

Weak type $(1, 1)$ of Riesz transform on some direct product manifolds with exponential volume growth

Hong-Quan Li^{1,*} & Jie-Xiang Zhu²

¹*School of Mathematical Sciences, Fudan University, Shanghai 200433, China;*

²*Center for Applied Mathematics, Tianjin University, Tianjin 300072, China*

Email: hongquan.li@fudan.edu.cn, 15110840006@fudan.edu.cn

Received November 21, 2022; accepted April 4, 2023; published online 2023

Abstract In this paper we are concerned with the Riesz transform on the direct product manifold $\mathbb{H}^n \times M$, where $\mathbb{H}^n$ is the n -dimensional real hyperbolic space and M is a connected complete non-compact Riemannian manifold satisfying the volume doubling property and generalized Gaussian or sub-Gaussian upper estimates for the heat kernel. We establish its weak type $(1, 1)$ property. In addition, we obtain the weak type $(1, 1)$ of the heat maximal operator in the same setting. Our arguments also work for a large class of direct product manifolds with exponential volume growth. Particularly, we provide a simpler proof of weak type $(1, 1)$ boundedness of some operators considered in the work of Li, Sjögren and Wu [27].

Keywords Riesz transform, heat kernel, direct product, real hyperbolic space

MSC(2020) 42B20, 58J35

Citation: Hong-Quan Li, Jie-Xiang Zhu. Weak type $(1, 1)$ of Riesz transform on some direct product manifolds with exponential volume growth. *Sci China Math*, 2022, 65, <https://doi.org/10.1007/s11425-016-5135-4>

1 Introduction and main results

The weak $(1, 1)$ and L^p boundedness of Riesz transform are one of the classic topics in harmonic analysis. In the last several decades, they have been investigated extensively in the setting of Riemannian manifolds. To formulate these problems, we first list the following terminologies and notations that will be used throughout the whole paper. Let (M, g) be a connected complete non-compact Riemannian manifold. Let us denote by $d\mu_M$ the Riemannian measure and by d_M the geodesic distance. Denote by $B(x, R)$ the open ball of center x and of radius $R > 0$. Set $V_M(x, R) = \mu_M(B(x, R))$. Let ∇_M be the Riemannian gradient, Δ_M be the Laplace-Beltrami operator on M . Denote by $(e^{t\Delta_M})_{t>0}$ the heat semigroup associated with Δ_M and $p_t^M(x, y)$ the heat kernel. By spectral analysis,

$$(-\Delta_M)^{-\frac{1}{2}} = \frac{1}{\sqrt{\pi}} \int_0^\infty e^{t\Delta_M} \frac{dt}{\sqrt{t}}.$$

Then the Riesz transform $\mathcal{R}$ on M is defined by $\nabla_M(-\Delta_M)^{-\frac{1}{2}}$. In the sequel, we use the notation $A \lesssim B$ if there exists a universal constant $C > 0$ such that $A \leq CB$, where the constant C may only depend on

* Corresponding author

the properties of the underlying manifolds. We write $A \sim B$, if $A \lesssim B$ and $B \lesssim A$. We also use $A \lesssim_\varepsilon B$ to denote $A \leq C_\varepsilon B$, where the constant C_ε depends on the parameter ε . In the present work, we shall consider the weak type $(1, 1)$ property of the Riesz transform on M , namely

$$\mu_M \left(\left\{ x; |\nabla_M (-\Delta_M)^{-\frac{1}{2}} f(x)| > \lambda \right\} \right) \lesssim \frac{\|f\|_1}{\lambda}, \quad \forall \lambda > 0, f \in C_0^\infty(M); \quad (1.1)$$

and its L^p boundedness properties, namely, for which $1 < p < +\infty$,

$$\|\nabla_M (-\Delta_M)^{-\frac{1}{2}} f\|_p \lesssim \|f\|_p, \quad \forall f \in C_0^\infty(M). \quad (1.2)$$

Since the case $p = 2$ is a direct result of integration by parts, then (1.2) for $1 < p < 2$ follows from (1.1) by interpolation. The classical singular integral operator theory says that on the Euclidean space, (1.1) and (1.2) are valid for all $1 < p < +\infty$ (see e.g. [32, Chapter III]). In contrast with the classical result, for any given $p_0 > 2$, there exist stochastically complete Riemannian manifolds, on which (1.2) is valid iff $1 < p < p_0$, and even iff $1 < p \leq 2$ (see e.g. [13, 24] for more details). It is worthwhile to mention that there are various recent counterexamples in the literature. The appendix below contains a brief discussion of (1.2) for $p > 2$ in our setting. In this paper we mainly focus on (1.1).

1.1 Volume doubling property and heat kernel upper estimates

For the further purpose, it is worthwhile discussing some basic assumptions on M and p_t^M . We say that M satisfies the volume doubling property if there exists $D > 0$, such that

$$V_M(x, 2R) \leq D V_M(x, R), \quad \forall x \in M, R > 0. \quad (D)$$

Given $m \geq 2$. Suppose that there exists some $c \in (0, 1)$, for all $x, y \in M$,

$$p_t^M(x, y) \leq c^{-1} \begin{cases} \frac{1}{V_M(x, \sqrt{t})} \exp \left(-c \frac{d_M(x, y)^2}{t} \right), & \text{if } 0 < t < 1; \\ \frac{1}{V_M(x, t^{\frac{1}{m}})} \exp \left(-c \left(\frac{d_M(x, y)^m}{t} \right)^{\frac{1}{m-1}} \right), & \text{if } t \geq 1. \end{cases} \quad (UE_m)$$

The heat kernel estimate (UE_2) (respectively, (UE_m) for $m > 2$) is called Gaussian upper estimate (respectively, sub-Gaussian upper estimate with the exponent m). (UE_2) is also known as Li-Yau upper estimate. If M satisfies (D) , then (UE_2) is equivalent to the relative Faber-Krahn inequality (see e.g. [22, Chapter 15] for more details). While (UE_m) for $m > 2$ holds on some fractal-like manifolds (see [11]).

Several different generalized forms of (UE_m) have been considered by other authors (see, e.g., [8, 9, 19]). However, we need to introduce a new one in our context. As we know, (UE_m) is unfortunately not stable under products, i.e. provided that M_1 and M_2 are connected complete non-compact Riemannian manifolds satisfying (UE_{m_1}) and (UE_{m_2}) with $m_1 \neq m_2$ respectively, then (UE_m) does not hold for any $m \geq 2$ on $M_1 \times M_2$ in general (see [28, §2]). On the other hand, when we investigate Riesz transform on product manifolds, sometimes the geodesic distance does not directly play a role so that we have to define other distances (or even quasi-distances) and work within the new framework (see [28]). Let $\tilde{d}$ denote a quasi-distance in the sense of Coifman-Weiss on M (see [12]). Besides we assume that $\tilde{d}$ induces the topology of M . Set

$$\tilde{B}(x, R) = \{y \in M; \tilde{d}(x, y) < R\}, \quad \tilde{V}_M(x, R) = \mu_M(\tilde{B}(x, R)), \quad \forall x \in M, R > 0.$$

We naturally hope our generalization is compatible with $\tilde{d}$. These facts lead us to the following definition. Given $0 < \theta \leq 1$, we say that M satisfies the (so-called) generalized heat kernel upper estimates with the exponent θ w.r.t. the quasi-distance $\tilde{d}$, if there exists a constant $c \in (0, 1)$ such that

$$p_t^M(x, y) \leq c^{-1} \frac{1}{\tilde{V}_M(x, \sqrt{t})} \exp \left(-c \left(\frac{\tilde{d}(x, y)^2}{t} \right)^\theta \right), \quad \forall t > 0, x, y \in M. \quad (GUE_\theta)$$

In the sequel, to avoid confusing of notations, we rewrite the volume doubling property of $(M, \tilde{d}, d\mu_M)$ as (for some constant $\tilde{D} > 0$)

$$\tilde{V}_M(x, 2R) \leq \tilde{D} \tilde{V}_M(x, R), \quad \forall x \in M, R > 0. \quad (ND)$$

Note that (ND) implies the following stronger property: there exist positive constants D' and v , such that

$$\tilde{V}_M(x, R_1) \leq D' \left(1 + \frac{R_1}{R_2}\right)^v \tilde{V}_M(x, R_2), \quad \forall x \in M, R_1, R_2 > 0. \quad (1.3)$$

In addition, the following assumptions on weighted estimates of $\nabla_M p_t^M(x, y)$ ¹⁾ will be requested. Given $1 \leq q \leq 2$ and $0 < \theta \leq 1$, we say that M satisfies weighted gradient integral estimates of the heat kernel with the exponent (q, θ) w.r.t. $\tilde{d}$, if there exist constants $c, c' > 0$ such that for all $t > 0, y \in M$,

$$\int_M |\nabla_M p_t^M(x, y)|^q \exp\left(c \left(\frac{\tilde{d}(x, y)^2}{t}\right)^\theta\right) d\mu_M(x) \leq \frac{c'}{t^{\frac{q}{2}} \tilde{V}_M(y, \sqrt{t})^{q-1}}. \quad (WGE_{q,\theta})$$

It is worth noting that in the case $\tilde{d} = d_M$, $(WGE_{2,1})$ is deduced from (ND) and (GUE_1) , see e.g. [13]. It has recently been pointed out implicitly in [11] that when $\tilde{d} = \max\{d_M, d_M^{\frac{m}{2}}\}$ for some $m > 2$, then $(WGE_{q, \frac{1}{m-1}})$ ($q \in (1, 2)$) is a corollary of (ND) and $(GUE_{\frac{1}{m-1}})$. But the above two conclusions are based on the fact that $|\nabla_M d_M| \leq 1$, which is obviously not valid for a general quasi-distance $\tilde{d}$. It should be reasonable to conjecture that, in general, $(WGE_{q,\theta})$ is not a direct consequence of the conjunction of (ND) and (GUE_θ) . That is the reason we list it here as an independent assumption. Section 2 contains a detailed discussion of (GUE_θ) and $(WGE_{q,\theta})$.

1.2 Some closely related known results about Riesz transform

In the past several decades, a great deal of mathematical effort has been devoted to (1.1) and (1.2) on various manifolds. Here we only mention a few early works [1, 2, 5, 6, 29, 30, 35]. Meanwhile, people try to establish (1.1) under as few assumptions as possible, and the volume doubling property (D) is among the most essential ones. Since (D) implies the weak $(1, 1)$ estimate of Hardy-Littlewood maximal operator and the Calderón-Zygmund decomposition of L^1 functions, which are convenient tools in harmonic analysis. In this regard, a seminal result has been obtained by T. Coulhon and X. T. Duong in [13], using singular integral theory developed in [19]. More precisely, they established (1.1) under (D) and (UE_2) . And recent progress can be found in [11] and [28]. To be more precise, (1.1) has been proved under the assumption that (D) and (UE_m) in [11], also under the weaker assumption that (ND) , (GUE_θ) and $(WGE_{1,\theta})$ in [28].

However, there is a wide variety of manifolds not satisfying the volume doubling property. For example, manifolds with both exponential volume growth and a positive bottom of the spectrum, on which we will concentrate. Compared with doubling manifolds, the main issue in this setting is the lack of an adequate singular integral theory. Nevertheless (1.1) has been obtained in the following few settings: non-compact symmetric spaces (see [2]), harmonic AN groups (see [3]), the Laplacian with drift on Euclidean spaces (see [27]) and on real hyperbolic spaces in the sense of [25], as well as the sub-Laplacian with drift on Heisenberg groups in the context of [26].

1.3 Real hyperbolic space

The real hyperbolic space $\mathbb{H}^n$ of dimension $n \geq 2$ is a typical manifold with exponential volume growth and a positive bottom of the spectrum. We hereafter adopt the upper half-space representation (see, e.g., [16, §5.7]): that is $\mathbb{H}^n = \mathbb{R}^{n-1} \times \mathbb{R}^+$ endowed with the metric $g_{\mathbb{H}^n} = u_n^{-2} \sum_{i=1}^n (du_i)^2$. The associated

¹⁾ Hereafter, for any smooth function $u(x, y)$ defined on $M \times M$, we adopt the shorthand notation $\nabla_M u(x, y) := \nabla_{M,x} u(x, y)$.

Riemannian measure is $d\mu_{\mathbb{H}^n}(u) = u_n^{-n} du$, where du is the Lebesgue measure. In this representation, the geodesic distance r between $u = (u_1, \dots, u_n), v = (v_1, \dots, v_n) \in \mathbb{R}^{n-1} \times \mathbb{R}^+$ can be computed as

$$r(u, v) = \operatorname{arccosh} \frac{\sum_{i=1}^{n-1} (u_i - v_i)^2 + u_n^2 + v_n^2}{2u_n v_n}.$$

The gradient of $f \in C^\infty(\mathbb{H}^n)$ is written as

$$\nabla_{\mathbb{H}^n} f = \left(u_n \frac{\partial f}{\partial u_1}, \dots, u_n \frac{\partial f}{\partial u_n} \right),$$

and the Laplace-Beltrami operator is given by

$$\Delta_{\mathbb{H}^n} f = u_n^2 \sum_{i=1}^n \frac{\partial^2 f}{\partial u_i^2} - (n-2)u_n \frac{\partial f}{\partial u_n}.$$

In Riemannian Geometry, it is well-known that the real hyperbolic space is homogeneous, the volume of geodesic ball of radius R is independent of its center, and

$$V_{\mathbb{H}^n}(u, R) \sim \begin{cases} R^n & \text{if } 0 < R \leq 1, \\ e^{(n-1)R} & \text{if } R > 1. \end{cases} \quad (1.4)$$

Consequently, $\mathbb{H}^n$ satisfies the following local volume doubling property in the sense that for any fixed $R_0 > 0$, there exists a positive constant D_{R_0} , such that

$$V_{\mathbb{H}^n}(u, 2R) \leq D_{R_0} V_{\mathbb{H}^n}(u, R), \quad \forall u \in \mathbb{H}^n, R \in (0, R_0). \quad (ND_{\text{loc}})$$

As for the behavior at infinity, its volume growth is at most exponential in the sense that there exist positive constants C and c such that

$$V_{\mathbb{H}^n}(u, \theta R) \leq C e^{c\theta} V_{\mathbb{H}^n}(u, R), \quad \forall u \in \mathbb{H}^n, \theta > 1, R \in (0, 1]. \quad (VE)$$

Obviously the properties (ND_{loc}) and (VE) can be defined on general (quasi-)metric measure spaces.

For the Riesz transform on $\mathbb{H}^n$, (1.2) for $1 < p < +\infty$ and (1.1) are known results.

1.4 Main results

We shall consider a class of product Riemannian manifolds $\mathcal{M}$ in this paper. $\mathcal{M}$ is obtained by taking the direct product of $\mathbb{H}^n$ and a connected complete non-compact Riemannian manifold M equipped with a quasi-distance $\tilde{d}$. Besides we always assume M satisfies (ND) , (GUE_θ) and $(WGE_{q,\theta})$ (with some $1 \leq q \leq 2$ and $0 < \theta \leq 1$) w.r.t. $\tilde{d}$.

It is well-known that the heat kernel $p_t^{\mathcal{M}}$ on $\mathcal{M}$ is given by (see e.g. [22])

$$p_t^{\mathcal{M}}(X, Y) = p_t^{\mathbb{H}^n}(u, v) p_t^M(x, y), \quad \forall t > 0, X = (u, x), Y = (v, y) \in \mathbb{H}^n \times M. \quad (1.5)$$

For our purpose, a quasi-distance d on $\mathcal{M}$ is defined as follows:

$$d(X, Y) = \max\{r(u, v), \tilde{d}(x, y)\}, \quad X = (u, x), Y = (v, y) \in \mathbb{H}^n \times M. \quad (1.6)$$

Note that d differs from the geodesic distance on $\mathcal{M}$ considered in differential geometry. In the sequel, for simplicity, we will omit coordinates of points in distance functions if they are self-evident. The Riemannian measure $d\mu_{\mathcal{M}}$ on $\mathcal{M}$ satisfies $d\mu_{\mathcal{M}} = d\mu_{\mathbb{H}^n} \otimes d\mu_M$. Then (ND_{loc}) of $\mathbb{H}^n$ and (ND) of M imply that $(\mathcal{M}, d, d\mu_{\mathcal{M}})$ satisfies (ND_{loc}) . As for the gradient of $f \in C^1(\mathcal{M})$, we have $\nabla_{\mathcal{M}} f = (\nabla_{\mathbb{H}^n} f, \nabla_M f)$.

The motivation for this work comes from two aspects. Firstly, recalling all the existing weak type $(1, 1)$ results on manifolds with exponential volume growth, whose bottom of the spectrum is positive (see [2, 3, 25–27]), approaches therein all require sharp pointwise estimates of spatial derivatives of the

heat kernel, which are highly non-trivial. Therefore it is natural to ask if the use of these estimates could be avoided when establishing (1.1).

Secondly, we are interested in the Riesz transform on direct product manifolds. We have showed in [28] that if M is the direct product manifold $M_1 \times \cdots \times M_n$, where M_i ($1 \leq i \leq n$) are connected complete non-compact Riemannian manifolds satisfying (D) and (UE_{m_i}) ($m_i \geq 2$), then (1.1) and (1.2) for $1 < p < 2$ remain valid on M . Hence it is reasonable to ask whether the corresponding results hold in the setting of direct product manifolds with exponential volume growth, such as $\mathbb{H}^n \times M$, where M is a connected complete non-compact Riemannian manifold satisfying (ND) , (GUE_θ) and $(WGE_{q,\theta})$. The aim of this paper is to give an affirmative answer to this question.

Now we can state our main result as follows:

Theorem 1.1. *Let M be a connected complete non-compact Riemannian manifold satisfying (ND) , (GUE_θ) and $(WGE_{q,\theta})$ (with $1 \leq q \leq 2$, $0 < \theta \leq 1$) w.r.t. the quasi-distance $\tilde{d}$. Moreover, suppose that for any given $R_0 > 0$, the ball $\tilde{B}(x, R_0)$ equipped with the induced quasi-distance $\tilde{d}$ and measure $d\mu_M$ satisfies (ND) , with a doubling constant independent of its center. Then, the Riesz transform on $\mathbb{H}^n \times M$ is of weak type $(1, 1)$ (so L^p bounded for $1 < p \leq 2$).*

Remark 1.2. The condition that $(\tilde{B}(x, R_0), \tilde{d}, d\mu_M)$ satisfies (ND) uniformly in $x \in M$ is a technical property that is only used in Section 3 for a localization argument. When $\tilde{d} = d_M$ and $(ND)_{\text{loc}}$ holds, it is automatically fulfilled, as shown in [4, Lemma 4.1] (cf. also [24, Lemme 4.3.1]). In addition, it can be observed that this property remains stable under products, provided that direct product manifolds are equipped with a quasi-distance of the type (1.6). Therefore, under the assumptions of Theorem 1.1, $\mathbb{H}^n \times M$ shares the same property.

The techniques developed here also apply to the study of the heat maximal operator on $\mathcal{M}$, which is defined by

$$\mathcal{H}(f)(X) := \sup_{t>0} |e^{t\Delta_{\mathcal{M}}} f(X)|, \quad \forall X \in \mathcal{M}, f \in L^p(\mathcal{M}), 1 \leq p \leq +\infty.$$

Since the heat semigroup $(e^{t\Delta_{\mathcal{M}}})_{t>0}$ is a diffusion semigroup symmetric w.r.t. $\mu_{\mathcal{M}}$, the L^p boundedness ($1 < p \leq +\infty$) of $\mathcal{H}$ can be found in [33, Section 3.3] or [7, Lemma 1.6.2]. In Section 5, we establish the weak type $(1, 1)$ property of $\mathcal{H}$ as follows:

Theorem 1.3. *Let M be a connected complete non-compact Riemannian manifold satisfying (ND) , (GUE_θ) w.r.t. the quasi-distance $\tilde{d}$ (with $0 < \theta \leq 1$). Then, the heat maximal operator on $\mathbb{H}^n \times M$ is of weak type $(1, 1)$ (so L^p bounded for $1 < p \leq +\infty$).*

For more results on the weak type $(1, 1)$ boundedness of $\mathcal{H}$ in the context of manifolds with exponential volume growth, we refer the reader to [2, 3, 15, 25–27].

Section 6 extends the above results to a large class of direct product manifolds with exponential volume growth. As a by-product, we generalize the main results in [27] and provide a much simpler proof.

2 Preliminaries

2.1 Some basic facts about (GUE_θ) and $(WGE_{q,\theta})$

To illustrate (GUE_θ) and $(WGE_{q,\theta})$, recall first that if M meets (D) and (UE_2) , then (GUE_1) and $(WGE_{2,1})$ are valid w.r.t. the geodesic distance (cf. e.g. [13] or [20]).

Next, suppose that M satisfies the volume doubling property and the sub-Gaussian heat kernel upper estimates with the exponent $m > 2$. Let

$$\tilde{d}(x, y) = \max\{d_M(x, y), d_M(x, y)^{\frac{m}{2}}\}, \quad \forall x, y \in M.$$

It is easy to check that $\tilde{d}$ defines a quasi-distance on M and that $(M, \tilde{d}, d\mu_M)$ satisfies (ND) , $(GUE_{\frac{1}{m-1}})$. It follows from [11, Lemma 2.2] that $(WGE_{q, \frac{1}{m-1}})$ is available for all $1 < q < 2$ with regard to $\tilde{d}$.

Fix a quasi-distance $\tilde{d}$ on M . Notice that given $0 < \theta' < \theta$, $s^{\theta'} \leq 1 + s^\theta$ holds for all $s \geq 0$. Hence

$$(GUE_\theta) \implies (GUE_{\theta'}), \quad (WGE_{q,\theta}) \implies (WGE_{q,\theta'}). \quad (2.1)$$

Assume now that (ND) and (GUE_θ) are valid. The standard method of decomposition in annuli implies that

$$\begin{aligned} \int_M \exp \left(-\gamma \left(\frac{\tilde{d}(x,y)^2}{t} \right)^\theta \right) d\mu_M(x) &= \int_{\tilde{B}(y,\sqrt{t})} + \sum_{i=0}^{+\infty} \int_{2^i\sqrt{t} \leq \tilde{d}(y,x) < 2^{i+1}\sqrt{t}} \\ &\leq C(\gamma, \theta) \tilde{V}_M(y, \sqrt{t}), \quad \forall t > 0, y \in M, \end{aligned} \quad (2.2)$$

where we have used (1.3) in the last line. Combining this with Hölder inequality, we obtain the following relation

$$(WGE_{q,\theta}) \implies (WGE_{q',\theta}), \quad \forall 1 \leq q' < q. \quad (2.3)$$

To finish this section, we point out an invariance of (GUE_θ) and $(WGE_{q,\theta})$ under direct product in the following sense:

Proposition 2.1. *Suppose that M_i ($i = 1, 2$) are connected complete non-compact Riemannian manifolds satisfying (ND) , (GUE_{θ_i}) and (WGE_{q_i,θ_i}) (with $1 \leq q_i \leq 2$, $0 < \theta_i \leq 1$) w.r.t. the quasi-distance $\tilde{d}_i$. Let $\theta = \min\{\theta_1, \theta_2\}$, $q = \min\{q_1, q_2\}$ and*

$$\tilde{d}(x, y) = \max_{i=1,2} \{\tilde{d}_i(x_i, y_i)\}, \quad \forall x = (x_1, x_2), y = (y_1, y_2) \in M_1 \times M_2.$$

Then on the direct product manifold $M = M_1 \times M_2$, we have (ND) , (GUE_θ) and $(WGE_{q,\theta})$ w.r.t. the quasi-distance $\tilde{d}$.

Proof. This proof is an easy task. It suffices to use properties of product manifolds and (2.2). So we omit it. $\square$

Remark 2.2. The result remains valid if one of M_i is a compact Riemannian manifold or a Lie group with polynomial growth.

2.2 Some further properties of $\mathbb{H}^n$

In this short section, we collect basic properties of $\mathbb{H}^n$ which will be used throughout this work. We first present some estimates of the heat kernel and its gradient, which can be found in [16, 18]:

The heat kernel $p_t^{\mathbb{H}^n}(u, v)$ depends only on t and $r = r(u, v)$. More precisely, it is known that for all $r \geq 0$ and $t > 0$,

$$p_t^{\mathbb{H}^n}(u, v) = p_t^{\mathbb{H}^n}(r) \sim t^{-\frac{n}{2}} (1 + r + t)^{\frac{n-3}{2}} (1 + r) \exp \left(-\frac{r^2}{4t} - \frac{n-1}{2}r - \frac{(n-1)^2}{4}t \right). \quad (2.4)$$

In particular, combining this with (1.4), $p_t^{\mathbb{H}^n}$ has the following small-time upper bound:

$$p_t^{\mathbb{H}^n}(u, u) \leq \frac{C}{V_{\mathbb{H}^n}(u, \sqrt{t})}, \quad \forall u \in \mathbb{H}^n, 0 < t \leq 1, \quad (2.5)$$

and the following local estimate: given $r_0 > 0$, for all $u, v \in \mathbb{H}^n$ with $r = r(u, v) \leq r_0$,

$$p_t^{\mathbb{H}^n}(u, v) \lesssim_{r_0} t^{-\frac{n}{2}} \exp \left(-\frac{r^2}{Ct} \right), \quad \forall t > 0. \quad (2.6)$$

As for the spatial derivative of $p_t^{\mathbb{H}^n}$, it is known that for $u, v \in \mathbb{H}^n$ and $t > 0$,

$$|\nabla_{\mathbb{H}^n} p_t^{\mathbb{H}^n}(u, v)| \sim t^{-\frac{n+2}{2}} (1 + r + t)^{\frac{n-1}{2}} r \exp \left(-\frac{r^2}{4t} - \frac{n-1}{2}r - \frac{(n-1)^2}{4}t \right). \quad (2.7)$$

We will make use of the following lemma, which asserts that on $\mathbb{H}^n$, a specific decay of the kernel toward infinity implies the weak type $(1, 1)$ property of the associated integral operator. For its proof, see [36].

Lemma 2.3. *The integral operator defined on $\mathbb{H}^n$ with the kernel*

$$S(u, v) = e^{-(n-1)r}$$

is of weak type $(1, 1)$.

2.3 Weak type $(1, 1)$ of sub-linear operators on product measure spaces

Now we present a simple criterion of weak type $(1, 1)$, which plays a role in our proof.

Lemma 2.4. *Let (X, μ) and (Y, ν) be measure spaces. Let $\mathcal{A}$ and $\mathcal{B}$ be sub-linear operators defined on $L^1(X, \mu)$ and $L^1(Y, \nu)$ respectively, which are of the forms*

$$\mathcal{A}(f_1)(x_1) = \int_X a(x_1, y_1) |f_1(y_1)| d\mu(y_1) \quad \text{and} \quad \mathcal{B}(f_2)(x_2) = \int_Y b(x_2, y_2) |f_2(y_2)| d\nu(y_2),$$

where a and b are non-negative measurable functions. Suppose that $\|\mathcal{A}\|_{L^1 \rightarrow L^{1,\infty}} \leq C$, i.e. it holds that

$$\mu(\{x_1 \in X; |\mathcal{A}(f_1)(x_1)| > \lambda\}) \leq C \frac{\|f_1\|_1}{\lambda}, \quad \forall \lambda > 0, \quad f_1 \in L^1(X, \mu).$$

Moreover, assume that there exists a constant $M > 0$ such that

$$\sup_{y_2 \in Y} \int_Y b(x_2, y_2) d\nu(y_2) \leq M.$$

Then, for the following tensor product type operator $\mathcal{T}$, defined on $L^1(X \times Y, \mu \otimes \nu)$,

$$\mathcal{T}(f)(x_1, x_2) = \iint_{X \times Y} a(x_1, y_1) b(x_2, y_2) |f(y_1, y_2)| d\mu(y_1) d\nu(y_2),$$

we have $\|\mathcal{T}\|_{L^1 \rightarrow L^{1,\infty}} \leq CM$.

Proof. Given $f \in L^1(X \times Y, \mu \otimes \nu)$, notice that $\int_Y b(x_2, y_2) |f(\cdot, y_2)| d\nu(y_2) \in L^1(X)$ for a.e.- ν $x_2 \in Y$. Therefore the weak type $(1, 1)$ estimate of $\mathcal{A}$ implies that

$$\mu(\{x_1 \in X; |\mathcal{T}(f)(x_1, x_2)| > \lambda\}) \leq \frac{C}{\lambda} \int_X \left(\int_Y b(x_2, y_2) |f(y_1, y_2)| d\nu(y_2) \right) d\mu(y_1)$$

holds for a.e.- ν $x_2 \in Y$. Combining this with Fubini's theorem, we yield that

$$\begin{aligned} & (\mu \otimes \nu)(\{(x_1, x_2) \in X \times Y; |\mathcal{T}(f)(x_1, x_2)| > \lambda\}) \\ &= \int_Y \mu(\{x_1 \in X; |\mathcal{T}(f)(x_1, x_2)| > \lambda\}) d\nu(x_2) \\ &\leq \frac{C}{\lambda} \int_Y \left(\int_X \left(\int_Y b(x_2, y_2) |f(y_1, y_2)| d\nu(y_2) \right) d\mu(y_1) \right) d\nu(x_2). \end{aligned}$$

Then changing the order of integration, by the assumption of b , this lemma is proved. $\square$

3 Standard method of localization for the Riesz transforms

In this section, we will study the local part of the Riesz transform $\mathcal{R} = \nabla_{\mathcal{M}}(-\Delta_{\mathcal{M}})^{-\frac{1}{2}}$. To begin with, we represent $\mathcal{R}$ as the sum of the local part $\mathcal{R}_{\text{loc}}$ and the part at infinity $\mathcal{R}_{\infty}$. This step proceeds along the standard line, which has been widely adopted by other authors. In previous works mentioned in the end of Section 1.2, the Calderón-Zygmund theory has been successfully applied to analyze $\mathcal{R}_{\text{loc}}$, mainly due to the availability of upper bounds for $|\nabla p_t^{\mathcal{M}}|$ and $|\nabla \nabla p_t^{\mathcal{M}}|$. However, in our setting, we do not have a (complete) pointwise upper bound of $|\nabla_{\mathcal{M}} p_t^{\mathcal{M}}|$, which motives the adoption of the method developed in [13].

3.1 Weak $(1, 1)$ of $\mathcal{R}_{\text{loc}}$ on general weighted Riemannian manifolds

The localization procedure could be done on a general non-compact complete, weighted Riemannian manifold $\mathcal{N}$ in the sense of [22] satisfying (ND_{loc}) w.r.t. a quasi-distance ρ . By the definition of ρ , there exists a constant $\omega \geq 1$ such that:

$$\rho(x, z) \leq \omega (\rho(x, y) + \rho(y, z)), \quad \forall x, y, z \in \mathcal{N}. \quad (3.1)$$

Selecting a $\frac{1}{4\omega^2}$ -maximal separated subset of $\mathcal{N}$, we get a countable family of balls $\{B^j = B(x^j, \frac{1}{2\omega}); x^j \in \mathcal{N}, j \in \mathbb{N}\}$. It follows from (ND_{loc}) that this family satisfies the following properties:

- (a) $\mathcal{N} = \bigcup_{j \in \mathbb{N}} B^j$;
- (b) the balls $B(x^j, \frac{1}{4\omega^2})$ are pairwise disjoint;
- (c) there exists $N \in \mathbb{N}$, such that every point of $\mathcal{N}$ is contained in at most N balls $(2\omega^2 + \omega)B^j := B(x^j, \omega + \frac{1}{2})$.

Fix a C^∞ partition of unity $(\varphi_j)_{j \in \mathbb{N}}$ such that $\text{supp } \varphi_j \subseteq B^j$. Given any $1 \leq p < +\infty$, the definition of partition of unity and the property (c) imply that $\frac{1}{N^p} \leq \sum_j \varphi_j^p \leq 1$. Therefore for all $f \in C_0^\infty(\mathcal{N})$,

$$\|f\|_p^p \sim \sum_j \|f\varphi_j\|_p^p. \quad (3.2)$$

For $f \in C_0^\infty(\mathcal{N})$, we write

$$\mathcal{R}(f) = \sum_j \chi_{(2\omega^2 + \omega)B^j} \mathcal{R}(f\varphi_j) + \sum_j (1 - \chi_{(2\omega^2 + \omega)B^j}) \mathcal{R}(f\varphi_j). \quad (3.3)$$

Let $R(x, y) = \frac{1}{\sqrt{\pi}} \int_0^\infty \nabla_{\mathcal{N}} p_t^{\mathcal{N}}(x, y) \frac{dt}{\sqrt{t}}$, which is the kernel of the Riesz transform. Then from the construction, we have for $x \in \mathcal{N}$,

$$\begin{aligned} \left| \sum_j (1 - \chi_{(2\omega^2 + \omega)B^j}) \mathcal{R}(f\varphi_j)(x) \right| &\leq \sum_j \int_{\{y; \rho > 1\}} |R(x, y)| |f(y)| \varphi_j(y) d\mu_{\mathcal{N}}(y) \\ &\leq \int_{\{y; \rho > 1\}} |R(x, y)| |f(y)| d\mu_{\mathcal{N}}(y). \end{aligned} \quad (3.4)$$

We define the local part of the Riesz transform

$$\mathcal{R}_{\text{loc}}(f)(x) := \sum_j \chi_{(2\omega^2 + \omega)B^j} \mathcal{R}(f\varphi_j)(x), \quad (3.5)$$

and the part at infinity

$$\mathcal{R}_\infty(f)(x) := \int_{\{y; \rho > 1\}} |R(x, y)| |f(y)| d\mu_{\mathcal{N}}(y). \quad (3.6)$$

Hence, in order to prove (1.1) on $\mathcal{N}$, it is sufficient to prove that $\mathcal{R}_{\text{loc}}$ and $\mathcal{R}_\infty$ are both of weak type $(1, 1)$. In this section we focus on $\mathcal{R}_{\text{loc}}$.

For the further developments, we address here a rather general theorem on the weak $(1, 1)$ of $\mathcal{R}_{\text{loc}}$. Indeed, it is an analogy of [13, Theorem 1.2] (see also [28, Proposition 3]).

Theorem 3.1. *Let $\mathcal{N}$ be a connected complete non-compact weighted Riemannian manifold satisfying (ND_{loc}) w.r.t. the quasi-distance ρ . Moreover, suppose that for any given $R_0 > 0$, the ball $B(x, R_0)$ equipped with the induced quasi-distance ρ and measure $d\mu_{\mathcal{N}}$ satisfies (ND) , with a doubling constant independent of its center. Assume that the following small-time heat kernel upper bound is valid: for some $\theta \in (0, 1]$, there exists a constant $c \in (0, 1)$ such that*

$$p_t^{\mathcal{N}}(x, y) \leq c^{-1} \frac{1}{V_{\mathcal{N}}(x, \sqrt{t})} \exp \left(-c \left(\frac{\rho(x, y)^2}{t} \right)^\theta \right), \quad \forall 0 < t \leq 1, x, y \in \mathcal{N}. \quad (GUE'_\theta)$$

Additionally, suppose that there exist some constant $a \geq \frac{1}{2}(2\omega^2 + \omega + 1)$, where $\omega \geq 1$ is the constant in (3.1), and a bounded function $F : [0, +\infty) \rightarrow (0, +\infty)$, which satisfies

$$\int_0^1 F\left(\frac{1}{u}\right) \frac{du}{u} < +\infty,$$

such that for any $0 < t \leq 1$, $s > 0$ and $y \in \mathcal{N}$,

$$\int_{\{x; \sqrt{t} \leq \rho \leq a\}} |\nabla_{\mathcal{N}} p_s^{\mathcal{N}}(x, y)| d\mu_{\mathcal{N}}(x) \leq \frac{1}{\sqrt{s}} F\left(\frac{t}{s}\right), \quad (WGE'_1)$$

Furthermore we assume that $\mathcal{R}_{\infty}$ is of weak type $(1, 1)$. Then $\mathcal{R}_{\text{loc}}$ is of weak type $(1, 1)$.

Remark 3.2. Before turning to the proof, let us make a few comments on the requirement $a \geq \frac{1}{2}(2\omega^2 + \omega + 1)$. First, it is purely technical and could be relaxed to $a > 0$. To see this, it suffices to suitably modify the numerical constants initially chosen in the localization procedure (involving the radius of balls used in the construction and the definition of $\mathcal{R}_{\infty}$) to adapt the value of a . While the remaining argument of the proof is unchanged. Second, in concrete settings, this requirement is usually inessential. For instance, on the product manifolds to be considered later, we can even achieve a stronger version of (WGE'_1) as follows. For any fixed $a > 0$, there exists a bounded function $F_a : [0, +\infty) \rightarrow (0, +\infty)$, depending on a , which satisfies

$$\int_0^1 F_a\left(\frac{1}{u}\right) \frac{du}{u} < +\infty,$$

such that for any $0 < t \leq 1$, $s > 0$ and $y \in \mathcal{N}$,

$$\int_{\{x; \sqrt{t} \leq \rho \leq a\}} |\nabla_{\mathcal{N}} p_s^{\mathcal{N}}(x, y)| d\mu_{\mathcal{N}}(x) \leq \frac{1}{\sqrt{s}} F_a\left(\frac{t}{s}\right).$$

Proof. This proof is actually a straightforward adaption of the argument used in [13, §4]. We include it for the sake of clarity. Assume for simplicity of exposition that $\omega = 1$ (the general case $\omega \geq 1$ can be handled accordingly but some notational inconvenience).

In view of the property (c), for any fixed $x \in \mathcal{N}$, there are at most N non-zero terms in the sum $\sum_j \chi_{3B^j} \mathcal{R}(f\varphi_j)(x)$. Consequently, for $\lambda > 0$,

$$\begin{aligned} \mu_{\mathcal{N}}(\{x \in \mathcal{N}; |\mathcal{R}_{\text{loc}}(f)(x)| > \lambda\}) &= \mu_{\mathcal{N}}(\{x \in \mathcal{N}; |\sum_j \chi_{3B^j} \mathcal{R}(f\varphi_j)(x)| > \lambda\}) \\ &\leq \sum_j \mu_{\mathcal{N}}(\{x \in 3B^j; |\mathcal{R}(f\varphi_j)(x)| > \lambda/N\}). \end{aligned} \quad (3.7)$$

According to (3.2) for $p = 1$, it suffices to show that there exists a universal constant $c > 0$, such that for all $\lambda > 0$ and $j \in \mathbb{N}$

$$\mu_{\mathcal{N}}(\{x \in 3B^j; |\mathcal{R}(f)(x)| > \lambda\}) \leq c \frac{\|f\|_1}{\lambda}, \quad \forall f \in C_0^\infty(B^j). \quad (3.8)$$

By assumption, each B^j satisfies (ND) with a constant independent of j . From now on we drop the subscripts of B^j . We can apply the classical Calderón-Zygmund decomposition. There exists a constant $\kappa > 0$ (only depends on the constant in (ND) , therefore independent of j) such that, given $f \in C_0^\infty(B)$ and $\lambda > \frac{\|f\|_1}{\mu_{\mathcal{M}}(B)}$, we can decompose f as

$$f = g + b = g + \sum_i b_i,$$

and find a sequence of balls $B_i = B(x_i, r_i) \subseteq B$, satisfying

$$(i) \quad \|g\|_2^2 \leq \kappa \lambda \|f\|_1;$$

(ii) $\text{supp } b_i \subseteq B_i$ and $\|b_i\|_1 \leq \kappa \lambda \mu_{\mathcal{N}}(B_i)$;

(iii) $\sum_i \mu_{\mathcal{N}}(2B_i) \leq \kappa \lambda^{-1} \|f\|_1$.

It is clear that $r_i \leq \text{diam}(B) \leq 1$.

Notice that

$$\begin{aligned} & \mu_{\mathcal{N}}(\{x \in 3B; |\mathcal{R}(f)(x)| > \lambda\}) \\ & \leq \mu_{\mathcal{N}}(\{x \in 3B; |\mathcal{R}(g)(x)| > \lambda/2\}) + \mu_{\mathcal{N}}(\{x \in 3B; |\mathcal{R}(b)(x)| > \lambda/2\}). \end{aligned} \quad (3.9)$$

Combining the L^2 boundedness of $\mathcal{R}$ with (i), the first term is easy to estimate. Thus we only need to treat the second one. We write

$$\mathcal{R}(b) = \sum_i \mathcal{R}(\chi_{5B} e^{t_i \Delta_{\mathcal{N}}} b_i) + \mathcal{R}((1 - \chi_{5B}) \sum_i e^{t_i \Delta_{\mathcal{N}}} b_i) + \sum_i \mathcal{R}((I - e^{t_i \Delta_{\mathcal{N}}}) b_i), \quad (3.10)$$

where $t_i = r_i^2$. Then

$$\begin{aligned} & \mu_{\mathcal{N}}(\{x \in 3B; |\mathcal{R}(b)(x)| > \lambda/2\}) \leq \mu_{\mathcal{N}}(\{x \in 3B; |\mathcal{R}(\sum_i \chi_{5B} e^{t_i \Delta_{\mathcal{N}}} b_i)(x)| > \lambda/6\}) \\ & + \mu_{\mathcal{N}}(\{x \in 3B; |\mathcal{R}((1 - \chi_{5B}) \sum_i e^{t_i \Delta_{\mathcal{N}}} b_i)(x)| > \lambda/6\}) \\ & + \mu_{\mathcal{N}}(\{x \in 3B; |\sum_i \mathcal{R}((I - e^{t_i \Delta_{\mathcal{N}}}) b_i)(x)| > \lambda/6\}). \end{aligned} \quad (3.11)$$

It follows from our assumption and (ND) that the centered Hardy-Littlewood maximal operator is bounded on $L^2(5B)$. On the other hand, by construction, each $t_i \leq 1$. Therefore we can apply (GUE'_θ) . By the duality method used in [13, §3], we obtain

$$\left\| \sum_i \chi_{5B} e^{t_i \Delta_{\mathcal{N}}} b_i \right\|_2^2 \lesssim \lambda \|f\|_1.$$

Hence the L^2 boundedness of $\mathcal{R}$ and Chebyshev's inequality imply that

$$\mu_{\mathcal{N}}(\{x \in 3B; |\sum_i \mathcal{R}(\chi_{5B} e^{t_i \Delta_{\mathcal{N}}} b_i)(x)| > \lambda/6\}) \lesssim \frac{\|f\|_1}{\lambda}. \quad (3.12)$$

As for $\mathcal{R}((1 - \chi_{5B}) \sum_i e^{t_i \Delta_{\mathcal{N}}} b_i)$, we shall use $\mathcal{R}_\infty$ to dominate it. In fact, for $x \in 3B$,

$$\left| \mathcal{R}((1 - \chi_{5B}) \sum_i e^{t_i \Delta_{\mathcal{N}}} b_i)(x) \right| \leq \mathcal{R}_\infty((1 - \chi_{5B}) \sum_i e^{t_i \Delta_{\mathcal{N}}} b_i)(x) \leq \mathcal{R}_\infty(\sum_i |e^{t_i \Delta_{\mathcal{N}}} b_i|)(x).$$

Combining this with the weak type $(1, 1)$ estimate of $\mathcal{R}_\infty$, using the contractivity of the heat semigroup $(e^{t \Delta_{\mathcal{N}}})_{t>0}$, we obtain

$$\mu_{\mathcal{N}}(\{x \in 3B; |\mathcal{R}((1 - \chi_{5B}) \sum_i e^{t_i \Delta_{\mathcal{N}}} b_i)(x)| > \lambda/6\}) \lesssim \frac{\sum_i \|b_i\|_1}{\lambda} \lesssim \frac{\|f\|_1}{\lambda}, \quad (3.13)$$

where we have used the properties (ii) and (iii) in the last inequality.

Therefore, by (3.9)-(3.13), to get (3.8), it remains to establish the following estimate:

$$\mu_{\mathcal{N}}(\{x \in 3B; |\sum_i \mathcal{R}((I - e^{t_i \Delta_{\mathcal{N}}}) b_i)(x)| > \lambda/6\}) \lesssim \frac{\|f\|_1}{\lambda},$$

which can be proved by the standard techniques. More precisely, the left-hand side of the above inequality is majorized by

$$\mu_{\mathcal{N}}(\cup_i 2B_i) + \mu_{\mathcal{N}}(\{x \in 3B \setminus (\cup_i 2B_i); |\sum_i \mathcal{R}((I - e^{t_i \Delta_{\mathcal{N}}}) b_i)(x)| > \lambda/6\}).$$

Using the properties (ii) and (iii) again, via Chebyshev's inequality, it is enough to prove that

$$\int_{3B \setminus 2B_i} |\mathcal{R}((I - e^{t_i \Delta_{\mathcal{N}}}) b_i)(x)| d\mu_{\mathcal{N}}(x) \lesssim \|b_i\|_1, \quad \forall i. \quad (3.14)$$

Denote by k_t the kernel of $\mathcal{R}(I - e^{t \Delta_{\mathcal{N}}})$. Since $\text{supp } b_i \subseteq B_i \subseteq B$, we have

$$\begin{aligned} & \int_{3B \setminus 2B_i} |\mathcal{R}((I - e^{t_i \Delta_{\mathcal{N}}}) b_i)(x)| d\mu_{\mathcal{N}}(x) \\ & \leq \int_{3B \setminus 2B_i} \left(\int_{B_i} |k_{t_i}(x, y)| |b_i(y)| d\mu_{\mathcal{N}}(y) \right) d\mu_{\mathcal{N}}(x) \\ & \leq \int_{B_i} \left(\int_{\{x; \sqrt{t_i} \leq \rho \leq 2\}} |k_{t_i}(x, y)| d\mu_{\mathcal{N}}(x) \right) |b_i(y)| d\mu_{\mathcal{N}}(y), \end{aligned}$$

where

$$k_t(x, y) = \frac{1}{\sqrt{\pi}} \int_0^\infty \nabla_{\mathcal{N}} p_s^{\mathcal{N}}(x, y) \left(\frac{1}{\sqrt{s}} - \frac{\chi_{\{s > t\}}}{\sqrt{s-t}} \right) ds.$$

Combining (WGE'_1) with the equality above, we yield that

$$\begin{aligned} \int_{\{x; \sqrt{t} \leq \rho \leq 2\}} |k_t(x, y)| d\mu_{\mathcal{N}}(x) & \lesssim \int_0^\infty \left| \frac{1}{\sqrt{s}} - \frac{\chi_{\{s > t\}}}{\sqrt{s-t}} \right| F\left(\frac{t}{s}\right) \frac{ds}{\sqrt{s}} \\ & \lesssim \int_0^t F\left(\frac{t}{s}\right) \frac{ds}{s} + \int_t^\infty \left| \frac{1}{\sqrt{s}} - \frac{\chi_{\{s > t\}}}{\sqrt{s-t}} \right| F\left(\frac{t}{s}\right) \frac{ds}{\sqrt{s}}. \end{aligned}$$

By our assumptions on F , the above integral is bounded from above uniformly in $t > 0$. Thus (3.14) is proved. This ends the proof of Theorem 3.1. $\square$

3.2 Consequences of Theorem 3.1

As an application of Theorem 3.1, we establish the weak type $(1, 1)$ boundedness of $\mathcal{R}_{\text{loc}}$ defined on $\mathcal{M} = \mathbb{H}^n \times M$ in this subsection. Actually, we can extend this result to general direct product manifolds with exponential volume growth. Let $\mathcal{E}$ be a connected complete non-compact weighted Riemannian manifold. Slightly abusing notation, we denote by r the geodesic distance on $\mathcal{E}$.

Corollary 3.3. *Let M be a connected complete non-compact weighted Riemannian manifolds satisfying (ND) , (GUE_θ) and $(WGE_{q,\theta})$ (with $1 \leq q \leq 2$, $0 < \theta \leq 1$) w.r.t. the quasi-distance $\tilde{d}$. Moreover, suppose that for any given $R_0 > 0$, the ball $\tilde{B}(x, R_0)$ equipped with the induced quasi-distance $\tilde{d}$ and measure $d\mu_M$ satisfies (ND) , with a doubling constant independent of its center. Assume that $\mathcal{E}$ satisfies (ND_{loc}) and (VE) , whose heat kernel $p_t^\mathcal{E}$ has the small-time upper bound*

$$p_t^\mathcal{E}(u, u) \leq \frac{C}{V_\mathcal{E}(u, \sqrt{t})}, \quad \forall u \in \mathcal{E}, \quad 0 < t \leq 1.$$

Furthermore, we assume that $\mathcal{R}_\infty$ defined on $\mathcal{E} \times M$ is of weak type $(1, 1)$. Then $\mathcal{R}_{\text{loc}}$ defined on $\mathcal{E} \times M$ is of weak type $(1, 1)$.

Proof. We denote $\mathcal{E} \times M$ briefly by $\mathcal{W}$. A quasi-distance d defined on $\mathcal{W}$ is also given by (1.6). According to Theorem 3.1, it is enough to verify the conditions (GUE'_θ) and (WGE'_1) on $\mathcal{W}$ because the other assumptions hold obviously, taking into account Remark 1.2.

We start with (GUE'_θ) . It follows from [21, Theorem 1.1] (cf. also [22, Chapter 16]) that the small-time upper bound of $p_t^\mathcal{E}$ and (ND_{loc}) of $\mathcal{E}$ imply the following off-diagonal estimate: given $\gamma > 4$,

$$p_t^\mathcal{E}(u, v) \lesssim_\gamma \frac{1}{\sqrt{V_\mathcal{E}(u, \sqrt{t})V_\mathcal{E}(v, \sqrt{t})}} \exp\left(-\frac{r^2}{\gamma t}\right), \quad \forall t \in (0, 1], \quad u, v \in \mathcal{E}.$$

And it is clear that by (VE),

$$V_{\mathcal{E}}(u, \sqrt{t}) \leq V_{\mathcal{E}}(v, \sqrt{t} + r) \leq C e^{c \frac{r}{\sqrt{t}}} V_{\mathcal{E}}(v, \sqrt{t}), \quad \forall t \in (0, 1], u, v \in \mathcal{E}.$$

Thus we obtain the following classical small time upper bound of $p_t^{\mathcal{E}}$: given $\gamma > 4$, we have

$$p_t^{\mathcal{E}}(u, v) \lesssim_{\gamma} \frac{1}{V_{\mathcal{E}}(u, \sqrt{t})} \exp\left(-\frac{r^2}{\gamma t}\right), \quad \forall t \in (0, 1], u, v \in \mathcal{E}. \quad (3.15)$$

Moreover, by the trivial inequality $s^{\theta} \leq 1 + s$ for all $s \geq 0$ and $0 < \theta \leq 1$, we get

$$p_t^{\mathcal{E}}(u, v) \lesssim_{\gamma} \frac{1}{V_{\mathcal{E}}(u, \sqrt{t})} \exp\left(-\frac{1}{\gamma} \left(\frac{r^2}{t}\right)^{\theta}\right), \quad \forall t \in (0, 1], u, v \in \mathcal{E}. \quad (3.16)$$

Now, for $t \in (0, 1]$ and $X = (u, x), Y = (v, y) \in \mathcal{E} \times M = \mathcal{W}$, the basic property of the heat kernel on direct product manifolds says that $p_t^{\mathcal{W}}(X, Y) = p_t^{\mathcal{E}}(u, v) p_t^M(x, y)$. Combining this with (3.16), the assumption (GUE_{θ}) on M implies that there exists a constant $c \in (0, 1)$ such that

$$\begin{aligned} p_t^{\mathcal{W}}(X, Y) &\lesssim \frac{1}{V_{\mathcal{E}}(u, \sqrt{t}) \tilde{V}_M(x, \sqrt{t})} \exp\left\{-c \left[\left(\frac{r(u, v)^2}{t}\right)^{\theta} + \left(\frac{\tilde{d}(x, y)^2}{t}\right)^{\theta}\right]\right\} \\ &\lesssim \frac{1}{V_{\mathcal{W}}(X, \sqrt{t})} \exp\left(-c \left(\frac{d(X, Y)^2}{t}\right)^{\theta}\right), \end{aligned} \quad (3.17)$$

where we have used in the last line the following trivial observation:

$$\{Y; d(X, Y) < s\} = \{v; r(u, v) < s\} \times \{y; \tilde{d}(x, y) < s\}, \quad \forall s > 0.$$

In other terms, (GUE'_{θ}) is valid on $\mathcal{W}$.

We turn to (WGE'_1) . Indeed, the requirement $a \geq \frac{1}{2}(2\omega^2 + \omega + 1)$ in Theorem 3.1 is only a technical assumption. The following proof also works for any given $a > 0$ in our setting (see Remark 3.2). Since (WGE'_1) holds trivially if $t > a^2$, in the following we assume $t \leq a^2$. Recalling the definition of d , we obtain for all $t \in [0, 1]$ and $Y = (v, y) \in \mathcal{W}$,

$$\begin{aligned} &\int_{\{X; \sqrt{t} \leq d \leq a\}} |\nabla_{\mathcal{W}} p_s^{\mathcal{W}}(X, Y)| d\mu_{\mathcal{W}}(X) \\ &\leq \int_{\{X; \sqrt{t} \leq r \leq a, \tilde{d} \leq a\}} |\nabla_{\mathcal{W}} p_s^{\mathcal{W}}(X, Y)| d\mu_{\mathcal{W}}(X) + \int_{\{X; r \leq a, \sqrt{t} \leq \tilde{d} \leq a\}} |\nabla_{\mathcal{W}} p_s^{\mathcal{W}}(X, Y)| d\mu_{\mathcal{W}}(X) \\ &\leq \int_{\{u; \sqrt{t} \leq r \leq a\}} p_s^{\mathcal{E}} d\mu_{\mathcal{E}} \int_{\{x; \tilde{d} \leq a\}} |\nabla_M p_s^M| d\mu_M + \int_{\{u; \sqrt{t} \leq r \leq a\}} |\nabla_{\mathcal{E}} p_s^{\mathcal{E}}| d\mu_{\mathcal{E}} \int_{\{x; \tilde{d} \leq a\}} p_s^M d\mu_M \\ &+ \int_{\{u; r \leq a\}} p_s^{\mathcal{E}} d\mu_{\mathcal{E}} \int_{\{x; \sqrt{t} \leq \tilde{d} \leq a\}} |\nabla_M p_s^M| d\mu_M + \int_{\{u; r \leq a\}} |\nabla_{\mathcal{E}} p_s^{\mathcal{E}}| d\mu_{\mathcal{E}} \int_{\{x; \sqrt{t} \leq \tilde{d} \leq a\}} p_s^M d\mu_M. \end{aligned} \quad (3.18)$$

This inequality implies that, in order to verify (WGE'_1) on $\mathcal{W}$, we only need to estimate the integrals on $\mathcal{E}$ and M separately. We first consider the terms involving $\mathcal{E}$. Our proof is mainly based on the weighted L^2 -norm estimate, which has become a standard tool in heat kernel theory. We refer the reader to the monograph [22] for more details. Given $D > 2$, we define the weighted integrals of the heat kernel and its gradient as follows: for $s > 0$ and $v \in \mathcal{E}$,

$$E(v, s) := \int_{\mathcal{E}} p_s^{\mathcal{E}}(u, v)^2 \exp\left(\frac{r^2}{Ds}\right) d\mu_{\mathcal{E}}(u),$$

and

$$E_1(v, s) := \int_{\mathcal{E}} |\nabla_{\mathcal{E}} p_s^{\mathcal{E}}(u, v)|^2 \exp\left(\frac{r^2}{Ds}\right) d\mu_{\mathcal{E}}(u).$$

Recall that a basic fact about the heat kernel is that $p_s^\mathcal{E}(v, v)$ is decreasing w.r.t. $s > 0$ for any given $v \in \mathcal{E}$. By [22, Theorem 16.3] (see also [21, Theorem 2.1]), it follows from the small-time upper bound of $p_t^\mathcal{E}$ and (ND_{loc}) that

$$E(v, s) \lesssim \frac{1}{V_\mathcal{E}(v, \sqrt{s})}, \quad \forall s \in (0, 1], v \in \mathcal{E}. \quad (3.19)$$

Moreover, let $\lambda_1 \geq 0$ denote the bottom of the spectrum of $(-\Delta_\mathcal{E})$. Combining (3.19) with the integral maximum principle (see e.g. [22, Theorem 12.1]), we obtain (see also [22, (15.25)])

$$E(v, s) \lesssim \frac{e^{-2\lambda_1 s}}{V_\mathcal{E}(v, \min\{\sqrt{s}, 1\})}, \quad \forall s > 0, v \in \mathcal{E}. \quad (3.20)$$

Then [20, Theorem 1.1] (notice that the main results and proofs therein remain valid in the setting of weighted Riemannian manifolds without modifications) implies that

$$E_1(v, s) \lesssim \frac{e^{-2\lambda_1 s}}{sV_\mathcal{E}(v, \min\{\sqrt{s}, 1\})}, \quad \forall s > 0, v \in \mathcal{E}. \quad (3.21)$$

On the other hand, by the method of decomposition in annuli, (VE) implies that

$$\int_{\{u; \sqrt{t} \leq r \leq a\}} e^{-\frac{r^2}{4Ds}} d\mu_\mathcal{E}(u) \lesssim_a e^{-\frac{t}{4Ds}} V_\mathcal{E}(v, \min\{\sqrt{s}, 1\}), \quad \forall s > 0, t \in [0, 1]. \quad (3.22)$$

Here and in (3.23)-(3.24) below, the implicit constants may increase exponentially w.r.t. a . Combining this with (3.20) and (3.21), by Cauchy-Schwarz inequality, we obtain for all $s > 0, t \in [0, 1]$ and $v \in \mathcal{E}$,

$$\int_{\{u; \sqrt{t} \leq r \leq a\}} p_s^\mathcal{E}(u, v) d\mu_\mathcal{E}(u) \lesssim_a \exp\left(-\frac{t}{4Ds} - \lambda_1 s\right), \quad (3.23)$$

and

$$\int_{\{u; \sqrt{t} \leq r \leq a\}} |\nabla_\mathcal{E} p_s^\mathcal{E}(u, v)| d\mu_\mathcal{E}(u) \lesssim_a \frac{1}{\sqrt{s}} \exp\left(-\frac{t}{4Ds} - \lambda_1 s\right). \quad (3.24)$$

Next, we turn to the analysis of the terms in (3.18) involving M , which is much simpler. A direct computation shows that there exists $c > 0$ such that for all $s > 0, t \in [0, 1]$ and $y \in M$,

$$\int_{\{x; \sqrt{t} \leq \tilde{d} \leq a\}} p_s^M(x, y) d\mu_M(x) \lesssim \exp\left(-c \left(\frac{t}{s}\right)^\theta\right), \quad (3.25)$$

and

$$\int_{\{x; \sqrt{t} \leq \tilde{d} \leq a\}} |\nabla_M p_s^M(x, y)| d\mu_M(x) \lesssim \frac{1}{\sqrt{s}} \exp\left(-c \left(\frac{t}{s}\right)^\theta\right). \quad (3.26)$$

In conclusion, collecting (3.23)-(3.26), from (3.18), (WGE'_1) holds on $\mathcal{W}$ for any $a > 0$ and $F_a(h) = c_a^{-1} e^{-c_a h^\theta}$, which completes our proof. $\square$

Remark 3.4. The weighted norm $E(\cdot, t)$ is an important tool to establish pointwise upper bounds of the heat kernel, based on the following inequality (see [20])

$$p_t^\mathcal{E}(u, v) \leq \sqrt{E(u, t/2)E(v, t/2)} \exp\left(-\frac{r^2}{2Dt}\right).$$

Combining this with (3.20) and (ND_{loc}) , the heat kernel $p_t^\mathcal{E}$ satisfies the following local estimate: given $r_0 > 0$, for all $t > 0$ and $u, v \in \mathcal{E}$ with $r = r(u, v) \leq r_0$,

$$p_t^\mathcal{E}(u, v) \lesssim_{r_0} \frac{1}{V_\mathcal{E}(u, \min\{\sqrt{t}, 1\})} \exp\left(-\frac{r^2}{2Dt} - \lambda_1 t\right), \quad (3.27)$$

Then by the similar argument used in (3.17), there exists $c > 0$ such that for all $t > 0$ and $X = (u, x)$, $Y = (v, y) \in \mathcal{W}$ with $d = d(X, Y) \leq 1$,

$$p_t^{\mathcal{W}}(X, Y) \lesssim \frac{1}{V_{\mathcal{W}}(X, \min\{\sqrt{t}, 1\})} \exp\left(-c \left(\frac{d^2}{t}\right)^\theta\right). \quad (3.28)$$

In Section 5, this estimate will be used in the study of the heat maximal operator.

It is clear that $\mathbb{H}^n$ is an example of such $\mathcal{E}$ (see (2.5)). Combining Corollary 3.3 with Theorems 4.1 and 4.2 in the next section, we conclude that

Corollary 3.5. $\mathcal{R}_{\text{loc}}$ defined on $\mathbb{H}^n \times M$ is of weak type $(1, 1)$.

The weak $(1, 1)$ property of $\mathcal{R}_\infty$ on a general $\mathcal{W}$ is a challenging problem. In Section 6, we shall give a sufficient condition, whose put-forward is inspired by the case $\mathcal{M} = \mathbb{H}^n \times M$.

4 Proof of Theorem 1.1: the part at infinity

In this section, we shall prove the weak type $(1, 1)$ estimate of the part at infinity $\mathcal{R}_\infty$. For this purpose, set in the sequel for $X = (u, x)$, $Y = (v, y) \in \mathbb{H}^n \times M$,

$$\begin{aligned} R_1(X, Y) &:= \frac{1}{\sqrt{\pi}} \int_0^\infty \left| \nabla_{\mathbb{H}^n} p_t^{\mathbb{H}^n}(u, v) \right| p_t^M(x, y) \frac{dt}{\sqrt{t}}, \\ R_2(X, Y) &:= \frac{1}{\sqrt{\pi}} \int_0^\infty p_t^{\mathbb{H}^n}(u, v) \left| \nabla_M p_t^M(x, y) \right| \frac{dt}{\sqrt{t}}. \end{aligned}$$

Now we define $\mathcal{R}_{i,\infty}$ ($i = 1, 2$) as follows:

$$\mathcal{R}_{i,\infty} f(X) := \int_{\{Y; d(X, Y) \geq 1\}} R_i(X, Y) |f(Y)| d\mu_{\mathcal{M}}(Y). \quad (4.1)$$

Notice that $\mathcal{R}_\infty(f) \leq \mathcal{R}_{1,\infty}(f) + \mathcal{R}_{2,\infty}(f)$. Hence it is sufficient to prove the operators $\mathcal{R}_{i,\infty}$ ($i = 1, 2$) are of weak type $(1, 1)$.

4.1 Weak $(1, 1)$ of $\mathcal{R}_{1,\infty}$

We start with $\mathcal{R}_{1,\infty}$, which is easier to handle for the reason that $R_1(X, Y)$ has an explicit pointwise upper bound.

Theorem 4.1. $\mathcal{R}_{1,\infty}$ defined on $\mathcal{M}$ is of weak type $(1, 1)$.

Proof. Since $\{(X, Y) \in \mathcal{M}^2; d > 1\} \subseteq S_1 \cup S_2$ with

$$S_1 := \{(X, Y) \in \mathcal{M}^2; \tilde{d} > 1, r \leq 1\}, \quad S_2 := \{(X, Y) \in \mathcal{M}^2; r > 1\}, \quad (4.2)$$

it suffices to show that the integral operators with kernels $R_1(X, Y) \chi_{S_j}$ ($j = 1, 2$) are of weak type $(1, 1)$.

We start with the kernel $R_1(X, Y) \chi_{S_1}$. And we shall prove that

$$\sup_{Y \in \mathcal{M}} \int_{\mathcal{M}} R_1(X, Y) \chi_{S_1}(X, Y) d\mu_{\mathcal{M}}(X) < +\infty, \quad (4.3)$$

then the integral operator with the kernel $R_1(X, Y) \chi_{S_1}$ is bounded on $L^1(\mathcal{M})$. Indeed, we can write

$$\begin{aligned} 0 < R_1(X, Y) &\lesssim \int_0^\infty \left| \nabla_{\mathbb{H}^n} p_t^{\mathbb{H}^n}(u, v) \right| p_t^M(x, y) \frac{dt}{\sqrt{t}} \\ &\leq \int_0^\infty \left| \nabla_{\mathbb{H}^n} p_t^{\mathbb{H}^n}(u, v) \right| dt \cdot \sup_{t>0} \frac{1}{\sqrt{t}} p_t^M(x, y). \end{aligned} \quad (4.4)$$

As a consequence of (2.7), for $r \in (0, 1]$, we have

$$\int_0^\infty \left| \nabla_{\mathbb{H}^n} p_t^{\mathbb{H}^n}(u, v) \right| dt \lesssim r \int_0^\infty t^{-\frac{n+2}{2}} \exp\left(-\frac{r^2}{Ct}\right) dt \lesssim r^{-(n-1)} \lesssim \frac{r}{V_{\mathbb{H}^n}(v, r)}, \quad (4.5)$$

where we have used (1.4) in the last line.

On the other hand, using the symmetric property of the heat kernel, (GUE_θ) and (1.3) imply that

$$\begin{aligned} \sup_{t>0} \frac{1}{\sqrt{t}} p_t^M(x, y) &= \sup_{t>0} \frac{1}{\sqrt{t}} p_t^M(y, x) \\ &\lesssim \sup_{t>0} \frac{1}{\sqrt{t}} \frac{1}{\widetilde{V}_M(y, \sqrt{t})} e^{-c\left(\frac{\widetilde{d}^2}{t}\right)^\theta} \\ &\lesssim \sup_{t>0} \frac{\widetilde{d}}{\sqrt{t}} \left(1 + \frac{\widetilde{d}}{\sqrt{t}}\right)^v e^{-c\left(\frac{\widetilde{d}^2}{t}\right)^\theta} \cdot \frac{1}{\widetilde{d}} \frac{1}{\widetilde{V}_M(y, \widetilde{d})} \\ &\lesssim \frac{1}{\widetilde{d}} \frac{1}{\widetilde{V}_M(y, \widetilde{d})}. \end{aligned} \quad (4.6)$$

Then by the standard method of decomposition in annuli, the volume doubling property implies that

$$\begin{aligned} &\sup_{y \in M} \int_{\{x; \widetilde{d} > 1\}} \frac{1}{\widetilde{d}} \frac{1}{\widetilde{V}_M(y, \widetilde{d})} d\mu_M(x) \\ &= \sup_{y \in M} \sum_{j=0}^{+\infty} \int_{\{x; 2^j < \widetilde{d} \leq 2^{j+1}\}} \frac{1}{\widetilde{d}} \frac{1}{\widetilde{V}_M(y, \widetilde{d})} d\mu_M(x) < +\infty. \end{aligned} \quad (4.7)$$

Similarly, the local volume doubling property on $\mathbb{H}^n$ (cf. (ND_{loc})) implies that

$$\sup_{v \in \mathbb{H}^n} \int_{\{u; r \leq 1\}} \frac{r}{V_{\mathbb{H}^n}(v, r)} d\mu_{\mathbb{H}^n}(u) < +\infty. \quad (4.8)$$

Hence, (4.3) is a direct consequence of (4.4)-(4.8).

We now turn to study $R_1(X, Y) \chi_{S_2}$. For a fixed $\varepsilon \in (0, \frac{1}{64})$, we write

$$\begin{aligned} 0 < R_1(X, Y) &\lesssim \int_0^\infty \left| \nabla_{\mathbb{H}^n} p_t^{\mathbb{H}^n}(u, v) \right| p_t^M(x, y) \frac{dt}{\sqrt{t}} \\ &\leq \int_0^\infty \left| \nabla_{\mathbb{H}^n} p_t^{\mathbb{H}^n}(u, v) \right| e^{\varepsilon \frac{r}{t}} dt \cdot \sup_{t>0} \frac{1}{\sqrt{t}} p_t^M(x, y) e^{-\varepsilon \frac{r}{t}}. \end{aligned}$$

For $r > 1$, the gradient estimate (2.7) implies that

$$\begin{aligned} &\int_0^\infty \left| \nabla_{\mathbb{H}^n} p_t^{\mathbb{H}^n}(u, v) \right| e^{\varepsilon \frac{r}{t}} dt \\ &\lesssim e^{-(n-1)r} \sqrt{r} \int_0^\infty \left(\frac{r}{t}\right)^{\frac{1}{2}} \left(1 + \frac{r}{t}\right)^{\frac{n-1}{2}} \exp\left\{-\frac{r}{4} \left((n-1)\sqrt{\frac{t}{r}} - \sqrt{\frac{r}{t}}\right)^2 + \varepsilon \frac{r}{t}\right\} \frac{dt}{t}. \end{aligned}$$

Changing variables by $t = r\lambda$, the last integral is equal to

$$\begin{aligned} &\int_0^\infty \lambda^{-\frac{1}{2}} (1 + \lambda^{-1})^{\frac{n-1}{2}} \exp\left\{-\frac{r}{4} \frac{((n-1)\lambda - 1)^2}{\lambda} + \frac{\varepsilon}{\lambda}\right\} \frac{d\lambda}{\lambda} \\ &= \int_{(n-1)\lambda - 1 \geq \frac{1}{2}} + \int_{-1 < (n-1)\lambda - 1 \leq -\frac{1}{2}} + \int_{-\frac{1}{2} < (n-1)\lambda - 1 < \frac{1}{2}}; \end{aligned}$$

and the standard Laplace's method says that the main contribution comes from the third term, which is bounded by a multiple of $r^{-\frac{1}{2}}$. We conclude that when $r > 1$,

$$\int_0^\infty \left| \nabla_{\mathbb{H}^n} p_t^{\mathbb{H}^n}(u, v) \right| e^{\varepsilon \frac{r}{t}} dt \lesssim_\varepsilon e^{-(n-1)r}. \quad (4.9)$$

According to Lemma 2.3, the integral operator on $\mathbb{H}^n$, with the kernel defined by the left-hand side of (4.9), is of weak type $(1, 1)$.

On the other hand, by a suitable modification to the proof of (4.6), we can show that for $r > 1$,

$$\sup_{t>0} \frac{1}{\sqrt{t}} p_t^M(x, y) e^{-\varepsilon \frac{r}{t}} \lesssim_{\varepsilon} \min \left\{ \frac{1}{\tilde{V}_M(y, 1)}, \frac{1}{\tilde{d}} \frac{1}{\tilde{V}_M(y, \tilde{d})} \right\}. \quad (4.10)$$

We claim that the integral operator on M , with the kernel defined by the left-hand side of the last inequality, is bounded on L^1 . Indeed, it can be easily verified that

$$\sup_{y \in M} \int_M \min \left\{ \frac{1}{\tilde{V}_M(y, 1)}, \frac{1}{\tilde{d}} \frac{1}{\tilde{V}_M(y, \tilde{d})} \right\} d\mu_M(x) < +\infty. \quad (4.11)$$

Hence, by Lemma 2.4, the integral operator with the kernel $R_1(X, Y) \chi_{S_2}$ is of weak type $(1, 1)$. In conclusion, the weak type $(1, 1)$ estimate of $\mathcal{R}_{1, \infty}$ is established. $\square$

4.2 Weak $(1, 1)$ of $\mathcal{R}_{2, \infty}$

In this section, we turn to the analysis of $\mathcal{R}_{2, \infty}$. Compared with $\mathcal{R}_{1, \infty}$, the main difficulty here is the lack of a pointwise upper bound of $|\nabla_M p_t^M|$. By (2.3), without loss of generality, we only assume $(WGE_{1, \theta})$ on M , i.e. there exist constants $c, c' > 0$ and $0 < \theta \leq 1$ such that for all $t > 0$, $y \in M$,

$$\int_M |\nabla_M p_t^M(x, y)| \exp \left(c \left(\frac{\tilde{d}(x, y)^2}{t} \right)^{\theta} \right) d\mu_M(x) \leq \frac{c'}{\sqrt{t}}.$$

In particular, as its corollary

$$\int_M |\nabla_M p_t^M(x, y)| d\mu_M(x) \lesssim \frac{1}{\sqrt{t}}. \quad (4.12)$$

Theorem 4.2. $\mathcal{R}_{2, \infty}$ defined on $\mathcal{M}$ is of weak type $(1, 1)$.

Proof. As presented in the preceding subsection, it suffices to show that the integral operators with kernels $R_2(X, Y) \chi_{S_j}$ ($j = 1, 2$) are of weak type $(1, 1)$.

We start with the analysis of $R_2(X, Y) \chi_{S_1}$. Notice that

$$\begin{aligned} 0 < R_2(X, Y) &\lesssim \int_0^\infty p_t^{\mathbb{H}^n}(u, v) |\nabla_M p_t^M(x, y)| \frac{dt}{\sqrt{t}} \\ &\leq \sup_{t>0} \sqrt{t} p_t^{\mathbb{H}^n}(u, v) \cdot \left(\int_0^\infty |\nabla_M p_t^M(x, y)| \frac{dt}{t} \right). \end{aligned} \quad (4.13)$$

The heat kernel estimate on $\mathbb{H}^n$ (see (2.6)) implies that when $r \leq 1$,

$$\sup_{t>0} \sqrt{t} p_t^{\mathbb{H}^n}(u, v) \lesssim \sup_{t>0} t^{-\frac{n}{2} + \frac{1}{2}} e^{-\frac{r^2}{Ct}} \lesssim r^{-(n-1)} \lesssim \frac{r}{V_{\mathbb{H}^n}(v, r)}. \quad (4.14)$$

On the other hand, the assumption $(WGE_{1, \theta})$ on M implies that there exists a constant $c > 0$, such that for all $y \in M$ and $t > 0$,

$$\int_{\{x; \tilde{d}>1\}} |\nabla_M p_t^M(x, y)| d\mu_M(x) \lesssim \frac{e^{-ct^{-\theta}}}{\sqrt{t}}.$$

Hence

$$\sup_{y \in M} \int_{\{x; \tilde{d}>1\}} \left(\int_0^\infty |\nabla_M p_t^M(x, y)| \frac{dt}{t} \right) d\mu_M(x) \lesssim \int_0^\infty \frac{e^{-ct^{-\theta}}}{t^{3/2}} dt < +\infty. \quad (4.15)$$

It deduces from (4.13)-(4.15) and (4.8) that

$$\sup_{Y \in \mathcal{M}} \int_{\mathcal{M}} R_2(X, Y) \chi_{S_1}(X, Y) d\mu_{\mathcal{M}}(X) < +\infty.$$

Therefore the integral operator with the kernel $R_2(X, Y) \chi_{S_1}$ is bounded on L^1 .

We turn to study $R_2(X, Y) \chi_{S_2}$. For a fixed $\varepsilon \in (0, \frac{1}{64})$, we write

$$\begin{aligned} 0 < R_2(X, Y) &\lesssim \int_0^\infty p_t^{\mathbb{H}^n}(u, v) |\nabla_M p_t^M(x, y)| \frac{dt}{\sqrt{t}} \\ &\leq \sup_{t>0} \sqrt{t} p_t^{\mathbb{H}^n}(u, v) e^{\varepsilon \frac{r}{t}} \cdot \left(\int_0^\infty |\nabla_M p_t^M(x, y)| e^{-\varepsilon \frac{r}{t}} \frac{dt}{t} \right). \end{aligned}$$

The heat kernel estimate (2.4) implies that when $r > 1$,

$$\begin{aligned} &\sup_{t>0} \sqrt{t} p_t^{\mathbb{H}^n}(u, v) e^{\varepsilon \frac{r}{t}} \\ &\lesssim \sup_{t>0} \frac{r}{t} \left(1 + \frac{r}{t}\right)^{\frac{n-3}{2}} \exp \left\{ -\frac{r}{4} \left((n-1) \sqrt{\frac{t}{r}} - \sqrt{\frac{r}{t}} \right)^2 + \varepsilon \frac{r}{t} \right\} \cdot e^{-(n-1)r} \\ &\lesssim e^{-(n-1)r} \sup_{\lambda>0} \lambda^{-1} (1 + \lambda^{-1})^{\frac{n-3}{2}} \exp \left\{ -\frac{1}{4} \frac{((n-1)\lambda - 1)^2}{\lambda} + \frac{\varepsilon}{\lambda} \right\} \\ &\lesssim_\varepsilon e^{-(n-1)r}. \end{aligned} \tag{4.16}$$

In the last step we have used the following simple observation: the continuous function

$$g_\varepsilon(\lambda) := \lambda^{-1} (1 + \lambda^{-1})^{\frac{n-3}{2}} \exp \left\{ -\frac{1}{4} \frac{((n-1)\lambda - 1)^2}{\lambda} + \frac{\varepsilon}{\lambda} \right\},$$

which is defined on $(0, +\infty)$, satisfies $\lim_{\lambda \rightarrow 0^+} g_\varepsilon(\lambda) = \lim_{\lambda \rightarrow +\infty} g_\varepsilon(\lambda) = 0$.

Then we set

$$\begin{aligned} T(X, Y) &:= e^{-(n-1)r} \cdot \left(\int_0^\infty |\nabla_M p_t^M(x, y)| e^{-\varepsilon \frac{r}{t}} \frac{dt}{t} \right) \chi_{\{r>1\}}, \\ \mathcal{T}f(X) &:= \int_{\mathcal{M}} T(X, Y) |f(Y)| d\mu_{\mathcal{M}}(Y). \end{aligned}$$

It is enough to prove that the operator $\mathcal{T}$ is of weak type $(1, 1)$.

We define for $j \in \mathbb{N}$,

$$T_j(X, Y) := T(X, Y) \chi_{\{2^j < r \leq 2^{j+1}\}}, \quad \mathcal{T}_j f(X) := \int_{\mathcal{M}} T_j(X, Y) |f(Y)| d\mu_{\mathcal{M}}(Y).$$

Then it is clear that

$$T_j(X, Y) \leq e^{-(n-1)r} \cdot \left(\int_0^\infty |\nabla_M p_t^M(x, y)| e^{-\varepsilon \frac{2^j}{t}} \frac{dt}{t} \right).$$

We claim that $\|\mathcal{T}_j\|_{L^1 \rightarrow L^{1,\infty}} \lesssim_\varepsilon 2^{-j/2}$. Combining Lemma 2.3 and Lemma 2.4, it remains to show that

$$\sup_{y \in M} \int_M \left(\int_0^\infty |\nabla_M p_t^M(x, y)| e^{-\varepsilon \frac{2^j}{t}} \frac{dt}{t} \right) d\mu_M(x) \lesssim_\varepsilon 2^{-j/2}, \quad \forall j \in \mathbb{N}.$$

In fact, changing the order of integration and using (4.12) lead to

$$\begin{aligned} &\int_M \left(\int_0^\infty |\nabla_M p_t^M(x, y)| e^{-\varepsilon \frac{2^j}{t}} \frac{dt}{t} \right) d\mu_M(x) \\ &= \int_0^\infty \left(\int_M |\nabla_M p_t^M(x, y)| d\mu_M(x) \right) e^{-\varepsilon \frac{2^j}{t}} \frac{dt}{t} \end{aligned}$$

$$\lesssim \int_0^\infty \frac{e^{-\varepsilon \frac{2^j}{t}}}{t^{3/2}} dt = 2^{-j/2} \int_0^\infty \frac{e^{-\varepsilon s^{-1}}}{s^{3/2}} ds \lesssim_\varepsilon 2^{-j/2}.$$

Then [34, Lemma 2.3] implies the weak type $(1, 1)$ of $\mathcal{T} = \sum_{j=0}^\infty \mathcal{T}_j$. Hence the integral operator with the kernel $R_2(X, Y) \chi_{S_2}$ is of weak type $(1, 1)$. The proof is completed. $\square$

5 The heat maximal operator on $\mathcal{M}$

In this section, we shall establish the weak type $(1, 1)$ property of the heat maximal operator on $\mathcal{M}$, which is defined by

$$\mathcal{H}(f)(X) = \sup_{t>0} |e^{t\Delta_{\mathcal{M}}} f(X)|, \quad \forall f \in L^p(\mathcal{M}), 1 \leq p \leq +\infty.$$

The basic strategy is the same as that used in the study of Riesz transform. We first split $\mathcal{H}$ as the sum of the local part and the part at infinity. Then we prove they are of weak type $(1, 1)$ separately.

Let us first write $\mathcal{H}$ as

$$\begin{aligned} \mathcal{H}(f)(X) &\leq \sup_{t>0} \left| \int_{\{Y; d \leq 1\}} p_t^{\mathcal{M}}(X, Y) f(Y) d\mu_{\mathcal{M}}(Y) \right| + \sup_{t>0} \left| \int_{\{Y; d > 1\}} p_t^{\mathcal{M}}(X, Y) f(Y) d\mu_{\mathcal{M}}(Y) \right| \\ &\leq \sup_{t>0} \int_{\{Y; d \leq 1\}} p_t^{\mathcal{M}}(X, Y) |f(Y)| d\mu_{\mathcal{M}}(Y) + \int_{\{Y; d > 1\}} \sup_{t>0} p_t^{\mathcal{M}}(X, Y) |f(Y)| d\mu_{\mathcal{M}}(Y) \\ &:= \mathcal{H}_{\text{loc}}(f)(X) + \mathcal{H}_{\infty}(f)(X). \end{aligned}$$

The local part $\mathcal{H}_{\text{loc}}$ is easier to handle. Combining the local heat kernel estimate (3.28) and $(ND)_{\text{loc}}$, $\mathcal{H}_{\text{loc}}$ can be dominated by a constant multiple of the local centered Hardy-Littlewood maximal operator M_{loc} , which is defined by

$$M_{\text{loc}}(f)(X) = \sup_{0 < R \leq 1} \frac{1}{V_{\mathcal{M}}(X, R)} \int_{B(X, R)} |f(Y)| d\mu_{\mathcal{M}}(Y),$$

whose weak $(1, 1)$ property follows from $(ND)_{\text{loc}}$. Hence we need only consider the part at infinity $\mathcal{H}_{\infty}$.

Theorem 5.1. $\mathcal{H}_{\infty}$ defined on $\mathcal{M} = \mathbb{H}^n \times M$ is of weak type $(1, 1)$, where M is as described in Theorem 1.3.

Proof. It suffices to show that the integral operators with the kernels $\sup_{t>0} p_t^{\mathcal{M}}(X, Y) \chi_{S_j}$ ($j = 1, 2$) are of weak type $(1, 1)$.

We start with the kernel $\sup_{t>0} p_t^{\mathcal{M}}(X, Y) \chi_{S_1}$. By (1.5), we obtain for $X = (u, x), Y = (v, y) \in \mathbb{H}^n \times M$,

$$\sup_{t>0} p_t^{\mathcal{M}}(X, Y) \leq \sup_{t>0} \sqrt{t} p_t^{\mathbb{H}^n}(u, v) \cdot \sup_{t>0} \frac{1}{\sqrt{t}} p_t^M(x, y).$$

It deduced from (4.6), (4.7), (4.14) and (4.8) that

$$\sup_{Y \in \mathcal{M}} \int_{\mathcal{M}} \sup_{t>0} p_t^{\mathcal{M}}(X, Y) \chi_{S_1}(X, Y) d\mu_{\mathcal{M}}(X) < +\infty.$$

Therefore the integral operator with the kernel $\sup_{t>0} p_t^{\mathcal{M}}(X, Y) \chi_{S_1}$ is bounded on L^1 .

Then we turn to the analysis of $\sup_{t>0} p_t^{\mathcal{M}}(X, Y) \chi_{S_2}$. Taking an $\varepsilon \in (0, \frac{1}{64})$, by (1.5), we can write for $r > 1$

$$\sup_{t>0} p_t^{\mathcal{M}}(X, Y) \leq \sup_{t>0} \sqrt{t} p_t^{\mathbb{H}^n}(u, v) e^{\varepsilon \frac{r}{t}} \cdot \sup_{t>0} \frac{1}{\sqrt{t}} p_t^M(x, y) e^{-\varepsilon \frac{r}{t}}.$$

Combining this with (4.10), (4.11) and (4.16), it follows from Lemmas 2.3 and 2.4 that the integral operator with the kernel $\sup_{t>0} p_t^{\mathcal{M}}(X, Y) \chi_{S_2}$ is of weak type $(1, 1)$. This finishes the proof. $\square$

Remark 5.2. Recall that the first-order heat maximal operator on $\mathcal{M}$ is defined by

$$\mathcal{H}_1(f)(X) = \sup_{t>0} \left| t \frac{\partial}{\partial t} e^{t\Delta_{\mathcal{M}}} f(X) \right|, \quad \forall f \in L^p(\mathcal{M}), 1 \leq p < +\infty.$$

It follows from [17, Theorem 4] that $|t \frac{\partial}{\partial t} p_t^M(x, y)|$ also satisfies the upper bound of type (GUE_{θ}) . Using the time derivative estimates of the heat kernel on real hyperbolic spaces (see [25, Section 5]), a similar argument allows us to obtain the weak $(1, 1)$ property of $\mathcal{H}_1$.

6 Some Generalizations

Note that our method works without major difficulties for a large class of direct product manifolds. In this section, we discuss some generalizations.

Consider the direct product manifold $\mathcal{E} \times M$, where M is the same as in Theorem 1.1 or 1.3, and $\mathcal{E}$ is a connected complete non-compact weighted Riemannian manifold whose volume growth at infinity is at most exponential. We also denote by r the geodesic distance on $\mathcal{E}$. Given $\sigma \in [0, \frac{1}{2})$, $\tau \in (0, \frac{1}{2}]$ and $r_0, \varepsilon > 0$, we define integral operators $\mathcal{K}_i$ ($= \mathcal{K}_{i,\text{loc}} + \mathcal{K}_{i,\infty}$) ($i = 1, 2$) on $\mathcal{E}$, with the following kernels respectively,

$$\begin{aligned} K_1(u, v) &= K_{1,\text{loc}}(u, v) + K_{1,\infty}(u, v) \\ &:= \int_0^\infty |\nabla_{\mathcal{E}} p_t^{\mathcal{E}}(u, v)| \frac{dt}{t^\sigma} \chi_{\{r \leq r_0\}} + \int_0^\infty |\nabla_{\mathcal{E}} p_t^{\mathcal{E}}(u, v)| e^{\varepsilon \frac{r}{t}} \frac{dt}{t^\sigma} \chi_{\{r > r_0\}}, \end{aligned} \quad (6.1)$$

and

$$\begin{aligned} K_2(u, v) &= K_{2,\text{loc}}(u, v) + K_{2,\infty}(u, v) \\ &:= \sup_{t>0} t^\tau p_t^{\mathcal{E}}(u, v) \chi_{\{r \leq r_0\}} + \sup_{t>0} t^\tau p_t^{\mathcal{E}}(u, v) e^{\varepsilon \frac{r}{t}} \chi_{\{r > r_0\}}. \end{aligned} \quad (6.2)$$

We are thus led to the following generalization of Theorems 1.1 and 1.3.

Theorem 6.1. Assume that $\mathcal{E}$ satisfies (ND_{loc}) and (VE) , whose heat kernel $p_t^{\mathcal{E}}$ has the small-time upper bound

$$p_t^{\mathcal{E}}(u, u) \leq \frac{C}{V_{\mathcal{E}}(u, \sqrt{t})}, \quad \forall u \in \mathcal{E}, 0 < t \leq 1.$$

Furthermore, suppose that there exist $\sigma \in [0, \frac{1}{2})$, $\tau \in (0, \frac{1}{2}]$ and $r_0, \varepsilon > 0$, such that the integral operators $\mathcal{K}_i$ ($i = 1, 2$) are of weak type $(1, 1)$. Then the Riesz transform defined on $\mathcal{E} \times M$ is of weak type $(1, 1)$, where M is as described in Theorem 1.1.

Theorem 6.2. Assume that $\mathcal{E}$ satisfies (ND_{loc}) and (VE) , whose heat kernel $p_t^{\mathcal{E}}$ has the small-time upper bound

$$p_t^{\mathcal{E}}(u, u) \leq \frac{C}{V_{\mathcal{E}}(u, \sqrt{t})}, \quad \forall u \in \mathcal{E}, 0 < t \leq 1.$$

Furthermore, suppose that there exist $\tau \in (0, \frac{1}{2}]$ and $r_0, \varepsilon > 0$, such that the integral operator $\mathcal{K}_2$ is of weak type $(1, 1)$. Then the heat maximal operator defined on $\mathcal{E} \times M$ is of weak type $(1, 1)$, where M is as described in Theorem 1.3.

The proofs of Theorems 1.1 and 1.3 are organized in such a way that they could be applied to this more general setting via a slight modification. So we omit the proofs of Theorems 6.1 and 6.2. It is worth noting that the assumptions of Theorem 6.1 actually imply the weak type $(1, 1)$ estimate of the Riesz transform defined on $\mathcal{E}$. In fact, it follows from spectral analysis that the weak type $(1, 1)$ estimate of $\mathcal{K}_{1,\infty}$ can deduce that of $\nabla_{\mathcal{E}}(-\Delta_{\mathcal{E}})^{-s} \chi_{\{r > r_0\}}$ for $s \in (0, 1 - \sigma]$. While by Theorem 3.1, the assumptions on $\mathcal{E}$ imply the weak type $(1, 1)$ estimate of $\nabla_{\mathcal{E}}(-\Delta_{\mathcal{E}})^{-1/2} \chi_{\{r \leq r_0\}}$. Similarly, the weak type $(1, 1)$ of the heat maximal operator defined on $\mathcal{E}$ follows.

Remark 6.3. (1) It is worth pointing out that the factor $e^{\varepsilon \frac{r}{t}}$ introduced in the definition of $\mathcal{K}_{i,\infty}$ ($i = 1, 2$) is not unique. For example, we can replace it by $(1 + \frac{r}{t})^\varsigma$ with $\varsigma \geq \frac{v+1}{2}$, where $v > 0$ is the constant in (1.3), with a slight modification of the proof. The advantage of using $e^{\varepsilon \frac{r}{t}}$ lies in the fact that it only depends on properties of $\mathcal{E}$.

(2) Theorems 6.1 and 6.2 remain valid if M is a Lie group with polynomial growth.

6.1 Sufficient conditions for L^1 boundedness of $\mathcal{K}_{i,\text{loc}}$

Since $\mathcal{E}$ satisfies (ND_{loc}) , even more is expected to be said about the local operators $\mathcal{K}_{i,\text{loc}}$ ($i = 1, 2$). In fact, they will be bounded on $L^1(\mathcal{E})$ under some additional assumptions, particularly for $\mathcal{E}$ with a positive bottom of the spectrum, such as $\mathbb{H}^n$. Before going further, we first recall that the estimates established in Section 3.2, particularly (3.24) and (3.27), remain valid on $\mathcal{E}$. Hence they will be used directly.

We begin by analysing $\mathcal{K}_{1,\text{loc}}$.

$$\begin{aligned} \int_{\mathcal{E}} K_{1,\text{loc}}(u, v) d\mu_{\mathcal{E}}(u) &= \int_0^\infty \left(\int_{\{u; r \leq r_0\}} |\nabla_{\mathcal{E}} p_t^{\mathcal{E}}(u, v)| d\mu_{\mathcal{E}}(u) \right) \frac{dt}{t^\sigma} \\ &= \left(\int_0^1 + \int_1^\infty \right) \left(\int_{\{u; r \leq r_0\}} |\nabla_{\mathcal{E}} p_t^{\mathcal{E}}(u, v)| d\mu_{\mathcal{E}}(u) \right) \frac{dt}{t^\sigma}. \end{aligned}$$

By (3.24), we have

$$\sup_{v \in \mathcal{E}} \int_0^1 \left(\int_{\{u; r \leq r_0\}} |\nabla_{\mathcal{E}} p_t^{\mathcal{E}}(u, v)| d\mu_{\mathcal{E}}(u) \right) \frac{dt}{t^\sigma} \lesssim \int_0^1 \frac{dt}{t^{\sigma+1/2}} < +\infty.$$

On the other hand, when $t \geq 1$, by Cauchy-Schwarz inequality, we get

$$\int_{\{u; r \leq r_0\}} |\nabla_{\mathcal{E}} p_t^{\mathcal{E}}(u, v)| d\mu_{\mathcal{E}}(u) \leq V_{\mathcal{E}}(v, r_0)^{1/2} \left(\int_{\mathcal{E}} |\nabla_{\mathcal{E}} p_t^{\mathcal{E}}(u, v)|^2 d\mu_{\mathcal{E}}(u) \right)^{1/2}. \quad (6.3)$$

Using integration by parts, Cauchy-Schwarz inequality implies that

$$\begin{aligned} \int_{\mathcal{E}} |\nabla_{\mathcal{E}} e^{t\Delta_{\mathcal{E}}/2} g(u)|^2 d\mu_{\mathcal{E}}(u) &= - \int_{\mathcal{E}} g(u) (\Delta_{\mathcal{E}} e^{t\Delta_{\mathcal{E}}} g)(u) d\mu_{\mathcal{E}}(u) \\ &\leq \|g\|_2 \|\Delta_{\mathcal{E}} e^{t\Delta_{\mathcal{E}}} g\|_2 \leq t^{-1} \|g\|_2^2, \quad \forall t > 0, g \in C_0^\infty(\mathcal{E}), \end{aligned} \quad (6.4)$$

where we have used in the last inequality the analyticity of $(e^{t\Delta_{\mathcal{E}}})_{t>0}$ on L^2 . It follows from the semigroup property that $p_t^{\mathcal{E}}(u, v) = e^{t\Delta_{\mathcal{E}}/2}(p_{t/2}^{\mathcal{E}}(\cdot, v))(u)$. Therefore combining (6.3) and (6.4) gives

$$\int_{\{u; r \leq r_0\}} |\nabla_{\mathcal{E}} p_t^{\mathcal{E}}(u, v)| d\mu_{\mathcal{E}}(u) \leq \frac{V_{\mathcal{E}}(v, r_0)^{1/2} \|p_{t/2}^{\mathcal{E}}(\cdot, v)\|_2}{t^{1/2}} = \frac{\sqrt{V_{\mathcal{E}}(v, r_0) p_t^{\mathcal{E}}(v, v)}}{t^{1/2}}.$$

Finally, we conclude that if

$$\sup_{v \in \mathcal{E}} \int_1^\infty \frac{\sqrt{V_{\mathcal{E}}(v, r_0) p_t^{\mathcal{E}}(v, v)}}{t^{\sigma+1/2}} dt < +\infty \quad (6.5)$$

holds, then $\mathcal{K}_{1,\text{loc}}$ is bounded on $L^1(\mathcal{E})$.

Now, we turn to the integral operator $\mathcal{K}_{2,\text{loc}}$. It follows from (3.27) that there exists a constant $C > 0$, such that for all $u, v \in \mathcal{E}$ with $r = r(u, v) \leq r_0$,

$$\begin{aligned} K_{2,\text{loc}}(u, v) &\leq \sup_{0 < t \leq 1} t^\tau p_t^{\mathcal{E}}(u, v) + \sup_{t > 1} t^\tau p_t^{\mathcal{E}}(u, v) \\ &\lesssim \sup_{0 < t \leq 1} \frac{t^\tau}{V_{\mathcal{E}}(u, \sqrt{t})} e^{-\frac{r^2}{Ct}} + \sup_{t > 1} t^\tau p_t^{\mathcal{E}}(u, v). \end{aligned} \quad (6.6)$$

First, we claim that there exist positive constants v_1 and v_2 such that

$$\frac{V_{\mathcal{E}}(u, R_1)}{V_{\mathcal{E}}(u, R_2)} \lesssim_{r_0} \left(\frac{R_1}{R_2}\right)^{v_1} + \left(\frac{R_1}{R_2}\right)^{v_2}, \quad \forall u \in \mathcal{E}, R_1 \in (0, r_0], R_2 \in (0, 1]. \quad (6.7)$$

Indeed, when $R_1 \geq R_2$, by iterating (ND_{loc}) , there exists $v_1 > 0$ such that

$$\frac{V_{\mathcal{E}}(u, R_1)}{V_{\mathcal{E}}(u, R_2)} \lesssim_{r_0} \left(\frac{R_1}{R_2}\right)^{v_1}, \quad 0 < R_1 \leq r_0.$$

When $R_1 < R_2$, the so-called (local) reversed volume doubling property (see [22, page 412]), which is a direct consequence of (ND_{loc}) and the connectedness of $\mathcal{E}$, implies that there exists $v_2 > 0$ such that

$$\frac{V_{\mathcal{E}}(u, R_1)}{V_{\mathcal{E}}(u, R_2)} \lesssim_{r_0} \left(\frac{R_1}{R_2}\right)^{v_2}.$$

In both cases, (6.7) is true.

Then applying (6.7) to the first term in (6.6), we have

$$\sup_{0 < t \leq 1} \frac{t^\tau}{V_{\mathcal{E}}(u, \sqrt{t})} e^{-\frac{r^2}{Ct}} \lesssim_{r_0} \sup_{0 < t \leq 1} \left(\left(\frac{r}{\sqrt{t}}\right)^{v_1-2\tau} + \left(\frac{r}{\sqrt{t}}\right)^{v_2-2\tau} \right) e^{-\frac{r^2}{Ct}} \cdot \frac{r^{2\tau}}{V_{\mathcal{E}}(u, r)}. \quad (6.8)$$

By the same argument used in (4.8), we deduce that for any $\tau > 0$,

$$\sup_{v \in \mathcal{E}} \int_{\{u; r \leq r_0\}} \frac{r^{2\tau}}{V_{\mathcal{E}}(u, r)} d\mu_{\mathcal{E}}(u) < +\infty.$$

Set $v' = \min\{v_1, v_2\}$. Therefore if $\tau \in (0, \frac{v'}{2}]$, the integral operator with the kernel $\sup_{0 < t \leq 1} t^\tau p_t^\mathcal{E} \chi_{\{r \leq r_0\}}$ is bounded on $L^1(\mathcal{E})$.

In conclusion, $\mathcal{K}_{2, \text{loc}}$ is bounded on $L^1(\mathcal{E})$ whenever there exists $\tau \in (0, \frac{v'}{2}]$ such that

$$\sup_{v \in \mathcal{E}} \int_{\{u; r \leq r_0\}} \sup_{t > 1} t^\tau p_t^\mathcal{E}(u, v) d\mu_{\mathcal{E}}(u) < +\infty. \quad (6.9)$$

In particular, if $\mathcal{E}$ has a positive bottom of the spectrum λ_1 , then by (3.27) and (ND_{loc}) , the conditions (6.5) and (6.9) can be verified easily.

6.2 More specific examples

To end this section, we present more specific examples of $\mathcal{E}$. We start with the following simple but instructive model. Given $\alpha > 0$, for non-negative numbers a_1 and a_2 that are not all zero, the function

$$h(u) = a_1 e^{\alpha u} + a_2 e^{-\alpha u}$$

is a positive α^2 -harmonic function defined on $\mathbb{R}$. Moreover, all the positive eigenfunctions of $\Delta_{\mathbb{R}}$ corresponding to a positive eigenvalue take this form. Let μ be the measure on $\mathbb{R}$ defined by $d\mu = h^2 du$, where du is the Lebesgue measure. We consider the weighted Riemannian manifold $\mathcal{E} = (\mathbb{R}, d\mu)$, equipped with the canonical Euclidean metric. It is easily seen that for all $u \in \mathcal{E}$,

$$V_{\mathcal{E}}(u, R) \sim \begin{cases} h^2(u)R & \text{if } 0 < R \leq 1, \\ h^2(u)e^{2\alpha R} & \text{if } R > 1. \end{cases}$$

The (weighted) Laplace operator on $\mathcal{E}$ is as follows:

$$\Delta_{\mathcal{E}} = \frac{d^2}{du^2} + 2 \frac{h'}{h} \frac{d}{du} = \frac{d^2}{du^2} + 2\alpha \frac{a_1 e^{2\alpha u} - a_2}{a_1 e^{2\alpha u} + a_2} \frac{d}{du}.$$

By Doob's h -transform (see [22]), the heat kernel on $\mathcal{E}$ is explicitly given by

$$p_t^{\mathcal{E}}(u, v) = \frac{1}{\sqrt{4\pi t}} \exp\left(-\frac{r^2}{4t} - \alpha^2 t\right) \frac{1}{h(u)h(v)},$$

and a direct computation gives rise to

$$|\nabla_{\mathcal{E}} p_t^{\mathcal{E}}(u, v)| \lesssim t^{-\frac{3}{2}} (r + t) \exp\left(-\frac{r^2}{4t} - \alpha^2 t\right) \frac{1}{h(u)h(v)},$$

where $r = r(u, v) = |u - v|$.

To apply Theorems 6.1 and 6.2, it only remains to verify the weak type $(1, 1)$ of $\mathcal{K}_i$ ($i = 1, 2$), since other assumptions on $\mathcal{E}$ are satisfied trivially. To do this, take $\sigma = 0$, $\tau = \frac{1}{2}$, $r_0 = 1$ and $\varepsilon \in (0, \frac{1}{8})$ in the expressions of $\mathcal{K}_i$. Remark that the local operators $\mathcal{K}_{i,\text{loc}}$ ($i = 1, 2$) are bounded on L^1 because $\lambda_1 = \alpha^2$ in this setting.

On the other hand, we have

$$K_{1,\infty}(u, v) \lesssim \int_0^\infty t^{-\frac{3}{2}} (r + t) \exp\left(-\frac{r^2}{4t} - \alpha^2 t + \frac{\varepsilon r}{t}\right) dt \cdot \frac{1}{h(u)h(v)} \lesssim \frac{e^{-\alpha r}}{h(u)h(v)},$$

and

$$K_{2,\infty}(u, v) \lesssim \sup_{t>0} \exp\left(-\frac{r^2}{4t} - \alpha^2 t + \frac{\varepsilon r}{t}\right) \cdot \frac{1}{h(u)h(v)} \lesssim \frac{e^{-\alpha r}}{h(u)h(v)}.$$

Notice that $\frac{e^{-\alpha r}}{h(u)h(v)} \leq \frac{1}{h^2(u)} \in L^{1,\infty}(\mathcal{E})$, the weak type $(1, 1)$ of $\mathcal{K}_i$ follows. Therefore $(\mathbb{R}, d\mu)$ is an example of $\mathcal{E}$ in our theorems. In particular, it is known that the Euclidean space $\mathbb{R}^n$ endowed with the drifted Laplacian can be viewed as a direct product manifold in essence. Then if we take $\mathcal{E} = (\mathbb{R}, e^{2u} du)$ and $M = (\mathbb{R}^{n-1}, dx)$, applying Theorems 6.1 and 6.2 gives an alternative proof of Theorems 1 and 2 in [27].

We emphasize that the assumptions on $\mathcal{E}$ could be verified on some manifolds with a positive bottom of the spectrum. As illustrated in Section 6.1, the main difficulty in this context is the weak $(1, 1)$ of $\mathcal{K}_{i,\infty}$ ($i = 1, 2$), which usually follows from the analogues of Lemma 2.3, involving sharp estimates of heat kernel and its gradient. According to the known results on the weak type $(1, 1)$, Theorems 6.1 and 6.2 are valid when $\mathcal{E}$ is taken as the following manifolds mentioned in Section 1.2: non-compact symmetric spaces, harmonic AN groups, the Laplacian with drift on real hyperbolic spaces, as well as the sub-Laplacian with drift on Heisenberg groups. As for specific properties of guaranteeing the validity of our assumptions, we refer the reader to [2, 36], [3], [25] and [26]. The details are left to the reader.

Acknowledgements This work is partially supported by NSF of China (Grants No. 12271102 No.11625102 and No. 11831004). The second author is also partially supported by the National Key R&D Program of China (No. 2022YFA1006000, 2020YFA0712900) and NNSFC (11921001). The authors would like to thank the anonymous referees for their many useful suggestions and valuable remarks which improve the writing of the paper.

References

- 1 Alexopoulos G. An application of homogenization theory to harmonic analysis: Harnack inequalities and Riesz transforms on Lie groups of polynomial growth. *Canad J Math*, 1992, 44: 691–727
- 2 Anker J-P. Sharp estimates for some functions of the Laplacian on noncompact symmetric spaces. *Duke Math J*, 1992, 65: 257–297
- 3 Anker J-P, Damek E, Yacoub C. Spherical analysis on harmonic AN groups. *Ann Scuola Norm Sup Pisa Cl Sci* (4), 1996, 23: 643–679
- 4 Auscher P, Coulhon T, Duong X T, Hofmann S. Riesz transform on manifolds and heat kernel regularity. *Ann Sci École Norm Sup* (4), 2004, 37: 911–957
- 5 Bakry D. Transformations de Riesz pour les semi-groupes symétriques. II. Étude sous la condition $\Gamma_2 \geq 0$. In: *Séminaire de Probabilités XIX, Lecture Notes in Math*, 1123. Berlin: Springer, 1985, 145–174

- 6 Bakry D. Étude des transformations de Riesz sur les variétés riemanniennes à courbure de Ricci minorée. In: Séminaire de Probabilités XXI, Lecture Notes in Math, 1247. Berlin: Springer, 1987, 137–172
- 7 Bakry D, Gentil I, Ledoux M. Analysis and geometry of Markov diffusion operators. Springer Cham, 2014
- 8 Blunck S, Kunstmann P C. Calderón-Zygmund theory for non-integral operators and the H^∞ -functional calculus. *Rev Mat Iberoam*, 2003, 19: 919–942
- 9 Blunck S, Kunstmann P C. Weak type (p, p) estimates for Riesz transform. *Math Z*, 2004, 247: 137–148
- 10 Carron G, Coulhon T, Hassell A. Riesz transform and L^p -cohomology for manifolds with Euclidean ends. *Duke Math J*, 2006, 133: 59–93
- 11 Chen L, Coulhon T, Feneuil J, Russ E. Riesz transform for $1 \leq p \leq 2$ without Gaussian heat kernel bound. *J Geom Anal*, 2017, 27: 1489–1514
- 12 Coifman R R, Weiss G. Analyse harmonique non-commutative sur certains espaces homogènes. Berlin: Springer-Verlag, 1971
- 13 Coulhon T, Duong X T. Riesz transforms for $1 \leq p \leq 2$. *Trans Amer Math Soc*, 1999, 351: 1151–1169
- 14 Coulhon T, Li H-Q. Estimations inférieures du noyau de la chaleur sur les variétés coniques et transformée de Riesz. *Arch Math*, 2004, 83: 229–242
- 15 Cowling M, Gaudry G, Giulini S, Mauceri G. Weak type $(1, 1)$ estimates for heat kernel maximal functions on Lie groups. *Trans Amer Math Soc*, 1991, 323: 637–649
- 16 Davies E B. Heat kernels and spectral theory. Cambridge: Cambridge University Press, 1989
- 17 Davies E B. Non-Gaussian aspects of heat kernel behaviour. *J Lond Math Soc*, 1997, 55: 105–125
- 18 Davies E B, Mandouvalos N. Heat kernel bounds on hyperbolic space and Kleinian groups. *Proc Lond Math Soc*, 1988, 57: 182–208
- 19 Duong X T, McIntosh A. Singular integral operators with non-smooth kernels on irregular domains. *Rev. Mat. Iberoam*, 1999, 15: 233–265
- 20 Grigor'yan A. Upper bounds of derivatives of the heat kernel on an arbitrary complete manifold. *J Funct Anal*, 1995, 127: 363–389
- 21 Grigor'yan A. Gaussian upper bounds for the heat kernel on arbitrary manifolds. *J Diff Geom*, 1997, 45: 33–52
- 22 Grigor'yan A. Heat kernel and analysis on manifolds. Boston: International Press, 2009
- 23 Hassell A, Sikora A. Riesz transforms on a class of non-doubling manifolds. *Comm Partial Differential Equations*, 2019, 44: 1072–1099
- 24 Li H-Q. La transformation de Riesz sur les variétés coniques. *J Funct Anal*, 1999, 168: 145–238
- 25 Li H-Q, Sjögren P. Weak type $(1, 1)$ for some operators related to the Laplacian with drift on real hyperbolic spaces. *Potential Anal*, 2017, 46: 463–484
- 26 Li H-Q, Sjögren P. Estimates for operators for the sub-Laplacian with drift in Heisenberg groups. *J Fourier Anal Appl*, 2022, 28: Paper No. 10
- 27 Li H-Q, Sjögren P, Wu Y-R. Weak type $(1, 1)$ for some operators for the Laplacian with drift. *Math Z*, 2016, 282: 623–633
- 28 Li H-Q, Zhu J-X. A note on “Riesz transform for $1 \leq p \leq 2$ without Gaussian heat kernel bound”. *J Geom Anal*, 2018, 28: 1597–1609
- 29 Lohoué N. Comparaison des champs de vecteurs et des puissances du laplacien sur une variété riemannienne à courbure non positive. *J Funct Anal*, 1985, 61: 164–201
- 30 Lohoué N. Transformées de Riesz et fonctions de Littlewood-Paley sur les groupes non moyennables. *C R Acad Sci Paris*, 1988, 306: 327–330
- 31 Saloff-Coste L. Aspects of Sobolev-type inequalities. Cambridge: Cambridge University Press, 2002
- 32 Stein E M. Singular integrals and differentiability properties of functions. Princeton: Princeton University Press, 1970.
- 33 Stein E M. Topics in Harmonic Analysis Related to the Littlewood-Paley Theory. Princeton: Princeton University Press; Tokyo: University of Tokyo Press, 1970
- 34 Stein E M, Weiss N J. On the convergence of Poisson integrals. *Trans Amer Math Soc*, 1969, 140: 35–54
- 35 Strichartz R S. Analysis of the Laplacian on the complete Riemannian manifold. *J Funct Anal*, 1983, 52: 48–79
- 36 Strömberg J-O, Weak type L^1 estimates for maximal functions on noncompact symmetric spaces. *Ann of Math (2)*, 1981, 114: 115–126

Appendix A Two counterexamples for $p > 2$

As already mentioned in our introduction, the L^p boundedness ($p > 2$) of the Riesz transform on $\mathcal{M} = \mathbb{H}^n \times M$ is a quite subtle problem. In general, (ND) , (GUE_θ) together with $(WGE_{q,\theta})$ of M is not sufficient. To illustrate this, we shall present two typical counterexamples. In this appendix, we always assume M satisfies (D) and (UE_2) , i.e. the quasi-distance $\tilde{d}$ is the geodesic distance d_M . Let us recall

that $(WGE_{2,1})$ is a corollary of our assumptions.

Example A.1. For $k \geq 2$, let $M = \mathbb{R}^k \# \mathbb{R}^k$ be the connected sum of two copies of $\mathbb{R}^k$, where the Riesz transform is bounded on L^p iff $1 < p \leq 2$ when $k = 2$ and $1 < p < k$ when $k \geq 3$ (see [10, 13]). By its construction, the Ricci curvature Ric_M equals zero identically outside a compact set, hence is bounded from below. On the other hand, it is well-known that $\text{Ric}_{\mathbb{H}^n} \equiv -(n-1)$. Therefore the Ricci curvature is bounded from below on $\mathcal{M} = \mathbb{H}^n \times M$. Then applying the main theorem of [6], we obtain for any given $1 < p < +\infty$,

$$\|\nabla_{\mathcal{M}} f\|_p \lesssim \|(-\Delta_{\mathcal{M}})^{\frac{1}{2}} f\|_p + \|f\|_p, \quad \forall f \in C_0^\infty(\mathcal{M}).$$

And one easily sees that $\mathcal{M}$ has a positive bottom of the spectrum. By interpolation,

$$\|f\|_p \lesssim \|(-\Delta_{\mathcal{M}})^{\frac{1}{2}} f\|_p, \quad \forall f \in C_0^\infty(\mathcal{M}).$$

Combining with the above inequalities, the L^p boundedness of the Riesz transform on $\mathcal{M}$ holds for any $1 < p < +\infty$. It is a bit surprising since taking direct product enlarges the range of p for the L^p boundedness of the Riesz transform. In addition, this argument does not use (D) and (UE_2) , so the same phenomenon happens when M is taken as the connected sums considered in [23].

Example A.2. Let M be a k -dimensional conical manifold with a compact boundaryless basis. The Riesz transform on M has been studied in [24]. Let λ_1 be the smallest strictly positive eigenvalue of the basis. Then the L^p boundedness of the Riesz transform is valid iff $1 < p < p_0$, where

$$p_0 = \begin{cases} k \left(\frac{k}{2} - \sqrt{\left(\frac{k-2}{2}\right)^2 + \lambda_1} \right)^{-1} & \text{if } \lambda_1 < k-1, \\ +\infty & \text{if } \lambda_1 \geq k-1. \end{cases}$$

We shall only focus on the case $\lambda_1 < k-1$. For any fixed $p \geq p_0$, suppose that the Riesz transform on $\mathcal{M} = \mathbb{H}^n \times M$ is bounded on L^p , i.e.

$$\|\nabla_{\mathcal{M}} F\|_p \lesssim \|(-\Delta_{\mathcal{M}})^{\frac{1}{2}} F\|_p, \quad \forall F \in D((-\Delta_{\mathcal{M}})^{\frac{1}{2}}). \quad (\text{A.1})$$

By the results of [14, p. 239], there exists $g \in C_0^\infty(M)$ and $t_0 > 0$, such that

$$\|\nabla_M e^{t_0 \Delta_M} g\|_p = +\infty.$$

Now we choose an $f \in C_0^\infty(\mathbb{H}^n)$ that is not identically 0, and set

$$\begin{aligned} G(u, x) &:= e^{t_0 \Delta_{\mathbb{H}^n}} f(u) \cdot e^{t_0 \Delta_M} g(x) \\ &= e^{t_0 \Delta_{\mathcal{M}}} (f \otimes g) \in D((-\Delta_{\mathcal{M}})^{\frac{1}{2}}). \end{aligned}$$

Then it is easily checked that

$$\begin{aligned} \|\nabla_{\mathcal{M}} G\|_p &\sim \|\nabla_{\mathbb{H}^n} e^{t_0 \Delta_{\mathbb{H}^n}} f\|_p \cdot \|e^{t_0 \Delta_M} g\|_p + \|e^{t_0 \Delta_{\mathbb{H}^n}} f\|_p \cdot \|\nabla_M e^{t_0 \Delta_M} g\|_p \\ &= +\infty. \end{aligned}$$

On the other hand, by the analyticity of $(e^{t \Delta_{\mathcal{M}}})_{t>0}$ on L^p , we obtain

$$\begin{aligned} \|(-\Delta_{\mathcal{M}})^{\frac{1}{2}} G\|_p &= \|(-\Delta_{\mathcal{M}})^{\frac{1}{2}} e^{t_0 \Delta_{\mathcal{M}}} (f \otimes g)\|_p \\ &\leq t_0^{-\frac{1}{2}} \|f \otimes g\|_p < +\infty, \end{aligned}$$

which contradicts (A.1). So when $p_0 \leq p < +\infty$, the Riesz transform on $\mathcal{M}$ is not L^p bounded.

As for the L^p boundedness ($p > 2$) of the Riesz transform on $\mathcal{M} = \mathbb{H}^n \times M$ for a general M , we can apply the criterions developed in [4]. Indeed, by [31, Theorem 5.5.3], the conjunction of (ND_{loc}) and the local Poincaré inequality is equivalent to the two-sided Gaussian heat kernel estimate for small time. Accordingly, [4, Theorem 1.8] asserts the following result:

Assume M is a connected stochastically complete non-compact Riemannian manifold satisfying (D), whose heat kernel has the two-sided Gaussian estimate for small time: there exist constants C and c , such that for all $x, y \in M$ and $0 < t < 1$,

$$\frac{c}{V_M(x, \sqrt{t})} \exp\left(-C \frac{d_M(x, y)^2}{t}\right) \leq p_t^M(x, y) \leq \frac{C}{V_M(x, \sqrt{t})} \exp\left(-c \frac{d_M(x, y)^2}{t}\right).$$

If for some $p_0 \in (2, +\infty]$ and $\alpha \geq 0$, it holds for all $t > 0$ that

$$\|\|\nabla_M e^{t\Delta_M}\|\|_{p_0 \rightarrow p_0} \lesssim \frac{e^{\alpha t}}{\sqrt{t}},$$

then the Riesz transform on $\mathcal{M}$ is bounded on L^p for all $p \in [2, p_0)$.