dy_2 \frac{dt_1 dt_2}{t_1 t_2}
\]
and \( \varphi^{(1)}, \varphi^{(2)} \) are the function mentioned in Lemma 5.1. Then it follows from Lemma 5.1 that the integral kernel \( K_{(t_1^2 \Delta^{(1)})^k \varphi^{(1)}(t_1 \sqrt{\Delta^{(1)}})} \) of the operator \( (t_1^2 \Delta^{(1)})^k \varphi^{(1)}(t_1 \sqrt{\Delta^{(1)}}) \) satisfy
\[
(5.16) \quad \text{supp } K_{(t_1^2 \Delta^{(1)})^k \varphi^{(1)}(t_1 \sqrt{\Delta^{(1)}})} \subset \{ (x, y) \in \mathbb{R}^{n+m} \times \mathbb{R}^{n+m} : |x - y| < t_1 \}
\]
and
\[
(5.17) \quad \text{supp } K_{(t_2^2 \Delta^{(2)})^k \varphi^{(2)}(t_2 \sqrt{\Delta^{(2)}})} \subset \{ (u, v) \in \mathbb{R}^{m} \times \mathbb{R}^{m} : |u - v| < t_2 \}.
\]
We now consider the support of \( (\Delta^{(1)})^{k_1} \otimes_2 \Delta^{(2)}{k_2} b_{\ell, \bar{R}} \). From the definition of \( b_{\ell, \bar{R}} \) as in (5.15), we have that
\[
(\Delta^{(1)})^{k_1} \otimes_2 \Delta^{(2)}{k_2} (b_{\ell, \bar{R}})(x_1, x_2)
\]
\[
:= \frac{1}{\lambda_\ell} \sum_{R \in B_\ell, R \subset \bar{R}} \int_{T(R)} t_1^{2M-2k} t_2^{2M-2k} dt_1 dt_2
\]
\[
\int_{\mathbb{R}^{2m}} K_{(t_1^2 \Delta^{(1)})^k \varphi^{(1)}(t_1 \sqrt{\Delta^{(1)}})} (x_1, y_1, x_2, z_2) K_{(t_2^2 \Delta^{(2)})^k \varphi^{(2)}(t_2 \sqrt{\Delta^{(2)}})} (z_2, y_2) d\z_2
\]
\[
(t_1^2 \Delta^{(1)} e^{-t_1^2 \Delta^{(1)}} t_2^2 \Delta^{(2)} e^{-t_2^2 \Delta^{(2)}}) (f)(y_1, y_2) dy_1 dy_2 \frac{dt_1 dt_2}{t_1 t_2}.
\]
Now from the following term on the right-hand side of the above equality

\[
\int_{\mathbb{R}^m} K(t_1^2 \Delta^{(1)})_{t_1 \sqrt{\Delta^{(1)}}} (x_1, y_1, x_2, z_2) K(t_2^2 \Delta^{(2)})_{t_2 \sqrt{\Delta^{(2)}}} (z_2, y_2) dz_2
\] (5.18)

and from the support conditions (5.16) and (5.17), we obtain that for \((x_1, x_2)\) in (5.18),

\[
|x_1 - y_1| \leq 3\ell(I_R), \quad |x_2 - z_2| \leq 3\ell(I_R), \quad \text{and} \quad |z_2 - y_2| \leq 3\ell(J_R)
\]

since \((y_1, y_2, t_1, t_2) \in T(R)\). Hence we obtain that

\[
|x_1 - y_1| \leq 3\ell(I_R) \quad \text{and} \quad |x_2 - y_2| \leq 3\ell(I_R) + 3\ell(J_R).
\]

As a consequence, we have that for every \(k_1, k_2 = 0, 1, \ldots, M,\)

\[
\text{supp} (\Delta^{(1)k_1} \otimes \Delta^{(2)k_2}) b_{\ell, R} \subseteq 10R,
\] (5.19)

where \(\ell(I_R) = \ell(I_R)\) and \(\ell(J_R) = \ell(I_R) + \ell(J_R)\).

Then based on the support condition above and on the definition of \(\tilde{\Omega}_\ell\), we obtain that

\[
\text{supp} a_\ell \subseteq \tilde{\Omega}_\ell.
\]

Next we estimate \(\|a_\ell\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}\). Taking \(g \in L^2(\mathbb{R}^n \times \mathbb{R}^m)\) with \(\|g\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)} = 1\), from the definition of \(a_\ell\), we have

\[
\left| \int_{\mathbb{R}^n \times \mathbb{R}^m} a_\ell(x_1, x_2)g(x_1, x_2) dx_1 dx_2 \right|
= \left| \left( \frac{1}{\lambda_\ell} \sum_{R \in B_{\ell, R} \text{max}} \sum_{R \in B_{\ell, R} \subset R} \int_{T(R)} \psi(t_1 \sqrt{\Delta^{(1)}}) \psi(t_2 \sqrt{\Delta^{(2)}})(g)(y_1, y_2) \right. \right.
\]

\[
\left. \left( t_1^2 \Delta^{(1)} e^{-t_2^2 \Delta^{(1)}} t_2^2 \Delta^{(2)} e^{-t_2^2 \Delta^{(2)}} \right) (f)(y_1, y_2) dy_1 dy_2 \frac{dt_1 dt_2}{t_1 t_2} \right| \leq \frac{1}{\lambda_\ell} \int_{\mathbb{R}^n \times \mathbb{R}^m} \left( \sum_{R \in B_{\ell}} \int_0^\infty \int_0^\infty |\psi(t_1 \sqrt{\Delta^{(1)}}) \psi(t_2 \sqrt{\Delta^{(2)}})(g)(y_1, y_2)|^2 \chi_{T(R)} \frac{dt_1 dt_2}{t_1 t_2} \right)^{\frac{1}{2}} dy_1 dy_2
\]

\[
\leq \frac{1}{\lambda_\ell} \|g\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)} \left\| \left( \sum_{R \in B_{\ell}} \int_0^\infty \int_0^\infty \left| t_1^2 \Delta^{(1)} e^{-t_2^2 \Delta^{(1)}} t_2^2 \Delta^{(2)} e^{-t_2^2 \Delta^{(2)}} \right|^2 \chi_{T(R)} \frac{dt_1 dt_2}{t_1 t_2} \right)^{\frac{1}{2}} \right\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)} \leq |\tilde{\Omega}_\ell|^{-\frac{1}{2}},
\]
and hence, we have $\|a\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)} \leq C|\Omega|^\frac{1}{2}$.

A similar argument to that above shows that for every $0 \leq k_1, k_2 \leq M$,

$$\sum_{R \in B_t, R \text{ max}} \ell(I_R)^{-4M} \ell(J_R)^{-4M} \|\ell(I_R)^2 \Delta^{(1)}\|^{k_1} \otimes_2 \ell(J_R)^2 \Delta^{(2)}\|^{k_2} \beta_{\ell,R}^2 \leq C|\Omega|^{-1}.$$ 

Combining all the estimates above, we can see that $a$ is a $(1,2,M)$-atom as in Definition 1.9 up to some constant depending only on $M, \psi$.

To see that the atomic decomposition $\sum_{\ell} \lambda_\ell a_\ell$ converges to $f$ in the $L^2(\mathbb{R}^n \times \mathbb{R}^m)$ norm, we only need to show that $\|\sum_{|\ell|>G} \lambda_\ell a_\ell\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)} \to 0$ as $G$ tends to infinity. To see this, first note that

$$\left\| \sum_{|\ell|>G} \lambda_\ell a_\ell \right\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)} = \sup_{h: \|h\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}=1} \left| \left\langle \sum_{|\ell|>G} \lambda_\ell a_\ell, h \right\rangle_{L^2(\mathbb{R}^n \times \mathbb{R}^m)} \right|.$$

Next, we have

$$\left| \left\langle \sum_{|\ell|>G} \lambda_\ell a_\ell, h \right\rangle_{L^2(\mathbb{R}^n \times \mathbb{R}^m)} \right| = \left| \sum_{|\ell|>G} \sum_{R \in B_t} \int_{T(R)} \psi(t_1 \sqrt{\Delta^{(1)}}) \psi(t_2 \sqrt{\Delta^{(2)}}) (h)(y_1, y_2) \right|$$

$$\times \left| (t_1^2 \Delta^{(1)} - t_2^2 \Delta^{(1)}, t_2^2 \Delta^{(2)} e^{-t_2^2 \Delta^{(2)}}) (f)(y_1, y_2) dy_1 dy_2 dt_1 dt_2 \right|$$

$$\leq \int_{\mathbb{R}^n \times \mathbb{R}^m} \left( \sum_{|\ell|>G} \sum_{R \in B_t} \int_0^\infty \int_0^\infty |\psi(t_1 \sqrt{\Delta^{(1)}}) \psi(t_2 \sqrt{\Delta^{(2)}}) (h)(y_1, y_2)|^2 \chi_{T(R)} dt_1 dt_2 \right)^{\frac{1}{2}}$$

$$\times \left( \sum_{|\ell|>G} \sum_{R \in B_t} \int_0^\infty \int_0^\infty \left| (t_1^2 \Delta^{(1)} - t_2^2 \Delta^{(1)}, t_2^2 \Delta^{(2)} e^{-t_2^2 \Delta^{(2)}}) (f)(y_1, y_2)|^2 \chi_{T(R)} \right| dt_1 dt_2 \right)^{\frac{1}{2}} dy_1 dy_2$$

$$\leq C \|h\|_{L^2}$$

$$\times \left\| \left( \sum_{|\ell|>G} \sum_{R \in B_t} \int_0^\infty \int_0^\infty \left| (t_1^2 \Delta^{(1)} - t_2^2 \Delta^{(1)}, t_2^2 \Delta^{(2)} e^{-t_2^2 \Delta^{(2)}}) (f)|^2 \chi_{T(R)} \right| dt_1 dt_2 \right)^{\frac{1}{2}} \right\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}$$

$$\to 0$$

as $G$ tends to $\infty$, since $\|S_{F, \Delta^{(1)}, \Delta^{(2)} f}\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)} < \infty$. This implies that $f = \sum_\ell \lambda_\ell a_\ell$ in the sense of $L^2(\mathbb{R}^n \times \mathbb{R}^m)$.

Next, we verify the estimate for the series $\sum_\ell |\lambda_\ell|$. To deal with this, we claim that for each $\ell \in \mathbb{Z}$,

$$\sum_{R \in B_t} \int_{T(R)} \left| (t_1^2 \Delta^{(1)} - t_2^2 \Delta^{(1)}, t_2^2 \Delta^{(2)} e^{-t_2^2 \Delta^{(2)}}) (f)(y_1, y_2)|^2 dy_1 dy_2 \frac{dt_1 dt_2}{t_1 t_2} \leq C 2^{2(\ell+1)} |\Omega_\ell|.$$ 

First we note that

$$\int_{\Omega_\ell \setminus \Omega_{\ell+1}} S_{F, \Delta^{(1)}, \Delta^{(2)} f}^2 (x_1, x_2) dx_1 dx_2 \leq 2^{2(\ell+1)} |\Omega_\ell|.$$
Also we point out that

\[
\int_{\Omega_{t+1}} S_{F,\triangle(1),\triangle(2)} (f)^2 dx_1 dx_2
= \int_{\Omega_{t+1}} \int_{\mathbb{R}^n_+} \int_{\mathbb{R}^m_+} \chi_t \lambda_t (x_1 - y_1, x_2 - y_2) 
   \left| \left( t_1^2 \triangle(1) e^{-t_2^2 \triangle(1)} \otimes_2 t_2^2 \triangle(2) e^{-t_2^2 \triangle(2)} \right) f(y_1, y_2) \right|^2 
   \frac{dy_1 dy_2 dt_1 dt_2}{t_1 t_2 t_1^m t_2^m} dx_1 dx_2
\]

\[
= \int_{\mathbb{R}^n_+} \int_{\mathbb{R}^m_+} \left| \left( t_1^2 \triangle(1) e^{-t_2^2 \triangle(1)} \otimes_2 t_2^2 \triangle(2) e^{-t_2^2 \triangle(2)} \right) f(y_1, y_2) \right|^2 
   \times \left| \left\{ (x_1, x_2) \in \tilde{\Omega}_t \setminus \Omega_{t+1} : |x_1 - y_1| < t_1, |x_2 - y_2| < t_1 + t_2 \right\} \right| 
   \frac{dy_1 dy_2 dt_1 dt_2}{t_1^m t_2^m},
\]

where the last inequality follows from the definition of \( B_\ell \). This shows that the claim holds.

As a consequence, we have

\[
\sum_{\ell} |\lambda_\ell| 
\leq C \sum_{\ell} \left\| \left( \sum_{R \in B_\ell} \int_0^\infty \int_0^\infty \left| \left( t_1^2 \triangle(1) e^{-t_2^2 \triangle(1)} \otimes_2 t_2^2 \triangle(2) e^{-t_2^2 \triangle(2)} \right) f \right|^2 \chi_{T(R)} \right| \frac{dt_1 dt_2}{t_1 t_2} \right\|_{L^2(\mathbb{R}_+ \times \mathbb{R}^m)} 
\times |\tilde{\Omega}_t|^{1/2}
\leq C \sum_{\ell} \left( \sum_{R \in B_\ell} \int_{T(R)} \left| \left( t_1^2 \triangle(1) e^{-t_2^2 \triangle(1)} \otimes_2 t_2^2 \triangle(2) e^{-t_2^2 \triangle(2)} \right) f(y_1, y_2) \right|^2 \frac{dy_1 dy_2 dt_1 dt_2}{t_1 t_2} \right)^{1/2} 
\times |\tilde{\Omega}_t|^{1/2}
\leq C \sum_{\ell} 2^{\ell+1} |\tilde{\Omega}_t| 
\leq C \sum_{\ell} 2^\ell |\tilde{\Omega}_t|
\leq C \| S_{F,\triangle(1),\triangle(2)} (f) \|_{L^1(\mathbb{R}_+ \times \mathbb{R}^m)}
\leq C \| f \|_{H^{1}_{F,\triangle(1),\triangle(2)}(\mathbb{R}_+ \times \mathbb{R}^m)}.
\]

This completes the proof of Proposition 5.7. \( \Box \)
6 The estimate $\sum_{j=1}^{n+m} \sum_{k=1}^{m} \|R_{j,k}(f)\|_1 + \|f\|_1 \lesssim \|S_F(f)\|_1$ via atomic decomposition

Based on the estimate in (6.3), to prove the estimate $\sum_{j=1}^{n+m} \sum_{k=1}^{m} \|R_{j,k}(f)\|_1 + \|f\|_1 \lesssim \|S_F(f)\|_1$, it suffices to prove that there exists a positive constant $C$ such that for $j = 0, 1, \ldots, n + m, k = 0, \ldots, m$

$$\|R_{j,k}(f)\|_1 \leq C\|S_{F,(\Omega)}(f)\|_1.$$ (6.1)

Indeed, as mentioned, $R_{j,k}$ is the composition of $R_j$ and $R_k$, and hence $R_{j,k}$ is bounded on $L^p(\mathbb{R}^{n+m}), 1 < p < \infty$. The flag Riesz transform can also be defined by $T := \nabla^{(1)\Delta(1)^{-1/2}} \otimes_2 \nabla^{(2)\Delta(2)^{-1/2}}$ as follows,

$$Tf(x_1, x_2) = \frac{1}{4\pi} \int_0^\infty \int_0^\infty (\nabla^{(1)} e^{-t_1 \Delta(1)} \otimes_2 \nabla^{(2)} e^{-t_2 \Delta(2)}) f(x_1, x_2) \frac{dt_1 dt_2}{\sqrt{t_1 t_2}}.$$ (6.2)

The estimate $\sum_{j=1}^{n+m} \sum_{k=1}^{m} \|R_{j,k}(f)\|_1 + \|f\|_1 \lesssim \|S_F(f)\|_1$ follows from the following theorem.

**Theorem 6.1.** The flag Riesz transform $\nabla^{(1)\Delta(1)^{-1/2}} \otimes_2 \nabla^{(2)\Delta(2)^{-1/2}}$ extends to a bounded operator from $H^1_{F,(\Omega)}(\mathbb{R}^n \times \mathbb{R}^m)$ to $L^1(\mathbb{R}^n \times \mathbb{R}^m)$.

**Proof.** Let $T := \nabla^{(1)(\Omega)^{-1/2}} \otimes_2 \nabla^{(2)(\Omega)^{-1/2}}$. It suffices to show that $T$ is uniformly bounded on each $(1, 2, M)$ atom $a$ with $M > \max\{n, m\}/2$, and there exists a constant $C > 0$ independent of $a$ such that

$$\|T(a)\|_{L^1(\mathbb{R}^n \times \mathbb{R}^m)} \leq C.$$ (6.3)

From the definition of $(1, 2, M)$ atom, it follows that $a$ is supported in some $\Omega \subset \mathbb{R}^n \times \mathbb{R}^m$ and $a$ can be further decomposed into $a = \sum_{R \in \mathcal{E}(\Omega)} a_R$. For any $R = I \times J \subset \Omega$, let $l$ be the biggest dyadic cube containing $I$, so that $l \times J \subset \Omega$, where $\Omega = \{x \in \mathbb{R}^n \times \mathbb{R}^m : M_s(\chi_{\Omega})(x) > 1/2\}$. Next, let $Q$ be the biggest dyadic cube containing $J$, so that $l \times Q \subset \Omega$, where $\Omega = \{x \in \mathbb{R}^n \times \mathbb{R}^m : M_s(\chi_{\Omega})(x) > 1/2\}$. Now let $\tilde{R}$ be the 100-fold dilate of $l \times Q$ concentric with $l \times Q$. Clearly, an application of the strong maximal function theorem shows that $| \bigcup_{R \in \Omega} \tilde{R} | \leq C|\Omega| \leq C|\Omega| \leq C|\Omega|$. From (iii) in the definition of $(1, 2, M)$ atom, we can obtain that

$$\int_{\bigcup_{l \times Q} |T(a)(x)|dx \leq \bigcup_{l \times Q} \tilde{R} |1/2 |T(a)||_{L^2} \leq C|\Omega|^{1/2} |a|_{L^2} \leq C|\Omega|^{1/2} |\Omega|^{-1/2} \leq C.$$ (6.4)

Therefore, the proof of (6.3) reduces to showing that

$$\int_{(\bigcup_{l \times Q} \tilde{R})} |T(a)(x)|dx \leq C.$$ (6.4)
Since \( a = \sum_{R \in m(\Omega)} a_R \), we have
\[
\int_{(\cup R)^c} |T(a)(x)| \, dx \\
\leq \sum_{R \in m(\Omega)} \int_{R^c} |T(a_R)(x)| \, dx \\
\leq \sum_{R \in m(\Omega)} \int_{(100)^c \times \mathbb{R}^n} |T(a_R)(x)| \, dx + \sum_{R \in m(\Omega)} \int_{\mathbb{R}^n \times (100)^c} |T(a_R)(x)| \, dx \\
= I + II.
\]

For term I, we observe that
\[
\int_{(100)^c \times \mathbb{R}^n} |T(a_R)(x_1, x_2)| \, dx_1 \, dx_2 \\
= \left( \int_{(100)^c \times 100J} + \int_{(100)^c \times (100)J^c} \right) |T(a_R)(x_1, x_2)| \, dx_1 \, dx_2 \\
= I_1 + I_2.
\]

Let us first estimate the term \( I_1 \). Then
\[
I_1 \leq \int_{(100)^c \times 100J} \left| \nabla^{(1)} \triangle^{(1)-1/2} \otimes_2 \nabla^{(2)} \triangle^{(2)-1/2} a_R(x_1, x_2) \right| \, dx_1 \, dx_2 \\
= \int_{(100)^c \times 100J} \left| \frac{1}{2\sqrt{\pi}} \int_0^\infty \nabla^{(1)} e^{-t_1 \triangle^{(1)}} \left( \nabla^{(2)} \triangle^{(2)-1/2} a_R(x_1, x_2) \right) \, dt_1 \right| \, dx_1 \, dx_2 \\
= \int_{(100)^c \times 100J} \left| \frac{1}{2\sqrt{\pi}} \int_0^{\ell(I)\ell(J)} \nabla^{(1)} e^{-t_1 \triangle^{(1)}} \left( \nabla^{(2)} \triangle^{(2)-1/2} a_R(x_1, x_2) \right) \, dt_1 \right| \, dx_1 \, dx_2 \\
+ \int_{(100)^c \times 100J} \left| \frac{1}{2\sqrt{\pi}} \int_{\ell(I)\ell(J)}^\infty \nabla^{(1)} e^{-t_1 \triangle^{(1)}} \left( \nabla^{(2)} \triangle^{(2)-1/2} a_R(x_1, x_2) \right) \, dt_1 \right| \, dx_1 \, dx_2 \\
=: I_{11} + I_{12}.
\]

We first consider \( I_{11} \). We write
\[
I_{11} \leq \int_{(100)^c \times 100J} \left| \frac{1}{2\sqrt{\pi}} \int_0^{\ell(I)\ell(J)} \nabla^{(1)} e^{-t_1 \triangle^{(1)}} \left( \nabla^{(2)} \triangle^{(2)-1/2} a_R(x_1, x_2) \right) \, dt_1 \right| \, dx_1 \, dx_2 \\
\leq C \sum_{j_1 = j_2 = 0}^{\infty} \sum_{|x_1 - x_j|} \int_{|x_1 - x_j| \approx 2^{j_1} \ell(I)} \int_{100J} \int_0^{\ell(I)\ell(J)} \int_{|y_2 - y_j| \approx 2^{j_2} \ell(J)} \frac{1}{\ell_1 \ell_2} e^{-\frac{|(x_1, x_2) - (y_1, y_2)|^2}{\ell_1 \ell_2}} \left| \left( \nabla^{(2)} \triangle^{(2)-1/2} a_R(y_1, y_2) \right) \right| \, dy_1 dy_2 \, dt_1 \, dx_1 \, dx_2 \\
\leq C \sum_{j_1 = j_2 = 0}^{\infty} \sum_{|x_1 - x_j|} \int_{|x_1 - x_j| \approx 2^{j_1} \ell(I)} \int_{100J} \int_0^{\ell(I)\ell(J)} \int_{|y_2 - y_j| \approx 2^{j_2} \ell(J)} \frac{\ell_1 \ell_2}{\ell_1 \ell_2} \left| \left( \nabla^{(2)} \triangle^{(2)-1/2} a_R(y_1, y_2) \right) \right| \, dy_1 dy_2 \, dt_1 \, dx_1 \, dx_2.
\]
\[
\frac{1}{\ell_1^{n+m}} \left( \frac{t_1}{|x_1 - x_2|^2} \right)^{\alpha_1} \left( \frac{t_1}{|y_2 - y_1|^2} \right)^{\alpha_2} \left| \left( \nabla^{(2)} \Delta^{(2)} \right)^{-1/2} a_R(y_1, y_2) \right| dy_1 dy_2 \frac{dt_1}{t_1} dx_1 dx_2,
\]

where we choose \(2\alpha_2 - m > 2\alpha_1 - n > 1\) and \(\tilde{j}\) is the smallest integer such that \(2^{\tilde{j}} I \cap (100l)^c \neq \emptyset\). Hence,

\[
I_{11} \leq C \sum_{j_1=j}^{\infty} \sum_{j_2=6}^{\infty} (2^{j_2} \ell(I))^{n} \ell(J)^m \frac{1}{(2^{j_1} \ell(I))^{2\alpha_1}} \frac{1}{(2^{j_2} \ell(J))^{2\alpha_2}} \int_0^{\ell(I) \ell(J)} \frac{dt_1}{t_1^{\alpha_1 + \alpha_2 - \frac{n+m}{2} - 1}} dt_1
\]

Using the fact that \(\ell(I) \leq \ell(J)\), we have

\[
I_{11} \leq C |I|^{\frac{1}{2}} |J|^{\frac{1}{2}} \sum_{j_1=j}^{\infty} \sum_{j_2=1}^{\infty} 2^{-j_1(2\alpha_1 - n)} 2^{-j_2(2\alpha_2 - \frac{m}{2})} \ell(I)^{\frac{\alpha_1}{2} - \alpha_1 + \alpha_2 - \frac{m}{2}} \ell(J)^{\frac{\alpha_2}{2} - \alpha_2} \|a_R\|_{L^2(\mathbb{R}^{n+m})}
\]

To consider \(I_{12}\), we choose \(2\alpha_1 - n > 0, 4\alpha_2 - m > 0, 2\alpha_2 - m > 2\alpha_1 - n\) and \(2\alpha_1 + 2\alpha_2 < n + m\). Hence

\[
I_{12} \leq \int_{(100l)^c \times 100l} \left| \frac{1}{2^{\frac{1}{2}\ell(I)}} \int_{\ell(I) \ell(J)}^{\infty} \nabla^{(1)} e^{-t_1 \Delta^{(1)}} \left( \nabla^{(2)} \Delta^{(2)} \right)^{-1/2} a_R(x_1, x_2) \frac{dt_1}{t_1^{\frac{1}{2}}} \right| dx_1 dx_2
\]

\[
\leq C \sum_{j_1=j}^{\infty} \sum_{j_2=6}^{\infty} \frac{1}{(2^{j_1} \ell(I))^{n}} \frac{1}{(2^{j_2} \ell(J))^{m}} \int_{x_1-x_2=0}^{2^{j_2} \ell(I)} \int_0^{\ell(I) \ell(J)} \int_{|y_2-y_1|=2^{j_2} \ell(J)} \frac{1}{(x_1 - x_2)^{\alpha_1}} \left( \frac{t_1}{|y_2 - y_1|^2} \right)^{\alpha_2} \left| \left( \nabla^{(2)} \Delta^{(2)} \right)^{-1/2} a_R(y_1, y_2) \right| dy_1 dy_2 \frac{dt_1}{t_1^{\frac{1}{2}}} dx_1 dx_2
\]

\[
\leq C \sum_{j_1=j}^{\infty} \sum_{j_2=6}^{\infty} (2^{j_1} \ell(I))^n \ell(J)^m \frac{1}{(2^{j_1} \ell(I))^{2\alpha_1}} \frac{1}{(2^{j_2} \ell(J))^{2\alpha_2}} \int_0^{\ell(I) \ell(J)} \frac{1}{t_1^{n+m - \alpha_1 - \alpha_2 + 1}} dt_1
\]

\[
|I|^{\frac{1}{2}} (2^{j_2} \ell(J))^{\frac{m}{2}} \left( \int_{|y_2-y_1|=2^{j_2} \ell(J)} \left| \left( \nabla^{(2)} \Delta^{(2)} \right)^{-1/2} a_R(y_1, y_2) \right| dy_1 dy_2 \right)^{\frac{1}{2}}
\]
where again the last inequality follows from the fact that 

$$\alpha > \frac{m}{2}.$$ 

Since $2\alpha_2 - m > 2\alpha_1 - n$ and $\ell(I) \leq \ell(J)$, we get

$$I_{12} \leq C|I|^{\frac{1}{2}}\ell(J)^{\frac{m}{2}} \sum_{j_1=j}^\infty \sum_{j_2=6}^\infty 2^{-j_1(2\alpha_1-n)} \ell(I)^{\frac{n}{2} - \alpha_1 + \alpha_2 - \frac{m}{2} \alpha_2 - \frac{m}{2}} \|a_R\|_{L^2(\mathbb{R}^{n+m})}.$$

We now consider $I_2$. We write

$$I_2 \leq \int_{(100)^c \times (100)^c} \left| \nabla^{(1)}(1)^{-1/2} \otimes_2 \nabla^{(2)}(2)^{-1/2} a_R(x_1, x_2) \right| \, dx_1 \, dx_2$$

$$= \int_{(100)^c \times (100)^c} \left| \frac{1}{2\sqrt{\pi}} \int_0^\infty \nabla^{(1)} e^{-t_1 \Delta^{(1)}} \nabla^{(2)} e^{-t_2 \Delta^{(2)}} \, \frac{dt_1}{\sqrt{t_1}} \right| \, dx_1 \, dx_2$$

$$= \int_{(100)^c \times (100)^c} \left| \frac{1}{2\sqrt{\pi}} \int_0^{\ell(I)^2} \nabla^{(1)} e^{-t_1 \Delta^{(1)}} \nabla^{(2)} e^{-t_2 \Delta^{(2)}} \, \frac{dt_1}{\sqrt{t_1}} \right| \, dx_1 \, dx_2$$

$$+ \int_{(100)^c \times (100)^c} \left| \frac{1}{2\sqrt{\pi}} \int_0^\infty \nabla^{(1)} e^{-t_1 \Delta^{(1)}} \nabla^{(2)} e^{-t_2 \Delta^{(2)}} \, \frac{dt_1}{\sqrt{t_1}} \right| \, dx_1 \, dx_2$$

$$= I_{21} + I_{22}.$$

We first consider $I_{21}$. From the heat kernel estimate and the support condition of $a_R$, it is clear that

$$I_{21} \leq C \int_{(100)^c \times (100)^c} \left| \frac{1}{2\sqrt{\pi}} \int_0^{\ell(I)^2} \int_{\mathbb{R}^{n+m}} q_{t_1}^{(1)}(x_1 - y_1, x_2 - y_2) \nabla^{(2)} e^{-t_2 \Delta^{(2)}} \, \frac{dt_1}{t_1} \right| \, dx_1 \, dx_2$$

$$\leq C \int_{(100)^c \times (100)^c} \left| \frac{1}{2\sqrt{\pi}} \int_0^{\ell(I)^2} \int_{1}^{\infty} q_{t_1}^{(1)}(x_1 - y_1, x_2 - y_2) \nabla^{(2)} e^{-t_2 \Delta^{(2)}} \, \frac{dt_1}{t_1} \right| \, dx_1 \, dx_2$$

$$+ C \int_{(100)^c \times (100)^c} \left| \frac{1}{2\sqrt{\pi}} \int_0^{\ell(I)^2} \sum_{k_2=0}^\infty \int_{|y_2 - y_1| \approx 2^{k_2} \ell(J)} q_{t_1}^{(1)}(x_1 - y_1, x_2 - y_2) \nabla^{(2)} e^{-t_2 \Delta^{(2)}} \, \frac{dt_1}{t_1} \right| \, dx_1 \, dx_2.$$
\[ I_{211} = I_{212} + I_{212}. \]

For the term \( I_{211} \), Hölder’s inequality gives

\[
I_{211} \leq C \int_{(100)^c \times (100)^c} \int_0^{\ell(I)^2} \left( \int_I \left( \int |q_1^{(1)}(x_1 - y_1, x_2 - y_2)|^2 dy_2 \right) \frac{1}{t_1} \frac{dt_1}{t_1} \right) \frac{dt_1}{t_1} \frac{dy_1}{dy_1} dx_1 dx_2
\]

\[
\leq C \int_{(100)^c \times (100)^c} \int_0^{\ell(I)^2} \left( \int_I \left( \int \left( \frac{1}{t_1} |e^{\frac{|(x_1, x_2)-(y_1, y_2)|^2}{t_1}}| \right)^2 dy_2 \right) \frac{1}{t_1} \frac{dt_1}{t_1} \right) \frac{dt_1}{t_1} \frac{dy_1}{dy_1} dx_1 dx_2.
\]

Choosing \( 2\alpha_1 - n > 0 \) and \( 2\alpha_2 - m > 0 \), the fact \( \ell(I) \leq \ell(J) \) implies

\[
I_{211} \leq C \sum_{j_1 = j}^{\infty} \sum_{j_2 = 0}^{\infty} \frac{1}{|x_1 - x_j|^2} \int_{|x_1 - x_j| \approx 2^{j_1} \ell(I)} \int_{|x_2 - y_j| \approx 2^{j_2} \ell(J)} \int_I \left( \int |a_R(y_1, y_2)|^2 dy_2 \right) \frac{1}{t_1} \frac{dt_1}{t_1} \frac{dy_1}{dy_1} dx_1 dx_2
\]

\[
\leq C |R|^{1/2} \sum_{j_1 = j}^{\infty} \sum_{j_2 = 0}^{\infty} \ell(I)^{2(\alpha_1 + \alpha_2 - \frac{n+m}{2})} \frac{1}{(2^{j_1} \ell(I))^{2\alpha_1 - n}} \frac{1}{(2^{j_2} \ell(J))^{2\alpha_2 - m}} \|a_R\|_{L^2(R^{n+m})} dx_1 dx_2
\]

\[
\leq C |R|^{1/2} \gamma_1(R)^{-(2\alpha_1 - n)} \|a_R\|_{L^2(R^{n+m})},
\]

where again the last inequality follows from the fact that

\[ 2^j \approx \ell(I) \]

\[ \frac{\ell(I)}{\ell(J)}. \]

To estimate \( I_{212} \), we have

\[
I_{212} \leq C \int_{(100)^c \times (100)^c} \int_0^{\ell(I)^2} \int_0^{\infty} \sum_{k_2 = 0}^{\infty} \frac{1}{t_1} \frac{dt_1}{t_1} \frac{dt_2}{t_2} \left| dx_1 dx_2 \right|
\]

\[
\int_I \int_{|y_2 - y_2| \approx 2^{k_2} \ell(J)} \int q_1^{(1)}(x_1 - y_1, x_2 - y_2) q_1^{(2)}(y_2 - z_2) a_R(y_1, z_2) dz_2 dy_1 dy_2
\]

\[
\leq C \int_{(100)^c \times (100)^c} \int_0^{\ell(I)^2} \int_0^{\infty} \sum_{k_2 = 0}^{\infty} \frac{1}{t_1} \frac{dt_1}{t_1} e^{-\frac{|(x_1, x_2)-(y_1, y_2)|^2}{t_1}} \int dx_1 dx_2.
\]
\[
\int J \int_{|y_2-z_2| \approx 2^{k_2} t(J)} \int J \frac{1}{t_2^2} e^{-\frac{|y_2-z_2|^2}{t_2}} dy_2 |a_R(y_1, z_2)| dz_2 dy_1 \frac{dt_1}{t_1} \frac{dt_2}{t_2} dx_1 dx_2 
\]
\[
\leq C \int_{(100)^c \times (100)^c} (\int J \int_{|y_2-z_2| \approx 2^{k_2} t(J)} (\int J \int_{|y_2-z_2| \approx 2^{k_2} t(J)} \left( (\int_{y_1}^t + \int_t^\infty) \frac{1}{t_2} e^{-\frac{\ell(J)^2}{2t_2}} dt_2 \right) \int J \int_{|y_2-z_2| \approx 2^{k_2} t(J)} e^{-\frac{|y_2-z_2|^2}{2t_2}} dy_2 |a_R(y_1, z_2)| dz_2 dy_1 \frac{dt_1}{t_1} \frac{dt_2}{t_2} dx_1 dx_2 
\]
\[
\leq C \int_{(100)^c \times (100)^c} \int J \int_{|y_2-z_2| \approx 2^{k_2} t(J)} \left( (\int_{y_1}^t + \int_t^\infty) \frac{1}{t_2} e^{-\frac{\ell(J)^2}{2t_2}} dt_2 \right) \int J \int_{|y_2-z_2| \approx 2^{k_2} t(J)} e^{-\frac{|y_2-z_2|^2}{2t_2}} dy_2 |a_R(y_1, z_2)| dz_2 dy_1 \frac{dt_1}{t_1} \frac{dt_2}{t_2} dx_1 dx_2 
\]
\[
=: I_{2121} + I_{2122}. 
\]

By the heat kernel estimate, we choose \(2\alpha_1 - n > 0\) and \(2\alpha_2 - m > 0\) to obtain

\[
I_{2121} \leq C \int_{(100)^c \times (100)^c} \int J \int_{|y_2-z_2| \approx 2^{k_2} t(J)} \left( (\int_{y_1}^t + \int_t^\infty) \frac{1}{t_2} e^{-\frac{\ell(J)^2}{2t_2}} dt_2 \right) \int J \int_{|y_2-z_2| \approx 2^{k_2} t(J)} e^{-\frac{|y_2-z_2|^2}{2t_2}} dy_2 |a_R(y_1, z_2)| dz_2 dy_1 \frac{dt_1}{t_1} \frac{dt_2}{t_2} dx_1 dx_2 
\]
\[
\leq C \int_{(100)^c \times (100)^c} \int J \int_{|y_2-z_2| \approx 2^{k_2} t(J)} \left( (\int_{y_1}^t + \int_t^\infty) \frac{1}{t_2} e^{-\frac{\ell(J)^2}{2t_2}} dt_2 \right) \int J \int_{|y_2-z_2| \approx 2^{k_2} t(J)} e^{-\frac{|y_2-z_2|^2}{2t_2}} dy_2 |a_R(y_1, z_2)| dz_2 dy_1 \frac{dt_1}{t_1} \frac{dt_2}{t_2} dx_1 dx_2 
\]
\[
\leq C \sum_{j_1=j}^{\infty} \int_{|x_1-x_t| \approx 2^{k_2} t(J)} \left( (\int_{y_1}^t + \int_t^\infty) \frac{1}{t_2} e^{-\frac{\ell(J)^2}{2t_2}} dt_2 \right) \int J \int_{|y_2-z_2| \approx 2^{k_2} t(J)} e^{-\frac{|y_2-z_2|^2}{2t_2}} dy_2 |a_R(y_1, z_2)| dz_2 dy_1 \frac{dt_1}{t_1} \frac{dt_2}{t_2} dx_1 dx_2 
\]

Hölder’s inequality gives

\[
I_{2121} \leq C |R|^{1/2} \sum_{j_1=j}^{\infty} \sum_{j_2=6}^{\infty} \left( (\int_{y_1}^t + \int_t^\infty) \frac{1}{t_2} e^{-\frac{\ell(J)^2}{2t_2}} dt_2 \right) \int J \int_{|y_2-z_2| \approx 2^{k_2} t(J)} e^{-\frac{|y_2-z_2|^2}{2t_2}} dy_2 |a_R(y_1, z_2)| dz_2 dy_1 \frac{dt_1}{t_1} \frac{dt_2}{t_2} dx_1 dx_2 
\]
\[
\leq C |R|^{1/2} \gamma_1(R)^{-2\alpha_1-n} \|a_R\|_{L^2(\mathbb{R}^{n+m})} \leq C |R|^{1/2} \gamma_1(R)^{-2\alpha_1-n} \|a_R\|_{L^2(\mathbb{R}^{n+m})}. 
\]

Choosing \(2\alpha_1 > n + m\) and \(2\alpha_2 > m\), we get

\[
I_{2122} \leq C \int_{(100)^c \times (100)^c} \int J \int_{|y_2-z_2| \approx 2^{k_2} t(J)} \left( (\int_{y_1}^t + \int_t^\infty) \frac{1}{t_2} e^{-\frac{\ell(J)^2}{2t_2}} dt_2 \right) \int J \int_{|y_2-z_2| \approx 2^{k_2} t(J)} e^{-\frac{|y_2-z_2|^2}{2t_2}} dy_2 |a_R(y_1, z_2)| dz_2 dy_1 \frac{dt_1}{t_1} \frac{dt_2}{t_2} dx_1 dx_2 
\]
Let \( I_{22} \). From the heat kernel estimate and the support condition of \( a_R \), it is clear that

\[
I_{22} \leq C \int_{(100)^c \times (100)^c} \left| \frac{1}{2\sqrt{\pi}} \int_0^\infty \ell(I)^2 \right| dx_1 dx_2
\]

\[
\times \int_{\mathbb{R}^{n+m}} \langle x_1 - y_1, x_2 - y_2 \rangle \nabla^{(2)} \Delta (2)^{-1/2} a_R(y_1, y_2) dy_1 dy_2 \frac{dt_1}{t_1} \vert \ dx_1 dx_2
\]

\[
\leq C \int_{(100)^c \times (100)^c} \left| \frac{1}{2\sqrt{\pi}} \int_0^\infty \ell(I)^2 \right| dx_1 dx_2
\]

\[
\times \int_{I} \int_{|x_1 - y_1| \geq 2y_2 \ell(J)} \langle x_1 - y_1, x_2 - y_2 \rangle \nabla^{(2)} \Delta (2)^{-1/2} a_R(y_1, y_2) dy_1 dy_2 \frac{dt_1}{t_1} \vert \ dx_1 dx_2
\]

\[
+ C \int_{(100)^c \times (100)^c} \left| \frac{1}{2\sqrt{\pi}} \int_0^\infty \ell(I)^2 \right| \sum_{k_2 = 0}^\infty \int_{I} \int_{|x_2 - y_2| \geq 2y_2 \ell(J)} \langle x_1 - y_1, x_2 - y_2 \rangle \nabla^{(2)} \Delta (2)^{-1/2} a_R(y_1, y_2) dy_1 dy_2 \frac{dt_1}{t_1} \vert \ dx_1 dx_2
\]

\[=: I_{221} + I_{222}. \]

Let \( a_R = (\Delta (1)^M \otimes_2 \mathbb{I}_2) a_{R,2} \) where \( a_{R,2} = (\mathbb{I}_1 \otimes_2 \Delta^{(2)} M) b_R \). For the term \( I_{221} \), Hölder’s inequality gives

\[
I_{221} = C \int_{(100)^c \times (100)^c} \left| \frac{1}{2\sqrt{\pi}} \int_0^\infty \ell(I)^2 \right| \int_{J} \int_{|x_1 - y_1| \geq 2y_2 \ell(J)} \langle x_1 - y_1, x_2 - y_2 \rangle \nabla^{(2)} \Delta (2)^{-1/2} a_{R,2}(y_1, y_2) dy_1 dy_2 \frac{dt_1}{t_1} \vert \ dx_1 dx_2
\]

\[
\leq C \int_{(100)^c \times (100)^c} \left| \frac{1}{2\sqrt{\pi}} \int_0^\infty \ell(I)^2 \right| \int_{J} \int_{|x_1 - y_1| \geq 2y_2 \ell(J)} (\int_{x_2 - y_2} |t_1 \Delta (1)^M q_{t_1}^{(1)}(x_1 - y_1, x_2 - y_2)|^2 dy_2)^{1/2} \ dx_1 dx_2
\]
\[
\left( \int_J |\nabla^2 \triangle^2|^{-1/2} a_R(y_1, y_2)^2 dy_2 \right)^{1/2} \frac{dt_1}{t_1^{1+M}} dx_1 dx_2 \\
\leq C|J|^{1/2} \int_{(100)^c \times (100)^c} \int_{t(J)^2} \int_{|x_2-y_J| \approx 2^j t(J)} \frac{1}{n+1} \frac{e^{-(|x_1-x_J| - |x_2-y_J|)^2}}{t_1} dt_1 dx_1 dx_2,
\]

Hence,

\[
I_{221} \leq C|R|^{1/2} \sum_{j_1=j=6}^{\infty} \int_{|x_1-x_J| \approx 2^j t(J)} \int_{|x_2-y_J| \approx 2^j t(J)} \int_{\ell(J)^2} \frac{1}{n+1} \frac{e^{-(|x_1-x_J| - |x_2-y_J|)^2}}{t_1} t_1^{\alpha_1 + \alpha_2 - \frac{n+m}{2} - 1 - M} dt_1 dx_1 dx_2
\]

where \(2\alpha_1 > n, 2\alpha_2 > m\) and \(2\alpha_1 + 2\alpha_2 < n + m + 2M\).

Let \(a_R = (\triangle^2)^M \otimes b_R\) where \(a_R = (\mathbb{I}_1 \otimes \triangle^2)^M b_R\). To estimate \(I_{222}\), we have

\[
I_{222} \leq C \int_{(100)^c \times (100)^c} \int_{t(J)^2} \int_{0}^{\infty} \sum_{k_2=0}^{\infty} \int_{|y_2-z_2| \approx 2^k t(J)} (t_1 \triangle^2)^M q^2_{t_1}(x_1 - y_1, x_2 - y_2) q^2_{t_1}(y_2 - z_2) a_R(y_1, z_2) d\tau_1 dy_1 dy_2 \frac{dt_1}{t_1^{1+M}} \frac{dt_2}{t_2} dx_1 dx_2,
\]

where \(2\alpha_1 > n, 2\alpha_2 > m\) and \(2\alpha_1 + 2\alpha_2 < n + m + 2M\).
Characterizations of flag Hardy spaces

By the heat kernel estimate, we choose $2\alpha_1 - n > 0, 2\alpha_2 - m > 0$ and $\alpha_1 + \alpha_2 < \frac{n}{2} + M$ to obtain

$$I_{2221} \leq C \int_{(100)^c \times (100)^c} \int_{\ell(I)^2} 1 \frac{1}{t_1 + m} \frac{n}{t_1^2} e^{-\frac{|x_1-y_1|^2}{t_1}} \int_0^{t_1} \frac{1}{t_2} e^{-\frac{\ell(J)^2}{2t_2}} \frac{dt_2}{t_2}$$

$$\int_I \sum_{k_2=0}^{\infty} \int_{|y_2-z_2| \approx 2^{k_2} \ell(J)} e^{-\frac{|y_2-z_2|^2}{2t_2}} e^{-\frac{|y_2-z_2|^2}{2t_2}} dy_2 |a_{R,2}(y_1, z_2)| dz_2 dy_1 \frac{dt_1}{t_1 + M} dx_1 dx_2$$

$$\leq C \sum_{j_1=j} \int_{|x_1-x_j| \approx 2^{j_1} \ell(I)} \int_{j_2=6}^{\infty} \int_{|x_2-y_j| \approx 2^{j_2} \ell(J)} \int_{\ell(I)^2} \int_I \frac{1}{t_1 + M} \frac{dt_1}{t_1 + M} dx_1 dx_2$$

$$\int_I \sum_{j_1=j} \int_{|x_1-x_j| \approx 2^{j_1} \ell(I)} \int_{j_2=6}^{\infty} \int_{|x_2-y_j| \approx 2^{j_2} \ell(J)} \int_{\ell(I)^2} \int_I \frac{1}{t_1 + M} \frac{dt_1}{t_1 + M} dx_1 dx_2$$

Choosing $n < 2\alpha_1 < n + m + 2M$ and $2\alpha_2 > m$, we get

$$I_{2222} \leq C \int_{(100)^c \times (100)^c} \int_{\ell(I)^2} 1 \frac{1}{t_1 + m} \frac{n}{t_1^2} e^{-\frac{|x_1-y_1|^2}{t_1}} \int_0^{t_1} \frac{1}{t_2} e^{-\frac{\ell(J)^2}{2t_2}} \frac{dt_2}{t_2}$$

$$\int_I \sum_{k_2=0}^{\infty} \int_{|y_2-z_2| \approx 2^{k_2} \ell(J)} e^{-\frac{|y_2-z_2|^2}{2t_2}} e^{-\frac{|y_2-z_2|^2}{2t_2}} dy_2 |a_{R,2}(y_1, z_2)| dz_2 dy_1 \frac{dt_1}{t_1 + M} \frac{dt_2}{t_2} dx_1 dx_2$$

$$\leq C \sum_{j_1=j} \int_{|x_1-x_j| \approx 2^{j_1} \ell(I)} \int_{j_2=6}^{\infty} \int_{|x_2-y_j| \approx 2^{j_2} \ell(J)} \int_{\ell(I)^2} \int_I \frac{1}{t_1 + M} \frac{dt_1}{t_1 + M} \frac{dt_2}{t_2} \frac{e^{-\frac{\ell(J)^2}{2t_2}}}{t_2}$$
Combining the above estimates, there exists a positive constant \( \delta_1 \) such that
\[
\int_{(100)^c \times \mathbb{R}^n} |T(a_R)(x_1, x_2)| dx_1 dx_2 \\
\leq C |R|^{1/2} \gamma_1(R)^{-\delta_1} \left( \|a_R\|_{L^2(\mathbb{R}^{n+1})} + \ell(I)^{-2M} \|a_R, 2\|_{L^2(\mathbb{R}^{n+1})} \right) \\
= C |R|^{1/2} \gamma_1(R)^{-\delta_1} \left( \|a_R\|_{L^2(\mathbb{R}^{n+1})} + \ell(I)^{-2M} \|(\mathbb{1}_1 \otimes_2 \Delta(2)^M) b_R\|_{L^2(\mathbb{R}^{n+1})} \right).
\]

Using Hölder’s inequality, Journé’s covering lemma and the properties of flag atoms, we have
\[
I := \sum_{R \in m(\Omega)} \int_{(100)^c \times \mathbb{R}^n} |T(a_R)(x_1, x_2)| dx_1 dx_2 \\
\leq C \left( \sum_{R \in m(\Omega)} |R|^{1/2} \gamma_1(R)^{-\delta_1} \ell(I)^{-2M} \ell(J)^{-2M} \right)^{1/2} \\
\times \left( \|((\ell(I)^2 \Delta(1)^M) \otimes_2 (\ell(J)^2 \Delta(2)^M) b_R\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)} + \|(\mathbb{1}_1 \otimes_2 \Delta(2)^M) b_R\|_{L^2(\mathbb{R}^{n+1})} \right) \\
\leq C \left( \sum_{R \in m(\Omega)} |R|^{1/2} \gamma_1(R)^{-2\delta_1} \right)^{1/2} \left( \sum_{R \in m(\Omega)} \ell(I)^{-4M} \ell(J)^{-4M} \right)^{1/2} \\
\times \left( \|((\ell(I)^2 \Delta(1)^M) \otimes_2 (\ell(J)^2 \Delta(2)^M) b_R\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)} + \|(\mathbb{1}_1 \otimes_2 \Delta(2)^M) b_R\|_{L^2(\mathbb{R}^{n+1})} \right)^{1/2} \\
\leq C |\Omega|^{\frac{2}{3}} |\Omega|^{-\frac{2}{3}} \leq C.
\]

For term II, we observe that
\[
\int_{\mathbb{R}^n \times (100S)^c} |T(a_R)(x_1, x_2)| dx_1 dx_2 \\
= \left( \int_{100I \times (100S)^c} + \int_{(100)^c \times (100S)^c} \right) |T(a_R)(x_1, x_2)| dx_1 dx_2 \\
= II_1 + II_2.
\]

Let us first estimate the term II_1. Then
\[
II_1 \leq \int_{100I \times (100S)^c} \left| \nabla^{(1)} \Delta^{(1)} \otimes_2 \nabla^{(2)} \Delta^{(2)} a_R(x_1, x_2) \right| dx_1 dx_2
\]
where the last inequality follows from the fact that

Characterizations of flag Hardy spaces

For the term of II

By the \(L^2(\mathbb{R}^{n+m})\) boundedness of \(\nabla^{(1)}(1)\),

\[
\Pi_{111} \leq C|R|^{1/2} \sum_{j=0}^{\infty} \int_{|x_j-y_j| \approx 2^{j} \epsilon} \left( \int_{|x_j-y_j| \approx 2^{j} \epsilon} a_R^{(1)} dx \right) a_R^{(1)} L^2(\mathbb{R}^{n+m})
\]

where the last inequality follows from the fact that

\[
2^j \approx \frac{j}{\ell(J)}.
\]

For the term of II_{12}, the heat kernel estimate gives

\[
\Pi_{112} \leq C \sum_{j=0}^{\infty} \int_{|x_j-y_j| \approx 2^{j} \epsilon} \left( \int_{|x_j-y_j| \approx 2^{j} \epsilon} a_R^{(1)} dx \right) a_R^{(1)} L^2(\mathbb{R}^{n+m})
\]
\[
\left( \int_0^{t_1} + \int_0^{t_2} \right) \int \int_1^{t_2} \frac{n+m}{2} e^{-\frac{|y-x|^2}{t_1}} |a_R(z_1, z_2)| dz_1 dz_2 \frac{dt_1}{t_1} \frac{dt_2}{t_2}
\]

\[=: \Pi_{1121} + \Pi_{1122}.\]

We do the integral for the variable \(y\) to get

\[
\Pi_{1121} \leq C|I| \sum_{j_2 = j_0}^{\infty} \int_{|x-y| = 2\ell(J)} \int_0^{\ell(J)+2} e^{-\frac{t}{2t_2}} \int_0^{t_2} \frac{n+m}{2} e^{-\frac{|y-x|^2}{t_1}} |a_R(z_1, z_2)| dz_1 dz_2 dt_1 dt_2 dx_2
\]

\[\leq C|I^{\gamma_2}(R)| \sum_{j_2 = j_0}^{\infty} \int_{|x-y| = 2\ell(J)} \int_0^{\ell(J)+2} e^{-\frac{t}{2t_2}} \int_0^{t_2} \frac{n+m}{2} e^{-\frac{|y-x|^2}{t_1}} |a_R(z_1, z_2)| dz_1 dz_2 dt_1 dt_2 dx_2
\]

where \(2\alpha_1 > n + m\) and \(2\alpha_2 > m\). Similarly,

\[
\Pi_{1122} \leq C|I^{\gamma_2}(R)| \sum_{j_2 = j_0}^{\infty} \int_{|x-y| = 2\ell(J)} \int_0^{\ell(J)+2} e^{-\frac{t}{2t_2}} \int_0^{t_2} \frac{n+m}{2} e^{-\frac{|y-x|^2}{t_1}} |a_R(z_1, z_2)| dz_1 dz_2 dt_1 dt_2 dx_2
\]

\[\leq C|I^{\gamma_2}(R)| \sum_{j_2 = j_0}^{\infty} \int_{|x-y| = 2\ell(J)} \int_0^{\ell(J)+2} e^{-\frac{t}{2t_2}} \int_0^{t_2} \frac{n+m}{2} e^{-\frac{|y-x|^2}{t_1}} |a_R(z_1, z_2)| dz_1 dz_2 dt_1 dt_2 dx_2
\]

where \(m < 2\alpha_1 < n + m\).

Let \(a_R = (I \otimes \Delta^{(2)} M) a_{R,1}\), where \(a_{R,1} = (\Delta^{(1)} M) \otimes \Pi_2 b_R\) We now consider \(\Pi_{12}\) and write

\[
\Pi_{12} \leq \int_{100 \times 100} \left| \frac{1}{2\pi} \int_{\ell'J} \nabla^{(2)}(t_2 \Delta^{(2)}) M e^{-t_2 \Delta^{(2)}} \left( \nabla^{(1)} \Delta^{(1)} \right)^{-1/2} a_{R,1}(x_1, x_2) \right| dt_2 \frac{dt_2}{t_2^{m+1}} dx_1 dx_2
\]

\[\leq C \sum_{j_2 = j_0}^{\infty} \int_{100 \times 100} \left| \frac{1}{2\pi} \int_{\ell'J} \nabla^{(2)}(t_2 \Delta^{(2)}) M e^{-t_2 \Delta^{(2)}} \left( \nabla^{(1)} \Delta^{(1)} \right)^{-1/2} a_{R,1}(x_1, x_2) \right| dt_2 \frac{dt_2}{t_2^{m+1}} dx_1 dx_2
\]

\[=: \Pi_{121} + \Pi_{122}.\]
By the $L^2(\mathbb{R}^{n+m})$ boundedness of $\nabla^{(1)} \triangle^{(1)} - 1/2$,

$$\Pi_{121} \leq C|R|^{1/2} \sum_{j=3}^\infty \int_{|x_2 - y_j| \approx 2^{2j} \ell(J)} \int_{\ell(J)^2} \int_{|y_2 - y_j| \approx 2^{2j} \ell(J)} t_2^{-\frac{m}{2} - M} e^{-\frac{|x_2 - y_j|^2}{t_2^2}} \frac{dt_2}{t_2} dx_2 \|a_{R,1}\|_{L^2(\mathbb{R}^{n+m})}$$

$$\leq C|R|^{1/2} \sum_{j=3}^\infty (2^{2j+1} \ell(J))^{m-2\alpha_2} \int_{\ell(J)^2} \int_{|y_2 - y_j| \approx 2^{2j} \ell(J)} \alpha_2 - \frac{m}{2} - M \frac{dt_2}{t_2} \|a_{R,1}\|_{L^2(\mathbb{R}^{n+m})}$$

$$\leq C|R|^{1/2} \gamma_2(R)^{-2M} \|a_{R,1}\|_{L^2(\mathbb{R}^{n+m})},$$

where $0 < 2\alpha_2 - m < M$.

For the term of $\Pi_{122}$, the heat kernel estimate gives

$$\Pi_{122} \leq C \sum_{j=3}^\infty \sum_{k=0}^\infty \int_{|x_2 - y_j| \approx 2^{2j} \ell(J)} \int_{\ell(J)^2} \int_{|y_2 - y_j| \approx 2^{2j} \ell(J)} t_2^{-\frac{m}{2} - M} e^{-\frac{|x_2 - y_j|^2}{t_2^2}}$$

$$\int_0^\infty \int_1 \int_{J_1} \frac{n+m}{2} e^{-\frac{|z_1 - y_j|^2}{t_2^2}} |a_{R,1}(z_1, z_2)| dz_1 dz_2 dt_1 \frac{dy_2}{t_2} dx_2$$

$$\leq C |\ell(J)| \sum_{j=3}^\infty \sum_{k=0}^\infty \int_{|x_2 - y_j| \approx 2^{2j} \ell(J)} \int_{\ell(J)^2} \int_{|y_2 - y_j| \approx 2^{2j} \ell(J)} t_2^{-\frac{m}{2} - M} e^{-\frac{|x_2 - y_j|^2}{t_2^2}}$$

$$\left( \int_0^{t_2} + \int_{t_2}^\infty \right) \int_1 \int_{J_1} \frac{n+m}{2} e^{-\frac{|z_1 - y_j|^2}{t_2^2}} |a_{R,1}(z_1, z_2)| dz_1 dz_2 dt_1 \frac{dy_2}{t_2} dx_2$$

$$=: \Pi_{1221} + \Pi_{1222}.$$
where $m < 2\alpha_1 < n + m$.

Combining the above estimates, there exists a positive constant $\delta_2$ such that

$$
\int_{\mathbb{R}^n \times (100S)^c} |T(a_R)(x_1, x_2)| \, dx_1 \, dx_2 
\leq C |R|^{1/2} |\gamma_2(R) - (2\alpha_1 - m)\ell(\mathcal{J}) - 2M| a_{R,1} \, ||_{L^2(\mathbb{R}^{n+m})}
$$

$$
\leq C |R|^{1/2} |\gamma_2(R) - (2\alpha_1 - m)\ell(\mathcal{J}) - 2M| a_{R,1} \, ||_{L^2(\mathbb{R}^{n+m})}
$$

$$
\leq C |R|^{1/2} |\gamma_2(R) - (2\alpha_1 - m)\ell(\mathcal{J}) - 2M| a_{R,1} \, ||_{L^2(\mathbb{R}^{n+m})},
$$

Using the Hölder’s inequality, Journé’s covering lemma and the properties of flag atoms, we have

$$
\Pi := \sum_{R \in \mathcal{M}(\Omega)} \int_{\mathbb{R}^n \times (100S)^c} \, dx_1 \, dx_2 
\leq C \sum_{R \in \mathcal{M}(\Omega)} |R|^{1/2} |\gamma_2(R) - (2\alpha_1 - m)\ell(I) - 2M| \ell(I) - 2M
$$

$$
\times \left( \left| (\ell(I)^2 \Delta^{(1)})^M \otimes_2 (\ell(I)^2 \Delta^{(2)})^M \right| a_R \right|_{L^2(\mathbb{R}^{n+m})} + \left| (\Delta^{(1)})^M \otimes_2 \mathbb{I}_2 \right| a_R \right|_{L^2(\mathbb{R}^{n+m})}
$$

$$
\leq C \left( \sum_{R \in \mathcal{M}(\Omega)} |R|^{\delta_2} |\gamma_2(R) - (2\alpha_1 - m)\ell(I) - 2M| \ell(I) - 2M
$$

$$
\times \left( \left| (\ell(I)^2 \Delta^{(1)})^M \otimes_2 (\ell(I)^2 \Delta^{(2)})^M \right| a_R \right|_{L^2(\mathbb{R}^{n+m})} + \left| (\Delta^{(1)})^M \otimes_2 \mathbb{I}_2 \right| a_R \right|_{L^2(\mathbb{R}^{n+m})} \right)^{1/2}
$$

$$
\leq C |\Omega|^{\frac{1}{2}} |\Omega|^{-\frac{1}{2}} \leq C.
$$

Therefore,

$$
\int_{(\cup R)^c} |T(a)(x)| \, dx 
\leq \sum_{R \in \mathcal{M}(\Omega)} \int_{(100)^c \times \mathbb{R}^n} |T(a_R)(x)| \, dx + \sum_{R \in \mathcal{M}(\Omega)} \int_{\mathbb{R}^n \times (100S)^c} |T(a_R)(x)| \, dx 
\leq C.
$$

The inequality (6.4) is done and the proof is completed. $\Box$

Based on the result above, we already showed that $\sum_{j=1}^{n+m} \sum_{k=1}^{m} \|R_{j,k}(f)\|_1 + \|f\|_1 \leq \|S_F(f)\|_1$. Now we conclude, based on all estimates provided from Section 2 to Section 4,
that
\[ \|S_F(f)\|_1 \lesssim \|S_F(u)\|_1 \lesssim \|u^*\|_1 \lesssim \|M_\Phi^*(f)\|_1 \lesssim \|u^*\|_1 \]
\[ \lesssim \|u^+\|_1 \lesssim \|M_\Phi^+(f)\|_1 \lesssim \|u^+\|_1 \]
\[ \lesssim \sum_{j=1}^{n+m} \sum_{k=1}^m \|R_{j,k}(f)\|_1 + \|f\|_1 \]
\[ \lesssim \|S_F(f)\|_1. \]

**Acknowledgement:** BDW’s research supported in part by National Science Foundation grants DMS # 1560955 and DMS # 1800057. JL supported by ARC DP 160100153 and Macquarie University New Staff Grant. MYL is supported by Ministry of Science and Technology, R.O.C. under Grant #MOST 106-2115-M-008-003-MY2 as well as supported by National Center for Theoretical Sciences of Taiwan. This paper started in 2007 when YSH visited JL at Sun Yat-Sen University, and subsequent work was done when MYL visited JL at Sun Yat-Sen University in 2012 and 2013, and JL visited MYL at National Central University, Taiwan in 2014 and when JL visited BDW at Washington University in STL in 2015.
References

[1] S-Y. A. Chang and R. Fefferman, Some recent developments in Fourier analysis and $H^p$ theory on product domains, Bull. A. M. S. 12 (1985), 1-43.

[2] S-Y. A. Chang and R. Fefferman, The Calderón-Zygmund decomposition on product domains, Amer. J. Math. 104 (1982), 455-468.

[3] S-Y. A. Chang and R. Fefferman, A continuous version of duality of $H^1$ with $BMO$ on the bidisc, Annals of math. 112 (1980), 179-201.

[4] T. Coulhon and A. Sikora, Gaussian heat kernel upper bounds via Phragmén-Lindelöf theorem, Proc. Lond. Math. 96 (2008), 507–544.

[5] X. Duong, J. Li, Y. Ou, J. Pipher and B. D. Wick, Commutators of multi-parameter flag singular integrals and applications, to appear in Analysis and PDE.

[6] C. Fefferman and E. M. Stein, $H^p$ spaces of several variables, Acta Math. 129 (1972), 137-194.

[7] R. Fefferman, Multi-parameter Fourier analysis, Study 112, Beijing Lectures in Harmonic Analysis, Edited by E. M. Stein, 47-130. Annals of Mathematics Studies Princeton University Press.

[8] R. Fefferman, Harmonic Analysis on product spaces, Ann. of Math. 126 (1987), 109-130.

[9] R. Fefferman, Multiparameter Calderón-Zygmund theory, Harmonic analysis and partial differential equations (Chicago, IL, 1996), 207-221, Chicago Lectures in Math., Univ. Chicago Press, Chicago, IL, 1999.

[10] R. Fefferman and E. M. Stein, Singular integrals on product spaces, Adv. Math. 45 (1982), 117-143.

[11] M. Frazier and B. Jawerth, A discrete transform and decomposition of distribution, J. Func. Anal. 93 (1990), 34–170.

[12] P. Glowacki, The Melin calculus for general homogeneous groups, Ark. Mat. 45 (2007), 31–48.

[13] P. Glowacki, Composition and $L^2$-boundedness of flag kernels, Colloq. Math. 118 (2010), 581–585.

[14] P. Glowacki, Correction to “Composition and $L^2$-boundedness of flag kernels”, Colloq. Math. 120 (2010), 331.
[15] R. Gundy and E. M. Stein, $H^p$ theory for the polydisk, Proc. Nat. Acad. Sci., 76 (1979),

[16] Y. Han, Calderón-type reproducing formula and the $Tb$ theorem, Rev. Mat. Iberoamericana 10 (1994), 51-91.

[17] Y. Han, Plancherel-Pólya type inequality on space of homogeneous type and its applications, Proc. Amer. Math. Soc. 126 (1998), 3315-3327.

[18] Y. Han, M.-Y. Lee, C.-C. Lin and Y.-C. Lin, Calderón-Zygmund operators on product Hardy spaces, J. Funct. Anal. 258 (2010), 2834–2861.

[19] Y. Han, Y.S. Han, J. Li and C. Tan, Marcinkiewicz multipliers and Lipschitz spaces on Heisenberg groups, to appear in Canadian Journal of Mathematics.

[20] Y. Han, G. Lu and E. Sawyer, Flag Hardy spaces and Marcinkiewicz multipliers on the Heisenberg group, Anal. PDE 7 (2014), 1465–1534.

[21] S. Hofmann, G. Z. Lu, D. Mitrea, M. Mitrea and L. X. Yan, Hardy spaces associated to nonnegative self-adjoint operators satisfying Davies-Gaffney estimates, Memoirs Amer. Math. Soc. 214 (2011), no. 1007.

[22] J. L. Journé, Calderón-Zygmund operators on product spaces, Rev. Mat. Iberoamericana 1 (1985), 55-92.

[23] K. Merryfield, $H^p$ spaces on poly-half spaces, Ph. D. Thesis, Univ. of Chicago, 1980.

[24] D. Müller, F. Ricci, and E. M. Stein, Marcinkiewicz multipliers and multi-parameter structure on Heisenberg(-type) groups, I, Invent. math. 119 (1995), 119-233.

[25] D. Müller, F. Ricci, and E. M. Stein, Marcinkiewicz multipliers and multi-parameter structure on Heisenberg(-type) groups, II, Math. Z. 221 (1996), 267-291.

[26] A. Nagel and E. M. Stein, On the product theory of singular integrals, Rev. Mat. Iberoamericana 20 (2004), 531-561.

[27] A. Nagel, F. Ricci, and E. M. Stein, Singular integrals with flag kernels and analysis on quadratic CR manifolds, J. Func. Anal. 181 (2001), 29-118.

[28] A. Nagel, F. Ricci, E. M. Stein, and S. Wainger, Singular integrals with flag kernels on homogeneous groups: I, Rev. Mat. Iberoamericana 28 (2012), 631-722.

[29] A. Nagel, F. Ricci, E. M. Stein, and S. Wainger, Algebras of singular integral operators with kernels controlled by multiple norms, arxiv:1511.05702.
[30] J. Pipher, Journé’s covering lemma and its extension to higher dimensions, Duke Mathematical Journal, 53 (1986), 683-690.

[31] A. Sikora, Riesz transform, Gaussian bounds and the method of wave equation, Math. Z. 247 (2004), 643–662.

[32] E.M. Stein, Harmonic analysis: Real variable methods, orthogonality and oscillatory integrals, Princeton Univ. Press, Princeton, NJ, (1993).

[33] E.M. Stein and G. Weiss, On the theory of harmonic functions of several variables, I: The theory of $H^p$ spaces, Acta Math. 103: 25-62.

[34] K. Yosida, Functional Analysis (fifth edition), Spring-Verlag, Berlin, 1978.

Department of Mathematics, Auburn University, Auburn, AL 36849-5310, U.S.A.

E-mail: hanyong@auburn.edu

Department of Mathematics, National Central University, Chung-Li 320, Taiwan.

&

National Center for Theoretical Sciences, 1 Roosevelt Road, Sec. 4, National Taiwan University, Taipei 106, Taiwan

E-mail: mylee@math.ncu.edu.tw

Department of Mathematics, Macquarie University, NSW, 2109, Australia

E-mail: ji.li@mq.edu.au

Department of Mathematics, Washington University–St. Louis, St. Louis, MO 63130-4899, U.S.A.

E-mail: wick@math.wustl.edu