On Fefferman–Stein type inequality on Shilov boundaries and applications

Ji Li

Abstract: In this paper, we establish the Fefferman–Stein type inequality for area integral and non-tangential maximal function on the Shilov boundary studied by Nagel and Stein [30]. The technique here is inspired by Fefferman–Stein [14] and Merryfield [22] but we bypass the use of Fourier or group structure as these were not available on the polynomial domains of finite type. Direct applications include the maximal function characterisation of product Hardy space and the weak type endpoint estimate for product Calderón–Zygmund operators (such as the Cauchy–Szegő projection) on the Shilov boundary.

1 Introduction

In this paper, we establish the Fefferman–Stein type good-λ inequality for area integral and non-tangential maximal function on a typical product space: the Shilov boundary of tensor product domains studied by Nagel and Stein [30].

Let $u(x,t)$ be a harmonic function in $\mathbb{R}^n \times (0,\infty)$. The non-tangential maximal function $u^*(x) = \sup_{|x-y|<t} |u(y,t)|$ and area integral $S(u)(x)^2 = \int_{|x-y|<t} |\nabla u(y,t)|^2 t^{1-n} dydt$ are two fundamental tools in the theory of singular integrals and the related function spaces. Fefferman and Stein [14] first showed that $\|u^*\|_{L^p(\mathbb{R}^n)} \approx \|S(u)\|_{L^p(\mathbb{R}^n)}$, $0 < p \le 1$, when $u(x,t) \to 0$ as $t \to \infty$ (for $1 < p < \infty$ this was known in [35]). The key objects in their proof are the following inequality ([14, (7.2)])

$$|\{x \in \mathbb{R}^n : S(u)(x) > \lambda\}| \lesssim |\{x \in \mathbb{R}^n : u^*(x) > \lambda\}| + \frac{1}{\lambda^2} \int_0^\lambda s|\{x \in \mathbb{R}^n : u^*(x) > s\}|ds \quad (1.1)$$

and the other inequality of the same type but with u^* and $S(u)$ interchanged (here the implicit constant is independent of λ). When u is given by the Poisson integral of f, a different proof of the L^p norm, $0 < p \le 1$, of u^* and $S(u)$ was given via atomic decomposition. Later, Gundy and Stein [16] established this result in the bi-disc for characterising the product Hardy space H^p, $0 < p \le 1$, via using holomorphic function and martingales. The key step mirrors (1.1), applied to the area integral and non-tangential maximal function in the bi-disc. It is natural to explore whether this is also true on the product spaces $\mathbb{R}^n \times \mathbb{R}^m$, noting that in the higher dimensional space, the analyticity is not available. Unlike the one-parameter setting, the equivalence $\|u^*\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)} \approx \|S(u)\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}$ (for $0 < p \le 1$) does not follow from atomic

Mathematics Subject Classification (2020): Primary 32A50,32T25; Secondary 43A85, 32W30
Key words: Shilov boundary, finite type domains, good-λ inequality.

decomposition directly. It is still not clear whether one can construct the atomic decomposition of f directly from the assumption $\|u^*\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)} < \infty$ (u is given by the double Poisson integral of f).

To overcome this, Merryfield [22] provided a new proof of (1.1), which bypassed the use of surface approximation ([14]) or analyticity ([16]), and hence the inequality $\|S(u)\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)} \lesssim \|u^*\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)}$ holds (the reverse can be done by atomic decomposition directly). However, one key ingredient in [22] is the Cauchy–Riemann equation for constructing a test function (for Littlewood–Paley estimate) from the given test function (for maximal function). Thus, although we have studied the Hardy spaces via area function and atomic decomposition in the general product spaces of homogeneous type (see for example [19, 6, 20]), the maximal function characterisation was only known in a few cases via establishing an analogy of (1.1): (1) the Muckenhoupt–Stein Bessel operator setting ([13]), where we exploited the Cauchy–Riemann type equation associated with the Bessel operators; (2) the multi-parameter flag setting of Nagel–Ricci–Stein [23]. In [17] we extended [22] to the flag Euclidean setting, (3) the product stratified Lie groups [10] and the flag setting of Heisenberg group [5], where we used the group structure and explicit pointwise upper and lower bound of Poisson kernel.

Besides the bi-disc and product Lie groups, one fundamental model domain is the Shilov boundary $\widetilde{M} = M_1 \times M_2$ studied by Nagel and Stein [30] which links to the $\bar\partial$-Neumann problem on decoupled boundaries (to ease the burden of notational complexity we consider the tensor product of two domains). Here each M_j is an unbounded polynomial domain of finite type m_j defined as follows.

Let M (for simplicity we first drop the subscript j) be given as

$$M := \big\{(z,w) \in \mathbb{C}^2 : \ \mathrm{Im}(w) = \mathcal{P}(z)\big\}, \tag{1.2}$$

where $\mathcal{P}(z)$ is a real, subharmonic, non-harmonic polynomial of degree m. We note that ([30]) M can be identified with $\mathbb{C}\times\mathbb{R} = \{(z,\mathfrak{t}) : z \in \mathbb{C}, \mathfrak{t} \in \mathbb{R}\}$. The basic (0,1) Levi vector field is then $\bar Z = \frac{\partial}{\partial \bar z} - i\frac{\partial \mathcal{P}}{\partial \bar z}\frac{\partial}{\partial \mathfrak{t}}$, where $i^2 = -1$, and we write $\bar Z = X_1 + iX_2$. The real vector fields $\{X_1, X_2\}$ and their commutators of orders up to m span the tangent space at each point. Associated with the domain M is the natural Carnot–Carathéodory metric $d(\mathbf{x},\mathbf{y})$ on M (for every $\mathbf{x},\mathbf{y} \in M$) and the measure μ of the nonisotropic ball $B(\mathbf{x}, r) = \{\mathbf{y} \in M : d(\mathbf{x},\mathbf{y}) < r\}$, which made (M, d, μ) as a space of homogeneous type in the sense of Coifman and Weiss [9]. Also denote $V_r(\mathbf{x}) = \mu(B(\mathbf{x}, r))$. Details are given in Section 2.

The typical example that we have in mind (regarding further development and applications [29, 24, 25]) is $\mathcal{P}(z) = \frac{1}{2k}|z|^{2k}$ for $z \in \mathbb{C}$ and k to be positive integers. When $k = 1$, M is the boundary of the Siegel domain in $\mathbb{C}^2$ and it is CR-diffeomorphic to the first Heisenberg group, from which it inherits the group structure. However, when $k \geq 2$, M does not have a group structure. Thus, we are more interested in the cases $k \geq 2$. The $\bar\partial$-problem, sub-Laplacian, Kohn-Laplacian and the related geometry, singular integrals have been intensively studied, see for example [7, 3, 12, 27, 30, 26, 34]. Very recent progresses on such model domains M ($k \geq 2$) include the explicit pointwise upper and lower bound for the Cauchy–Szegő kernel and its applications to boundedness and compactness of commutators ([2]), and the construction that for each $k \geq 2$, M can be lifted to a Lie group G ([1]), which provided an explicit and optimal lifting Lie algebra comparing to the result of Rothschild and Stein [33].

We now state our result in detail. Let $\widetilde{M} = M_1 \times M_2$, where each M_j is the example domain as above with $\mathcal{P}_j(z) = \frac{1}{2k_j}|z|^{2k_j}$, $j = 1, 2$. Let d_j be the Carnot–Carathéodory metric on M_j and μ_j be the corresponding measure. For the sake of simplicity, throughout the paper we denote $d\mu_j(\mathbf{x}_j) = d\mathbf{x}_j$ and $|A|$ represents the measure of the set A. Let $\mathcal{L}_j$ be the sub-Laplacian of M_j and $P^{[j]}_{t_j}$ be the Poisson semigroup $e^{-t_j\sqrt{\mathcal{L}_j}}$, $j = 1, 2$. Consider the non-tangential maximal function

$$\mathcal{N}^\beta_P(f)(\mathbf{x}_1, \mathbf{x}_2) := \sup_{\substack{(\mathbf{y}_1,\mathbf{y}_2)\in\widetilde{M}, t_1>0, t_2>0,\\ d_1(\mathbf{x}_1,\mathbf{y}_1)<\beta t_1,\\ d_2(\mathbf{x}_2,\mathbf{y}_2)<\beta t_2}} |P^{[1]}_{t_1} P^{[2]}_{t_2}(f)(\mathbf{y}_1, \mathbf{y}_2)|$$

with the constant $\beta > 0$. Consider also the Littlewood–Paley area function. Let

$$\nabla_{t_1,M_1} := (\partial_{t_1}, X_{1,1}, X_{1,2}), \quad \nabla_{t_2,M_2} := (\partial_{t_2}, X_{2,1}, X_{2,2}).$$

Then for any fixed $\beta \in (0, \infty)$, the Littlewood–Paley area function $S^\beta_P(f)(\mathbf{x}_1, \mathbf{x}_2)$ is defined as

$$S^\beta_P(f)(\mathbf{x}_1, \mathbf{x}_2) := \left(\iint_{\Gamma^\beta(\mathbf{x}_1,\mathbf{x}_2)} |t_1\nabla_{t_1,M_1} P^{[1]}_{t_1}\ t_2\nabla_{t_2,M_2} P^{[2]}_{t_2}(f)(\mathbf{y}_1, \mathbf{y}_2)|^2 \frac{d\mathbf{y}_1 d\mathbf{y}_2 dt_1 dt_2}{t_1 V_{t_1}(\mathbf{x}_1) t_2 V_{t_2}(\mathbf{x}_2)} \right)^{1/2},$$

where $\Gamma^\beta(\mathbf{x}_1, \mathbf{x}_2) = \Gamma^\beta_1(\mathbf{x}_1) \otimes \Gamma^\beta_2(\mathbf{x}_2)$ and $\Gamma^\beta_j(\mathbf{x}_j) = \{(\mathbf{y}_j, t_j) \in M_j \times \mathbb{R}_+ : d_j(\mathbf{x}_j, \mathbf{y}_j) < \beta t_j\}$, $j = 1, 2$. For simplicity, we denote $S^1_P(f)(\mathbf{x}_1, \mathbf{x}_2)$ by $S_P(f)(\mathbf{x}_1, \mathbf{x}_2)$.

The main result of this paper is the following.

Theorem 1.1. *There exist $C > 0$ and $\beta > 1$ such that for all $f \in C^\infty_0(\widetilde{M})$ and for all $\lambda > 0$,*

$$\begin{aligned} &|\{(\mathbf{x}_1, \mathbf{x}_2) \in \widetilde{M} :\ S_P(f)(\mathbf{x}_1, \mathbf{x}_2) > \lambda\}| \qquad (1.3)\\ &\leq C\big|\{(\mathbf{x}_1, \mathbf{x}_2) \in \widetilde{M} : \mathcal{N}^\beta_P(f)(\mathbf{x}_1, \mathbf{x}_2) > \lambda\}\big| + \frac{C}{\lambda^2}\int_0^\lambda s|\{(\mathbf{x}_1, \mathbf{x}_2) \in \widetilde{M} : \mathcal{N}^\beta_P(f)(\mathbf{x}_1, \mathbf{x}_2) > s\}|ds. \end{aligned}$$

The main step here is to establish a modified version of good-λ inequality, that is, to prove that $|\{(\mathbf{x}_1, \mathbf{x}_2) \in A_\beta(\lambda) :\ S_P(f)(\mathbf{x}_1, \mathbf{x}_2) > \lambda\}|$ is bounded by the right-hand side of (3.4), where $A_\beta(\lambda)$ is a suitable subset of $E_\beta(\lambda) := \big\{(\mathbf{x}_1, \mathbf{x}_2) \in \widetilde{M} : \mathcal{N}^\beta_P(f)(\mathbf{x}_1, \mathbf{x}_2) \leq \lambda\big\}$. The key idea that is different from [22] is to inherit the property of the double Poisson integral $u = P^{[1]}_{t_1} P^{[2]}_{t_2}(\chi_{E_\beta(\lambda)})$ via a smooth function which captures the range of u. Thus, we bypass the restriction on a smooth function which is even, and has compact support. We only rely on the pointwise upper bound of the Poisson kernel as well as the conservation property. The proof will be given in Section 3 after some necessary preliminaries in Section 2.

We further note that the above result also holds for $\widetilde{M} = M_1 \times \cdots \times M_n$ for general $n > 2$. It suffices to repeat the proof by induction.

As applications, we point out that: (1) Theorem 1.1 passes the endpoint weak type estimate ($L \log L(\widetilde{M})$ to weak $L^1(\widetilde{M})$) from the strong maximal function $\mathcal{M}_s$ to the Littlewood–Paley area function S^β_P. Thus, through the approach of R. Fefferman [15] (local version) and the recent study [11] (global version), we obtain the $L \log L(\widetilde{M})$ to weak $L^1(\widetilde{M})$ for product Calderón–Zygmund operators on $\widetilde{M}$; (2) Theorem 1.1 also gives rise to the characterisation of product Hardy space on $\widetilde{M}$, which was established via Littlewood–Paley area function and characterised equivalently by atomic decomposition [19, 20]. Details are in Section 4.

2 Notation and preliminaries

In this section, we recall the basic geometry of the Shilov boundary $\widetilde{M} = M_1 \times M_2$ [30] with each M_j given in (1.2), where $\mathcal{P}(z) = \frac{1}{2k}|z|^{2k}$, $k \geq 2$. It is clear that the degree m of $\mathcal{P}(z)$ is given by $m = 2k$.

2.1 Basic geometry of Carnot–Carathéodory space M

We first recall the *control metric on M* given in [30] (see also [31, 27, 28, 34]). Note that we write the complex (0,1) vector field $\overline{Z} = X_1 + iX_2$, where $\{X_1, X_2\}$ are real vector fields on M. Define the metric d on M as follows. If $\mathbf{x}, \mathbf{y} \in M$ and $\delta > 0$, let $AC(\mathbf{x}, \mathbf{y}, \delta)$ denote the set of absolutely continuous mapping $\gamma : [0, 1] \to M$ such that $\gamma(0) = \mathbf{x}$ and $\gamma(1) = \mathbf{y}$, and such that for almost all $t \in [0, 1]$ we have $\gamma'(t) = \alpha_1(t)X_1(\gamma(t)) + \alpha_2(t)X_2(\gamma(t))$ with $|\alpha_1(t)|^2 + |\alpha_2(t)|^2 < \delta^2$. Then we define

$$d(\mathbf{x}, \mathbf{y}) = \inf\{\delta > 0 : \ AC(\mathbf{x}, \mathbf{y}, \delta) \neq \emptyset\}.$$

The corresponding nonisotropic ball is defined as $B(\mathbf{x}, \delta) = \{\mathbf{y} \in M : \ d(\mathbf{x}, \mathbf{y}) < \delta\}$, and let $V_\delta(\mathbf{x})$ denotes its volume. From [30] we know that there is a positive constant C_d such that for every $\mathbf{x} \in M$, $\lambda \geq 1$ and $\delta > 0$,

$$V_{\lambda\delta}(\mathbf{x}) \leq C_d \lambda^m V_\delta(\mathbf{x}). \tag{2.1}$$

We also set $V(\mathbf{x}, \mathbf{y}) = V_{d(\mathbf{x},\mathbf{y})}(\mathbf{x})$. From the doubling property we observe that $V(\mathbf{x}, \mathbf{y}) \approx V(\mathbf{y}, \mathbf{x})$ where the implicit constants are independent of $\mathbf{x}$ and $\mathbf{y}$.

2.2 Sub-Laplacian on M

Consider the sub-Laplacian $\mathcal{L}$ on M in self-adjoint form, given by $\mathcal{L} = \sum_{j=1}^2 X_j^* X_j$. Here $(X_j^*\varphi, \psi) = (\varphi, X_j\psi)$, where $(\varphi, \psi) = \int_M \varphi(\mathbf{x})\bar{\psi}(\mathbf{x})d\mathbf{x}$, and $\varphi, \psi \in C_0^\infty(M)$, the space of C^∞ functions on M with compact support. In general, we have $X_j^* = -X_j + a_j$, where $a_j \in C^\infty(M)$. In our particular setting, we see that $a_j = 0$. That is

$$\mathcal{L} = -\sum_{j=1}^{2} X_j^2.$$

The solution of the following initial value problem for the heat equation,

$$\frac{\partial u}{\partial s}(\mathbf{x}, s) + \mathcal{L}_{\mathbf{x}} u(\mathbf{x}, s) = 0$$

with $u(\mathbf{x}, 0) = f(\mathbf{x})$, is given by $u(\mathbf{x}, s) = H_s(f)(\mathbf{x})$, where H_s is the operator given via the spectral theorem by $H_s = e^{-s\mathcal{L}}$, and an appropriate self-adjoint extension of the non-negative operator $\mathcal{L}$ initially defined on $C_0^\infty(M)$. For $f \in L^2(M)$,

$$H_s(f)(\mathbf{x}) = \int_M H(s, \mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y}.$$

Moreover, $H(s, \mathbf{x}, \mathbf{y})$ has some nice properties (see Proposition 2.3.1 in [30] and Theorem 2.3.1 in [27]). We restate them as follows:

(1) $H(s, \mathbf{x}, \mathbf{y}) \in C^\infty\big([0,\infty) \times M \times M \backslash \{s = 0 \text{ and } \mathbf{x} = \mathbf{y}\}\big)$.

(2) For every integer $N \geq 0$,

$$|\partial_s^j \partial_X^L \partial_Y^K H(s, \mathbf{x}, \mathbf{y})| \lesssim \frac{1}{(d(\mathbf{x},\mathbf{y}) + \sqrt{s})^{2j+K+L}} \frac{1}{V(\mathbf{x},\mathbf{y}) + V_{\sqrt{s}}(\mathbf{x}) + V_{\sqrt{s}}(\mathbf{y})} \left(\frac{\sqrt{s}}{d(\mathbf{x},\mathbf{y}) + \sqrt{s}}\right)^N,$$

(3) For each integer $L \geq 0$ there exist an integer N_L and a constant C_L so that if $\varphi \in C_0^\infty(B(\mathbf{x}_0, \delta))$, then for all $s \in (0, \infty)$,

$$|\partial_X^L H_s[\varphi](\mathbf{x}_0)| \leq C_L \delta^{-L} \sup_{\mathbf{x}} \sum_{|J| \leq N_L} \delta^{|J|} |\partial_X^J \varphi(\mathbf{x})|.$$

(4) For all $(s, \mathbf{x}, \mathbf{y}) \in (0, \infty) \times M \times M$, $H(s, \mathbf{x}, \mathbf{y}) = H(s, \mathbf{y}, \mathbf{x})$, $H(s, \mathbf{x}, \mathbf{y}) \geq 0$.

(5) For all $(s, \mathbf{x}) \in (0, \infty) \times M$, $\int_M H(s, \mathbf{x}, \mathbf{y}) d\mathbf{y} = 1$.

(6) For $1 \leq p \leq \infty$, $\|H_s[f]\|_{L^p(M)} \leq \|f\|_{L^p(M)}$.

(7) For every $\varphi \in C_0^\infty(M)$ and $1 \leq p < \infty$, $\lim_{s \to 0} \|H_s[\varphi] - \varphi\|_{L^p(M)} = 0$.

2.3 Poisson kernel estimate on M

From (2) in Section 2.2, we see that there is a positive constant C_H such that for all $s > 0$, $\mathbf{x}, \mathbf{y} \in M$,

$$|H(s, \mathbf{x}, \mathbf{y})| \leq C_H \frac{1}{V_{\sqrt{s}}(\mathbf{x},\mathbf{y}) + V_{\sqrt{s}}(\mathbf{x}) + V_{\sqrt{s}}(\mathbf{y})} \left(\frac{\sqrt{s}}{d(\mathbf{x},\mathbf{y}) + \sqrt{s}}\right)^N.$$

Let $P_t(\mathbf{x}, \mathbf{y})$ be the kernel of the Poisson semigroup $\mathrm{e}^{-t\sqrt{\mathcal{L}}}$. The estimates for $P_t(\mathbf{x}, \mathbf{y})$ follow from the subordination formula

$$\mathrm{e}^{-t\sqrt{\mathcal{L}}} = \frac{1}{2\sqrt{\pi}} \int_0^\infty \frac{t\mathrm{e}^{\frac{-t^2}{4s}}}{\sqrt{s}} \mathrm{e}^{-s\mathcal{L}} \frac{ds}{s},$$

the doubling property of the measure and the estimate of $|H(s, \mathbf{x}, \mathbf{y})|$ as above. We have that for each $\mathbf{x} \in M$ and $t > 0$

$$\int_M P_t(\mathbf{x}, \mathbf{y}) d\mathbf{y} = 1, \tag{2.2}$$

and that there exists $C_P > 0$ such that for each $\mathbf{x}, \mathbf{y} \in M$ and $t > 0$,

$$|P_t(\mathbf{x}, \mathbf{y})| \leq C_P \frac{1}{V(\mathbf{x},\mathbf{y}) + V_t(\mathbf{x}) + V_t(\mathbf{y})} \cdot \frac{t}{t + d(\mathbf{x},\mathbf{y})}. \tag{2.3}$$

Here C_P depends on C_H and the upper dimension m as in (2.1). From this size estimate we see that for every $f \in L^2(M)$,

$$|P_t(f)(\mathbf{x})| \leq C_P (C_d 2^m + 1) \mathcal{M}(f)(\mathbf{x}), \tag{2.4}$$

where C_d and m are the constants from (2.1) and $\mathcal{M}$ is the Hardy–Littlewood maximal operator such that

$$\mathcal{M}(f)(\mathbf{x}) = \sup_{B \ni x} \frac{1}{|B|} \int_B |f(\mathbf{y})| d\mathbf{y},$$

where B runs over all metric balls in M.

2.4 Basic geometry of Shilov boundary $\widetilde{M} = M_1 \times M_2$

Consider $\widetilde{M} = M_1 \times M_2$ such that $M_j = \{(z_j, w_j) \in \mathbb{C}^2 : \ \mathrm{Im}(w_j) = \mathcal{P}_j(z_j)\}$ with the vector fields $X_{j,1}$ and $X_{j,2}$, $j = 1, 2$. We denote $\vec{\mathbf{x}} = (\mathbf{x}_1, \mathbf{x}_2) \in M_1 \times M_2$.

The nonisotropic distance d_j on M_j can be regarded as a function on M_j which depends only on the variables (z_j, t_j), where $t_j = \mathrm{Re}(w_j)$. In addition, there is a nonisotropic metric $d_{\sum}$ on $\widetilde{M}$ induced by all real vector fields $\{X_{1,1}, X_{1,2}, X_{2,1}, X_{2,2}\}$. If $\vec{\mathbf{x}}, \vec{\mathbf{y}} \in \widetilde{M}$ and $\delta > 0$, let $AC(\vec{\mathbf{x}}, \vec{\mathbf{y}}, \delta)$ denote the set of absolutely continuous mappings $\gamma : [0, 1] \to \widetilde{M}$ such that $\gamma(0) = \vec{\mathbf{x}}$ and $\gamma(1) = \vec{\mathbf{y}}$, and such that for almost every $t \in [0, 1]$ we have $\gamma'(t) = \sum_{j=1}^{2}(\alpha_{j,1}(t)X_{j,1}(\gamma(t)) + \alpha_{j,2}(t)X_{j,2}(\gamma(t)))$ with $\sum_{j=1}^{2}(|\alpha_{j,1}(t)|^2 + |\alpha_{j,2}(t)|^2) < \delta^2$. Then

$$d_{\sum}(\vec{\mathbf{x}}, \vec{\mathbf{y}}) = \inf\{\delta > 0 \mid AC(\vec{\mathbf{x}}, \vec{\mathbf{y}}, \delta) \neq \emptyset\}.$$

Similar to (2.4), we also have that for every $f \in L^2(\widetilde{M})$,

$$|P_{t_1}^{[1]} P_{t_2}^{[2]}(f)(\mathbf{x}_1, \mathbf{x}_2)| \leq C_0 \mathcal{M}_S(f)(\mathbf{x}_1, \mathbf{x}_2), \tag{2.5}$$

where C_0 depends on the constant in (2.4), $P_{t_j}^{[j]}$ denotes the Poisson semigroup on M_j, and $\mathcal{M}_S$ is the strong maximal function on $\widetilde{M}$ such that

$$\mathcal{M}_S(f)(\mathbf{x}_1, \mathbf{x}_2) = \sup_{B_1 \times B_2 \ni (\mathbf{x}_1, \mathbf{x}_2)} \frac{1}{|B_1 \times B_2|} \int_{B_1 \times B_2} |f(\mathbf{y}_1, \mathbf{y}_2)| d\mathbf{y}_1 d\mathbf{y}_2,$$

where B_j runs over all metric balls in M_j for $j = 1, 2$.

3 Proof of Theorem 1.1: Fefferman–Stein type inequality on $\widetilde{M}$

For $j = 1, 2$, let M_j be the model domain as defined in Section 2, with the vector fields $X_{j,1}$ and $X_{j,2}$ and the sub-Laplacian $\mathcal{L}_j$.

We aim to prove that there exist constants $C > 0$ and $\beta > 1$ such that for $f \in C_0^\infty(\widetilde{M})$ and for all $\alpha > 0$,

$$\begin{aligned} &|\{(\mathbf{x}_1, \mathbf{x}_2) \in \widetilde{M} : \ S_P(f)(\mathbf{x}_1, \mathbf{x}_2) > \alpha\}| \\ &\leq C\big|\{(\mathbf{x}_1, \mathbf{x}_2) \in \widetilde{M} : \ \mathcal{N}_P^\beta(f)(\mathbf{x}_1, \mathbf{x}_2) > \alpha\}\big| + \frac{C}{\alpha^2} \iint_{E_\beta(\alpha)} \mathcal{N}_P^\beta(f)(\mathbf{x}_1, \mathbf{x}_2)^2 d\mathbf{x}_1 d\mathbf{x}_2, \end{aligned} \tag{3.1}$$

where

$$E_\beta(\alpha) = \{(\mathbf{x}_1, \mathbf{x}_2) \in \widetilde{M} : \ \mathcal{N}_P^\beta(f)(\mathbf{x}_1, \mathbf{x}_2) \leq \alpha\}.$$

We denote

$$E_\beta(\alpha)^c = \widetilde{M} \backslash E_\beta(\alpha) = \{(\mathbf{x}_1, \mathbf{x}_2) \in \widetilde{M} : \ \mathcal{N}_P^\beta(f)(\mathbf{x}_1, \mathbf{x}_2) > \alpha\}.$$

For $j = 1, 2$, we note that for any $f \in L^2(M_j)$, $u_j(\mathbf{x}_j, t_j) := e^{-t_j\sqrt{\mathcal{L}_j}}(f)(\mathbf{x}_j)$, $(\mathbf{x}_j, t_j) \in M_j \times \mathbb{R}_+$, is harmonic, in the sense that

$$\Delta_{t_j, M_j} u_j(\mathbf{x}_j, t_j) = \mathcal{L}_j u_j(\mathbf{x}_j, t_j) - \partial_{t_j}^2 u_j(\mathbf{x}_j, t_j) = 0,$$

where $\Delta_{t_j,M_j} := \mathcal{L}_j - \partial_{t_j}^2$ and we use the fact that $\partial_{t_j}^2 u_j(\mathbf{x}_j,t_j) = \partial_{t_j}^2 e^{-t_j\sqrt{\mathcal{L}_j}}(f)(\mathbf{x}_j) = \mathcal{L}_j u_j(\mathbf{x}_j,t_j)$. Consequently, for $j=1,2$, for the gradient $\nabla_{t_j,M_j} = (\partial_{t_j},\nabla_{M_j}) = (\partial_{t_j}, X_{j,1}, X_{j,2})$, the following formula holds: for every $(\mathbf{x}_j,t_j)\in M_j\times\mathbb{R}_+$,

$$\begin{aligned} 2|\nabla_{t_j,M_j}u_j(\mathbf{x}_j,t_j)|^2 &= 2|\nabla_{t_j,M_j}u_j(\mathbf{x}_j,t_j)|^2 - 2u_j(\mathbf{x}_j,t_j)\Delta_{t_j,M_j}u_j(\mathbf{x}_j,t_j) \\ &= -\Delta_{t_j,M_j}\left(u_j^2(\mathbf{x}_j,t_j)\right). \end{aligned} \tag{3.2}$$

Next, for all $\alpha>0$ and $f\in L^1(\widetilde{M})$ satisfying $\mathcal{N}_P^\beta(f)\in L^1(\widetilde{M})$, define

$$A_\beta(\alpha) := \left\{(\mathbf{x}_1,\mathbf{x}_2)\in\widetilde{M} : \mathcal{M}_S(\chi_{E_\beta(\alpha)^c})(\mathbf{x}_1,\mathbf{x}_2) \le \frac{1}{10C_0}\right\},$$

where C_0 is the constant in (2.5).

From the definition, it is direct to see that

$$E_\beta(\alpha)^c \subset A_\beta(\alpha)^c = \widetilde{M}\backslash A_\beta(\alpha)$$

and hence $A_\beta(\alpha)\subset E_\beta(\alpha)$.

Next, from the L^2-boundedness of the strong maximal function $\mathcal{M}_S$, we see that

$$|A_\beta(\alpha)^c| \le C\left\|\mathcal{M}_S(\chi_{E_\beta(\alpha)^c})\right\|_{L^2(\widetilde{M})}^2 \le C|E_\beta(\alpha)^c|, \tag{3.3}$$

where the constant C is independent of α and β. Then we split

$$\begin{aligned} &|\{(\mathbf{x}_1,\mathbf{x}_2)\in\widetilde{M} :\ S_P(f)(\mathbf{x}_1,\mathbf{x}_2)>\alpha\}| \\ &\le |\{(\mathbf{x}_1,\mathbf{x}_2)\in A_\beta(\alpha)^c :\ S_P(f)(\mathbf{x}_1,\mathbf{x}_2)>\alpha\}| + |\{(\mathbf{x}_1,\mathbf{x}_2)\in A_\beta(\alpha) :\ S_P(f)(\mathbf{x}_1,\mathbf{x}_2)>\alpha\}| \\ &\le C\left|E_\beta(\alpha)^c\right| + \frac{1}{\alpha^2}\iint_{A_\beta(\alpha)} S_P(f)(\mathbf{x}_1,\mathbf{x}_2)^2 d\mathbf{x}_1 d\mathbf{x}_2, \end{aligned} \tag{3.4}$$

where the last inequality follows from (3.3) and from Chebyshev's inequality. Now it suffices to estimate the second term in the right-hand side of last inequality above.

Let

$$g(\mathbf{x}_1,\mathbf{x}_2) := \chi_{E_\beta(\alpha)}(\mathbf{x}_1,\mathbf{x}_2) \quad\text{and}\quad W_\beta := \bigcup_{(\mathbf{x}_1,\mathbf{x}_2)\in A_\beta(\alpha)} \Gamma(\mathbf{x}_1,\mathbf{x}_2).$$

We first note that for $(\mathbf{y}_1,\mathbf{y}_2)\in\widetilde{M}$ and $t_1,t_2>0$,

$$\begin{aligned} P_{t_1}^{[1]}P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2) &= P_{t_1}^{[1]}P_{t_2}^{[2]}\left(1-\chi_{E_\beta(\alpha)^c}\right)(\mathbf{y}_1,\mathbf{y}_2) \\ &= 1 - P_{t_1}^{[1]}P_{t_2}^{[2]}\left(\chi_{E_\beta(\alpha)^c}\right)(\mathbf{y}_1,\mathbf{y}_2), \end{aligned}$$

where we use the fact that $P_{t_1}^{[1]}P_{t_2}^{[2]}(1)=1$. Next, note that for $(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2)\in W_\beta$, there is $(\mathbf{x}_1,\mathbf{x}_2)\in A_\beta(\alpha)$ such that $(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2)\in\Gamma(\mathbf{x}_1,\mathbf{x}_2)$, and hence

$$P_{t_1}^{[1]}P_{t_2}^{[2]}\left(\chi_{E_\beta(\alpha)^c}\right)(\mathbf{y}_1,\mathbf{y}_2) \le C_0\mathcal{M}_S(\chi_{E_\beta(\alpha)^c})(\mathbf{x}_1,\mathbf{x}_2) < \frac{1}{10},$$

where the first inequality follows from (2.5) and the last inequality follows from the fact that $(\mathbf{x}_1,\mathbf{x}_2)\in A_\beta(\alpha)$.

Then we see that for every $(\mathbf{y}_1, \mathbf{y}_2, t_1, t_2) \in W_\beta$, we obtain that

$$P_{t_1}^{[1]} P_{t_2}^{[2]}(g)(\mathbf{y}_1, \mathbf{y}_2) > \frac{9}{10}.$$

Next, we claim that if β is chosen sufficient large, then there is a constant $C_1 \in (0, \frac{9}{10})$, such that for any $(\mathbf{y}_1, \mathbf{y}_2, t_1, t_2) \in \big(\widetilde{W}_\beta\big)^c := \Big(\widetilde{M} \times [0,\infty) \times [0,\infty)\Big) \Big\backslash \widetilde{W}_\beta$,

$$P_{t_1}^{[1]} P_{t_2}^{[2]}(g)(\mathbf{y}_1, \mathbf{y}_2) \leq C_1, \tag{3.5}$$

where

$$\widetilde{W}_\beta := \bigcup_{(\mathbf{x}_1, \mathbf{x}_2) \in E_\beta(\alpha)} \Gamma^\beta(\mathbf{x}_1, \mathbf{x}_2).$$

In fact, for every $(\mathbf{y}_1, \mathbf{y}_2, t_1, t_2) \in \big(\widetilde{W}_\beta\big)^c$, we see that for any $(\mathbf{z}_1, \mathbf{z}_2) \in E_\beta(\alpha)$, we have either $d_1(\mathbf{y}_1, \mathbf{z}_1) \geq \beta t_1$, or $d_2(\mathbf{y}_2, \mathbf{z}_2) \geq \beta t_2$, or both. Hence, we have

$$\begin{aligned}
0 \leq P_{t_1}^{[1]} P_{t_2}^{[2]}(g)(\mathbf{y}_1, \mathbf{y}_2) &= \iint_{\widetilde{M}} \chi_{E_\beta(\alpha)}(\mathbf{z}_1, \mathbf{z}_2) P_{t_1}^{[1]}(\mathbf{y}_1, \mathbf{z}_1) P_{t_2}^{[2]}(\mathbf{y}_2, \mathbf{z}_2) d\mathbf{z}_1 d\mathbf{z}_2 \\
&\leq \int_{d_1(\mathbf{y}_1, \mathbf{z}_1) \geq \beta t_1} P_{t_1}^{[1]}(\mathbf{y}_1, \mathbf{z}_1) d\mathbf{z}_1 \int_{M_2} P_{t_2}^{[2]}(\mathbf{y}_2, \mathbf{z}_2) d\mathbf{z}_2 \\
&\quad + \int_{M_1} P_{t_1}^{[1]}(\mathbf{y}_1, \mathbf{z}_1) d\mathbf{z}_1 \int_{d_2(\mathbf{y}_2, \mathbf{z}_2) \geq \beta t_2} P_{t_2}^{[2]}(\mathbf{y}_2, \mathbf{z}_2) d\mathbf{z}_2 \\
&\leq \int_{d_1(\mathbf{y}_1, \mathbf{z}_1) \geq \beta t_1} P_{t_1}^{[1]}(\mathbf{y}_1, \mathbf{z}_1) d\mathbf{z}_1 + \int_{d_2(\mathbf{y}_2, \mathbf{z}_2) \geq \beta t_2} P_{t_2}^{[2]}(\mathbf{y}_2, \mathbf{z}_2) d\mathbf{z}_2,
\end{aligned}$$

where the last inequality follows from the conservation property (2.2). To continue, by decomposing $\{z_j \in M_j : d_j(\mathbf{y}_j, \mathbf{z}_j) \geq \beta t_j\}$, $j = 1, 2$, into annuli and using the size estimate (2.3), we have

$$0 \leq P_{t_1}^{[1]} P_{t_2}^{[2]}(g)(\mathbf{y}_1, \mathbf{y}_2) \leq \frac{C}{\beta} \to 0, \ \ (\text{as } \beta \to \infty),$$

where the constant C depends on the constant C_d and m in (2.1) and on C_P in (2.3). Thus, there is some $\beta > 1$ such that our claim (3.5) holds.

This also shows that if $P_{t_1}^{[1]} P_{t_2}^{[2]}(g)(\mathbf{y}_1, \mathbf{y}_2) > C_1$, then $(\mathbf{y}_1, \mathbf{y}_2, t_1, t_2) \in \widetilde{W}_\beta$.

To continue, we now choose a smooth cut-off function $\varphi(t) \in C^\infty(\mathbb{R})$ such that $\varphi(t) = 1$ when $t \geq \frac{9}{10}$ and $\varphi(t) = 0$, when $t \leq C_1$.

Besides, for simplicity, we denote $v_{\mathbf{y}_2, t_2}(\mathbf{y}_1) := \nabla_{t_2, M_2} P_{t_2}^{[2]}(f)(\mathbf{y}_1, \mathbf{y}_2)$. Then,

$$\begin{aligned}
&\iint_{A_\beta(\alpha)} S_P(f)(\mathbf{x}_1, \mathbf{x}_2)^2 d\mathbf{x}_1 d\mathbf{x}_2 \\
&= \iint_{A_\beta(\alpha)} \iint_{\Gamma(\mathbf{x}_1, \mathbf{x}_2)} |t_1 \nabla_{t_1, M_1} P_{t_1}^{[1]} \ t_2 \nabla_{t_2, M_2} P_{t_2}^{[2]}(f)(\mathbf{y}_1, \mathbf{y}_2)|^2 \frac{d\mathbf{y}_1 d\mathbf{y}_2 dt_1 dt_2}{t_1 V_{t_1}(\mathbf{x}_1) t_2 V_{t_2}(\mathbf{x}_2)} d\mathbf{x}_1 d\mathbf{x}_2 \\
&= \iint_{A_\beta(\alpha)} \iint_{\Gamma(\mathbf{x}_1, \mathbf{x}_2)} \Big| t_1 \nabla_{t_1, M_1} P_{t_1}^{[1]} t_2 v_{\mathbf{y}_2, t_2}(\mathbf{y}_1) \Big|^2 \frac{d\mathbf{y}_1 d\mathbf{y}_2 dt_1 dt_2}{t_1 V_{t_1}(\mathbf{x}_1) t_2 V_{t_2}(\mathbf{x}_2)} d\mathbf{x}_1 d\mathbf{x}_2 \\
&\leq \iiiint_{W_\beta} \Big| \nabla_{t_1, M_1} P_{t_1}^{[1]} v_{\mathbf{y}_2, t_2}(\mathbf{y}_1) \Big|^2 t_1 t_2 dt_1 dt_2 d\mathbf{y}_1 d\mathbf{y}_2
\end{aligned}$$

$$\leq \iiiint_{\widetilde{M}\times\mathbb{R}_+\times\mathbb{R}_+} \left|\nabla_{t_1,M_1} P^{[1]}_{t_1} v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)\right|^2 \times \left|\varphi\Big(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big)\right|^2 t_1t_2dt_1dt_2d\mathbf{y}_1d\mathbf{y}_2. \tag{3.6}$$

To continue, we note that $P^{[1]}_{t_1} v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)$ as a function of $(\mathbf{y}_1,t_1)$ is harmonic in $M_1\times\mathbb{R}_+$. Moreover, $P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)$ as a function of $(\mathbf{y}_1,t_1)$ is also harmonic in $M_1\times\mathbb{R}_+$. Hence, by using (3.2) the following equality holds:

$$\begin{aligned}
&\left|\nabla_{t_1,M_1} P^{[1]}_{t_1} v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)\right|^2 \left|\varphi\Big(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big)\right|^2 \qquad (3.7)\\
&= -\frac{1}{2}\Delta_{t_1,M_1}\left(\Big(P^{[1]}_{t_1} v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)\Big)^2\cdot\varphi\Big(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big)^2\right)\\
&\quad - 4P^{[1]}_{t_1} v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)\nabla_{t_1,M_1}P^{[1]}_{t_1} v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)\\
&\qquad \times\varphi\Big(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big)\varphi'\Big(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big)\nabla_{t_1,M_1}P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)\\
&\quad - |P^{[1]}_{t_1} v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)|^2\varphi'\Big(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big)^2\left|\nabla_{t_1,M_1}P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)\right|^2\\
&\quad - |P^{[1]}_{t_1} v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)|^2\varphi\Big(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big)\varphi''\Big(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big)\left|\nabla_{t_1,M_1}P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)\right|^2\\
&=: f_1(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2)+f_2(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2)+f_3(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2)+f_4(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2).
\end{aligned}$$

We note that by Young's inequality,

$$\begin{aligned}
|f_2(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2)| &\leq \frac{1}{10}\left|\nabla_{t_1,M_1} P^{[1]}_{t_1} v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)\right|^2 \left|\varphi\Big(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big)\right|^2\\
&\quad + 40|P^{[1]}_{t_1} v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)|^2\left|\varphi'\Big(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big)\right|^2|\nabla_{t_1,M_1}P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)|^2\\
&=: f_{21}(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2)+f_{22}(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2).
\end{aligned}$$

We can see that the integral

$$\iiiint_{\widetilde{M}\times\mathbb{R}_+\times\mathbb{R}_+} f_{21}(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2)t_1t_2dt_1dt_2d\mathbf{y}_1d\mathbf{y}_2.$$

can be absorbed by the right-hand side of (3.6), while $f_{22}(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2)$ is quite similar to $f_3(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2)$ and $f_4(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2)$. Hence, we further have

$$\begin{aligned}
&f_{22}(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2)+f_3(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2)+f_4(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2) \qquad (3.8)\\
&\leq 40|P^{[1]}_{t_1} v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)|^2\left|\Phi'\Big(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big)\right|^2\left|\nabla_{t_1,M_1}P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)\right|^2\\
&=: \mathfrak{f}_2(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2),
\end{aligned}$$

where we choose $\Phi(t)$ such that

$$\Phi'(t)=((\varphi'(t))^4+(\varphi(t)\varphi''(t))^2)^{1/4}, \qquad \text{and} \qquad \Phi(C_1)=0.$$

Note that $\Phi'(t) \geq 0$ and $\Phi'(t) = 0$ for $t \leq C_1$ or $t > \frac{9}{10}$. In addition, via assuming $\Phi(C_1) = 0$, we see that Φ exhibits behavior similar to ϕ.

To continue, note that the right-hand side of (3.6) is bounded by

$$\begin{aligned}
&\frac{10}{9}\left|\iiiint_{\widetilde{M}\times\mathbb{R}_+\times\mathbb{R}_+} f_1(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2)t_1t_2dt_1dt_2d\mathbf{y}_1d\mathbf{y}_2\right| \\
&\qquad + \frac{10}{9}\left|\iiiint_{\widetilde{M}\times\mathbb{R}_+\times\mathbb{R}_+} \mathfrak{f}_2(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2)t_1t_2dt_1dt_2d\mathbf{y}_1d\mathbf{y}_2\right| \\
&=: \mathrm{I}_1 + \mathrm{I}_2.
\end{aligned}$$

For the term I_1, integration by parts yields that

$$\begin{aligned}
&\left|\iint_{M_1\times\mathbb{R}_+} f_1(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2)t_1dt_1d\mathbf{y}_1\right| \\
&= \frac{1}{2}\left|\iint_{M_1\times\mathbb{R}_+} \mathcal{L}_1\left(P^{[1]}_{t_1}v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)^2\cdot\varphi(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))^2\right)t_1dt_1d\mathbf{y}_1\right. \\
&\qquad \left.-\iint_{M_1\times\mathbb{R}_+}\partial^2_{t_1}\left(P^{[1]}_{t_1}v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)^2\cdot\varphi(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))^2\right)t_1dt_1d\mathbf{y}_1\right| \\
&= \frac{1}{2}\left|\int_{\mathbb{R}_+} t_1\nabla_{M_1}\left(P^{[1]}_{t_1}v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)^2\cdot\varphi(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))^2\right)\Big|_{d_1(\mathbf{0},\mathbf{y}_1)=\infty}dt_1\right. \\
&\qquad -\int_{M_1}\partial_{t_1}\left(P^{[1]}_{t_1}v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)^2\cdot\varphi(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))^2\right)t_1\Big|_{t_1=0}^{t_1=\infty}d\mathbf{y}_1 \\
&\qquad \left.+\iint_{M_1\times\mathbb{R}_+}\partial_{t_1}\left(P^{[1]}_{t_1}v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)^2\cdot\varphi(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))^2\right)dt_1d\mathbf{y}_1\right| \\
&\leq \mathrm{I}_{11}+\mathrm{I}_{12}+\mathrm{I}_{13}.
\end{aligned}$$

For I_{11}, note that

$$\begin{aligned}
&t_1\nabla_{M_1}\left(P^{[1]}_{t_1}v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)^2\cdot\varphi(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))^2\right) \\
&= 2P^{[1]}_{t_1}v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)t_1\nabla_{M_1}P^{[1]}_{t_1}v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)\cdot\varphi(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))^2 \\
&\qquad + P^{[1]}_{t_1}v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)^2\cdot 2\varphi(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))t_1\nabla_{M_1}\varphi(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)).
\end{aligned}$$

The size condition of the Poisson kernel in (2.3) yields that

$$t_1\nabla_{M_1}\left(P^{[1]}_{t_1}v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)^2\cdot\varphi(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))^2\right)\Big|_{d_1(\mathbf{0},\mathbf{y}_1)=\infty} = 0.$$

Thus, $\mathrm{I}_{11} = 0$.

For I_{12}, we have

$$\begin{aligned}
&\partial_{t_1}\left(P^{[1]}_{t_1}v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)^2\cdot\varphi(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))^2\right)t_1 \\
&= 2P^{[1]}_{t_1}v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)t_1\partial_{t_1}P^{[1]}_{t_1}v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)\cdot\varphi(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))^2 \\
&\quad + P^{[1]}_{t_1}v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)^2\cdot 2\varphi(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))\varphi'(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))\,t_1\partial_{t_1}P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2).
\end{aligned}$$

For $t_1 \to \infty$, from the decay of Poisson kernel in (2.3) and the functional calculus we see that the kernel of $t_1\partial_{t_1} P_{t_1}^{[1]}$ also satisfies a similar size condition as in (2.3). Thus, we have that

$$\partial_{t_1}\left(P_{t_1}^{[1]} v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)^2 \cdot \varphi(P_{t_1}^{[1]}P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2))^2\right) t_1 \to 0. \tag{3.9}$$

On the other hand, when $t_1 \to 0^+$, we see that the term $t_1\partial_{t_1} P_{t_1}^{[1]} v_{\mathbf{y}_2,t_2}(\mathbf{y}_1) \to 0$, since $t_1\partial_{t_1} P_{t_1}^{[1]}$ has integration zero. Thus, (3.9) also holds when $t_1 \to 0^+$. This gives that $\mathrm{I}_{12} = 0$.

For I_{13}, we have

$$\begin{aligned}\mathrm{I}_{13} &= \left|\int_{M_1}\left(|P_{t_1}^{[1]} v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)|^2 \cdot \varphi(P_{t_1}^{[1]}P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2))^2\right)\Big|_{t_1=0}^{t_1=\infty} d\mathbf{y}_1\right| \\ &\le \int_{M_1} |v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)|^2 |\varphi(P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2))|^2 d\mathbf{y}_1,\end{aligned}$$

where the last inequality follows from the fact that $|P_{t_1}^{[1]} v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)|^2 \cdot \varphi(P_{t_1}^{[1]}P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2))^2 \to 0$ as $t_1 \to \infty$ and that $P_{t_1}^{[1]} \to$ identity as $t_1 \to 0^+$.

Therefore,

$$\begin{aligned}\mathrm{I}_1 &\le \frac{5}{9}\iiint_{\widetilde{M}\times\mathbb{R}_+} |v_{\mathbf{y}_2,t_2}(\mathbf{y}_1)|^2 |\varphi(P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2))|^2 t_2\, dt_2\, d\mathbf{y}_2\, d\mathbf{y}_1 \\ &= \frac{5}{9}\int_{M_1}\iint_{M_2\times\mathbb{R}_+} |\nabla_{t_2,M_2} P_{t_2}^{[2]}(f)(\mathbf{y}_1,\mathbf{y}_2)|^2 |\varphi(P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2))|^2 t_2\, dt_2\, d\mathbf{y}_2 \quad d\mathbf{y}_1.\end{aligned} \tag{3.10}$$

Using the same argument as in (3.7), we have

$$\begin{aligned}&\left|\nabla_{t_2,M_2} P_{t_2}^{[2]}(f)(\mathbf{y}_1,\mathbf{y}_2)\right|^2 \left|\varphi\Big(P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big)\right|^2 \\ &= -\frac{1}{2}\Delta_{t_2,M_2}\left(P_{t_2}^{[2]}(f)(\mathbf{y}_1,\mathbf{y}_2)^2 \cdot \varphi\Big(P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big)^2\right) \\ &\quad - 4P_{t_2}^{[2]}(f)(\mathbf{y}_1,\mathbf{y}_2)\nabla_{t_2,M_2} P_{t_2}^{[2]}(f)(\mathbf{y}_1,\mathbf{y}_2) \\ &\qquad \times \varphi\Big(P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big)\varphi'\Big(P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big)\nabla_{t_2,M_2} P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2) \\ &\quad - |P_{t_2}^{[2]}(f)(\mathbf{y}_1,\mathbf{y}_2)|^2 \varphi'\Big(P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big)^2 \left|\nabla_{t_2,M_2} P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2)\right|^2 \\ &\quad - |P_{t_2}^{[2]}(f)(\mathbf{y}_1,\mathbf{y}_2)|^2 \varphi\Big(P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big)\varphi''\Big(P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big) \left|\nabla_{t_2,M_2} P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2)\right|^2 \\ &=: h_1(\mathbf{y}_1,\mathbf{y}_2,t_2) + h_2(\mathbf{y}_1,\mathbf{y}_2,t_2) + h_3(\mathbf{y}_1,\mathbf{y}_2,t_2) + h_4(\mathbf{y}_1,\mathbf{y}_2,t_2).\end{aligned} \tag{3.11}$$

Then, the term h_2 can be handled by using the same estimate as we did for f_2, that is, we dominate h_2 by

$$\begin{aligned}&\frac{1}{10}\left|\nabla_{t_2,M_2} P_{t_2}^{[2]}(f)(\mathbf{y}_1,\mathbf{y}_2)\right|^2 \left|\varphi\Big(P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big)\right|^2 \\ &\qquad + 40\left|P_{t_2}^{[2]}(f)(\mathbf{y}_1,\mathbf{y}_2)\right|^2 \left|\varphi'\Big(P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big)\right|^2 \left|\nabla_{t_2,M_2} P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2)\right|^2 \\ &=: h_{21}(\mathbf{y}_1,\mathbf{y}_2,t_2) + h_{22}(\mathbf{y}_1,\mathbf{y}_2,t_2).\end{aligned}$$

Again, we see that

$$\frac{5}{9}\int_{M_1}\iint_{M_2\times\mathbb{R}_+} h_{21}(\mathbf{y}_1,\mathbf{y}_2,t_2)t_2dt_2d\mathbf{y}_2\ d\mathbf{y}_1$$

can be absorbed by the right-hand side of (3.10), and we further have

$$\begin{aligned}&h_{22}(\mathbf{y}_1,\mathbf{y}_2,t_2)+h_3(\mathbf{y}_1,\mathbf{y}_2,t_2)+h_4(\mathbf{y}_1,\mathbf{y}_2,t_2)\\&\le 40|P_{t_2}^{[2]}(f)(\mathbf{y}_1,\mathbf{y}_2)|^2\left|\Psi\left(P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2)\right)\right|^2\left|\nabla_{t_2,M_2}P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2)\right|^2\\&=:\mathfrak{h}_2(\mathbf{y}_1,\mathbf{y}_2,t_2),\end{aligned}\tag{3.12}$$

where we choose

$$\Psi(t)=((\varphi'(t))^4+(\varphi(t)\varphi''(t))^2)^{1/4}.$$

Then we further have

$$\begin{aligned}\mathrm{I}_1&\le C\int_{M_1}\iint_{M_2\times\mathbb{R}_+} h_1(\mathbf{y}_1,\mathbf{y}_2,t_2)t_2dt_2d\mathbf{y}_2\ \ d\mathbf{y}_1\\&\quad+C\int_{M_1}\iint_{M_2\times\mathbb{R}_+}\mathfrak{h}_2(\mathbf{y}_1,\mathbf{y}_2,t_2)t_2dt_2d\mathbf{y}_2\ \ d\mathbf{y}_1\\&=:\widetilde{\mathrm{I}_{11}}+\widetilde{\mathrm{I}_{12}}\end{aligned}$$

with an positive absolute constant C. By repeating a similar integration by parts as in the estimates for I_1, we have

$$\widetilde{\mathrm{I}_{11}}\le C\iint_{\widetilde{M}} f(\mathbf{y}_1,\mathbf{y}_2)^2\varphi(g(\mathbf{y}_1,\mathbf{y}_2))^2d\mathbf{y}_1d\mathbf{y}_2.$$

It follows from the definition of the non-tangential maximal function $\mathcal{N}_P^\beta(f)$ that $f(\mathbf{y}_1,\mathbf{y}_2)\le\mathcal{N}_P^\beta(f)(\mathbf{y}_1,\mathbf{y}_2)$. Besides, from the definition of the functions g and φ, we see that for $(\mathbf{y}_1,\mathbf{y}_2)$ with $\varphi(g(\mathbf{y}_1,\mathbf{y}_2))\ne 0$, we have that $g(\mathbf{y}_1,\mathbf{y}_2)>C_1$, which shows that $(\mathbf{y}_1,\mathbf{y}_2)\in E_\beta(\alpha)$. Hence, $\mathcal{N}_P^\beta(f)(\mathbf{y}_1,\mathbf{y}_2)\le\alpha$. Thus,

$$\widetilde{\mathrm{I}_{11}}\le C\iint_{E_\beta(\alpha)}\mathcal{N}_P^\beta(f)(\mathbf{y}_1,\mathbf{y}_2)^2d\mathbf{y}_1d\mathbf{y}_2.$$

For $\widetilde{\mathrm{I}_{12}}$, again, from the definitions of the functions g and Ψ (which inherit the support condition from φ), we see that for $(\mathbf{y}_1,\mathbf{y}_2,t_2)$ with $\Psi(P_{t_2}^{[2]}g(\mathbf{y}_1,\mathbf{y}_2))\ne 0$, we have that $P_{t_2}^{[2]}g(\mathbf{y}_1,\mathbf{y}_2)>C_1$. From (3.5) we see that $(\mathbf{y}_1,\mathbf{y}_2,0,t_2)\in\widetilde{W}_\beta$. Hence, there exists $\mathbf{z}_2$ such that $(\mathbf{y}_1,\mathbf{z}_2)\in E_\beta(\alpha)$ and that $|P_{t_2}^{[2]}(f)(\mathbf{y}_1,\mathbf{y}_2)|\le\mathcal{N}_P^\beta(f)(\mathbf{y}_1,\mathbf{z}_2)\le\alpha$. Thus,

$$\begin{aligned}\widetilde{\mathrm{I}_{12}}&\le C\iiint_{\widetilde{M}\times\mathbb{R}_+}|P_{t_2}^{[2]}(f)(\mathbf{y}_1,\mathbf{y}_2)|^2\left|\Psi\left(P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2)\right)\right|^2\left|\nabla_{t_2,M_2}P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2)\right|^2 t_2dt_2d\mathbf{y}_2\ \ d\mathbf{y}_1\\&\le C\alpha^2\iiint_{\widetilde{M}\times\mathbb{R}_+}\left|\nabla_{t_2,M_2}P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2)\right|^2 t_2dt_2d\mathbf{y}_2\ \ d\mathbf{y}_1\\&=C\alpha^2\iiint_{\widetilde{M}\times\mathbb{R}_+}\left|t_2\nabla_{t_2,M_2}P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2)\right|^2\frac{dt_2}{t_2}d\mathbf{y}_2\ \ d\mathbf{y}_1\end{aligned}$$

$$= C\alpha^2 \iiint_{\widetilde{M}\times\mathbb{R}_+} \left|t_2\nabla_{t_2,M_2}P^{[2]}_{t_2}(1-g)(\mathbf{y}_1,\mathbf{y}_2)\right|^2 \frac{dt_2}{t_2}d\mathbf{y}_2 \ \ d\mathbf{y}_1,$$

where the last equality follows from the fact that the kernel of $t_2\nabla_{t_2,M_2}P^{[2]}_{t_2}$ has integration zero. Therefore, by using the Littlewood–Paley estimate, we obtain that

$$\widetilde{\mathrm{I}_{12}} \le C\alpha^2\|1-g\|^2_{L^2(\widetilde{M})} = C\alpha^2\left|E_\beta(\alpha)^c\right|.$$

This finishes the estimate of the term I_1. We now turn to I_2. By noting that

$$\nabla_{t_1,M_1}\Phi\Big(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big) = \Phi'\Big(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big)\nabla_{t_1,M_1}P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2),$$

we have

$$\begin{aligned}\mathrm{I}_2 = \frac{400}{9}\iiiint_{\widetilde{M}\times\mathbb{R}_+\times\mathbb{R}_+} &\left|\nabla_{t_2,M_2}P^{[2]}_{t_2}P^{[1]}_{t_1}(f)(\mathbf{y}_1,\mathbf{y}_2)\right|^2 \\ &\times\left|\nabla_{t_1,M_1}\Phi\Big(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)\Big)\right|^2 t_1t_2dt_1dt_2d\mathbf{y}_1d\mathbf{y}_2.\end{aligned} \tag{3.13}$$

Observe that

$$\begin{aligned}&|\nabla_{t_2,M_2}P^{[1]}_{t_1}P^{[2]}_{t_2}(f)(\mathbf{y}_1,\mathbf{y}_2)|^2|\nabla_{t_1,M_1}\Phi(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))|^2\\
&= -\frac{1}{2}\Delta_{t_2,M_2}\left(P^{[1]}_{t_1}P^{[2]}_{t_2}(f)(\mathbf{y}_1,\mathbf{y}_2)^2|\nabla_{t_1,M_1}\Phi(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))|^2\right)\\
&\quad - 4P^{[1]}_{t_1}P^{[2]}_{t_2}(f)(\mathbf{y}_1,\mathbf{y}_2)\nabla_{t_2,M_2}P^{[1]}_{t_1}P^{[2]}_{t_2}(f)(\mathbf{y}_1,\mathbf{y}_2)\\
&\qquad\qquad \times\nabla_{t_1,M_1}\Phi(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))\nabla_{t_2,M_2}\nabla_{t_1,M_1}\Phi(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))\\
&\quad - |P^{[1]}_{t_1}P^{[2]}_{t_2}f(\mathbf{y}_1,\mathbf{y}_2)|^2|\nabla_{t_2,M_2}\nabla_{t_1,M_1}\Phi(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))|^2\\
&\quad - |P^{[1]}_{t_1}P^{[2]}_{t_2}(f)(\mathbf{y}_1,\mathbf{y}_2)|^2\\
&\qquad\qquad \times\nabla_{t_1,M_1}\Phi(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))\nabla_{t_2,M_2}\nabla_{t_2,M_2}\nabla_{t_1,M_1}\Phi(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))\\
&=: \mathfrak{F}_1(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2)+\mathfrak{F}_2(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2)+\mathfrak{F}_3(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2)+\mathfrak{F}_4(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2).\end{aligned}$$

Thus, the right-hand side of (3.13) is bounded by $\mathrm{II}_{21}+\mathrm{II}_{22}+\mathrm{II}_{23}+\mathrm{II}_{24}$, where

$$\mathrm{II}_{2j} := C\left|\iiiint_{\widetilde{M}\times\mathbb{R}_+\times\mathbb{R}_+}\mathfrak{F}_j(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2)t_1t_2dt_1dt_2d\mathbf{y}_1d\mathbf{y}_2\right|, \quad j=1,2,3,4.$$

To estimate the term II_{21}, we first let $\Phi_1(t)$ be a smooth function on $\mathbb{R}$ such that

$$\Phi_1'(t) = (\Phi'(t)^4 + (\Phi(t)\Phi''(t))^2)^{\frac{1}{4}}.$$

Before we move on, note that again, from the definition of the functions g and Φ_1' (which inherits the support condition from Φ), we see that for $(\mathbf{y}_1,\mathbf{y}_2,t_1)$ with $\Phi_1'(P^{[1]}_{t_1}g(\mathbf{y}_1,\mathbf{y}_2))\neq 0$, we have that $P^{[1]}_{t_1}g(\mathbf{y}_1,\mathbf{y}_2) > C_1$. From (3.5) we see that $(\mathbf{y}_1,\mathbf{y}_2,t_1,0)\in\widetilde{W}_\beta$. Hence, there exists $\mathbf{z}_1$ such that $(\mathbf{z}_1,\mathbf{y}_2)\in E_\beta(\alpha)$ and that $|P^{[1]}_{t_1}(f)(\mathbf{y}_1,\mathbf{y}_2)|\le\mathcal{N}^\beta_P(f)(\mathbf{z}_1,\mathbf{y}_2)\le\alpha$.

Next, by repeating a similar integration by parts, we have

$$
\begin{aligned}
\mathrm{II}_{21} &\le C \iint_{\widetilde{M}} f(\mathbf{y}_1,\mathbf{y}_2)^2 \Phi(g(\mathbf{y}_1,\mathbf{y}_2))^2 d\mathbf{y}_1 d\mathbf{y}_2 \\
&\quad + C \iiint_{\widetilde{M}\times\mathbb{R}_+} P_{t_1}^{[1]}(f)(\mathbf{y}_1,\mathbf{y}_2)^2 |\nabla_{t_1,M_1}\Phi_1(P_{t_1}^{[1]}(g)(\mathbf{y}_1,\mathbf{y}_2))|^2 t_1 dt_1 d\mathbf{y}_1 d\mathbf{y}_2 \\
&\le C \iint_{E_\beta(\alpha)} \mathcal{N}_P^\beta(f)(\mathbf{y}_1,\mathbf{y}_2)^2 d\mathbf{y}_1 d\mathbf{y}_2 \\
&\quad + C\alpha^2 \iiint_{\widetilde{M}\times\mathbb{R}_+} |\nabla_{t_1,M_1}(P_{t_1}^{[1]}(g)(\mathbf{y}_1,\mathbf{y}_2))|^2 t_1 dt_1 d\mathbf{y}_1 d\mathbf{y}_2 \\
&= C \iint_{E_\beta(\alpha)} \mathcal{N}_P^\beta(f)(\mathbf{y}_1,\mathbf{y}_2)^2 d\mathbf{y}_1 d\mathbf{y}_2 \\
&\quad + C\alpha^2 \iiint_{\widetilde{M}\times\mathbb{R}_+} |(t_1\nabla_{t_1,M_1}P_{t_1}^{[1]}(1-g)(\mathbf{y}_1,\mathbf{y}_2))|^2 \frac{dt_1 d\mathbf{y}_1 d\mathbf{y}_2}{t_1} \\
&\le C \iint_{E_\beta(\alpha)} \mathcal{N}_P^\beta(f)(\mathbf{y}_1,\mathbf{y}_2)^2 d\mathbf{y}_1 d\mathbf{y}_2 + C\alpha^2 \|1-g\|_{L^2(\widetilde{M})}^2 \\
&= C \iint_{E_\beta(\alpha)} \mathcal{N}_P^\beta(f)(\mathbf{y}_1,\mathbf{y}_2)^2 d\mathbf{y}_1 d\mathbf{y}_2 + C\alpha^2 |E_\beta(\alpha)^c|,
\end{aligned}
$$

where in the second inequality we used the chain rule

$$
\nabla_{t_1,M_1}\Phi_1(P_{t_1}^{[1]}(g)(\mathbf{y}_1,\mathbf{y}_2)) = \Phi_1'(P_{t_1}^{[1]}(g)(\mathbf{y}_1,\mathbf{y}_2))\nabla_{t_1,M_1}(P_{t_1}^{[1]}(g)(\mathbf{y}_1,\mathbf{y}_2))
$$

and the inequality $|\Phi_1'(P_{t_1}^{[1]}(g)(\mathbf{y}_1,\mathbf{y}_2))| \le C$ as well as the support condition on Φ_1'.

For the term II_{22}, we apply Young's inequality to see that

$$
\begin{aligned}
\mathrm{II}_{22} \le \frac{1}{10} &\iiiint_{\widetilde{M}\times\mathbb{R}_+\times\mathbb{R}_+} |\nabla_{t_2,M_2}P_{t_1}^{[1]}P_{t_2}^{[2]}(f)(\mathbf{y}_1,\mathbf{y}_2)|^2 \\
&\qquad\qquad \times |\nabla_{t_1,M_1}\Phi(P_{t_1}^{[1]}P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2))|^2 t_1 t_2 dt_1 dt_2 d\mathbf{y}_1 d\mathbf{y}_2 \\
&+ C \iiiint_{\widetilde{M}\times\mathbb{R}_+\times\mathbb{R}_+} |P_{t_1}^{[1]}P_{t_2}^{[2]}(f)(\mathbf{y}_1,\mathbf{y}_2)|^2 \\
&\qquad\qquad \times |\nabla_{t_2,M_2}\nabla_{t_1,M_1}\Phi(P_{t_1}^{[1]}P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2))|^2 t_1 t_2 dt_1 dt_2 d\mathbf{y}_1 d\mathbf{y}_2 \\
=:\ & \mathrm{II}_{221} + \mathrm{II}_{222}.
\end{aligned}
$$

Since II_{221} can be absorbed by I_2, it suffices to estimate the term II_{222}. By the chain rule,

$$
\begin{aligned}
&\nabla_{t_2,M_2}\nabla_{t_1,M_1}\Phi(P_{t_1}^{[1]}P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2)) \\
&= \Phi''(P_{t_1}^{[1]}P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2))\nabla_{t_2,M_2}P_{t_1}^{[1]}P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2)\nabla_{t_1,M_1}P_{t_1}^{[1]}P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2) \\
&\quad + \Phi'(P_{t_1}^{[1]}P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2))\nabla_{t_2,M_2}\nabla_{t_1,M_1}P_{t_1}^{[1]}P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2).
\end{aligned}
$$

Then II_{222} can be further bounded by II_{2221} and II_{2222} with the above two integrands respectively.

Denote $\mathcal{M}_1$ and $\mathcal{M}_2$ be the Hardy–Littlewood maximal functions on M_1 and M_2, respectively. Then by the support property of Φ'',

$$
\mathrm{II}_{2221} = \iiiint_{\widetilde{M}\times\mathbb{R}_+\times\mathbb{R}_+} |P_{t_1}^{[1]}P_{t_2}^{[2]}(f)(\mathbf{y}_1,\mathbf{y}_2)|^2 |\Phi''(P_{t_1}^{[1]}P_{t_2}^{[2]}(g)(\mathbf{y}_1,\mathbf{y}_2))|^2
$$

$$
\begin{aligned}
&\times |\nabla_{t_2,M_2} P^{[1]}_{t_1} P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)|^2 |\nabla_{t_1,M_1} P^{[1]}_{t_1} P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)|^2 t_1 t_2 dt_1 dt_2 d\mathbf{y}_1 d\mathbf{y}_2\\
\le{}& C\alpha^2 \iiiint_{\widetilde{M}\times\mathbb{R}_+\times\mathbb{R}_+} |\nabla_{t_2,M_2} P^{[1]}_{t_1} P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)|^2\\
&\qquad\times |\nabla_{t_1,M_1} P^{[1]}_{t_1} P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)|^2 t_1 t_2 dt_1 dt_2 d\mathbf{y}_1 d\mathbf{y}_2\\
\le{}& C\alpha^2 \iiiint_{\widetilde{M}\times\mathbb{R}_+\times\mathbb{R}_+} \left|\mathcal{M}_1\left(\left|t_2\nabla_{t_2,M_2} P^{[2]}_{t_2}(g)\right|\right)(\mathbf{y}_1,\mathbf{y}_2)\right|^2\\
&\qquad\times \left|\mathcal{M}_2\left(\left|t_1\nabla_{t_1,M_1} P^{[1]}_{t_1}(g)\right|\right)(\mathbf{y}_1,\mathbf{y}_2)\right|^2 \frac{dt_1 dt_2 d\mathbf{y}_1 d\mathbf{y}_2}{t_1 t_2}.
\end{aligned}
$$

Applying Hölder's inequality, we further have

$$
\begin{aligned}
\mathrm{II}_{2221} \le{}& C\alpha^2 \left(\iint_{\widetilde{M}} \left(\int_{\mathbb{R}_+} \left|\mathcal{M}_1\left(\left|t_2\nabla_{t_2,M_2} P^{[2]}_{t_2}(g)\right|\right)(\mathbf{y}_1,\mathbf{y}_2)\right|^2 \frac{dt_2}{t_2}\right)^2 d\mathbf{y}_1 d\mathbf{y}_2\right)^{1/2}\\
&\times \left(\iint_{\widetilde{M}} \left(\int_{\mathbb{R}_+} \left|\mathcal{M}_2\left(\left|t_1\nabla_{t_1,M_1} P^{[1]}_{t_1}(g)\right|\right)(\mathbf{y}_1,\mathbf{y}_2)\right|^2 \frac{dt_1}{t_1}\right)^2 d\mathbf{y}_1 d\mathbf{y}_2\right)^{1/2}\\
\le{}& C\alpha^2 \left(\iint_{\widetilde{M}} \left(\int_{\mathbb{R}_+} \left|t_2\nabla_{t_2,M_2} P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)\right|^2 \frac{dt_2}{t_2}\right)^2 d\mathbf{y}_1 d\mathbf{y}_2\right)^{1/2}\\
&\times \left(\iint_{\widetilde{M}} \left(\int_{\mathbb{R}_+} \left|t_1\nabla_{t_1,M_1} P^{[1]}_{t_1}(g)(\mathbf{y}_1,\mathbf{y}_2)\right|^2 \frac{dt_1}{t_1}\right)^2 d\mathbf{y}_1 d\mathbf{y}_2\right)^{1/2}\\
={}& C\alpha^2 \left(\iint_{\widetilde{M}} \left(\int_{\mathbb{R}_+} \left|t_2\nabla_{t_2,M_2} P^{[2]}_{t_2}(1-g)(\mathbf{y}_1,\mathbf{y}_2)\right|^2 \frac{dt_2}{t_2}\right)^2 d\mathbf{y}_1 d\mathbf{y}_2\right)^{1/2}\\
&\times \left(\iint_{\widetilde{M}} \left(\int_{\mathbb{R}_+} \left|t_1\nabla_{t_1,M_1} P^{[1]}_{t_1}(1-g)(\mathbf{y}_1,\mathbf{y}_2)\right|^2 \frac{dt_1}{t_1}\right)^2 d\mathbf{y}_1 d\mathbf{y}_2\right)^{1/2}\\
\le{}& C\alpha^2 \|1-g\|^2_{L^2(\widetilde{M})}\\
={}& C\alpha^2\, |E_\beta(\alpha)^c|,
\end{aligned}
$$

where the second inequality we used the vector-valued inequality for the Hardy–Littlewood maximal functions and the last inequality we applied the Littlewood–Paley theory. Next it follows from the support condition on Φ' that for $(\mathbf{y}_1,\mathbf{y}_2,t_1,t_2)$ satisfying $\Phi'(P^{[1]}_{t_1} P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)) \neq 0$, we have $|P^{[1]}_{t_1} P^{[2]}_{t_2}(f)(\mathbf{y}_1,\mathbf{y}_2)| \le \alpha$.

As a consequence, we have

$$
\begin{aligned}
\mathrm{II}_{2222} ={}& \iiiint_{\widetilde{M}\times\mathbb{R}_+\times\mathbb{R}_+} |P^{[1]}_{t_1} P^{[2]}_{t_2}(f)(\mathbf{y}_1,\mathbf{y}_2)|^2 |\Phi'(P^{[1]}_{t_1} P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))|^2\\
&\times |\nabla_{t_2,M_2}\nabla_{t_1,M_1} P^{[1]}_{t_1} P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)|^2 t_1 t_2 dt_1 dt_2 d\mathbf{y}_1 d\mathbf{y}_2\\
\le{}& C\alpha^2 \iiiint_{\widetilde{M}\times\mathbb{R}_+\times\mathbb{R}_+} |t_1\nabla_{t_1,M_1} P^{[1]}_{t_1} t_2\nabla_{t_2,M_2} P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)|^2 \frac{dt_1 dt_2 d\mathbf{y}_1 d\mathbf{y}_2}{t_1 t_2}\\
={}& C\alpha^2 \iiiint_{\widetilde{M}\times\mathbb{R}_+\times\mathbb{R}_+} |t_1\nabla_{t_1,M_1} P^{[1]}_{t_1} t_2\nabla_{t_2,M_2} P^{[2]}_{t_2}(1-g)(\mathbf{y}_1,\mathbf{y}_2)|^2 \frac{dt_1 dt_2 d\mathbf{y}_1 d\mathbf{y}_2}{t_1 t_2}
\end{aligned}
$$

$$
\begin{aligned}
&\le C\alpha^2\|1-g\|^2_{L^2(\widetilde{M})}\\
&= C\alpha^2\left|E_\beta(\alpha)^c\right|,
\end{aligned}
$$

where in the second inequality we applied the Littlewood–Paley theory again.

Similar to the estimate of II_{222}, we obtain that

$$
\mathrm{II}_{23}\le C\alpha^2\left|\{(\mathbf{y}_1,\mathbf{y}_2)\in\widetilde{M}:\ \mathcal{N}_P^\beta(f)(\mathbf{y}_1,\mathbf{y}_2)>\alpha\}\right|.
$$

Finally, we turn to estimate II_{24}. By the chain rule,

$$
\begin{aligned}
&\nabla_{t_1,M_1}\Phi(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))\nabla_{t_2,M_2}\nabla_{t_2,M_2}\nabla_{t_1,M_1}\Phi(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))\\
&=(\Phi'\Phi''')(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))|\nabla_{t_1,M_1}P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)|^2|\nabla_{t_2,M_2}P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)|^2\\
&\quad+2(\Phi'\Phi'')(P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2))\nabla_{t_1,M_1}P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)\nabla_{t_2,M_2}P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)\\
&\qquad\times\nabla_{t_1,M_1}\nabla_{t_2,M_2}P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2).
\end{aligned}
$$

Thus, II_{24} can be dominated by $\mathrm{II}_{241}+\mathrm{II}_{242}$ with respect to the above two terms in the integrand respectively. Using the same argument, we get that

$$
\mathrm{II}_{241}\le C\alpha^2\left|E_\beta(\alpha)^c\right|.
$$

Next, it follows from the support property of $\Phi'\Phi''$ and Hölder's inequality that

$$
\begin{aligned}
\mathrm{I}_{242}&\le C\alpha^2\iiiint_{\widetilde{M}\times\mathbb{R}_+\times\mathbb{R}_+}|\nabla_{t_1,M_1}P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)|\,|\nabla_{t_2,M_2}P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)|\\
&\qquad\times|\nabla_{t_1,M_1}\nabla_{t_2,M_2}P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)|t_1t_2dt_1dt_2d\mathbf{y}_1d\mathbf{y}_2\\
&\le C\alpha^2\bigg(\iiiint_{\widetilde{M}\times\mathbb{R}_+\times\mathbb{R}_+}|\nabla_{t_1,M_1}P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)|^2\\
&\qquad\qquad\times|\nabla_{t_2,M_2}P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)|^2t_1t_2dt_1dt_2d\mathbf{y}_1d\mathbf{y}_2\bigg)\\
&\quad+C\alpha^2\bigg(\iiiint_{\widetilde{M}\times\mathbb{R}_+\times\mathbb{R}_+}|\nabla_{t_1,M_1}\nabla_{t_2,M_2}P^{[1]}_{t_1}P^{[2]}_{t_2}(g)(\mathbf{y}_1,\mathbf{y}_2)|^2t_1t_2dt_1dt_2d\mathbf{y}_1d\mathbf{y}_2\bigg)\\
&\le C\alpha^2\left|E_\beta(\alpha)^c\right|,
\end{aligned}
$$

where the last inequality follows from the estimates of the terms II_{2221} and II_{2222}, that is, via taking the vector-valued Hardy–Littlewood maximal function estimate and then the Littlewood–Paley estimates.

Combining these estimates together, we conclude that

$$
\iint_{A_\beta(\alpha)}S_P(f)(\mathbf{x}_1,\mathbf{x}_2)^2d\mathbf{x}_1d\mathbf{x}_2\le C\iint_{E_\beta(\alpha)}\mathcal{N}_P^\beta(f)(\mathbf{y}_1,\mathbf{y}_2)^2d\mathbf{y}_1d\mathbf{y}_2+C\alpha^2\left|E_\beta(\alpha)^c\right|.
$$

Hence, combining (3.4) and the above estimate, we see that

$$
\begin{aligned}
&\left|\{(\mathbf{y}_1,\mathbf{y}_2)\in\widetilde{M}:\ S_P(f)(\mathbf{y}_1,\mathbf{y}_2)>\alpha\}\right|\\
&\le C\left|\{(\mathbf{y}_1,\mathbf{y}_2)\in\widetilde{M}:\ \mathcal{N}_P^\beta(f)(\mathbf{y}_1,\mathbf{y}_2)>\alpha\}\right|+\frac{C}{\alpha^2}\iint_{E_\beta(\alpha)}\mathcal{N}_P^\beta(f)(\mathbf{y}_1,\mathbf{y}_2)^2d\mathbf{y}_1d\mathbf{y}_2.
\end{aligned}
$$

The proof of Theorem 1.1 is complete.

4 Applications

We address two direct applications of Theorem 1.1.

4.1 Weak endpoint estimate for Cauchy–Szegő projection on $\widetilde{M}$

We recall that in the well-known result of Diaz [12] (see also recent result [2]), the explicit pointwise size estimate and regularity estimate of the Cauchy–Szegő kernel on M is given as follows.

Theorem A. *For $\mathbf{x}$ and $\mathbf{y}$ in M with $\mathbf{x} \neq \mathbf{y}$, the Cauchy–Szegő projection $\mathbf{S}$ associated with the kernel $S(\mathbf{x}, \mathbf{y})$ is a Calderón–Zygmund operator, i.e.,*

$$|S(\mathbf{x}, \mathbf{y})| \approx \frac{1}{V(\mathbf{x}, \mathbf{y})};$$

there is $\epsilon > 0$ such that for $\mathbf{x} \neq \mathbf{x}'$ and for $d(\mathbf{x}, \mathbf{x}') \leq cd(\mathbf{x}, \mathbf{y})$ with some small positive constant c,

$$|S(\mathbf{x}, \mathbf{y}) - S(\mathbf{x}', \mathbf{y})| \leq C_1 \frac{1}{V(\mathbf{x}, \mathbf{y})} \left(\frac{d(\mathbf{x}, \mathbf{x}')}{d(\mathbf{x}, \mathbf{y})}\right)^{\epsilon};$$

for $\mathbf{y} \neq \mathbf{y}'$ and for $d(\mathbf{y}, \mathbf{y}') \leq cd(\mathbf{x}, \mathbf{y})$ with some small positive constant c,

$$|S(\mathbf{x}, \mathbf{y}) - S(\mathbf{x}, \mathbf{y}')| \leq C_1 \frac{1}{V(\mathbf{x}, \mathbf{y})} \left(\frac{d(\mathbf{y}, \mathbf{y}')}{d(\mathbf{x}, \mathbf{y})}\right)^{\epsilon}$$

with some constant $C_1 > 0$.

Consider $\widetilde{M} = M_1 \times M_2$. The Cauchy–Szegő projection $\widetilde{\mathbf{S}} = \mathbf{S}_1 \circ \mathbf{S}_2$ is a product Calderón–Zygmund operator on $L^2(\widetilde{M})$, where $\mathbf{S}_j$ is the Cauchy–Szegő projection on M_j for $j = 1, 2$. The general framework of product Calderón–Zygmund operator on space of homogeneous type was studied in [18].

Following the framework in [11], we see that Theorem 1.1 gives rise to the weak endpoint estimate of $S_P(f)(\mathbf{x}_1, \mathbf{x}_2)$. To be more explicit, following [8] we first see that there is $C > 0$ such that for all $\lambda > 0$,

$$|\{(\mathbf{x}_1, \mathbf{x}_2) \in \widetilde{M} : \mathcal{M}(f)(\mathbf{x}_1, \mathbf{x}_2) > \lambda\}| \leq C \|\lambda^{-1} f\|_{L \log L(\widetilde{M})}, \tag{4.1}$$

where

$$\|f\|_{L \log L(\widetilde{M})} = \iint_{\widetilde{M}} |f(\mathbf{x}_1, \mathbf{x}_2)| \log(e + |f(\mathbf{x}_1, \mathbf{x}_2)|) d\mathbf{x}_1 d\mathbf{x}_2.$$

See also [11, Section 3] for the proof of (4.1) for strong maximal function on product Lie groups, where the arguments can be modified to our setting $\widetilde{M} = M_1 \times M_2$.

Then Theorem 1.1 yields that $|\{(\mathbf{x}_1, \mathbf{x}_2) \in \widetilde{M} : S_P(f)(\mathbf{x}_1, \mathbf{x}_2) > \lambda\}| \leq C \|\lambda^{-1} f\|_{L \log L(\widetilde{M})}$, and hence we have the atomic decomposition of $L \log L(\widetilde{M})$, which further gives

$$|\{(\mathbf{x}_1, \mathbf{x}_2) \in \widetilde{M} : \widetilde{\mathbf{S}}(f)(\mathbf{x}_1, \mathbf{x}_2) > \lambda\}| \leq C \|\lambda^{-1} f\|_{L \log L(\widetilde{M})}. \tag{4.2}$$

We also note that this weak type endpoint estimate (4.2) also holds for the general product Calderón–Zygmund operator on $\widetilde{M}$ (we refer to the full definition of the product non-isotropic smooth operators on $\widetilde{M}$ as introduced by Nagel and Stein [30], as well as the Journé type product Calderón–Zygmund operators [21, 18]).

Thus, (4.2) is also true for the Marcinkiewicz multiplier $m(\Box_b^{(1)}, \Box_b^{(2)})$ acting on the Shilov boundary $\widetilde{M}$, where $\Box_b^{(j)}$ is the Kohn Laplacian on M_j for $j = 1, 2$. (Note that in [4] we proved that under suitable assumptions on the multiplier function m, $m(\Box_b^{(1)}, \Box_b^{(2)})$ is a product Calderón–Zygmund operator of Journé type).

4.2 Maximal function characterisation for product Hardy space on $\widetilde{M}$

For the sake of simplicity, we take Hardy space $H^1(\widetilde{M})$ as an application. The same argument holds for $H^p(\widetilde{M})$ for $p < 1$ with a more complicated but very standard definition via distributions (see for example the test function space and distributions as in [19]).

We now define

$$H^1_{S_P}(\widetilde{M}) = \{f \in L^1(\widetilde{M}) : S_P(f) \in L^1(\widetilde{M})\}. \tag{4.3}$$

Following the Plancherel–Polya inequality in [19], we see that $H^1_{S_P}(\widetilde{M})$ is equivalent to the Hardy space $H^1(\widetilde{M})$ in [19] given via discrete square function (or discrete area functions) using the discrete reproducing formula via approximation to identity. Thus, $H^1_{S_P}(\widetilde{M})$ has atomic decomposition, which we refer to [20].

Next, we define

$$H^1_{max}(\widetilde{M}) = \{f \in L^1(\widetilde{M}) : \mathcal{N}_P^\beta(f) \in L^1(\widetilde{M})\}. \tag{4.4}$$

Then we have the following equivalence characterisation.

Proposition 4.1. *$H^1_{S_P}(\widetilde{M})$ coincides with $H^1_{max}(\widetilde{M})$ and they have equivalent norms.*

Proof. Suppose $f \in H^1_{S_P}(\widetilde{M}) \cap L^2(\widetilde{M})$, then following [20] f has an atomic decomposition $f = \sum_j \lambda_j a_j$ where each a_j is a product atom and $\sum_j |\lambda_j| \approx \|f\|_{H^1_{S_P}(\widetilde{M})}$. Thus, it suffices to show that $\|\mathcal{N}_P^\beta(a_j)\|_{L^1(\widetilde{M})} \leq C$ for every atom a_j, where C is an absolute constant. This is a standard argument, which follows from the size and cancellation of a_j, the size and regularity estimates of the Poisson kernels $P_{t_1}^{[1]}$ and $P_{t_2}^{[2]}$, and Journé's covering lemma [32]. Thus, we have $\|f\|_{H^1_{max}(\widetilde{M})} \leq C\|f\|_{H^1_{S_P}(\widetilde{M})}$. Then via the density of $H^1_{S_P}(\widetilde{M}) \cap L^2(\widetilde{M})$ in $H^1_{S_P}(\widetilde{M})$, we see that $H^1_{S_P}(\widetilde{M}) \subset H^1_{max}(\widetilde{M})$.

Then reverse direction $\|f\|_{H^1_{S_P}(\widetilde{M})} \leq C\|f\|_{H^1_{max}(\widetilde{M})}$ follows from Theorem 1.1 and hence $H^1_{max}(\widetilde{M}) \subset H^1_{S_P}(\widetilde{M})$.

The proof is complete. □

Proposition 4.1 provides the maximal function characterisation of $H^1(\widetilde{M})$.

Acknowledgement: Ji Li would like to thank Prof. Michael Cowling and Dr. Liangchuan Wu for suggestions. Ji Li is supported by ARC DP 220100285.

References

[1] D.-C. Chang, J. Li, A. Ottazzi and Q. Wu, *Optimal lifting of Levi-degenerate hypersurfaces and applications to the Cauchy–Szegő projection*, arXiv:2309.00525. 2

[2] D.-C. Chang, J. Li, J. Tie and Q. Wu, *The Kohn–Laplacian and Cauchy–Szegő projection on model domains*, Ann. Math. Sci. Appl., **8** (2023), no. 1, 111–155. 2, 17

[3] D.-C. Chang, A. Nagel and E.M. Stein, *Estimates for the $\bar{\partial}$-Neumann problem in pseudoconvex domains of finite type in $\mathbb{C}^2$*, Acta Math., **169** (1992), 153–228. 2

[4] P. Chen, M. Cowling, G. Hu and J. Li, *Marcinkiewicz multipliers associated with the Kohn Laplacian on the Shilov boundary of the product domain in $\mathbb{C}^{2n}$*, Math. Z., **300** (2022), no. 1, 347–376. 18

[5] P. Chen, M.G. Cowling, M.Y. Lee, J. Li and A. Ottazzi, *Flag Hardy space theory on Heisenberg groups and applications*, arXiv:2102.07371. 2

[6] P. Chen, X. Duong, J. Li, L.X. Yan and L. Ward, *Product Hardy spaces associated to operators with heat kernel bounds on spaces of homogeneous type*, Math. Z., **282** (2016), 1033–1065. 2

[7] M. Christ, *Regularity properties of the $\bar{\partial}_b$-equation on weakly pseudoconvex CR manifolds of dimension 3*, J. Amer. Math. Soc., **1** (1988), 587–646. 2

[8] A. Córdoba and R. Fefferman, *A geometric proof of the strong maximal theorem*, Ann. of Math., **102** (1975), 95–100. 17

[9] R. Coifman and G. Weiss, *Analyse harmonique non-commutative sur certains espaces homogènes*, Lecture Notes in Math. **242**, Springer, Berlin, 1971. 2

[10] M.G. Cowling, Z. Fan, J. Li and L. Yan, *Characterizations of product Hardy spaces on stratified groups by singular integrals and maximal functions*, arXiv:2210.01265. 2

[11] M.G. Cowling, M. Lee, J. Li and J. Pipher, *Endpoint estimate of product singular integral on stratified Lie groups*, arXiv:2312.15920. 3, 17

[12] K.P. Diaz, *The Sezgö kernel as a singular integral kernel on a family of weakly pseudoconvex domains*, Trans. Amer. Math. Soc., **304** (1987), 147–170. 2, 17

[13] X. Duong, J. Li, B.D. Wick and D. Yang, *Characterizations of product Hardy spaces in Bessel setting*, J. Fourier Anal. Appl., **27** (2021), no. 2, Paper No. 24, 65 pp. 2

[14] C. Fefferman and E.M. Stein, *H^p spaces of several variables*, Acta Math., **129** (1972), 137–193. 1, 2

[15] R. Fefferman, *A note on a lemma of Zo*, Proc. Amer. Math. Soc., **96** (1986), 241–246. 3

[16] R. Gundy and E. M. Stein, *H^p theory for the polydisc*, Proc. Nat. Acad. Sci., **76** (1979), 1026–1029. 1, 2

[17] Y. Han, M.Y. Lee, J. Li and B.D. Wick, *Maximal function, Littlewood-Paley theory, Riesz transforms and atomic decomposition in the multi parameter flag setting*, Mem. Amer. Math. Soc., **279** (2022), no. 1373. 2

[18] Y. Han, J. Li and C. Lin, *Criterion of the L^2 boundedness and sharp endpoint estimates for singular integral operators on product spaces of homogeneous type*, Ann. Sc. Norm. Super. Pisa Cl. Sci., (5) **16** (2016), no. 3, 845–907. 17, 18

[19] Y. Han, J. Li and G. Lu, *Multiparameter Hardy space theory on Carnot–Carathéodory spaces and product spaces of homogeneous type*, Trans. Amer. Math. Soc., **365** (2013), no. 1, 319–360. 2, 3, 18

[20] Y. Han, J. Li, C. Pereyra and L. Ward, *Atomic decomposition of product Hardy spaces via wavelet bases on spaces of homogeneous type*, New York J. Math., **27** (2021), 1173–1239. 2, 3, 18

[21] J. L. Journé, *Calderón–Zygmund operators on product space*, Rev. Mat. Iberoam. **1** (1985), 55–92. 18

[22] K. Merryfield, *On the area integral, Carleson measures and H^p in the polydisc*, Indiana Univ. Math. J., **34** (1985), 663–685. 1, 2, 3

[23] 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. 2

[24] A. Nagel, F. Ricci, E.M. Stein and S. Wainger, *Singular integrals with flag kernels on homogeneous groups, I,* Rev. Mat. Iberoam. **28** (2012), 631–722. 2

[25] A. Nagel, F. Ricci, E.M. Stein and S. Wainger, *Algebras of singular integral operators with kernels controlled by multiple norms*, Mem. Amer. Math. Soc. **256** (2018), no. 1230, vii+141 pp. 2

[26] A. Nagel, J. P. Rosay, E. M. Stein and S. Wainger, *Estimates for the Bergman and Szegö kernels in $\mathbb{C}^2$*, Ann. of Math., **129** (1989), 113–149. 2

[27] A. Nagel and E. M. Stein, *The $\Box_b$-Heat equation on pseudoconvex manifolds of finite type in $\mathbb{C}^2$*, Math. Z., **238** (2001), 37–88. 2, 4

[28] A. Nagel and E. M. Stein, *Differentiable control metrics and scaled bump functions*, J. Differential Geom., **57** (2001), 37–88. 4

[29] A. Nagel and E. M. Stein, *The $\bar{\partial}_b$-complex on decoupled boundaries in $\mathbb{C}^n$*, Ann. of Math., **164** (2006), 649–713. 2

[30] A. Nagel and E. M. Stein, *On the product theory of singular integrals*, Rev. Mat. Iberoam., **20** (2004), 531–561. 1, 2, 4, 18

[31] A. Nagel, E. M. Stein and S. Wainger, *Balls and metrics defined by vector fields I. Basic properties*, Acta Math., **155** (1985), 103–147. 4

[32] J. Pipher, *Journé's covering lemma and its extension to higher dimensions*, Duke Mathematical Journal, **53** (1986), 683–690. 18

[33] L. Rothschild and E.M. Stein, *Hypoelliptic differential operators and nilpotent groups*, Acta Math., **137** (1976), 247–320. 3

[34] B. Street, *The $\Box_b$ heat equation and multipliers via the wave equation*, Math. Z., **263** (2009), 861–886. 2, 4

[35] E.M. Stein, On the functions of Littlewood-Paley, Lusin and Marcinkiewicz, Trans. Amer. Math. Soc., **88** (1958), 430–466. 1

School of Mathematical and Physical Sciences, Macquarie University, NSW 2109, Australia
E-mail address: ji.li@mq.edu.au