Hessian heat kernel estimates and Calderón-Zygmund inequalities on complete Riemannian manifolds

Jun Cao[1], Li-Juan Cheng*[2], and Anton Thalmaier[3]

[1]Department of Applied Mathematics, Zhejiang University of Technology,
Hangzhou 310023, The People's Republic of China
caojun1860@zjut.edu.cn

[2]School of Mathematics, Hangzhou Normal University
Hangzhou 311123, The People's Republic of China
chenglj@zjut.edu.cn

[3]Department of Mathematics, University of Luxembourg, Maison du Nombre,
L-4364 Esch-sur-Alzette, Luxembourg
anton.thalmaier@uni.lu

January 10, 2022

Abstract

We address some fundamental questions about geometric analysis on Riemannian manifolds. The L^p-Calderón-Zygmund inequality is one of the cornerstones in the regularity theory of elliptic equations, and it has been asked under which geometric conditions it holds for a reasonable class of non-compact Riemannian manifolds, and to what extent assumptions on the derivative of curvature and on the injectivity radius of the manifold are necessary. In the present paper, for $1 < p < 2$, we give a positive answer for the validity of the L^p-Calderón-Zygmund inequality on a Riemannian manifold assuming only a lower bound on the Ricci curvature. It is well known that this alone is not sufficient for $p > 2$. In this case we complement the study of Güneysu-Pigola (2015) and derive sufficient geometric criteria for the validity of the Calderón-Zygmund inequality under additional Kato class bounds on the Riemann curvature tensor and the covariant derivative of Ricci curvature. Bounds in the Kato class are integral conditions and much weaker than pointwise bounds. Throughout the proofs, probabilistic tools, like Hessian formulas and Bismut type representations for heat semigroups, play a significant role.

Contents

* Corresponding author

2010 *Mathematics Subject Classification*. Primary: 35K08; Secondary: 35J10, 35J30, 47G40.

Key words and phrases. Brownian motion, Heat semigroup, Heat kernel, Bismut type Hessian representations, Hessian formula, Calderón-Zygmund inequalities, the Hardy-Littlewood maximal function, good-λ inequalities, Kato class.

This work has been project is supported by the National Natural Science Foundation of China (Grant No. 12071431). The third author has been supported by the Fonds National de la Recherche Luxembourg (project GEOMREV O14/7628746).

1 Introduction

The Hessian operator Hess, which contains all information of the second order derivatives, is a fundamental object in the second order smooth analysis. To study the Hessian operator, one usually needs the Calderón-Zygmund inequality that controls Hess by the simpler Laplacian operator Δ (see [6, 20, 35, 26, 36] and the references therein for its wide applications particularly in the regularity theory of elliptic equations).

In the Euclidean space $\mathbb{R}^d$, the Calderón-Zygmund inequality says that for any $p \in (1, \infty)$ and $u \in C_c^\infty(\mathbb{R}^n)$, it holds

$$\|\operatorname{Hess} u\|_{L^p(\mathbb{R}^d)} \le C\|\Delta u\|_{L^p(\mathbb{R}^d)}, \tag{1.1}$$

where $C = C(d, p) > 0$ is a constant depending only on d and p. The inequality (1.1) is first proved by Calderón and Zygmund via their seminal theory of singular integral operators based on the explicit representation of the Green kernel of the Laplacian. A remarkable consequence of inequality (1.1) is the fact that, in the Euclidean space, the Sobolev norms

$$\begin{aligned}\|u\|_{W^{2,p}(\mathbb{R}^d)} &= \|u\|_{L^p(\mathbb{R}^d)} + \|\nabla u\|_{L^p(\mathbb{R}^d)} + \|\operatorname{Hess} u\|_{L^p(\mathbb{R}^d)},\\ \|u\|_{\widetilde{W}^{2,p}(\mathbb{R}^d)} &= \|u\|_{L^p(\mathbb{R}^d)} + \|\Delta u\|_{L^p(\mathbb{R}^d)}\end{aligned}$$

are equivalent on $C_c^\infty(\mathbb{R}^n)$ (see [36]).

The Calderón-Zygmund inequality (1.1) extends to second order uniformly elliptic operators $L = -\operatorname{div} A\nabla$ with variable coefficients A on domains $\Omega \subset \mathbb{R}^d$ without any boundary conditions (see e.g. [20]). In this setting, we have the following local Calderón-Zygmund inequality that given any domains $\Omega_1 \Subset \Omega$, $p \in (1, \infty)$ and $u \in C_c^\infty(\Omega)$, it holds

$$\|\operatorname{Hess} u\|_{L^p(\Omega_1)} \le C\Big(\|u\|_{L^p(\Omega)} + \|Lu\|_{L^p(\Omega)}\Big), \tag{1.2}$$

where $C = C(\Omega_1, \Omega, d, p, A) > 0$ is a constant depending on Ω_1, Ω, p, d and the elliptic coefficients of A. The proof of (1.2) is based on (1.1) and a perturbation argument of A. In particular, if A is a constant matrix, then one can get rid of the term $\|u\|_{L^p(\Omega)}$ in (1.2).

A further step was taken by Güneysu and Pigola in [26] where they considered the following global Calderón-Zygmund inequality on Riemannian manifolds M of the form that for any $p \in (1, \infty)$ and $u \in C_c^\infty(M)$, it holds

$$\|\operatorname{Hess} u\|_{L^p(M)} \le C_1\|u\|_{L^p(M)} + C_2\|\Delta u\|_{L^p(M)}, \tag{$\mathbf{CZ}(p)$}$$

where Δ is the Laplace-Beltrami operator on M and C_1, C_2 are two positive constants. It is known that such an inequality $\mathbf{CZ}(p)$ can hold or fail depending on p and the geometry of M. We give a brief summary of the state of the art on this subject (see [36] for a more detailed survey, as well as [27]). For an example of a Riemannian manifold of positive sectional curvature where $\mathbf{CZ}(p)$ fails for large values of p, see [33], as well as [30] for a general result in this direction.

To begin with, we make some conventions on the notation. Throughout this paper, let (M, g) be a complete non-compact connected d-dimensional Riemannian manifold, ∇ the Levi-Civita covariant derivative, $\mathrm{Hess} = \nabla d$ the Hessian operator on functions, and μ the Riemannian volume measure on M. We denote by $|\cdot|$ the norm in the tangent space, and by $\|\cdot\|_p$ the norm in $L^p(M, \mu)$ for $1 \le p \le \infty$. The Laplace-Beltrami operator Δ acting on functions and forms, is understood as self-adjoint positive operator on $L^2(\mu)$.

If $p = 2$ and there is a bound $\mathrm{Ric} \ge -K$ of the Ricci curvature of M for some constant $K > 0$, then it is well-known that **CZ(2)** is a straightforward consequence of Bochner's identity, see the Appendix. The extension of **CZ**(p) from $p = 2$ to an arbitrary $p \in (1, \infty)$ is much more involved and an intriguing problem. Inspired by the proof of (1.1), a possible way to establish **CZ**(p) is a similar potential theoretical approach that represents u via the Green kernel of the Laplace-Beltrami operator Δ (see [35, 41, 34]). Although the method works pretty well for many related questions of the associated Poisson equation, it has the drawback that it applies only under some restrictive conditions such as compactness of M or nonnegative Ricci curvature.

To overcome this drawback, Güneysu and Pigola [26, 36] introduced two methods of proof for **CZ**(p) avoiding the use of Green kernel. The first method is based on a gluing procedure that connects the local consideration to the global result. To be precise, if M has bounded Ricci curvature and a strictly positive injectivity radius, then it is proved in [26] that the Calderón-Zygmund inequality **CZ**(p) holds for any $p \in (1, \infty)$ with implicit constants depending on p, $\|\mathrm{Ric}\|_{L^\infty}$, the dimension d and the injectivity radius. Moreover, if $p > \max\{2,\ d/2\}$ and M has bounded sectional curvature, then **CZ**(p) also holds (see [36, Theorem 5.18]). The second method, called the functional analytic method, uses boundedness results for the covariant Riesz transform for $1 < p < 2$ from [40].

Let us sketch the main idea of the second method in [26]. Inequality **CZ**(p) is usually reduced to the existence of positive constants C and σ such that

$$\big\| |\mathrm{Hess}(\Delta + \sigma)^{-1} u| \big\|_p \le C \|u\|_p,$$

which is equivalent to

$$\big\| |\nabla(\Delta_1 + \sigma)^{-1/2} \circ d(\Delta_0 + \sigma)^{-1/2} u| \big\|_p \le C \|u\|_p. \tag{1.3}$$

Here and hereafter, we write $\|\cdot\|_p := \|\cdot\|_{L^p(M)}$ for simplicity. The problem is thus reduced to the study of conditions for boundedness of the classical Riesz transform $d(\Delta_0 + \sigma)^{-1/2}$ on functions and boundedness of the covariant Riesz transform $\nabla(\Delta_1 + \sigma)^{-1/2}$ on one-forms. This approach in [26] however is restricted to $p \in (1, 2)$; the constants C_1, C_2 in **CZ**(p) depend on dimension d, p, $\|R\|_\infty$, $\|\nabla R\|_\infty$ and on the constants D, δ from the following local volume doubling assumption: there are constants $C > 0$, $0 \le \delta < 2$ such that

$$V(x, tr) \le C t^d \mathrm{e}^{Ctr} V(x, r) \tag{LD}$$

for all $x \in M$, $r > 0$ and $t \ge 1$, where $V(x, r) := \mu(B(x, r))$.

Very recently, Baumgarth, Devyver and Güneysu [4] studied the covariant Riesz transform on j-forms. Their results can be applied to **CZ**(p) when $1 < p < 2$ requiring the same curvature conditions as in [26] but without the local volume doubling assumption made in [26]. This comes from the fact that (**LD**) already holds if $\mathrm{Ric} \ge -K$ for some $K \ge 0$, as can be seen by the Bishop-Gromov comparison theorem and the well-known formula for the volume of balls in hyperbolic space. However, it seems difficult to establish **CZ**(p) for $p > 2$ in this way, as when trying to extend the machinery of [3] to L^p-boundedness of covariant Riesz transform for $p > 2$, the local Poincaré inequality, used explicitly in [3], does not make sense on differential forms.

The observations above raise the following questions:

1. *In the case $1 < p < 2$, is it possible to weaken the assumptions on the Riemann curvature tensor $\|R\|_\infty$ and $\|\nabla R\|_\infty$, e.g., replacing them by Ricci curvature bounds?*
2. *Without a lower control on the injectivity radius, under which conditions on the manifold M, the inequality* **CZ(p)** *holds on M for $p > 2$?*

In order to answer the above two questions affirmatively, we develop a new functional analytic method of proof for **CZ**(p) that works for all $p \in (1, \infty)$. Unlike the second method used in [26, 36] that reduces **CZ(p)** to the boundedness (1.3), our method makes use of the observation that **CZ**(p) is equivalent to the L^p-boundedness of the following operator:

$$\mathrm{Hess}(\Delta + \sigma)^{-1} = \int_0^\infty \mathrm{e}^{-\sigma t}\, \mathrm{Hess}\, P_t\, dt,$$

where P_t denotes the heat semigroup generated by $-\Delta$. Based on this observation, our strategy to solve the above two questions is first establishing some Hessian heat kernel estimates and then bridging from Hessian heat kernel estimates to **CZ(p)**.

However, references on Hessian heat kernel estimates are sparsely to find in the literature, especially when compared to the situation of heat kernel estimates and gradient estimates for the heat kernel. Therefore, in the following section, we invest some effort in deriving sharp Hessian heat kernel estimates first. It turns out that there is a big difference between the treatment of the cases $p < 2$ and $p > 2$. In particular, if $p \in (1, 2)$, we only assume that M satisfies the following curvature condition:

$$\mathrm{Ric} \geq -K \tag{\textbf{Ric}}$$

for some $K \geq 0$. The curvature condition (**Ric**) then give rise to the local Gaussian upper bounds of the heat kernel and the local volume doubling condition **(LD)** which enable us to deduce a series of L^2 weighted off-diagonal estimates of the Hessian heat kernel (see Section 2.1 below). These estimates together with the classical argument of Calderón-Zygmund decomposition (see [12]) allow to derive our first main result.

Theorem 1.1. *Let (M, g) be a complete Riemannian manifold satisfying* (**Ric**). *Let $1 < p < 2$ be fixed. Then there exists a constant $\sigma > 0$ such that the operator* $\mathrm{Hess}(\Delta + \sigma)^{-1}$ *is bounded in L^p, i.e.* **CZ(p)** *holds.*

Theorem 1.1 answers question 1 affirmatively. Comparing Theorem 1.1 with existing results on **CZ(p)**, it should be pointed out first that the result is valid without any injectivity radius assumptions and secondly, instead of boundedness of $\|R\|_\infty$ and $\|\nabla R\|_\infty$, only a lower bound of Ric is needed. In view of the argument for $p = 2$, the assumption (**Ric**) seems close to sharp.

For the case $p > 2$, it is well known [36, 33, 30] that the curvature condition (**Ric**) alone is not enough for **CZ(p)** to hold. Hence more geometric information about the manifold is required in this case for the validity of **CZ(p)**.

To formulate appropriate geometric conditions, we introduce some probabilistic quantities. Denote by X_t the diffusion process generated by $-\Delta$, which is assumed to be non-explosive (M is stochastically complete). Then, for $f \in \mathcal{B}_b(M)$,

$$\mathbb{E}^x[f(X_t)] = P_t f(x) = (\mathrm{e}^{-\Delta t} f)(x).$$

A Borel function $K\colon M \to \mathbb{R}$ is said to be in the *Kato class* $\mathcal{K}(M)$ of M, if

$$\lim_{t\to 0+} \sup_{x\in M} \int_0^t \mathbb{E}^x[|K(X_s)|]\, ds = 0. \tag{\textbf{K}}$$

Obviously, $\mathcal{K}(M)$ is a linear space and $\mathcal{K}(M) \subset L^1_{\mathrm{loc}}(M)$. The Kato class has been introduced in [31] on Euclidean space and then applied to investigate singular potentials, see for instance [1, 38, 22, 37]. Concerning criteria for functions to be in the Kato class the reader may consult [24, 23, 28]. Note that for dimension $d \geq 2$, if (**Ric**) holds, then the heat kernel has a local on-diagonal estimate which implies $L^p(M) + L^\infty(M) \subset \mathcal{K}(M)$ for $p > d/2$ by [28, Proposition 3.2].

To deal with the case $p > 2$, our conditions are given in terms of Kato bounds on the geometric quantities.

Condition (H) Ric $\geq -K$ *for some* $K \geq 0$ *and there exist* $K_1, K_2 \in \mathcal{K}(M)$ *such that*

$$|R|^2(x) \leq K_1(x) \quad \textit{and} \quad \left|\nabla \operatorname{Ric}^\sharp + d^* R\right|^2(x) \leq K_2(x), \quad x \in M, \tag{\textbf{H}}$$

where for $x \in M$ *and* $v_1, v_2, v_3 \in T_xM$,

$$|R|(x) = \sup\left\{|R^{\#,\#}(v_1, v_2)|_{\mathrm{HS}}(x)\colon\ v_1, v_2 \in T_xM,\ |v_1| \leq 1,\ |v_2| \leq 1\right\},$$

and

$$\langle d^* R(v_1, v_2), v_3\rangle = \langle (\nabla_{v_3} \operatorname{Ric}^\sharp)(v_1), v_2\rangle - \langle (\nabla_{v_2} \operatorname{Ric}^\sharp)(v_3), v_1\rangle,$$

with $R^{\#,\#}(v_1, v_2) = R(\cdot, v_1, v_2, \cdot)$*, the curvature tensor* R *and* $\operatorname{Ric}^\sharp(v) = \operatorname{Ric}(\cdot, v)^\sharp$ *for* $v \in T_xM$.

Under the condition (**H**), we are able to prove the following key pointwise inequalities for the Hessian of the semigroup that for any $t > 0$, $f \in C_c^\infty(M)$ and $x \in M$, it holds

$$|\operatorname{Hess} P_t f|(x) \leq \mathrm{e}^{2Kt} P_t |\operatorname{Hess} f|(x) + C\mathrm{e}^{(2K+\theta)t}(P_t|\nabla f|^2)^{1/2}(x), \tag{1.4}$$

and

$$t\,|\operatorname{Hess} P_t f|(x) \leq C(1 + \sqrt{t})\,\mathrm{e}^{(2K+\theta)t}\left(P_t|f|^2\right)^{1/2}(x) \tag{1.5}$$

for some constants $C, \theta > 0$ (see Propositions 2.7 and 2.8 below). Both of them are proved by using some probabilistic tools established in [18, 19, 32, 39, 2]. More precisely, inequality (1.4) is proved by a stochastic approach based on a second order derivative formula for the heat semigroup (see (2.21) below); inequality (1.5) is established by means of Bismut-type representation formulas for the Hessian of heat semigroups which were first proved by Elworthy and Li (see [18, 19]).

The inequalities (1.4) and (1.5) enable us to establish a series of pointwise Hessian heat kernel estimates. Based on these results, we can give an affirmative answer to the question 2 above.

Theorem 1.2. *Let* M *be a complete Riemannian manifold satisfying* (**H**)*. Let* $p > 2$ *be fixed. Then there exists a constant* $\sigma > 0$ *such that the operator* $\operatorname{Hess}(\Delta + \sigma)^{-1}$ *is bounded in* L^p*, i.e.* **CZ**(p) *holds.*

Theorem 1.2 gives in particular an answer to the open question in [26] about sufficient conditions for **CZ(**p**)** when $p > 2$ in the absence of control of the injectivity radius. It is worth mentioning that compared with the sufficient conditions even in the case $1 < p < 2$ ([26, Theorem D], [4, Corollary 1.8]), in Theorem 1.2 the geometric quantities $|R|$ and $|\nabla \operatorname{Ric}|$ do not need to be uniformly bounded on M; it is sufficient to have bounds in the Kato class which is a kind of integral condition. It should be mentioned that for $p > d/2$ in conjunction with $p \geq 2$ the validity of **CZ**(p) has been obtained under strong curvature assumption, namely uniform boundedness of the sectional curvature tensor, but without conditions involving the derivative of curvature, see Theorem 5.18 in the survey of Pigola [36].

As indicated above, the pointwise inequalities (1.4) and (1.5) play a key role in the proof of Theorem 1.2, which also reflects the difference between our approach and that of classical local Riesz transform (see [13, 14, 3]). In the latter case, the following domination property that for some positive constants c_1, c_2 and C

$$|\nabla P_t f| \leq C e^{c_1 t} P_{c_2 t} |\nabla f| \tag{1.6}$$

is indispensable for boundedness of the Riesz transform in case $p > 2$ (see [13, 14]). Unfortunately, the machinery of [13] cannot work well in our situation since a Hessian estimate of the type

$$|\operatorname{Hess} P_t f| \leq C e^{c_1 t} P_{c_2 t} |\operatorname{Hess} f|$$

would be required which however only holds in very specific cases (like flat manifolds). On the other hand, if we aim at using the techniques from [3] directly, the main difficulty to deal with is that there is no suitable Hessian replacement of the local Poincaré inequality which is heavily used throughout their proof.

To overcome these two obstacles, we take advantage of the techniques from [3], in particular, the sharp maximal function and good-λ inequalities. By means of these tools and inequality (1.4), we observe that if the additional term involving $(P_t|\nabla f|^2)^{1/2}$ is treated and controlled as an error term, then boundedness of $\operatorname{Hess}(\Delta + \sigma)$ in L^p can be established also for $p > 2$ with the help of pointwise Hessian estimates of the heat kernel. In conclusion, the crucial observation is that the pointwise inequality (1.4) may serve as a Hessian replacement of (1.6) circumventing the non-availability of the local Poincaré inequality.

The rest of the paper is organized as follows. In Section 2, we give various forms of estimates on the Hessian of the heat kernel and on the corresponding semigroups, which are used throughout the proof of **CZ**(p). In Sections 3 finally, we present proofs for Theorem 1.1 and 1.2 respectively. A suitable version of the localization techniques of [17] is included.

Acknowledgements The authors are indebted to Batu Güneysu, Stefano Pigola and Giona Veronelli for very helpful comments on the first draft of this paper.

2 Hessian heat kernel estimates

This section is divided into two parts: the first part is on L^2-estimates of the Hessian of heat kernels under the assumption of a lower Ricci curvature bound; the second part is on the estimates derived under condition **(H)** from the Hessian formula and Bismut type formulas for the Hessian of semigroups, which are established by means of the techniques from stochastic analysis.

2.1 L^2-estimates for the Hessian of heat kernel

In this section, we assume Ricci curvature to be bounded below, i.e. validity of condition (**Ric**). Then, in particular, the local doubling assumption (**LD**) with respect to μ holds. Then for x and $y \in M$, we obviously have $B(y, \sqrt{t}) \subset B(x, \sqrt{t} + \rho(x, y))$. Thus if (**LD**) holds, then

$$V(y, \sqrt{t}) \le V\Big(x, \sqrt{t} + \rho(x, y)\Big) \le C\Big(1 + \frac{\rho(x, y)}{\sqrt{t}}\Big)^d \exp\Big(C(\sqrt{t} + \rho(x, y))\Big) V(x, \sqrt{t}). \tag{2.1}$$

It is well-known that, under a lower Ricci curvature bound, the heat kernel allows an off-diagonal estimate [15]. The following lemma gives a pointwise off-diagonal estimate for the heat kernel and its time derivative.

Lemma 2.1. *Assume that* (**Ric**) *holds. Then for any $\alpha \in (0, \frac{1}{4})$, there exist constants C and $C_1 > 0$ depending on the dimension d and α such that for all $x, y \in M$ and $t > 0$,*

$$p_t(x, y) + \left|\frac{\partial p_t}{\partial t}(x, y)\right| \le \frac{C}{V(y, \sqrt{t})} \exp\left(-\alpha \frac{\rho^2(x, y)}{t} + C_1 K t\right). \tag{2.2}$$

Proof. The estimate of $p_t(x, y)$ is an easy consequence of (2.1) and [15, Theorem 2] or [42, Theorem 2.4.4], where it is proved that for all $x, y \in M$ and $t > 0$,

$$p_t(x, y) \le \frac{C}{V(y, \sqrt{t})} \exp\left(-\alpha \frac{\rho^2(x, y)}{t} + C_1 K t\right). \tag{2.3}$$

The estimate for $|\frac{\partial p_t}{\partial t}(x, y)|$ follows from that of $p_t(x, y)$ and the analytic property of the semigroup (see [16, Theorem 4] or [21, Corollary 3.3]). □

The following lemma gives weighted L^2-integral estimates for the heat kernel, its gradient and its Laplacian.

Lemma 2.2. *Assume that* (**Ric**) *holds. Let $\alpha \in (0, \frac{1}{4})$ be as in Lemma* 2.1. *For all $\gamma \in (0, 2\alpha)$, $s > 0$ and $y \in M$,*

$$\int_M \Big[|p_s(x, y)|^2 + s|\nabla_x p_s(x, y)|^2 + s^2|\Delta_x p_s(x, y)|^2\Big] \mathrm{e}^{\gamma \frac{\rho^2(x,y)}{s}} \mu(dx) \le \frac{C_\gamma}{V(y, \sqrt{s})} \mathrm{e}^{2C' s},$$

where $C_\gamma > 0$ depends on γ and $C' > 0$ on α and K.

Proof. By (**LD**), it is easy to see that for all $\gamma > 0$, $s, t > 0$ and $y \in M$, there exist two positive constants C_γ (depending on γ and the constants in (**LD**)) and $\tilde{C}$ such that

$$\begin{aligned}
\int_{\rho(x,y)\ge\sqrt{t}} \mathrm{e}^{-2\gamma\frac{\rho^2(x,y)}{s}} \mu(dx) &\le \mathrm{e}^{-\gamma t/s} \int_M \mathrm{e}^{-\gamma\frac{\rho^2(x,y)}{s}} \mu(dx) \\
&\le \mathrm{e}^{-\gamma t/s} \sum_{i=0}^{\infty} V(y, (i+1)\sqrt{s}) \mathrm{e}^{-\gamma i^2} \\
&\le C\mathrm{e}^{-\gamma t/s} V(y, \sqrt{s}) \sum_{i=0}^{\infty} (i+1)^d \mathrm{e}^{-\gamma i^2} \mathrm{e}^{C(i+1)\sqrt{s}} \\
&\le C\mathrm{e}^{-\gamma t/s} \mathrm{e}^{C\sqrt{s}} V(y, \sqrt{s}) \sum_{i=0}^{\infty} (i+1)^d \mathrm{e}^{-\gamma i^2} \mathrm{e}^{\gamma i^2/2 + C^2 s/2} \\
&\le C\mathrm{e}^{-\gamma t/s} \mathrm{e}^{C\sqrt{s} + C^2 s/2} V(y, \sqrt{s}) \sum_{i=0}^{\infty} (i+1)^d \mathrm{e}^{-\gamma i^2/2}
\end{aligned}$$

$$\le C_\gamma V(y,\sqrt{s})\, \mathrm{e}^{-\gamma t/s}\mathrm{e}^{\tilde{C}s}, \tag{2.4}$$

where the second inequality comes from condition (**LD**). The remainder of the proof follows from [12, Lemmas 2.1-2.3]. □

We now turn to the estimates for the Hessian of heat kernel. The following lemma shows that the Hessian of heat semigroup also satisfies an L^2-Gaffney off-diagonal estimate. Note that in the following discussion the constant C will be different in different lines without confusion.

Lemma 2.3. *Assume that* (**Ric**) *holds. There exist constants $C, C_2 > 0$ such that for all $t \in (0,\infty)$, all Borel subsets $E, F \subset M$ with compact closure, and all $f \in L^2(M)$ with* $\operatorname{supp} f \subset E$,

$$\left\|\mathbb{1}_F t\,|\operatorname{Hess} P_t f|\right\|_2 \le C(1+\sqrt{t})\exp\left(-\frac{C_2\,\rho^2(E,F)}{t}\right)\|f\|_2.$$

Proof. Recall the following L^2-Gaffney off-diagonal estimate on one-forms [4]. If (**Ric**) holds, then for $\alpha \in \Gamma_{L^2}(T^*M)$ with support $\operatorname{supp}(\alpha) \subset E$ and any $s \in (0,1)$, it holds

$$\begin{aligned}\left\|\mathbb{1}_F\sqrt{s}\,|\nabla \mathrm{e}^{-s\Delta^{(1)}}\alpha|\right\|_2 &\le C\,(1+\sqrt{s})\exp\left(-\frac{c_1\,\rho^2(E,F)}{s}\right)\|\mathbb{1}_E\alpha\|_2\\ &\le 2C\exp\left(-\frac{c_1\rho^2(E,F)}{s}\right)\|\mathbb{1}_E\alpha\|_2\end{aligned}$$

for some positive constants C and c_1. On the other hand, we have Gaffney's off-diagonal estimate for $|\nabla P_{t_2} f|$ (see [3, (3.1)]), i.e. for all $f \in L^2(M)$ with support in E and $u > 0$,

$$\left\|\mathbb{1}_F\sqrt{u}\,|\nabla P_u f|\right\|_2 \le C\exp\left(-\frac{c_2\rho^2(E,F)}{u}\right)\|\mathbb{1}_E f\|_2$$

for some positive constants C and c_2. For $t > 0$, denoting by $\Delta^{(1)}$ the Laplacian on one-forms, we may write

$$\operatorname{Hess} P_t f = \frac{\sqrt{\left(t-(\frac{t}{2}\wedge 1)\right)}\,\nabla \mathrm{e}^{-\left(t-(\frac{t}{2}\wedge 1)\right)\Delta^{(1)}}\left(\sqrt{\frac{t}{2}\wedge 1}\,dP_{\frac{t}{2}\wedge 1} f\right)}{\sqrt{\left(t-(\frac{t}{2}\wedge 1)\right)(\frac{t}{2}\wedge 1)}}$$

so that using the composition rule of Gaffney's off-diagonal estimate (see [29, Lemma 2.3]), we obtain

$$\left\|\mathbb{1}_E t\,|\operatorname{Hess} P_t f|\right\|_2 \le C\frac{t}{\sqrt{\left(t-(\frac{t}{2}\wedge 1)\right)(\frac{t}{2}\wedge 1)}}\exp\left(-\frac{c\,\rho^2(E,F)}{\max\left\{\frac{t}{2}\wedge 1, t-(\frac{t}{2}\wedge 1)\right\}}\right)\|\mathbb{1}_E f\|_2,$$

for some positive constants C and c. Note that for $0 < t/2 < 1$,

$$\max\left\{\frac{t}{2}\wedge 1, t-(\frac{t}{2}\wedge 1)\right\} = \max\left\{\frac{t}{2}, t-\frac{t}{2}\right\} = \frac{t}{2};\qquad \frac{t}{\sqrt{\left(t-(\frac{t}{2}\wedge 1)\right)(\frac{t}{2}\wedge 1)}} = 2,$$

and that for $t/2 \ge 1$,

$$\max\left\{\frac{t}{2}\wedge 1, t-(\frac{t}{2}\wedge 1)\right\} = \max\{1, t-1\} = t-1 \le t;\qquad \frac{t}{\sqrt{\left(t-(\frac{t}{2}\wedge 1)\right)(\frac{t}{2}\wedge 1)}} = \frac{t}{\sqrt{t-1}} \le \sqrt{2t}.$$

Therefore, we conclude that there exist positive constants C and C_2 such that

$$\left\|\mathbb{1}_F t\,|\operatorname{Hess} P_t f|\right\|_2 \le C\,(1+\sqrt{t})\exp\left(-\frac{C_2\rho^2(E,F)}{t}\right)\|\mathbb{1}_E f\|_2.$$ □

The following proposition gives L^2-weighted estimates for the Hessian of the heat kernel.

Proposition 2.4. *Assume that* (**Ric**) *holds. Fix* $\alpha \in (0, \frac{1}{4})$ *as in Lemma* 2.1*. Then for all* $\gamma \in (0, 2\alpha)$*,* $s > 0$ *and* $y \in M$*, there exists a constant* $C > 0$ *such that*

$$\int_M |\operatorname{Hess}_x p_s(x,y)|^2 \exp\left(\gamma \frac{\rho^2(x,y)}{s}\right) \mu(dx) \le \frac{C(1+Ks)\,\mathrm{e}^{2C's}}{s^2 V(y,\sqrt{s})}$$

where $C' > 0$ *is the same constant as in Lemma 2.2.*

Proof. We begin by integrating Bochner's identity (4.1) to obtain

$$\begin{aligned}\frac{1}{2}\int |\nabla p_s|^2\, \Delta \mathrm{e}^{\gamma\rho^2(x,y)/s}\, \mu(dx) &= \frac{1}{2}\int_M (\Delta|\nabla p_s|^2)\, \mathrm{e}^{\gamma\rho^2(x,y)/s}\, \mu(dx)\\ &= \int_M \Big(-|\operatorname{Hess} p_s|^2_{\mathrm{HS}} + g(\nabla\Delta p_s, \nabla p_s) - \operatorname{Ric}(\nabla p_s, \nabla p_s)\Big)\, \mathrm{e}^{\gamma\rho^2(x,y)/s}\, \mu(dx),\end{aligned}$$

which then implies

$$\begin{aligned}&\int_M |\operatorname{Hess} p_s|^2_{\mathrm{HS}}\, \mathrm{e}^{\gamma\rho^2(x,y)/s}\, \mu(dx) \qquad (2.5)\\ &= -\frac{1}{2}\int_M |\nabla p_s|^2 \Delta \mathrm{e}^{\gamma\rho^2(x,y)/s}\, \mu(dx) + \int_M g(\nabla\Delta p_s, \nabla p_s)\, \mathrm{e}^{\gamma\rho^2(x,y)/s}\, \mu(dx)\\ &\quad - \int_M \operatorname{Ric}(\nabla p_s, \nabla p_s)\, \mathrm{e}^{\gamma\rho^2(x,y)/s}\, \mu(dx)\\ &= -\frac{\gamma}{s}\int_M |\nabla p_s|^2 \operatorname{div}\Big(\mathrm{e}^{\gamma\rho^2(x,y)/s}\rho(x,y)\nabla\rho(\cdot,y)(x)\Big)\, \mu(dx) + \int \Delta p_s \operatorname{div}(\mathrm{e}^{\gamma\rho^2(x,y)/s}\nabla p_s)\, \mu(dx)\\ &\quad - \int \operatorname{Ric}(\nabla p_s, \nabla p_s)\, \mathrm{e}^{\gamma\rho^2(x,y)/s}\, \mu(dx) =: \mathrm{I}_1 + \mathrm{I}_2 + \mathrm{I}_3\end{aligned}$$

(with the sign convention $\Delta = \operatorname{div}\nabla$ for the divergence). Now for any $\beta \in (0, 2\alpha)$, let

$$E(s,y,\beta) := \int_M \Big[|p_s(x,y)|^2 + s|\nabla_x p_s(x,y)|^2 + s^2|\Delta_x p_s(x,y)|^2|\Big]\, \mathrm{e}^{\beta\rho^2(x,y)/s}\, \mu(dx).$$

Then by condition (**Ric**) and Lemma 2.2, we have

$$\mathrm{I}_3 \le K\int_M |\nabla p_s|^2\, \mathrm{e}^{\gamma\rho^2(x,y)/s}\, \mu(dx) \le \frac{K}{s}E(s,y,\gamma) \le \frac{K}{s}\frac{C_\gamma}{V(y,\sqrt{s})}\, \mathrm{e}^{2C's}. \qquad (2.6)$$

For I_2, using the fact $|\nabla\rho| \le 1$, we deduce from Lemma 2.2 again that

$$\begin{aligned}\mathrm{I}_2 &= \int_M (\Delta p_s)^2\, \mathrm{e}^{\gamma\rho^2(x,y)/s}\, \mu(dx) - \frac{2\gamma}{s}\int_M (\Delta p_s)\, \mathrm{e}^{\gamma\rho^2(x,y)/s}\rho(x,y)\langle\nabla\rho, \nabla p_s\rangle\, \mu(dx)\\ &\le 2\int_M (\Delta p_s)^2\, \mathrm{e}^{\gamma\rho^2(x,y)/s}\, \mu(dx) + \frac{\gamma^2}{s^2}\int_M \mathrm{e}^{\gamma\rho^2(x,y)/s}\rho^2(x,y)|\nabla p_s|^2\, \mu(dx)\\ &\le \frac{C_\gamma}{s^2}E(s,y,\gamma') \le \frac{1}{s^2}\frac{C_{\gamma'}}{V(y,\sqrt{s})}\, \mathrm{e}^{2C's}, \qquad (2.7)\end{aligned}$$

where $\gamma < \gamma' < 2\alpha$ with $\gamma' - \gamma$ sufficiently small.

To bound I_1, we write

$$\begin{aligned}\frac{s}{\gamma}\mathrm{I}_1 = &\int_M |\nabla p_s|^2\, \mathrm{e}^{\gamma\rho^2(x,y)/s}\,\mu(dx) + 2\int_M |\nabla p_s|^2\, \mathrm{e}^{\gamma\rho^2(x,y)/s}\frac{\gamma\rho^2(x,y)}{s}\,\mu(dx)\\ &- \int_M |\nabla p_s|^2\, \mathrm{e}^{\gamma\rho^2(x,y)/s}\rho(x,y)\Delta\rho(\cdot,y)(x)\,\mu(dx).\end{aligned} \tag{2.8}$$

By the Laplacian comparison theorem (see e.g. [8, p185]), we have

$$-\Delta\rho(\cdot,y)(x) \le \frac{d-1}{\rho(x,y)} + K\rho(x,y)$$

outside of the cut-locus. This yields in particular

$$\begin{aligned}&-\frac{\gamma}{s}\int_M |\nabla p_s|^2\, \mathrm{e}^{\gamma\rho^2(x,y)/s}\rho(x,y)\Delta\rho(\cdot,y)(x)\,\mu(dx)\\ &\le \frac{(d-1)\gamma}{s}\int_M |\nabla p_s|^2\, \mathrm{e}^{\gamma\rho^2(x,y)/s}\,\mu(dx) + K\int_M |\nabla p_s|^2\, \mathrm{e}^{\gamma\rho^2(x,y)/s}\,\frac{\gamma\rho^2(x,y)}{s}\,\mu(dx)\\ &\le \frac{(d-1)\gamma}{s}\int_M |\nabla p_s|^2\, \mathrm{e}^{\gamma\rho^2(x,y)/s}\,\mu(dx) + K\int_M |\nabla p_s|^2\, \mathrm{e}^{\frac{\gamma'\rho^2(x,y)}{s}}\,\mu(dx)\\ &\le \frac{C_\gamma(1+Ks)}{s^2}E(s,y,\gamma')\end{aligned}$$

with $\gamma < \gamma' < 2\alpha$ as in (2.7). Combining this estimate with (2.8), it follows

$$\mathrm{I}_1 \le C_\gamma\frac{(1+Ks)}{s^2}E(s,y,\gamma') \le \frac{c(1+Ks)}{s^2\sqrt{V(y,\sqrt{s})}}\,\mathrm{e}^{2C's}. \tag{2.9}$$

Altogether (2.5) through (2.9), we conclude that

$$\int_M |\operatorname{Hess} p_s|_{\mathrm{HS}}^2\, \mathrm{e}^{\gamma\rho^2(x,y)/s}\,\mu(dx) = \mathrm{I}_1 + \mathrm{I}_2 + \mathrm{I}_3 \le c\frac{(1+Ks)}{s^2\sqrt{V(y,\sqrt{s})}}\,\mathrm{e}^{2C's}.$$

which completes the proof of Proposition 2.4 by using the fact

$$|\operatorname{Hess} p_s|^2 \le |\operatorname{Hess} p_s|_{\mathrm{HS}}^2. \qquad \square$$

Using Proposition 2.4, we obtain the following L^2-integral estimates for Hessian of heat kernel.

Corollary 2.5. *Assume* (**Ric**) *holds. Fix* $\alpha \in (0, \frac{1}{4})$ *as in Lemma* 2.1. *There exist* $0 < \beta < \alpha$ *and* $C'' > C' > 0$ *with* C' *as in Lemma 2.2 such that*

$$\int_{\rho(x,y)\ge t^{1/2}} |\operatorname{Hess}_x p_s(x,y)|\,\mu(dx) \le C(1+\sqrt{s})\,\mathrm{e}^{C''s}\,\mathrm{e}^{-\beta t/s}s^{-1}$$

for all $y \in M$ *and* $s, t > 0$.

Proof. Let $0 < \beta < \alpha$. By Cauchy's inequality we obtain

$$\int_{\rho(x,y)\ge t^{1/2}} |\operatorname{Hess}_x p_s(x,y)|\,\mu(dx)$$

$$\leq \left(\int_M |\operatorname{Hess}_x p_s(x,y)|^2 \, \mathrm{e}^{2\beta\rho^2(x,y)/s} \, \mu(dx)\right)^{1/2} \left(\int_{\rho(x,y)\geq t^{1/2}} \mathrm{e}^{-2\beta\rho^2(x,y)/s} \, \mu(dx)\right)^{1/2}$$
$$\leq \frac{C\,\mathrm{e}^{C's}(1+\sqrt{s})}{s\sqrt{V(y,\sqrt{s})}} \sqrt{V(y,\sqrt{s})}\,\mathrm{e}^{-\beta t/s}\,\mathrm{e}^{\tilde{C}s}$$
$$= \frac{C(1+\sqrt{s})}{s}\,\mathrm{e}^{C's-\beta t/s}\,\mathrm{e}^{\tilde{C}s} \leq \frac{C(1+\sqrt{s})}{s}\,\mathrm{e}^{C''s-\beta t/s}$$

where the second inequality follows from Proposition 2.4 and inequality (2.4). This finishes the proof. □

2.2 Stochastic Hessian formulas and pointwise estimates for Hessian of heat kernel

In this subsection, we establish some pointwise and L^p-integral estimates for the Hessian of the heat kernel when $p > 2$. To this end, let us first introduce some necessary notations. For fixed $x \in M$, let B_t be the stochastic anti-development of $X_\cdot(x)$ which is a Brownian motion in T_xM. Let $//_t\colon T_xM \to T_{X_t(x)}M$ be the parallel transport and $Q_t\colon T_xM \to T_{X_t}M$ be the damped parallel transport defined as the solution to the following pathwise ordinary covariant differential equation along the trajectories of X_t,

$$DQ_t = -\operatorname{Ric}^\sharp Q_t\,dt, \quad Q_0 = \mathrm{id}_{T_xM} \tag{2.10}$$

with $DQ_t = //_t\, d\, //_t^{-1} Q_t$. For each $w \in T_xM$ define an operator-valued process $W_t(\cdot, w) : T_xM \to T_{X_t}M$ by

$$W_t(\cdot,w) = Q_t \int_0^t Q_r^{-1} R(//_r\, dB_r, Q_r(\cdot))Q_r(w) - Q_t \int_0^t Q_r^{-1}(\nabla \operatorname{Ric}^\sharp + d^*R)(Q_r(\cdot), Q_r(w))\,dr.$$

This means that the process $W_t(\cdot, w)$ is the solution to the following covariant Itô equation

$$\begin{cases} DW_t(\cdot,w) = R(//_t dB_t, Q_t(\cdot))Q_t(w) - (d^*R + \nabla \operatorname{Ric}^\sharp)(Q_t(\cdot), Q_t(w))\,dt - \operatorname{Ric}^\sharp(W_t(\cdot,w))\,dt, \\ W_0(\cdot, w) = 0. \end{cases}$$

First, we collect some easy curvature estimates for Riemannian manifolds M satisfying condition (**H**).

Lemma 2.6. *Assume that* (**H**) *holds. There exist constants* $C > 0$, $\theta > 0$ *such that for any* $t > 0$,

$$\sup_{x\in M} \mathbb{E}^x \left[\int_0^t |R|^2(X_s)\,ds + \int_0^t |\nabla \operatorname{Ric}^\sharp + d^*R|^2(X_s)\,ds\right] \leq C\mathrm{e}^{2\theta t}. \tag{2.11}$$

Proof. By means of the trivial inequality $s \leq \mathrm{e}^s$ for $s \geq 0$, we see that

$$\sup_{x\in M} \mathbb{E}^x \left[\int_0^t |R|^2(X_s)\,ds + \int_0^t |\nabla \operatorname{Ric}^\sharp + d^*R|^2(X_s)\,ds\right]$$
$$\leq \sup_{x\in M} \mathbb{E}^x \left[\int_0^t K_1(X_s)\,ds + \int_0^t K_2(X_s)\,ds\right]$$
$$\leq \sup_{x\in M} \mathbb{E}^x \left[\exp\left(\int_0^t K_1(X_s)\,ds\right)\right] + \sup_{x\in M} \mathbb{E}^x \left[\exp\left(\int_0^t K_2(X_s)\,ds\right)\right].$$

Recall in [25, Lemma 3.9] that for any $K \in \mathcal{K}(M)$, there exist positive constants c and C such

$$\sup_{x\in M} \mathbb{E}^x \left[\exp\left(\int_0^t |K(X_s)|\, ds \right) \right] \le C\mathrm{e}^{ct}.$$

This immediately completes the proof by substituting K_1 and K_2 as K. □

By this lemma and Bismut-type Hessian formula, we obtain the following Hessian estimates for heat semigroup.

Proposition 2.7. *Assume that* (**H**) *holds. Then, for any $p \ge 2$, there exist positive constants C and θ as in Lemma* 2.6 *such that, for every $f \in \mathcal{B}_b(M)$, $t > 0$,*

$$t\,|\operatorname{Hess} P_t f|(x) \le C(1+\sqrt{t})\,\mathrm{e}^{(2K+\theta)t}\,(P_t|f|^p)^{1/p}(x); \tag{2.12}$$

$$t\,\big\| |\operatorname{Hess} P_t f| \big\|_p \le C(1+\sqrt{t})\,\mathrm{e}^{(2K+\theta)t}\|f\|_p. \tag{2.13}$$

Proof. We start by recalling the following Bismut type formula for $\operatorname{Hess} P_t$ (see [2, 19, 9]). Let $x \in D$ with $v, w \in T_xM$, $f \in \mathcal{B}_b(M)$ and $0 < s < t$. Suppose that D_1 and D_2 are regular domains such that $x \in D_1$ and $\bar{D}_1 \subset D_2 \subset D$, and denote by σ, τ the exit times of $X.(x)$ from D_1 and D_2 respectively. Assume that k, ℓ are bounded adapted processes with paths in the Cameron-Martin space $L^2([0,t];[0,1])$ such that $k_r = 0$ for $r \ge \sigma \wedge s$, $k_0 = 1$, $\ell_r = 1$ for $r \le \sigma \wedge s$, $\ell_s = 0$ for $r \ge \tau \wedge t$. Then, for $f \in \mathcal{B}_b(M)$, we have

$$\begin{aligned}
(\operatorname{Hess} P_t f))(v,w) &= -\mathbb{E}^x \left[f(X_t) \int_0^t \langle W_s(\dot{k}_s v, w), //_s\, dB_s\rangle \right] \\
&\quad + \mathbb{E}\left[f(X_t) \int_s^t \langle Q_r(\dot{\ell}_r w), //_r\, dB_r\rangle \int_0^s \langle Q_r(\dot{k}_r v), //_r\, dB_r\rangle \right].
\end{aligned} \tag{2.14}$$

Note that under condition (**H**) one has $|Q_t|^2 \le \mathrm{e}^{2Kt}$ and according to the Itô formula, then for $t < \tau < \tau_D$ and $0 < \delta < 2K$,

$$\begin{aligned}
d|W_s(v,w)|^2 &= 2\langle R(//_s dB_s, Q_s(v))Q_s(w), W_s(v,w)\rangle + 2|R^{\sharp,\sharp}(Q_s(v), Q_s(w))|^2_{\mathrm{HS}}\, ds \\
&\quad - 2\langle (d^*R + \nabla \operatorname{Ric}^\sharp)(Q_s(v), Q_s(w)), W_s(v,w)\rangle\, ds \\
&\quad - 2\operatorname{Ric}(W_s(v,w), W_s(v,w))\, ds \\
&\le 2\langle R(//_s dB_s, Q_s(v))Q_s(w), W_s(v,w)\rangle + 2\mathrm{e}^{4Ks}K_1(X_s)\, ds \\
&\quad + 2\sqrt{K_2(X_s)}\,\mathrm{e}^{2Ks}|W_s(v,w)|\, ds + 2K|W_s(v,w)|^2\, ds \\
&\le 2\langle R(//_s dB_s, Q_s(v))Q_s(w), W_s(v,w)\rangle + 2\mathrm{e}^{4Ks}K_1(X_s)\, ds \\
&\quad + (2\delta)^{-1}\mathrm{e}^{4Ks}K_2(X_s)\, ds + (2K+\delta)|W_s(v,w)|^2\, ds.
\end{aligned}$$

From this and Lemma 2.6 there exist constants $C > 0$ and $\theta > 0$ such that

$$\begin{aligned}
\mathbb{E}|W_{t\wedge\tau_D}(v,w)|^2 &\le 2\mathrm{e}^{(2K+\delta)t}\left(\mathbb{E}\left[\int_0^t \mathrm{e}^{(2K-\delta)s}K_1(X_s)\, ds \right] + (4\delta)^{-1}\mathbb{E}\left[\int_0^t \mathrm{e}^{(2K-\delta)s}K_2(X_s)\, ds \right] \right) \\
&\le 2\max\{1, (4\delta)^{-1}\}\mathrm{e}^{4Kt}\mathbb{E}\left[\int_0^t K_1(X_s)\, ds + \int_0^t K_2(X_s)\, ds \right] \\
&\le C\,\mathrm{e}^{(4K+2\theta)t}.
\end{aligned} \tag{2.15}$$

Since the upper of $\mathbb{E}|W_{t\wedge\tau_D}(v,w)|^2$ is independent of the domain D, taking a sequence of domains D_n increasing to M, we conclude that $\mathbb{E}|W_t(v,w)|^2 \le C\,\mathrm{e}^{(4K+2\theta)t}$. All these estimates allow to verify that formula (2.14) holds for the deterministic functions $k_s = \frac{t-2s}{t} \vee 0$ and $\ell_s = 1 \wedge \frac{2(t-s)}{t}$ as well. Then, by estimate (2.15) and $|Q_s| \le \mathrm{e}^{Ks}$, we get for any $p \ge 2$,

$$\begin{aligned}
|\operatorname{Hess} P_t f| &\le \frac{2}{t}(P_t|f|^p)^{1/p}\left[\int_0^{t/2}\mathbb{E}|W_s(v,w)|^2\,ds\right]^{1/2}\\
&\quad+\frac{4}{t^2}(P_t|f|^p)^{1/p}\mathbb{E}\left[\Big(\int_0^{t/2}\langle Q_s(w),//_s dB_s\rangle\Big)^2\mathbb{E}\Big[\Big(\int_{t/2}^{t}\langle Q_s(v),//_s dB_s\rangle\Big)^2\Big|\mathscr{F}_{t/2}\Big]\right]^{1/2}\\
&\le \frac{2}{t}(P_t|f|^p)^{1/p}\left(\int_0^{t/2}\mathrm{e}^{(4K+2\theta)s}\,ds\right)^{1/2}\\
&\quad+\frac{4}{t^2}(P_t|f|^p)^{1/p}\mathbb{E}\left[\int_0^{t/2}\mathrm{e}^{2Ks}\,ds\right]^{1/2}\left(\int_{t/2}^{t}\mathrm{e}^{2Ks}\,ds\right)^{1/2}\\
&\le \frac{C}{\sqrt{t}}\,\mathrm{e}^{(2K+\theta)t}(P_t|f|^p)^{1/p}+\frac{C}{t}\,\mathrm{e}^{2Kt}(P_t|f|^p)^{1/p}\\
&\le \frac{C(1+\sqrt{t})}{t}\,\mathrm{e}^{(2K+\theta)t}(P_t|f|^p)^{1/p} \qquad (2.16)
\end{aligned}$$

which proves (2.12). Moreover, integration with respect to the measure μ yields

$$\mu(|\operatorname{Hess} P_t f|^p)^{1/p} \le \frac{C(1+\sqrt{t})}{t}\,\mathrm{e}^{(2K+\theta)t}\mu(|f|^p)^{1/p}.$$

This proves (2.13) and finishes the proof. □

The following proposition gives a pointwise estimate for the Hessian of heat semigroup.

Proposition 2.8. *Assume that* (**H**) *holds. Then there exist constants $C,\theta>0$ such that*

$$|\operatorname{Hess} P_t f| \le \mathrm{e}^{2Kt}\left(P_t|\operatorname{Hess} f|^2\right)^{1/2} + C\mathrm{e}^{(2K+\theta)t}\left(P_t|df|^2\right)^{1/2} \qquad (2.17)$$

for $t>0$ and $f\in C_0^\infty(M)$.

Proof. Recall the following relation

$$d\Delta f = \operatorname{tr}\operatorname{Hess}(df) - df(\mathrm{Ric}^\sharp) \qquad (2.18)$$

where $\operatorname{Hess}(df) = \nabla^2 df$. Moreover, using the differential Bianchi identity [10, p. 185], we have

$$\operatorname{Hess}(\Delta f) = \operatorname{tr}\operatorname{Hess}(\operatorname{Hess} f) - 2(\operatorname{Hess} f)(\mathrm{Ric}^\sharp\otimes\mathrm{id} - R^{\#,\#}) - df(d^*R + \nabla\,\mathrm{Ric}^\sharp). \qquad (2.19)$$

Now, for $s>0$ and $v,w\in T_xM$, let

$$N_s(v,w) := (\operatorname{Hess} P_{t-s}f)(Q_s(v),Q_s(w))(X_s) + (dP_{t-s}f)(W_s(v,w))(X_s).$$

We claim that N_s is a local martingale. Combining (2.18), (2.19) and applying Itô's formula, we obtain

$$dN_s(v,w) = (\nabla_{//_s dB_s}\operatorname{Hess} P_{t-s}f)(Q_s(v),Q_s(w)) + (\operatorname{Hess} P_{t-s}f)(\frac{D}{ds}Q_s(v),Q_s(w))\,ds$$

$$
\begin{aligned}
&+ (\operatorname{Hess} P_{t-s} f)(Q_s(v), \frac{D}{ds} Q_s(w))\, ds + \partial_s (\operatorname{Hess} P_{t-s} f)(Q_s(v), Q_s(w))\, ds \\
&+ \operatorname{tr} \operatorname{Hess}(\operatorname{Hess} P_{t-s} f)(Q_s(v), Q_s(w))\, ds + (\nabla_{//_s dB_s} dP_{t-s} f)(W_s(v,w)) \\
&+ (dP_{t-s} f)(DW_s(v,w)) + \langle d(dP_{t-s} f), DW_s(v,w)\rangle + \partial_s (dP_{t-s} f)(W_s(v,w))\, ds \\
&+ \operatorname{tr} \operatorname{Hess}(dP_{t-s} f)(W_s(v,w))\, ds \stackrel{\mathrm{m}}{=} 0,
\end{aligned}
$$

where $\stackrel{\mathrm{m}}{=}$ denotes equality modulo the differential of a local martingale, and here we used for the quadratic covariation the formula

$$
[d(dP_{t-s} f), DW_s(v,w)]_s = (\operatorname{Hess} P_{t-s} f)(R^{\#,\#}(Q_s(v), Q_s(w)))\, ds.
$$

Therefore, N_s is a local martingale. Recall that condition (**H**) implies

$$
|Q_t|^2 \le \mathrm{e}^{2Kt} \quad \text{and} \quad \mathbb{E}|W_t(v,w)|^2 \le C\mathrm{e}^{(4K+2\theta)t}
$$

for some positive constants. In view of estimate (2.16), $|\operatorname{Hess} P_{t-s} f|$ is bounded on $[0, t-\varepsilon] \times M$, and

$$
|\nabla P_{t-s} f| \le \mathrm{e}^{K(t-s)} \| |\nabla f| \|_\infty
$$

as a consequence of the derivative formula (see e.g. [42, Theorem 2.2.3]): for all $v \in T_x M$ and $x \in M$,

$$
\langle (\nabla P_t f)(x), v \rangle = \mathbb{E}^x \left[\langle (\nabla f)(X_t), Q_t v \rangle \right]. \tag{2.20}
$$

Thus N_s is a martingale on the time interval $[0, t-\varepsilon]$, and by taking expectation at time 0 and $t-\varepsilon$, we arrive at

$$
(\operatorname{Hess} P_t f)(v,w) = \mathbb{E}\Big[(\operatorname{Hess} P_\varepsilon f)(Q_{t-\varepsilon}(v), Q_{t-\varepsilon}(w))(X_t) + dP_\varepsilon f(W_{t-\varepsilon}(v,w))(X_{t-\varepsilon}) \Big]. \tag{2.21}
$$

Letting ε tend to 0, it then follows that

$$
\begin{aligned}
|\operatorname{Hess} P_t f(v,w)| &\le \Big| \mathbb{E}\Big[\operatorname{Hess} f(Q_t(v), Q_t(w))(X_t) \Big] \Big| + \Big| \mathbb{E}\Big[df(W_t(v,w))(X_t) \Big] \Big| \\
&\le P_t \Big[|\operatorname{Hess} f|\, |Q_t(v)|\, |Q_t(w)| \Big] + (P_t |df|^2)^{1/2} \mathbb{E}\Big[|W_t(v,w)|^2 \Big]^{1/2} \\
&\le (P_t |\operatorname{Hess} f|^2)^{1/2}\, \mathrm{e}^{2Kt} + C\, \mathrm{e}^{(2K+\theta)t}\, (P_t |df|^2)^{1/2}.
\end{aligned}
$$

□

Remark 2.9. Let

$$
\|R\|_\infty := \sup_{x \in M} |R|(x) \quad \text{and} \quad \|\nabla \operatorname{Ric}^\sharp + d^* R\|_\infty := \sup_{x \in M} |\nabla \operatorname{Ric}^\sharp + d^* R|(x).
$$

If assumption (**H**) is replaced by (**Ric**) with $K \ge 0$, $\|R\|_\infty < \infty$ and $\|\nabla \operatorname{Ric}^\sharp + d^* R\|_\infty < \infty$, then it is straightforward to see from (2.15) that for $0 < \delta < 2K$,

$$
\mathbb{E}|W_t(v,w)|^2 \le \mathrm{e}^{(2K+\delta)t} \mathbb{E}\left[\int_0^t \mathrm{e}^{(2K-\delta)s} \|R\|_\infty^2\, ds + \int_0^t \mathrm{e}^{(2K-\delta)s} \|\nabla \operatorname{Ric}^\sharp + d^* R\|_\infty^2\, ds \right] \le C\mathrm{e}^{4Kt} t
$$

for some explicit constant $C > 0$. Proceeding as in the proof of Proposition 2.8 above, we obtain in this case

$$
|\operatorname{Hess} P_t f| \le \mathrm{e}^{2Kt} (P_t |\operatorname{Hess} f|^2)^{1/2} + C\sqrt{t}\mathrm{e}^{2Kt} (P_t |df|^2)^{1/2}.
$$

Proposition 2.8 is used to establish pointwise Gaussian upper bounds for the Hessian of the heat kernel via inequality (2.17). Note that Carron [7] applied pointwise Gaussian upper bounds for the gradient of the heat kernel to the Riesz transform on complete manifolds whose Ricci curvature satisfies a quadratic decay control. Here, we shall use the following Hessian heat kernel estimate in the next section to establish the Calderón-Zygmund inequality.

Proposition 2.10. *Assume that* (**H**) *holds. Fix* $\alpha \in (0, 1/4)$ *as in Lemma* 2.1 *and let* $\theta > 0$ *be as in* (2.11). *Then there exist* $0 < \beta < 2\alpha$ *and* $C_3 > 0$ *such that*

$$|\operatorname{Hess}_x p_t(x,y)| \le \frac{C(1+\sqrt{t})}{tV(y,\sqrt{t})} \exp\left(-\beta \frac{\rho^2(x,y)}{t} + \frac{1}{2}(C_3+\theta)t\right).$$

for any $t > 0$ *and* $x, y \in M$.

Proof. For $t > 0$ and $x, y \in M$, write

$$\operatorname{Hess}_x p_{2t}(x,y) = \operatorname{Hess} P_t(p_t(\cdot,y))(x)$$

Applying Proposition 2.8 and taking $\gamma \in (0, 2\alpha)$, we have

$$\begin{aligned}\left|\operatorname{Hess} P_t(p_t(\cdot,y))(x)\right| &\le \mathrm{e}^{2Kt}\left(\int |\operatorname{Hess}_z p_t(z,y)|^2\, \mathrm{e}^{\gamma\rho^2(z,y)/t}\, \mathrm{e}^{-\gamma\rho^2(z,y)/t}\, |p_t(x,z)|\, \mu(dz)\right)^{1/2} \\ &\quad + C\mathrm{e}^{(2K+\theta)t}\left(\int |\nabla_z p_t(z,y)|^2\, \mathrm{e}^{\gamma\rho^2(z,y)/t}\, \mathrm{e}^{-\gamma\rho^2(z,y)/t}\, |p_t(x,z)|\, \mu(dz)\right)^{1/2}.\end{aligned}$$

Moreover, by Lemma 2.2 and Proposition 2.4, we know that there exists $C' > 0$ such that

$$\int_M \left(|\nabla_z p_t(z,y)|^2\, \mathrm{e}^{\gamma\rho^2(z,y)/t} + |\operatorname{Hess}_z p_t(z,y)|^2\, \mathrm{e}^{\gamma\rho^2(z,y)/t}\right) \mu(dz) \le \frac{C(1+Kt)}{t^2 V(y,\sqrt{t})}\, \mathrm{e}^{2C't}.$$

This further implies that there exists a constant $C_3 > 0$ such that

$$\begin{aligned}\left|\operatorname{Hess} P_t(p_t(\cdot,y))(x)\right| &\le \frac{C(1+\sqrt{t})}{t\sqrt{V(y,\sqrt{t})}}\, \mathrm{e}^{((C'+2K)+\theta)t} \sup_{z\in M}\left\{\mathrm{e}^{-\gamma\rho^2(z,y)/t}|p_t(x,z)|\right\}^{1/2} \\ &\le \frac{C(1+\sqrt{t})}{tV(y,\sqrt{t})}\, \mathrm{e}^{(C_3+\theta)t}\, \mathrm{e}^{-\beta\rho^2(x,y)/t}\end{aligned}$$

for β small enough, where in the last inequality we have used the estimate

$$\sup_{z\in M}\left\{\mathrm{e}^{-\gamma\rho^2(z,y)/t}|p_t(x,z)|\right\} \le \frac{\mathrm{e}^{C_1Kt}}{V(y,\sqrt{t})}\mathrm{e}^{-2\beta\rho^2(x,y)/t},$$

which can be deduced from Lemma 2.1. Thus, we conclude

$$|\operatorname{Hess}_x p_{2t}(x,y)| \le \frac{C(1+\sqrt{t})}{tV(y,\sqrt{t})}\, \mathrm{e}^{(C_3+\theta)t}\, \mathrm{e}^{-\beta\rho^2(x,y)/t}$$

which finishes the proof of Proposition 2.10. □

By interpolating the L^p-Hessian inequality with the L^2-Gaffney off-diagonal estimates, one obtains L^p-Gaffney off-diagonal estimates for any $p > 2$ as follows.

Proposition 2.11. *Assume that* (**H**) *hold. Then for $p > 2$ there exists constants $C, C_4 > 0$ such that for all $t \in (0, \infty)$, all Borel subsets $E, F \subset M$ with compact closure, and all $f \in L^p(M)$ with* $\operatorname{supp} f \subset E$,

$$\left\| \mathbb{1}_F t |\operatorname{Hess} P_t f| \right\|_p \le C(1 + \sqrt{t})\, \mathrm{e}^{(2K+\theta)t}\, \mathrm{e}^{-C_4 \rho^2(E,F)/t}\, \|f\|_p.$$

Proof. Let $p > 2$ and $t > 0$. For E, F and f as above, by inequality (2.12), there exist positive constants C and C_2 such that

$$\begin{aligned}\int_F t^p |\operatorname{Hess} P_t f|^p(x)\,\mu(dx) &\le C\mathrm{e}^{(2K+\theta)pt}(1 + \sqrt{t})^p \int_F (P_t |f|^{p/2})^2(x)\,\mu(dx) \\ &\le C\mathrm{e}^{-2C_2\rho^2(E,F)/t}\, \mathrm{e}^{(2K+\theta)pt}\, (1 + \sqrt{t})^p \int_E |f|^p\,\mu(dx)\end{aligned}$$

where the last inequality follows from the L^2-Gaffney off-diagonal estimates for $P_t f$ (Lemma 2.3), i.e. the existence of positive constants C and C_2 such that

$$\left\| \mathbb{1}_F |P_t f^{p/2}| \right\|_2 \le C\, \mathrm{e}^{-C_2\rho^2(E,F)/t} \left\| \mathbb{1}_E f^{p/2} \right\|_2$$

for all $f \in L^p(M)$ with $\operatorname{supp} f \subset E$. □

3 Calderón-Zygmund inequality for $p \neq 2$

In this section, we prove the main results of this paper. We first prove Theorem 1.1 in Section 3.1, namely **CZ(**p**)** for $p \in (1, 2)$ under the condition (**Ric**); then in Section 3.2, we prove Theorem 1.2, namely, **CZ(**p**)** for $p > 2$ under the condition (**H**). Recall that **CZ(2)** always holds under condition (**Ric**).

3.1 The case $p \in (1, 2)$

Let M be a complete Riemannian manifold satisfying (**Ric**). In this subsection, we prove Theorem 1.1 and show that the Calderón-Zygmund inequality holds for $p \in (1, 2)$. To be precise, let

$$T = \operatorname{Hess}(\Delta + \sigma)^{-1} = \int_0^\infty \mathrm{e}^{-\sigma t} \operatorname{Hess} P_t\, dt. \tag{3.1}$$

We show that T is bounded on $L^p(M)$ for any $p \in (1, 2)$, provided $\sigma > 0$ is large enough. Since T is already bounded on $L^2(M)$, using interpolation, it suffices to prove that T is of weak type $(1, 1)$, that is for some constant C,

$$\mu(\{x \in M : |T f(x)| > \lambda\}) \le C \frac{\|f\|_1}{\lambda} \tag{3.2}$$

for all $\lambda > 0$ and $f \in L^1(M)$. To this end, we need the following technical lemma from [3, Section 4] on the finite overlap property of M.

Lemma 3.1. [3] *Assume that* (**LD**) *holds. There exists a countable subset $C = \{x_j\}_{j \in \Lambda} \subset M$ such that*

(i) $M = \cup_{j\in\Lambda} B(x_j, 1)$;

(ii) $\{B(x_j, 1/2)\}_{j\in\Lambda}$ *are disjoint;*

(iii) *there exists* $N_0 \in \mathbb{N}$ *such that for any* $x \in M$*, there are at most* N_0 *balls* $B(x_j, 4)$ *containing* x*;*

(iv) *for any* $c_0 \geq 1$*, there exists* $C > 0$ *such that for any* $j \in \Lambda$*,* $x \in B(x_j, c_0)$ *and* $r \in (0, \infty)$*,*

$$\mu\left(B(x, 2r) \cap B(x_j, c_0)\right) \leq C\mu\left(B(x, r) \cap B(x_j, c_0)\right)$$

and

$$\mu(B(x, r)) \leq C\mu\left(B(x, r) \cap B(x_j, c_0)\right)$$

for any $x \in B(x_j, c_0)$ *and* $r \in (0, 2c_0]$.

The following lemma provides the localization argument to prove (3.2).

Lemma 3.2. *Assume that* (**Ric**) *holds. Let* $C = \{x_j\}_{j\in\Lambda}$ *be a countable subset of* M *having finite overlap property as in Lemma* 3.1*. Let* $\sigma > C''$ *where* C'' *is as in Corollary 2.5. Suppose that there exists a positive constant* C *such that*

$$\mu\left(\{x\colon \mathbb{1}_{B(x_j,2)}|Tf(x)| > \lambda\}\right) \leq \frac{C}{\lambda}\|f\|_1 \tag{3.3}$$

for any $j \in \Lambda$*,* $\lambda \in (0, \infty)$ *and* $f \in C_0^\infty(B(x_j, 1))$*. Then* (3.2) *holds for any* $f \in C_0^\infty(M)$*.*

Proof. For $j \in \Lambda$, let $B_j := B(x_j, 1)$ and let $\{\varphi_j\}_{j\in\Lambda}$ be a C^∞-partition of unity such that $0 \leq \varphi_j \leq 1$ and is supported in B_j. Then, for any $f \in C_0^\infty(M)$ and $x \in M$, write

$$Tf(x) = \sum_{j\in\Lambda} \mathbb{1}_{2B_j} T(f\varphi_j)(x) + \sum_{j\in\Lambda}(1 - \mathbb{1}_{2B_j})T(f\varphi_j)(x),$$

which yields that for any $\lambda > 0$,

$$\begin{aligned}\mu(\{x\colon |Tf(x)| > \lambda\}) &\leq \mu\left(\left\{x\colon \sum_{j\in\Lambda} \mathbb{1}_{2B_j}|T(f\varphi_j)(x)| > \frac{\lambda}{2}\right\}\right) + \mu\left(\left\{x\colon \sum_{j\in\Lambda}(1 - \mathbb{1}_{2B_j})|T(f\varphi_j)(x)| > \frac{\lambda}{2}\right\}\right)\\ &=: \mathrm{I}_1 + \mathrm{I}_2.\end{aligned}$$

For I_1, by Lemma 3.1(iii) and (3.3), we have

$$\mathrm{I}_1 \leq \sum_{j\in\Lambda}\mu\left(\left\{x : \mathbb{1}_{2B_j}|T(f\varphi_j)(x)| > \frac{\lambda}{2N_0}\right\}\right) \lesssim \frac{1}{\lambda}\|f\|_1 \tag{3.4}$$

as desired, where the notation $a \lesssim b$ means $a \leq Cb$ for some constant C.

To bound I_2, again by Lemma 3.1(iii), since φ_j is supported in B_j, it is easy to see that

$$\sum_{j\in\Lambda} |(1 - \mathbb{1}_{2B_j})(x)\varphi_j(y)| \leq N_0 \mathbb{1}_{\{\rho(x,y)\geq 1\}}.$$

Hence, according to the definition of T in (3.1) and Corollary 2.5, we get

$$\mathrm{I}_2 \leq \frac{2}{\lambda}\sum_{j\in\Lambda}\left\|\left|\left(1 - \mathbb{1}_{2B_j}\right)T\left(f\varphi_j\right)\right|\right\|_1$$

$$
\begin{aligned}
&\lesssim \frac{1}{\lambda}\int_M\left(\int_0^\infty \mathrm{e}^{-\sigma t}\int_M |\mathrm{Hess}_x\, p_t(x,y)|\sum_{j\in\Lambda}|(1-\mathbb{1}_{2B_j})(x)\varphi_j(y)||f(y)|\,\mu(dy)\,dt\right)\mu(dx)\\
&\lesssim \frac{1}{\lambda}\int_M\int_0^\infty \mathrm{e}^{-\sigma t}\left(\int_{\rho(x,y)\geq 1}|\,\mathrm{Hess}_x\, p_t(x,y)|\,\mu(dx)\right)dt\,|f(y)|\,\mu(dy)\\
&\leq \frac{1}{\lambda}\int_M |f(y)|\,\mu(dy)\int_0^\infty \mathrm{e}^{-\sigma t}(1+\sqrt{t})\,\mathrm{e}^{C''t}\,\mathrm{e}^{-\beta/t}t^{-1}\,dt,
\end{aligned}
$$

where $\beta\in(0,\alpha)$. Thus, since $\sigma > C''$, we obtain

$$
\mathrm{I}_2 \lesssim \frac{1}{\lambda}\int_0^\infty \mathrm{e}^{(C''-\sigma)t-\beta/t}\frac{(1+\sqrt{t})}{t}\,dt\,\|f\|_1 \lesssim \frac{1}{\lambda}\|f\|_1,
$$

which combined with the estimate about I_1 in (3.4) finishes the proof of Lemma 3.2. □

To prove property (3.3), we remove the subscript j and simply write B for each $B_j := B(x_j, 1)$. Let $c_0 \geq 1$. By Lemma 3.1(iv), we have that (c_0B, μ, ρ) is a metric measure subspace satisfying the *volume doubling property* that there exists $C_D \geq 1$ such that

$$
\mu\,(B(x,2r)\cap c_0B) \leq C_D\,\mu\,(B(x,r)\cap c_0B) \tag{\textbf{D}}
$$

for all $x\in c_0B$ and $r>0$.

We also need the following Calderón-Zygmund decomposition from [11].

Lemma 3.3 ([11])**.** *Let $(\mathcal{X},\nu,\rho)$ be a metric measure space satisfying* (**D**) *with c_0B replaced by $\mathcal{X}$. Let $f\in L^1(\mathcal{X})$ and $\lambda\in(0,\infty)$. Assume $\|f\|_{L^1} < \lambda\nu(\mathcal{X})$. Then f has a decomposition of the form*

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

such that

(a) *$g(x)\leq C\lambda$ for almost all $x\in M$;*

(b) *there exists a sequence of balls $B_i = B(x_i, r_i)$ so that the support of each b_i is contained in B_i:*

$$
\int_{\mathcal{X}} |b_i(x)|\,\nu(dx) \leq C\lambda\nu(B_i) \quad \textit{and} \quad \int_{\mathcal{X}} b_i(x)\,\nu(dx) = 0;
$$

(c) $\displaystyle\sum_i \nu(B_i) \leq \frac{C}{\lambda}\int_{\mathcal{X}}|f(x)|\,\nu(dx)$*;*

(d) *there exists $k_0\in\mathbb{N}^*$ such that each point of M is contained in at most k_0 balls B_i.*

Lemma 3.4. *Assume that* (**Ric**) *holds. Let $\lambda\in(0,\infty)$ and $f\in L^1(B)$ be as in Lemma* 3.3*. Assume that $\{b_i\}$ is the sequence of bad functions as in Lemma* 3.3 *and $\{P_t^\sigma\}_{t\geq 0}$ the heat semigroup associated to $-(\Delta+\sigma)$ with $\sigma>0$. Then there exist $\sigma>0$ and $C>0$ independent of f such that*

$$
\Big\|\sum_i P_{t_i}^\sigma b_i\Big\|_2^2 \leq C\lambda\|f\|_1
$$

where $t_i = r_i^2$ with r_i denoting the radius of the ball B_i as in Lemma 3.3(b).

Proof. Recall that $\operatorname{supp} b_i \subset B(x_i, \sqrt{t_i})$. Using the upper bound of the heat kernel in Lemma 2.2 and Lemma 3.3, we have for $x \in M$,

$$\begin{aligned}
|P_{t_i}^{\sigma} b_i(x)| &\leq \int_M \frac{\mathrm{e}^{-\sigma' t_i - \alpha \frac{\rho^2(x,y)}{t_i}}}{V(x, \sqrt{t_i})} |b_i(y)|\,\mu(dy) \\
&\leq \frac{C}{V(x, \sqrt{t_i})}\, \mathrm{e}^{-\sigma' t_i - \alpha' \frac{\rho^2(x,x_i)}{t_i}} \int_{B_i} |b_i(y)|\,\mu(dy) \\
&\leq C_2 \lambda \int_M \frac{\mathrm{e}^{-\sigma' t_i - \alpha'' \frac{\rho^2(x,y)}{t_i}}}{V(x, \sqrt{t_i})} \mathbb{1}_{B_i}(y)\,\mu(dy),
\end{aligned}$$

where $\sigma' = \sigma - C_1 K > 0$ and $0 < \alpha'' < \alpha' < \alpha$. It is therefore sufficient to verify that

$$\Big\| \sum_i \int_M \frac{\mathrm{e}^{-\sigma' t_i - \alpha'' \frac{\rho^2(\cdot,y)}{t_i}}}{V(\cdot, \sqrt{t_i})} \mathbb{1}_{B_i}(y)\,\mu(dy) \Big\|_2 \lesssim \Big\| \sum_i \mathbb{1}_{B_i} \Big\|_2, \tag{3.5}$$

since from this and Lemma 3.3 we obtain as consequence, as desired,

$$\Big\| \sum_i P_{t_i}^{\sigma} b_i \Big\|_2^2 \lesssim \lambda^2 \Big\| \sum_i \mathbb{1}_{B_i} \Big\|_2^2 \lesssim \lambda^2 \sum_i \mu(B_i) \lesssim \lambda \|f\|_1.$$

In order to prove (3.5), we write by duality

$$\begin{aligned}
&\Big\| \sum_i \int_M \frac{\mathrm{e}^{-\sigma' t_i - \alpha'' \frac{\rho^2(\cdot,y)}{t_i}}}{V(\cdot, \sqrt{t_i})} \mathbb{1}_{B_i}(y)\,\mu(dy) \Big\|_2 \qquad (3.6)\\
&\quad = \sup_{\|u\|_2 = 1} \left| \int_M \left(\sum_i \int_M \frac{\mathrm{e}^{-\sigma' t_i - \alpha'' \frac{\rho^2(x,y)}{t_i}}}{V(x, \sqrt{t_i})} \mathbb{1}_{B_i}(y)\,\mu(dy) \right) u(x)\,\mu(dx) \right| \\
&\quad \leq \sup_{\|u\|_2 = 1} \int_M \sum_i \left(\int_M \frac{\mathrm{e}^{-\sigma' t_i - \alpha'' \frac{\rho^2(x,y)}{t_i}}}{V(x, \sqrt{t_i})} |u(x)|\,\mu(dx) \right) \mathbb{1}_{B_i}(y)\,\mu(dy).
\end{aligned}$$

By (2.1), we have for any $x \in M$ and $y \in B_i$,

$$V(y, \sqrt{t_i}) \leq C \left(1 + \frac{\rho(x,y)}{\sqrt{t_i}} \right)^d \mathrm{e}^{C(t_i^{1/2} + \rho(x,y))} V(x, \sqrt{t_i})$$

with $\delta \in [0, 2)$. From this, we obtain that there exist $0 < \tilde{\alpha} < \alpha''' < \alpha''$ such that

$$\begin{aligned}
&\int_M \frac{\mathrm{e}^{-\sigma' t_i - \alpha'' \frac{\rho^2(x,y)}{t_i}}}{V(x, \sqrt{t_i})} |u(x)|\,\mu(dx) \\
&\quad \lesssim \frac{\mathrm{e}^{-\frac{1}{2}\sigma' t_i}}{V(y, \sqrt{t_i})} \int_M \mathrm{e}^{-\alpha''' \frac{\rho^2(x,y)}{t_i}} |u(x)|\,\mu(dx) \\
&\quad \lesssim \frac{1}{V(y, \sqrt{t_i})} \left(\int_{\rho(x,y) < \sqrt{t_i}} |u(x)|\,\mu(dx) + \sum_{k=0}^{\infty} \int_{2^k \sqrt{t_i} \leq \rho(x,y) < 2^{k+1} \sqrt{t_i}} \mathrm{e}^{-\alpha''' \frac{\rho^2(x,y)}{t_i}} |u(x)|\,\mu(dx) \right)
\end{aligned}$$

$$
\begin{aligned}
&\le \frac{1}{V(y,\sqrt{t_i})}\left(\int_{B(y,\sqrt{t_i})}|u(x)|\,\mu(dx)+\sum_{k=0}^{\infty}\mathrm{e}^{-\alpha'''2^{2k}}\int_{B(y,2^{k+1}\sqrt{t_i})}|u(x)|\,\mu(dx)\right)\\
&=\left(1+\sum_{k=0}^{\infty}\frac{V(y,2^{k+1}\sqrt{t_i})}{V(y,\sqrt{t_i})}\,\mathrm{e}^{-\alpha'''2^{2k}}\right)(\mathscr{M}u)(y)\\
&\le\left(1+C\sum_{k=0}^{\infty}2^{(k+1)d}\mathrm{e}^{C2^{k+1}\sqrt{t_i}}\mathrm{e}^{-\tilde{\alpha}2^{2k}}\right)(\mathscr{M}u)(y)\\
&\le\left(1+C\sum_{k=0}^{\infty}2^{(k+1)d}\mathrm{e}^{C2^{k+1}}\mathrm{e}^{-\tilde{\alpha}2^{2k}}\right)(\mathscr{M}u)(y)\lesssim(\mathscr{M}u)(y),
\end{aligned}
$$

where

$$
(\mathscr{M}u)(y):=\sup_{r>0}\frac{1}{V(y,r)}\int_{B(y,r)}|u(x)|\,\mu(dx)
$$

denotes the Hardy-Littlewood maximal function of u. This together with (3.6) and the L^2-boundedness of $\mathscr{M}$ gives

$$
\left\|\sum_i\int_M\frac{\mathrm{e}^{-\alpha''\frac{d^2(\cdot,y)}{t_i}}}{V(\cdot,\sqrt{t_i})}\mathbb{1}_{B_i}(y)\,\mu(dy)\right\|_2\lesssim\sup_{\|u\|_2=1}\int_M(\mathscr{M}u)(y)\sum_i\mathbb{1}_{B_i}(y)\,\mu(dy)\lesssim\left\|\sum_i\mathbb{1}_{B_i}\right\|_2,
$$

which shows that (3.5) holds true and finishes the proof of Lemma 3.4. □

With the help of Lemmas 3.2 through 3.4, we are now in position to the proof of Theorem 1.1.

Proof of Theorem 1.1. Recall that $T=\mathrm{Hess}(\Delta+\sigma)^{-1}$. We choose σ big enough such that $\sigma>\max\{C'',C_1K\}$, where $\tilde{\alpha}$ is as in proof of Lemma 3.4. By Lemma 3.2, it suffices to prove

$$
\mu(\{x\in 2B\colon |Tf(x)|>\lambda\})\lesssim\frac{\|f\|_1}{\lambda}\tag{3.7}
$$

for all $f\in C_0^\infty(B)$. By means of Lemma 3.3 with $\mathcal{X}=B$, we deduce that f has a decomposition

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

which implies

$$
\begin{aligned}
\mu(\{x\in 2B:|Tf(x)|>\lambda\})&\le\mu\left(\left\{x\in 2B:|Tg(x)|>\frac{\lambda}{2}\right\}\right)+\mu\left(\left\{x\in 2B:|Tb(x)|>\frac{\lambda}{2}\right\}\right)\\
&=:\mathrm{I}_1+\mathrm{I}_2.
\end{aligned}\tag{3.8}
$$

Using the facts that T is bounded on $L^2(M)$ and that $|g(x)|\le C\lambda$, we obtain as desired

$$
\mathrm{I}_1\lesssim\lambda^{-2}\|Tg\|_2^2\lesssim\lambda^{-2}\|g\|_2^2\lesssim\lambda^{-1}\|g\|_1\lesssim\lambda^{-1}\|f\|_1.\tag{3.9}
$$

We now turn to the estimate of I_2. Recall that $\{P_t^\sigma\}_{t\ge0}$ is the heat semigroup generated by $-(\Delta+\sigma)$, that is $P_t^\sigma=\mathrm{e}^{-t\sigma}P_t$. We write

$$
Tb_i=TP_{t_i}^\sigma b_i+T(I-P_{t_i}^\sigma)b_i,
$$

where $t_i = r_i^2$ with r_i the radius of B_i. By Lemma 3.4, we have

$$\Big\|\sum_i P^\sigma_{t_i} b_i\Big\|_2^2 \lesssim \lambda \|f\|_1.$$

This combined with the L^2-boundedness of T yields

$$\mu\left(\left\{x\in 2B\colon \left|T\left(\sum_i P^\sigma_{t_i} b_i\right)(x)\right| > \frac{\lambda}{2}\right\}\right) \lesssim \frac{1}{\lambda}\|f\|_1 \tag{3.10}$$

as desired. Consider now the term $T\sum_i (I-P^\sigma_{t_i})b_i$. We write

$$\begin{aligned}\mu&\left(\left\{x\in 2B\colon \left|T\left(\sum_i (I-P^\sigma_{t_i})b_i\right)(x)\right| > \frac{\lambda}{2}\right\}\right) \\ &\leq \sum_i \mu(2B_i) + \mu\left(\left\{x\in 2B\setminus \cup_i 2B_i\colon \left|T\left(\sum_i (I-P^\sigma_{t_i})b_i\right)\right|(x) > \frac{\lambda}{2}\right\}\right).\end{aligned} \tag{3.11}$$

From Lemma 3.3, it follows that

$$\sum_i \mu(2B_i) \lesssim \frac{\|f\|_1}{\lambda} \tag{3.12}$$

as desired. To estimate the second term, let $k^\sigma_{t_i}(x,y)$ denote the integral kernel of the operator $T(I-P^\sigma_{t_i})$. Note that

$$(\Delta+\sigma)^{-1}(I-P^\sigma_{t_i}) = \int_0^{+\infty} (P^\sigma_s - P^\sigma_{t_i+s})\,ds = \int_0^{t_i} P^\sigma_s\,ds$$

and

$$T(I-P^\sigma_{t_i}) = \mathrm{Hess}(\Delta+\sigma)^{-1}(I-P^\sigma_{t_i}) = \int_0^{t_i} \mathrm{Hess}\,P^\sigma_s\,ds.$$

Therefore,

$$k^\sigma_{t_i}(x,y) = \int_0^{t_i} \mathrm{Hess}_x\, p^\sigma_s(x,y)\,ds, \tag{3.13}$$

where $p^\sigma_s = \mathrm{e}^{-\sigma s} p_s$ is the kernel of P^σ_s with respect to μ. Since b_i is supported in B_i, we have

$$\begin{aligned}\int_{2B\setminus(2B_i)} \left|T\left((I-P^\sigma_{t_i})b_i\right)(x)\right|\mu(dx) &\leq \int_{2B\setminus(2B_i)}\left(\int_{B_i} |k^\sigma_{t_i}(x,y)||b_i(y)|\,\mu(dy)\right)\mu(dx) \\ &\leq \int_{B_i}\left(\int_{\rho(x,y)\geq t_i^{1/2}} |k^\sigma_{t_i}(x,y)|\,\mu(dx)\right)|b_i(y)|\,\mu(dy).\end{aligned} \tag{3.14}$$

Now by means of (3.13) and Corollary 2.5, we get

$$\int_{\rho(x,y)\geq t_i^{1/2}} |k^\sigma_{t_i}(x,y)|\,\mu(dx) \leq \int_0^{t_i}\left(\int_{\rho(x,y)\geq t_i^{1/2}} |\,\mathrm{Hess}_x\, p_s(x,y)|\,\mu(dx)\right)\mathrm{e}^{-s\sigma}\,ds$$

$$\leq C \int_0^{t_i} \mathrm{e}^{-\beta t_i/s} s^{-1} \mathrm{e}^{C''s} \mathrm{e}^{-s\sigma}(1+\sqrt{s})\, ds$$
$$\leq C \int_0^1 \frac{\mathrm{e}^{-\beta/u}}{u}\, du < \infty,$$

where for the last inequality we use the fact that

$$\mathrm{e}^{s(C''-\sigma)}(1+\sqrt{s}) < \infty, \quad s \in (0,\infty).$$

The estimate above together with (3.14) and Lemma 3.3 implies that

$$\mu\left(\left\{x \in 2B \setminus \cup_i 2B_i \colon \left|T\left(\sum_i (I - P_{t_i}^\sigma) b_i\right)(x)\right| > \frac{\lambda}{2}\right\}\right) \lesssim \frac{\|f\|_1}{\lambda}. \tag{3.15}$$

Altogether, combining (3.8) through (3.10), (3.12) and (3.15), we conclude that (3.7) holds which completes the proof of Theorem 1.1. □

3.2 The case $p \in (2, \infty)$

Let M be a complete Riemannian manifold satisfying (**H**). In this subsection, we prove Theorem 1.2 and show that the Calderón-Zygmund inequality $\mathbf{CZ}(p)$ holds for all $p \in (2,\infty)$, that is,

$$\|Tf\|_p \lesssim \|f\|_p \tag{3.16}$$

holds for any $f \in L^p(M)$ with T as in (3.1) and $\sigma \in (2K+\theta, \infty)$ where K, θ are respectively in (**Ric**) and (2.11).

To this end, let w be a C^∞ function on $[0,\infty)$ satisfying $0 \leq w \leq 1$ and

$$w(t) = \begin{cases} 1 & \text{on } [0, 3/4], \\ 0 & \text{on } [1, \infty), \end{cases}$$

and let $\widetilde{T}$ be an operator defined by

$$\widetilde{T} f := \int_0^\infty v(t) \operatorname{Hess} P_t f\, dt \tag{3.17}$$

with $v(t) := w(t)\mathrm{e}^{-\sigma t}$. We need the following lemma, which reduces (3.16) to a time and spatial localized version.

Lemma 3.5. *Assume that* (**H**) *holds. Let $p \in (2,\infty)$ and $\{x_j\}_{j\in\Lambda}$ be a countable subset of M having the finite overlap property as in Lemma* 3.1. *If there exists a positive constant C such that*

$$\left\| |\widetilde{T}(f)| \right\|_{L^p(B(x_j,4))} \leq C \|f\|_{L^p(B(x_j,1))} \tag{3.18}$$

holds for any $j \in \Lambda$ and $f \in C_0^\infty(B(x_j,1))$ with $\widetilde{T}$ defined as in (3.17)*, then* (3.16) *holds.*

Proof. By the fact that $w \equiv 1$ on $[0, 3/4]$, we obtain from Proposition 2.7 that if $\sigma > 2K + \theta$, then for any $g \in L^p(M)$,

$$\left\| \int_0^\infty (1 - w(t))\, \mathrm{e}^{-\sigma t} |\operatorname{Hess} P_t g|\, dt \right\|_p \lesssim \int_{3/4}^\infty \mathrm{e}^{(2K+\theta-\sigma)t} \frac{1+\sqrt{t}}{t} dt\, \|g\|_p \lesssim \|g\|_p.$$

This and (3.17) imply that to prove (3.16), it suffices to show that

$$\left\| |\widetilde{T}(g)| \right\|_p \lesssim \|g\|_p \tag{3.19}$$

for any $g \in C_0^\infty(M)$.

Let $(x_j)_{j\in\Lambda}$ be a countable subset of M having the finite overlap property as in Lemma 3.1. Let $\{\varphi\}_j$ be a corresponding C^∞ partition of unity such that $0 \le \varphi_j \le 1$ and φ_j is supported in $B_j := B(x_j, 1)$. Let χ_j be the characteristic function of the ball $4B_j$. For any $g \in C_0^\infty(M)$ and $x \in M$, write

$$\widetilde{T}g(x) \le \sum_{j\in\Lambda} \chi_j \widetilde{T}(g\varphi_j)(x) + \sum_{j\in\Lambda}(1-\chi_j)\widetilde{T}(g\varphi_j)(x) =: \mathrm{I}(x) + \mathrm{II}(x). \tag{3.20}$$

By Lemma 3.1, we know

$$\sum_{j\in\Lambda} \left|(1-\chi_j)(x)\varphi_j(y)\right| \le N_0 \mathbb{1}_{\{\rho(x,y)\ge 3\}}.$$

Hence, by Hölder's inequality, we have

$$\begin{aligned}
\mathrm{II}(x) &\le \int_0^1 \int_M |\mathrm{Hess}_x\, p_t(x,y)| \left(\sum_{j\in\Lambda} \left|(1-\chi_j)(x)\varphi_j(y)\right| \right) |g(y)|\, \mu(dy)\, dt \\
&\le N_0 \int_0^1 \int_{\rho(x,y)\ge 3} |\mathrm{Hess}_x\, p_t(x,y)||g(y)|\, \mu(dy)\, dt \\
&\lesssim \int_0^1 \left(\int_M \left| t\,\mathrm{Hess}_x\, p_t(x,y) \right|^p \mathrm{e}^{\gamma\rho^2(x,y)/t} \left(V(y,\sqrt{t})\right)^{p/p'} |g(y)|^p\, \mu(dy) \right)^{1/p} \frac{\mathrm{e}^{-c/t}}{t}\, dt,
\end{aligned}$$

where $c = \gamma p'/p$ and γ is a positive constant small enough so that in view of Proposition 2.10 and an argument similar to the proof of (2.4), it implies that for any $t \in (0,1)$,

$$\int_M |t\,\mathrm{Hess}_x\, p_t(x,y)|^p\, \mathrm{e}^{\gamma \frac{\rho^2(x,y)}{t}}\, \mu(dx) \lesssim \frac{1}{\left(V(y,\sqrt{t})\right)^{p-1}}.$$

This immediately implies

$$\begin{aligned}
&\int_M |\mathrm{II}(x)|^p\, \mu(dx) \qquad (3.21)\\
&\lesssim \int_M \left[\int_0^1 \left(\int_M |t\,\mathrm{Hess}_x\, p_t(x,y)|^p\, \mathrm{e}^{\gamma\rho^2(x,y)/t} V(y,\sqrt{t})^{p/p'} |f(y)|^p\, \mu(dy) \right)^{1/p} \frac{\mathrm{e}^{-c/t}}{t}\, dt \right]^p \mu(dx) \\
&\lesssim \int_M \int_0^1 \left(\int_M |t\,\mathrm{Hess}_x\, p_t(x,y)|^p\, \mathrm{e}^{\gamma\rho^2(x,y)/t}\, V(y,\sqrt{t})^{p/p'} |f(y)|^p\, \mu(dy) \right) dt\, \mu(dx) \\
&\qquad \times \left(\int_0^1 \frac{\mathrm{e}^{-cp'/t}}{t^{p'}}\, dt \right)^{p/p'} \\
&\lesssim \int_0^1 \left[\int_M [V(y,\sqrt{t})]^{p/p'} |f(y)|^p \left(\int_M |t\,\mathrm{Hess}_x\, p_t(x,y)|^p\, \mathrm{e}^{\gamma\rho^2(x,y)/t}\, \mu(dx) \right) \mu(dy) \right] dt \\
&\lesssim \int_M |f(y)|^p \mu(dy)
\end{aligned}$$

as desired.

Now we estimate I(x). Using Lemma 3.1, we know that the balls $\{4B_j\}_{j\in\Lambda}$ are of uniform overlap and hence

$$\sum_j \|h\chi_j\|_{p'}^{p'} \lesssim \|h\|_{p'}^{p'}$$

for all $h \in C_0^\infty(M)$. Since $g\varphi_j \in C_0^\infty(B(x_j,1))$ and using (3.18), we conclude that

$$\begin{aligned}\left|\int_M h(x)\mathrm{I}(x)\,\mu(dx)\right| &\le \int_M |h(x)|\left|\sum_j \chi_j \widetilde{T}(g\varphi_j)(x)\right|\mu(dx)\\ &\lesssim \sum_j \|g\varphi_j\|_p\|h\chi_j\|_{p'} \lesssim \|g\|_p\|h\|_{p'},\end{aligned}$$

which together with (3.20) and (3.21) implies (3.16), and hence finishes the proof of Lemma 3.5. □

From Lemma 3.5, it is easy to see that to prove T is of strong type (p,p) for $p>2$, it suffices to show that $\widetilde{T}$ is bounded from $L^p(B_j,\mu)$ to $L^p(4B_j,\mu)$ for each $j\in\Lambda$ as in (3.18). To this end, we need an L^p local bounded criterion from [3] via maximal functions. Recall that the *local maximal function* by

$$(\mathscr{M}_{\mathrm{loc}}f)(x) := \sup_{\substack{B\ni x\\ r(B)\le 32}} \frac{1}{\mu(B)}\int_B |f|\,d\mu, \quad x\in M, \tag{3.22}$$

for any locally integrable function f on M. From (**LD**), it follows that $\mathscr{M}_{\mathrm{loc}}$ is bounded on $L^p(M)$ for all $1<p\le\infty$. For a measurable subset $E\subset M$, the *maximal function relative to E* is defined by

$$(\mathscr{M}_E f)(x) := \sup_{B \text{ ball in } M,\, B\ni x} \frac{1}{\mu(B\cap E)}\int_{B\cap E} |f|\,d\mu, \quad x\in E, \tag{3.23}$$

for any locally integrable function f on M. If in particular E is a ball with radius r, it is enough to consider balls B with radii not exceeding $2r$. It is also easy to see $\mathscr{M}_E$ is of weak type $(1,1)$ and $L^p(M)$-bounded for $1<p\le\infty$ if E satisfies the *relative doubling property*, namely, if there exists a constant C_E such that for all $x\in E$ and $r>0$,

$$\mu(B(x,2r)\cap E) \le C_E\,\mu(B(x,r)\cap E). \tag{3.24}$$

Note that in Lemma 3.1 (iv), for any $j\in\Lambda$, the subsets $4B_j$ satisfy the relative doubling property (3.24) with a constant independent of j.

The following theorem is essential to the proof of Theorem 1.2.

Lemma 3.6. *Let $p\in(2,\infty)$ and assume that* (**LD**) *holds. For any ball B of M centered in $4B_j$ with radius less than 8, assume that*

(i) *there exists an integer n depending only on condition* (**LD**) *such that the map $f\to \mathscr{M}^{\#}_{4B_j,\tilde{T},n}f$ is bounded from $L^p(B_j,\mu)$ to $L^p(4B_j,\mu)$ with operator norm independent of j, where*

$$\mathscr{M}^{\#}_{4B_j,\tilde{T},n}f(x) := \sup_{B \text{ ball in } M,\, B\ni x}\left(\frac{1}{\mu(B\cap 4B_j)}\int_{B\cap 4B_j}\left|\widetilde{T}(I-P_{r^2})^n f(y)\right|^2\,\mu(dy)\right)^{1/2};$$

(ii) *for all $k \in \{1, 2, \ldots, n\}$, and all $f \in L^2(M, \mu)$ supported in B_j, there exists a sublinear operator S_j bounded from $L^p(B_j, \mu)$ to $L^p(4B_j, \mu)$ with operator norm independent of j such that for $x \in B \cap 4B_j$,*

$$\left(\frac{1}{\mu(B \cap 4B_j)} \int_{B \cap 4B_j} |\widetilde{T} P_{kr^2} f|^p \, d\mu \right)^{1/p} \leq C \left(\mathscr{M}_{4B_j}(|\widetilde{T} f|^2) + (S_j f)^2 \right)^{1/2} (x) \tag{3.25}$$

where $\mathscr{M}_{4B_j}$ is as in (3.23).

Then $\widetilde{T}$ is bounded from $L^p(B_j, \mu)$ to $L^p(4B_j, \mu)$, that is (3.18) *holds with a constant depending on p, the doubling constant C_D in* (**D**)*, the operator norm of $\widetilde{T}$ on $L^2(M, \mu)$, the operator norms of $\mathscr{M}^{\#}_{4B_j, \tilde{T}, n}$ and S_j on L^p, and the constant in* (3.25).

Proof. We use the following result [3, Theorem 2.4]:
Let $p_0 \in (2, \infty]$ and (M, μ, ρ) be a measured metric space. Suppose that T is a bounded sublinear operator which is bounded on $L^2(M, \mu)$, and let $\{A_r\}_{r>0}$, be a family of linear operators acting on $L^2(M, \mu)$. Let E_1 and E_2 be two subsets of M such that E_2 has the relative doubling property, $\mu(E_2) < \infty$ and $E_1 \subset E_2$. Assume that

(i) *the sharp maximal functional $\mathscr{M}^{\#}_{E_2, T, A}$ is bounded from $L^p(E_1, \mu)$ into $L^p(E_2, \mu)$ for all $p \in (2, p_0)$, where*

$$(\mathscr{M}^{\#}_{E_2, T, A} f)^2(x) = \sup_{B \text{ ball in } M,\ B \ni x} \frac{1}{\mu(B \cap E_2)} \int_{B \cap E_2} |T(I - A_{r(B)}) f|^2 \, d\mu$$

for $x \in E_2$;

(ii) *for some sublinear operator S bounded from $L^p(E_1, \mu)$ into $L^p(E_2, \mu)$ for all $p \in (2, p_0)$,*

$$\left(\frac{1}{\mu(B \cap E_2)} \int_{B \cap E_2} |T A_{r(B)} f|^{p_0} \, d\mu \right)^{1/p_0} \leq C \left(\mathscr{M}_{E_2}(|T f|^2) + (S(f)^2) \right)^{1/2} (x), \tag{3.26}$$

for all $f \in L^2(M, \mu)$ supported in E_1, all balls B in M and all $x \in B \cap E_2$, where $r(B)$ is the radius of B.

If $2 < p < p_0$ and $Tf \in L^p(E_2, \mu)$ whenever $f \in L^p(E_1, \mu)$, then T is bounded from $L^p(E_1, \mu)$ into $L^p(E_2, \mu)$ and its operator norm is bounded by a constant depending only on the operator norm of T on $L^2(M, \mu)$, C_{E_2} (see in (3.24) *for E_2), p and p_0, the operator norms of $\mathscr{M}^{\#}_{4B_j, \tilde{T}, n}$ and S on L^p, and the constant in* (3.26).

Here we may take E_1 and E_2 as B_j and $4B_j$ respectively, as the sets B_j and $4B_j$ have the relatively volume doubling property as in (3.24) with the constant C_E independent of j (see Lemma 3.1). Moreover, taking the operators $\{A_r\}_{r>0}$ such that

$$I - A_r = (I - P_{r^2})^n$$

for some integer n sufficiently large, then the result follows directly. □

Now it suffices to check (i) and (ii) of Lemma 3.6. We establish two technical lemmas which verify (i) and (ii) respectively. To this end, observe that (**LD**) implies: for all $r_0 > 0$ there exists C_{r_0} such that for all $x \in M$, $r \in (0, r_0)$,

$$V(x, 2r) \leq C_{r_0} V(x, r).$$

An easy consequence of the definition is that for all $y \in M$, $0 < r < 8$ and $s \geq 1$ satisfying $sr < 32$,

$$V(y, sr) \leq C s^{D_L} V(y, r), \tag{3.27}$$

for some constants C and $D_L > 0$. The following lemma plays an important role, when checking (i) of Lemma 3.6.

Lemma 3.7. *Assume that* (**H**) *hold. Then there exists an integer n such that the inequality*

$$\sup_{B\, ball\ in\ M,\, B\ni x} \left(\frac{1}{\mu(B \cap 4B_j)} \int_{B\cap 4B_j} \left| \widetilde{T}(I - P_{r^2})^n f(y) \right|^2 \mu(dy) \right)^{1/2} \leq C \left(\mathscr{M}_{\mathrm{loc}}(|f|^2)(x) \right)^{1/2} \tag{3.28}$$

holds for any $x \in 4B_j$, $f \in L^2(4B_j)$ satisfying $\operatorname{supp} f \subset 4B_j$*, where $\mathscr{M}_{\mathrm{loc}}$ is defined by* (3.22).

Proof. Viewing the left-hand side of (3.28) as the maximal function relative to $4B_j$, since the radius of $4B_j$ is 4, it is enough to consider balls B of radii not exceeding 8. Let $B = B(x_0, r)$ be an arbitrary ball containing x and satisfying $x_0 \in 4B_j$ and $r \in (0, 8)$. By Lemma 3.1, we know that

$$\mu(B) \lesssim \mu(B \cap 4B_j) \tag{3.29}$$

with an implicit constant independent of B and j. Hence, it is easy to see

$$\frac{1}{\mu(B \cap 4B_j)} \int_{B\cap 4B_j} \left| \widetilde{T}(I - P_{r^2})^n f \right|^2 d\mu \lesssim \frac{1}{\mu(B)} \int_B \left| \widetilde{T}(I - P_{r^2})^n f \right|^2 d\mu.$$

Thus, we only need to show that

$$\sup_{B \text{ ball in } M,\, B\ni x} \left(\frac{1}{\mu(B)} \int_B \left| \widetilde{T}(I - P_{r^2})^n f(y) \right|^2 \mu(dy) \right)^{1/2} \lesssim \left(\mathscr{M}_{\mathrm{loc}}(|f|^2)(x) \right)^{1/2}, \tag{3.30}$$

for any $x \in 4B_j$. Since $r < 8$, we choose $i_r \in \mathbb{Z}_+$ satisfying

$$2^{i_r} r \leq 8 < 2^{i_r+1} r. \tag{3.31}$$

Denote by

$$C_i \ \text{ the annulus } 2^{i+1}B \setminus (2^i B) \ \text{ if } i \geq 2 \text{ and } C_1 = 4B. \tag{3.32}$$

Using the fact $\operatorname{supp} f \subset 4B_j \subset 2^i B$ when $i > i_r$, we find that

$$f = \sum_{i=1}^{i_r} f \mathbb{1}_{C_i} =: \sum_{i=1}^{i_r} f_i$$

which then implies

$$\left\| |\widetilde{T}(I - P_{r^2})^n f| \right\|_{L^2(B)} \leq \sum_{i=1}^{i_r} \left\| |\widetilde{T}(I - P_{r^2})^n f_i| \right\|_{L^2(B)}. \tag{3.33}$$

For $i = 1$ we use the L^2-boundedness of $\widetilde{T}(I - P_{r^2})^n$ to obtain

$$\left\| |\widetilde{T}(I - P_{r^2})^n f_1| \right\|_{L^2(B)} \leq \|f\|_{L^2(4B)} \leq \mu(4B)^{1/2} (\mathscr{M}_{\mathrm{loc}}(|f|^2)(x))^{1/2} \tag{3.34}$$

as desired. For $i \geq 2$, we infer from (3.17) that

$$\begin{aligned}
\widetilde{T}(I - P_{r^2})^n f_i &= \int_0^\infty v(t)\,\mathrm{Hess}(P_t(I - P_{r^2})^n f_i)\,dt \\
&= \int_0^\infty v(t) \sum_{k=0}^n \binom{n}{k} (-1)^k\, \mathrm{Hess}\ P_{t+kr^2}(f_i)\,dt \\
&= \int_0^\infty \left(\sum_{k=0}^n \binom{n}{k} (-1)^k \mathbf{1}_{\{t>kr^2\}} v(t-kr^2) \right) \mathrm{Hess}\ P_t(f_i)\,dt \\
&=: \int_0^\infty g_r(t)\ \mathrm{Hess}\, P_t f_i\,dt.
\end{aligned}$$

For g_r, according to the definition $v(t) = w(t)\mathrm{e}^{-\sigma t}$ and an elementary calculation, we observe that

$$\begin{cases}
|g_r(t)| \lesssim 1 & \text{for } 0 < t \leq (1+nr^2) \wedge (1+n)r^2, \\
|g_r(t)| \lesssim r^{2n} & \text{for } (1+nr^2) \wedge (1+n)r^2 < t \leq 1 + nr^2, \\
g_r(t) = 0 & \text{for } t > 1 + nr^2,
\end{cases}$$

where the second estimate comes from $w^{(i)}(t) \lesssim 1$ for $t \in [0, 1]$ and $i \in \{1, \ldots, n\}$, along with

$$\left| \sum_{k=0}^n \binom{n}{k} (-1)^k w(t - kr^2) \right| \leq C_n \sup_{u \geq \frac{t}{n+1}} \left| (w(u)\,\mathrm{e}^{-\sigma u})^{(n)} \right| r^{2n} \lesssim r^{2n}$$

for all $t \geq (1+n)r^2$ and some constant C_n. Combined with Lemma 2.3, this gives

$$\left\| \left| \widetilde{T}(I - P_{r^2})^n f_i \right| \right\|_{L^2(B)} \lesssim \left(\int_0^\infty |g_r(t)|(1 + \sqrt{t})\,\mathrm{e}^{-\alpha' 4^i r^2/t} \frac{dt}{t} \right) \|f_i\|_{L^2(C_i)}$$

where using the fact that $0 < r < 8$, we know

$$\begin{aligned}
&\int_0^\infty (1 + \sqrt{t})|g_r(t)|\,\mathrm{e}^{-\frac{\alpha' 4^i r^2}{t}} \frac{dt}{t} \\
&\lesssim \int_0^{(1+nr^2)\wedge(1+n)r^2} (1 + \sqrt{t})\,\mathrm{e}^{-\frac{\alpha' 4^i r^2}{t}} \frac{dt}{t} + \int_{(1+nr^2)\wedge(1+n)r^2}^{1+nr^2} (1 + \sqrt{t}) r^{2n}\,\mathrm{e}^{-\frac{\alpha' 4^i r^2}{t}} \frac{dt}{t} \\
&\leq C_n \left(\int_0^{(n+1)r^2} \mathrm{e}^{-\frac{\alpha' 4^i r^2}{t}} \frac{dt}{t} + \int_{(1+nr^2)\wedge(1+n)r^2}^{1+nr^2} (1 + \sqrt{t}) r^{2n} \frac{t^{n-1}}{4^{in} r^{2n}}\,dt \right) \\
&\leq C_n \left(4^{-in} + 4^{-in} r^{2n} (1 + \sqrt{r}) \right) \leq C_n' 4^{-in}.
\end{aligned}$$

Now an easy consequence of the local doubling (3.27), since $r(2^i B) \leq 8$ when $1 \leq i \leq i_r$, is that

$$\mu(2^{i+1}B) \leq C 2^{(i+1)D_L} \mu(B),$$

with constants C and D_L independent of B and i. Therefore, as $C_i \subset 2^{i+1}B$,

$$\|f\|_{L^2(C_i)} \leq \mu(2^{i+1}B)^{1/2} (\mathscr{M}_{\mathrm{loc}}(|f|^2)(x))^{1/2} \leq C 2^{iD_L/2} \mu(B)^{1/2} (\mathscr{M}_{\mathrm{loc}}(|f|^2)(x))^{1/2}.$$

Using the definition of i_r, $r \leq 8$, and by choosing $2n > D_L/2$, we finally obtain

$$\left\| |\widetilde{T}(I - P_{r^2})^n f| \right\|_{L^2(B)} \leq C' \left(\sum_{i=1}^{i_r} 2^{i(D_L/2 - 2n)} \right) \mu(B)^{1/2} (\mathscr{M}_{\mathrm{loc}}(|f|^2)(x))^{1/2},$$

which proves Proposition 2.7. □

The following lemma is essential to the proof of part (ii) of Lemma 3.6.

Lemma 3.8. *Let $p \in (2,\infty)$. Assume that* (**H**) *holds. For a ball B with radius $r \in (0,8)$, let $i = i_r$ be an integer such that* (3.31) *holds. Then the following estimate hold: For any C^2-function f supported in C_i as in* (3.32)*, and each $k \in \{1,\dots,n\}$, where $n \in \mathbb{N}$ is chosen according to Lemma* 3.7*, one has*

$$\begin{aligned}\left(\frac{1}{\mu(B)}\int_B |\operatorname{Hess} P_{kr^2} f|^p\,d\mu\right)^{1/p} \qquad\qquad (3.35)\\ \leq C\mathrm{e}^{-\alpha_1 4^i}\left[\left(\frac{1}{\mu(2^{i+1}B)}\int_{C_i}|\nabla f|^2\,d\mu\right)^{1/2} + \left(\frac{1}{\mu(2^{i+1}B)}\int_{C_i}|\operatorname{Hess} f|^2\,d\mu\right)^{1/2}\right]\end{aligned}$$

and

$$\left(\frac{1}{\mu(B)}\int_B |\operatorname{Hess} P_{kr^2} f|^p\,d\mu\right)^{1/p} \leq \frac{C\mathrm{e}^{-\alpha_2 4^i}}{r^2}\left(\frac{1}{\mu(2^{i+1}B)}\int_{C_i}|f|^2\,d\mu\right)^{1/2} \qquad (3.36)$$

for some positive constants C and α_1, α_2 depending on K,θ, p and the constants in (**LD**).

Proof. We first observe that Lemma 2.8 yields

$$\begin{aligned}\left(\int_B |\operatorname{Hess} P_t f|^p\,d\mu\right)^{1/p} &\leq \mathrm{e}^{2Kt}\left(\int_B \left(P_t|\operatorname{Hess} f|^2\right)^{p/2}(x)\,\mu(dx)\right)^{1/p}\\ &\quad + C\mathrm{e}^{(2K+\theta)t}\left(\int_B \left(P_t|df|^2\right)^{p/2}(x)\,\mu(dx)\right)^{1/p}. \qquad (3.37)\end{aligned}$$

We substitute $t = kr^2$ in estimate (3.37). By Lemma 2.1, one has the upper bound of $p_t(x,y)$,

$$p_t(x,y) \leq \frac{C_\alpha}{V(y,\sqrt{t})}\exp\left(-\alpha\frac{\rho^2(x,y)}{t} + C_1Kt\right)$$

for all $x,y \in M$. Since $r \leq 8$, it follows from the above estimate that for all $x \in B$,

$$\begin{aligned}\left|P_{kr^2}(|df|^2)(x)\right| &\leq C\int_{C_i} V(y,\sqrt{k}r)^{-1}\exp\left(-\alpha\frac{\rho^2(x,y)}{kr^2} + C_1Kkr^2\right)|df|^2(y)\,\mu(dy)\\ &\leq C\mathrm{e}^{-\alpha 4^i/k}\int_{C_i} V(y,\sqrt{k}r)^{-1}|df|^2(y)\,\mu(dy).\end{aligned}$$

Moreover, since $y \in C_i$, we have $2^{i+1}B \subset B(y, 2^{i+2}r)$, and then by (**LD**), we know that

$$\frac{1}{V(y,\sqrt{k}r)} \leq \frac{2^{d(i+2)}\mathrm{e}^{C2^{i+2}}}{V(y,2^{i+2}r)} \leq \frac{2^{d(i+2)}\mathrm{e}^{C2^{i+2}}}{\mu(2^{i+1}B)}.$$

It then follows that

$$|P_{kr^2}|df|^2(x)| \leq C\mathrm{e}^{-c4^i}2^{d(i+2)}\mathrm{e}^{C2^{i+2}}\left(\frac{1}{\mu(2^{i+1}B)}\int_{C_i}|\nabla f|^2\,d\mu\right)$$

for all $x \in B$, and there exists $\alpha_1 < c$ such that

$$\left(\frac{1}{\mu(B)}\int_B (P_{kr^2}|\nabla f|^2)^{p/2}\,d\mu\right)^{1/p} \leq C\mathrm{e}^{-\alpha_1 4^i}\left(\frac{1}{\mu(2^{i+1}B)}\int_{C_i}|\nabla f|^2\,d\mu\right)^{1/2}. \qquad (3.38)$$

With similar arguments, we obtain

$$\left(\frac{1}{\mu(B)}\int_B (P_{kr^2}|\operatorname{Hess} f|^2)^{p/2}\,d\mu\right)^{1/p} \le C\mathrm{e}^{-\alpha_1 4^i}\left(\frac{1}{\mu(2^{i+1}B)}\int_{C_i}|\operatorname{Hess} f|^2\,d\mu\right)^{1/2}.$$

Altogether these yields

$$\begin{aligned}&\left(\frac{1}{\mu(B)}\int_B |\operatorname{Hess} P_{kr^2} f|^p\,d\mu\right)^{1/p}\\ &\le C\mathrm{e}^{-\alpha_1 4^i}\left(\left(\frac{1}{\mu(2^{i+1}B)}\int_{C_i}|\nabla f|^2\,d\mu\right)^{1/2}+\left(\frac{1}{\mu(2^{i+1}B)}\int_{C_i}|\operatorname{Hess} f|^2\,d\mu\right)^{1/2}\right),\end{aligned}$$

which completes the proof of (3.35).

Next we observe from Proposition 2.7 that

$$\left(\int_B |\operatorname{Hess} P_t f|^p\,d\mu\right)^{1/p} \le \frac{(1+\sqrt{t})}{t}\mathrm{e}^{(2K+\theta)t}\left(\int_B \left(P_t|f|^2\right)^{p/2}(x)\,\mu(dx)\right)^{1/p}. \tag{3.39}$$

Substituting $t = kr^2$ in (3.39), as $r \in (0, 8)$, there exists a constant C such that

$$\left(\int_B |\operatorname{Hess} P_{kr^2} f|^p\,d\mu\right)^{1/p} \le \frac{C}{r^2}\left(\int_B \left(P_{kr^2}|f|^2\right)^{p/2}(x)\,\mu(dx)\right)^{1/p}.$$

With similar arguments as for (3.38), we then complete the proof of (3.36). □

With the help of the Lemmas 3.6, 3.7 and 3.8, we are now in position to turn to the proof of Theorem 1.2.

Proof of Theorem 1.2. By Lemma 3.6, we only need to show (i) and (ii) of Lemma 3.6 hold true under our condition (**H**). We first verify (i) of Lemma 3.6. Observe from Lemma 3.7 that there is an integer n depending only on D_L as in (3.27) such that for all $f \in L^p(B_j)$ and $x \in 4B_j$,

$$\sup_{B \text{ ball in } M, B\ni x}\frac{1}{\mu(B\cap 4B_j)}\int_{B\cap 4B_j}|\widetilde{T}(I-P_{r^2})^n f|^2\,d\mu \le C\mathscr{M}_{\mathrm{loc}}(|f|^2)(x).$$

Recall that $\mathscr{M}_{\mathrm{loc}}$ is bounded on $L^p(M)$ for $1 < p \le \infty$; thus for all $p > 2$, the operator $\mathscr{M}^{\#}_{4B_j,\tilde{T},n}$ is bounded from $L^p(B_j,\mu)$ to $L^p(4B_j,\mu)$ uniformly in j, i.e. assertion (i) is proved.

Next, we prove (ii) of Lemma 3.6. Assume that $f \in L^2(B_j)$ and let $h = \int_0^\infty v(t)P_t f\,dt$ with v defined as in (3.17). Since $\widetilde{T} f = \operatorname{Hess} h$ and the inequality (3.29) for $B \cap 4B_j$, we have

$$\begin{aligned}\left(\frac{1}{\mu(B\cap 4B_j)}\int_{B\cap 4B_j}|\widetilde{T}P_{kr^2} f|^p\,d\mu\right)^{1/p} &= \left(\frac{1}{\mu(B\cap 4B_j)}\int_{B\cap 4B_j}|\operatorname{Hess} P_{kr^2} h|^p\,d\mu\right)^{1/p}\\ &\lesssim \left(\frac{1}{\mu(B)}\int_B |\operatorname{Hess} P_{kr^2} h|^p\,d\mu\right)^{1/p}.\end{aligned}$$

We write

$$\operatorname{Hess} P_{kr^2} h = \sum_{i=1}^{\infty}\operatorname{Hess} P_{kr^2}(h\mathbb{1}_{C_i}) = \sum_{i=1}^{\infty}\operatorname{Hess} P_{kr^2} g_i,$$

where $g_i = h\mathbb{1}_{C_i}$ and C_i is as in (3.32). Next, we distinguish the two regimes $i \le i_r$ and $i > i_r$ where i_r is defined as in (3.31). In the regime $i \le i_r$, by the inequality (3.35) in Lemma 3.8, we have

$$\begin{aligned}&\left(\frac{1}{\mu(B)}\int_B |\operatorname{Hess} P_{kr^2} g_i|^p \, d\mu\right)^{1/p}\\ &\le C\mathrm{e}^{-\alpha_1 4^i}\left(\left(\frac{1}{\mu(2^{i+1}B)}\int_{C_i}|\nabla h|^2\, d\mu\right)^{1/2} + \left(\frac{1}{\mu(2^{i+1}B)}\int_{C_i}|\operatorname{Hess} h|^2\, d\mu\right)^{1/2}\right)\\ &\le C\mathrm{e}^{-\alpha_1 4^i}\left(\left(\mathscr{M}_{\mathrm{loc}}(|\nabla h|^2)(x)\right)^{1/2} + \left(\frac{1}{\mu(2^{i+1}B)}\int_{C_i}|\operatorname{Hess} h|^2\, d\mu\right)^{1/2}\right).\end{aligned}$$

On the other hand, since $C_i \subset 2^{i+1}B$, we know

$$\begin{aligned}&\frac{1}{\mu(2^{i+1}B)}\int_{C_i}\left|\operatorname{Hess} h\right|^2 d\mu\\ &\le \frac{1}{\mu(2^{i+1}B\cap 4B_j)}\int_{(2^{i+1}B)\cap 4B_j}|\operatorname{Hess} h|^2 d\mu + \frac{1}{\mu(2^{i+1}B)}\int_{2^{i+1}B}\mathbb{1}_{M\setminus 4B_j}|\operatorname{Hess} h|^2\, d\mu\\ &\le \mathscr{M}_{4B_j}(|\operatorname{Hess} h|^2)(x) + \mathscr{M}_{\mathrm{loc}}(|\operatorname{Hess} h|^2\mathbb{1}_{M\setminus 4B_j})(x),\end{aligned}$$

for any $x \in B\cap 4B_j$. Hence in this case

$$\begin{aligned}&\left(\frac{1}{\mu(B)}\int_B \Big|\sum_{i=1}^{i_r}\operatorname{Hess} P_{kr^2} g_i\Big|^p \, d\mu\right)^{1/p} \qquad (3.40)\\ &\le \sum_{i=1}^{i_r} C\mathrm{e}^{-\alpha 4^i}\Big(\mathscr{M}_{\mathrm{loc}}(|\nabla h|^2) + \mathscr{M}_{4B_j}(|\operatorname{Hess} h|^2) + \mathscr{M}_{\mathrm{loc}}(|\operatorname{Hess} h|^2\mathbb{1}_{M\setminus 4B_j})\Big)^{1/2}(x).\end{aligned}$$

For the second regime $i > i_r$, we proceed with inequality (3.35) in Lemma 3.8 that there exist positive constants c_1 and c_2 such that

$$\left(\frac{1}{\mu(B)}\int_B |\operatorname{Hess} P_{kr^2} g_i|^p \, d\mu\right)^{1/p} \le \frac{c_1\mathrm{e}^{-c_2 4^i}}{r^2}\left(\frac{1}{\mu(2^{i+1}B)}\int_{C_i}|h|^2\, d\mu\right)^{1/2}. \qquad (3.41)$$

On the other hand, since $i > i_r$, it is easy to see $4B_j \subset 2^{i+1}B$, thus

$$\begin{aligned}\Big(\frac{1}{\mu(2^{i+1}B)}\int_{C_i}|h|^2\, d\mu\Big)^{1/2} &\le \left(\frac{1}{\mu(2^{i+1}B)}\int_0^1 v(t)\int_{C_i} P_t|f|^2\, d\mu\, dt\right)^{1/2} \qquad (3.42)\\ &\le C\left(\frac{1}{\mu(4B_j)}\int_{B_j}|f|^2\, d\mu\right)^{1/2} = C\left(\frac{1}{\mu(4B_j)}\int_{B_j}|f|^2\, d\mu\right)^{1/2}\\ &\le C(\mathscr{M}_{4B_j}(|f|^2)(x))^{1/2}.\end{aligned}$$

The contribution of the terms in the second regime $i > i_r$ is bounded by combining (3.41) and (3.42),

$$\sum_{i>i_r}\left(\frac{1}{\mu(B)}\int_B |\operatorname{Hess} P_{kr^2} g_i|^p \, d\mu\right)^{1/p} \le \sum_{i>i_r}\frac{c_1\mathrm{e}^{-c_2 4^i}}{r^2}(\mathscr{M}_{4B_j}(|f|^2)(x))^{1/2} \qquad (3.43)$$

and it remains to recall that $1/r^2 \le 4^i/8$ and $1/r \le 2^i/(2\sqrt{2})$ when $i > i_r$.

We then conclude from (3.40) and (3.43) that for any $p > 2$ and $k \in \{1, 2, \ldots, n\}$, there exists a constant C independent of j such that

$$\left(\frac{1}{\mu(B \cap 4B_j)} \int_{B\cap 4B_j} |\widetilde{T} P_{kr^2} f|^p \, d\mu\right)^{1/p} \le C(\mathscr{M}_{4B_j}(|\widetilde{T} f|^2) + (S_j f)^2)^{1/2}(x)$$

all $f \in L^2(M,\mu)$ supported in B_j, all balls B in M and all $x \in B \cap 4B_j$, where the radius r of B is less than 8, and where

$$(S_j f)^2 := \mathscr{M}_{\rm loc}(|\widetilde{T} f|^2 \mathbb{1}_{M\setminus 4B_j}) + \mathscr{M}_{\rm loc}(|\nabla h|^2) + \mathscr{M}_{4B_j}(|f|^2). \tag{3.44}$$

Our last step is to show that the operator S_j defined in (3.44) is bounded from $L^p(B_j)$ to $L^p(4B_j)$ for any $p \in (2, \infty)$ with operator norm independent of j. By (3.44), we only need to show that the operators

$$(\mathscr{M}_{\rm loc}(|\widetilde{T} f|^2 \mathbb{1}_{M\setminus 4B_j}))^{1/2}, \ \ (\mathscr{M}_{\rm loc}(|\nabla h|^2))^{1/2} \ \text{ and } \ (\mathscr{M}_{4B_j}(|f|^2))^{1/2}$$

are respectively bounded from $L^p(B_j)$ to $L^p(4B_j)$. Indeed, for any $f \in L^p(4B_j)$, by Lemma 3.1 we know that $4B_j$ satisfies the doubling property (**D**), which combined with $p > 2$ implies that $(\mathscr{M}_{4B_j}(|f|^2))^{1/2}$ is bounded from $L^p(B_j)$ to $L^p(4B_j)$ by a constant depending only on the doubling property (**D**). On the other hand, using the fact from Bismut's formula [5] that

$$\left\| \sqrt{t} \nabla P_t f \right\|_p \lesssim \mathrm{e}^{Kt} \|f\|_p,$$

we deduce from $p > 2$ and $\sigma > 2K + \theta$ that

$$\left\| \left(\mathscr{M}_{\rm loc}(|\nabla h|^2)\right)^{1/2} \right\|_p \lesssim \|\nabla h\|_p \le \left(\int_0^\infty \frac{v(t)\,\mathrm{e}^{Kt}}{\sqrt{t}} \, dt\right) \|f\|_{L^p(B_j)} \le C\|f\|_{L^p(B_j)}.$$

Finally, the L^p-boundedness of

$$(\mathscr{M}_{\rm loc}(|\widetilde{T} f|^2 \mathbb{1}_{M\setminus 4B_j}))^{1/2}$$

follows from the $L^{p/2}$-boundedness of $(\mathscr{M}_{\rm loc}(|\widetilde{T} f|^2)^{1/2}$ and an argument similar to (3.21). This implies that the operator S_j is bounded from $L^p(B_j)$ to $L^p(4B_j)$ with an upper bound independent of j.

We conclude that the requirements (i) and (ii) in Lemma 3.6 both hold under the condition (**H**). Thus, the operator $\widetilde{T}$ is bounded from $L^p(B_j, \mu)$ to $L^p(4B_j, \mu)$ for $p > 2$ with a constant independent of j. Therefore, by Lemma 3.5, the operator T is strong type (p, p) for $p > 2$. This proves Theorem 1.2. □

4 Appendix

In this appendix, we include the proof of **(CZ)**(2) in the case that the underlying manifold M has a lower bound Ricci curvature for reader's convenience (see [26, 36]). To be precise, let M be a complete Riemannian manifold satisfying the curvature condition (**Ric**). For any $u \in C_0^\infty(M)$, Bochner's formula gives

$$-\frac{1}{2}\Delta|\nabla u|^2 = |\operatorname{Hess} u|_{\rm HS}^2 - g(\nabla\Delta u, \nabla u) + \operatorname{Ric}(\nabla u, \nabla u). \tag{4.1}$$

Integration by parts, together with the lower Ricci bound, leads to

$$\begin{aligned}0 &= -\frac{1}{2}\int_M \Delta|\nabla u|^2\\ &= \int_M |\operatorname{Hess} u|_{\mathrm{HS}}^2 - \int_M g(\nabla\Delta u, \nabla u) + \int_M \operatorname{Ric}(\nabla u, \nabla u)\\ &\geq \int_M |\operatorname{Hess} u|_{\mathrm{HS}}^2 - \int_M (\Delta u)^2 - K\int_M |\nabla u|^2\\ &= \int_M |\operatorname{Hess} u|_{\mathrm{HS}}^2 - \int_M (\Delta u)^2 - K\int_M u\Delta u.\end{aligned}$$

Then, by estimating the last integral via Young's inequality, we obtain for any $\sigma > 0$,

$$\| |\operatorname{Hess} u| \|_{L^2}^2 \leq \| |\operatorname{Hess} u|_{\mathrm{HS}} \|_{L^2}^2 \leq \frac{K\sigma^2}{2}\|u\|_{L^2}^2 + \left(1 + \frac{K}{2\sigma^2}\right)\|\Delta u\|_{L^2}^2, \tag{4.2}$$

which establishes **CZ(2)**.

Remark 4.1. Inequality (4.2) extends from $u \in C_0^\infty(M)$ to $u \in H^{2,2}(M)$. Thus, in particular, if $u \in L^2(M)$ is a distributional solution to the Poisson equation

$$\Delta u = f \quad \text{on } M \tag{4.3}$$

for some $f \in L^2(M)$, then (4.2) provides an L^2-Hessian estimate of the solution u.

Remark 4.2. Recall that for L^2-gradient estimates of u in (4.3), the lower bound on the Ricci tensor is not needed. Indeed, integrating (4.3) by parts and using Young inequality, one obtains directly for every $\sigma > 0$,

$$\int_M |\nabla u|^2 = \int_M fu \leq \frac{\sigma^2}{2}\int_M u^2 + \frac{1}{2\sigma^2}\int_M f^2,$$

which is the L^2-gradient estimate

$$\|\nabla u\|_{L^2}^2 \leq \frac{\sigma^2}{2}\|u\|_{L^2}^2 + \frac{1}{2\sigma^2}\|f\|_{L^2}^2. \tag{4.4}$$

References

[1] Michael Aizenman and Barry Simon, *Brownian motion and Harnack inequality for Schrödinger operators*, Comm. Pure Appl. Math. **35** (1982), no. 2, 209–273. MR 644024

[2] Marc Arnaudon, Holger Plank, and Anton Thalmaier, *A Bismut type formula for the Hessian of heat semigroups*, C. R. Math. Acad. Sci. Paris **336** (2003), no. 8, 661–666. MR 1988128

[3] Pascal Auscher, Thierry Coulhon, Xuan Thinh Duong, and Steve Hofmann, *Riesz transform on manifolds and heat kernel regularity*, Ann. Sci. École Norm. Sup. (4) **37** (2004), no. 6, 911–957. MR 2119242

[4] Robert Baumgarth, Baptiste Devyver, and Batu Güneysu, *Estimates for the covariant derivative of the heat semigroup on differential forms, and covariant Riesz transforms*, arXiv:2107.00311 (2021).

[5] Jean-Michel Bismut, *Large deviations and the Malliavin calculus*, Progress in Mathematics, vol. 45, Birkhäuser Boston, Inc., Boston, MA, 1984. MR 755001

[6] Alberto P. Caldéron and Antoni S. Zygmund, *On the existence of certain singular integrals*, Acta Math. **88** (1952), 85–139. MR 52553

[7] Gilles Carron, *Riesz transform on manifolds with quadratic curvature decay*, Rev. Mat. Iberoam. **33** (2017), no. 3, 749–788. MR 3713030

[8] Isaac Chavel, *Eigenvalues in Riemannian geometry*, Pure and Applied Mathematics, vol. 115, Academic Press, Inc., Orlando, FL, 1984, Including a chapter by Burton Randol, With an appendix by Jozef Dodziuk. MR 768584

[9] Xin Chen, Xue-Mei Li, and Bo Wu, *Logarithmic heat kernels: estimates without curvature restrictions*, arXiv:2106.02746 (2021).

[10] Bennett Chow and Richard S. Hamilton, *Constrained and linear Harnack inequalities for parabolic equations*, Invent. Math. **129** (1997), no. 2, 213–238. MR 1465325

[11] Ronald R. Coifman and Guido Weiss, *Analyse harmonique non-commutative sur certains espaces homogènes*, Lecture Notes in Mathematics, Vol. 242, Springer-Verlag, Berlin-New York, 1971, Étude de certaines intégrales singulières. MR 0499948

[12] Thierry Coulhon and Xuan Thinh Duong, *Riesz transforms for* $1 \leq p \leq 2$, Trans. Amer. Math. Soc. **351** (1999), no. 3, 1151–1169. MR 1458299

[13] ———, *Riesz transforms for* $p > 2$, C. R. Acad. Sci. Paris Sér. I Math. **332** (2001), no. 11, 975–980. MR 1838122

[14] ———, *Riesz transform and related inequalities on noncompact Riemannian manifolds*, Comm. Pure Appl. Math. **56** (2003), no. 12, 1728–1751. MR 2001444

[15] E. Brian Davies, *The state of the art for heat kernel bounds on negatively curved manifolds*, Bull. London Math. Soc. **25** (1993), no. 3, 289–292. MR 1209255

[16] ———, *Non-Gaussian aspects of heat kernel behaviour*, J. London Math. Soc. (2) **55** (1997), no. 1, 105–125. MR 1423289

[17] Xuan T. Duong and Derek W. Robinson, *Semigroup kernels, Poisson bounds, and holomorphic functional calculus*, J. Funct. Anal. **142** (1996), no. 1, 89–128. MR 1419418

[18] K. David Elworthy and Xue-Mei Li, *Formulae for the derivatives of heat semigroups*, J. Funct. Anal. **125** (1994), no. 1, 252–286. MR 1297021

[19] ———, *Bismut type formulae for differential forms*, C. R. Acad. Sci. Paris Sér. I Math. **327** (1998), no. 1, 87–92. MR 1650216

[20] David Gilbarg and Neil S. Trudinger, *Elliptic partial differential equations of second order*, Classics in Mathematics, Springer-Verlag, Berlin, 2001, Reprint of the 1998 edition.

[21] Alexander Grigoryan, *Gaussian upper bounds for the heat kernel on arbitrary manifolds*, J. Differential Geom. **45** (1997), no. 1, 33–52. MR 1443330

[22] Batu Güneysu, *Nonrelativistic hydrogen type stability problems on nonparabolic 3-manifolds*, Ann. Henri Poincaré **13** (2012), no. 7, 1557–1573. MR 2982633

[23] ______, *Covariant Schrödinger semigroups on Riemannian manifolds*, Operator Theory: Advances and Applications, vol. 264, Birkhäuser/Springer, Cham, 2017. MR 3751359

[24] ______, *Heat kernels in the context of Kato potentials on arbitrary manifolds*, Potential Anal. **46** (2017), no. 1, 119–134. MR 3595965

[25] Batu Güneysu and Diego Pallara, *Functions with bounded variation on a class of Riemannian manifolds with Ricci curvature unbounded from below*, Math. Ann. **363** (2015), no. 3-4, 1307–1331. MR 3412360

[26] Batu Güneysu and Stefano Pigola, *The Calderón-Zygmund inequality and Sobolev spaces on noncompact Riemannian manifolds*, Adv. Math. **281** (2015), 353–393. MR 3366843

[27] ______, *L^p-interpolation inequalities and global Sobolev regularity results*, Ann. Mat. Pura Appl. (4) **198** (2019), no. 1, 83–96, With an appendix by Ognjen Milatovic. MR 3918620

[28] Batu Güneysu and Olaf Post, *Path integrals and the essential self-adjointness of differential operators on noncompact manifolds*, Math. Z. **275** (2013), no. 1-2, 331–348. MR 3101810

[29] Steve Hofmann and José María Martell, *L^p bounds for Riesz transforms and square roots associated to second order elliptic operators*, Publ. Mat. **47** (2003), no. 2, 497–515. MR 2006497

[30] Shouhei Honda, Luciano Mari, Michele Rimoldi, and Giona Veronelli, *Density and non-density of $C_c^\infty \hookrightarrow W^{k,p}$ on complete manifolds with curvature bounds*, Nonlinear Anal. **211** (2021), Paper No. 112429, 26. MR 4268755

[31] Tosio Kato, *Schrödinger operators with singular potentials*, Israel J. Math. **13** (1972), 135–148 (1973). MR 333833

[32] Xue-Mei Li, *Stochastic differential equations on noncompact manifolds*, University of Warwick, Thesis (1992).

[33] Ludovico Marini and Giona Veronelli, *The L^p-Calderón-Zygmund inequality on non-compact manifolds of positive curvature*, Ann. Global Anal. Geom. **60** (2021), no. 2, 253–267. MR 4291611

[34] Lei Ni, *The Poisson equation and Hermitian-Einstein metrics on holomorphic vector bundles over complete noncompact Kähler manifolds*, Indiana Univ. Math. J. **51** (2002), no. 3, 679–704. MR 1911050

[35] Lei Ni, Yuguang Shi, and Luen-Fai Tam, *Poisson equation, Poincaré-Lelong equation and curvature decay on complete Kähler manifolds*, J. Differential Geom. **57** (2001), no. 2, 339–388. MR 1879230

[36] Stefano Pigola, *Global Calderón-Zygmund inequalities on complete Riemannian manifolds*, arXiv:2011.03220v1 (2020).

[37] Christian Rose and Peter Stollmann, *Manifolds with Ricci curvature in the Kato class: Heat kernel bounds and applications*, Analysis and Geometry on Graphs and Manifolds, London Mathematical Society Lecture Note Series, vol. 461, Cambridge University Press, 2020, pp. 76–94.

[38] Barry Simon, *Schrödinger semigroups*, Bull. Amer. Math. Soc. (N.S.) **7** (1982), no. 3, 447–526. MR 670130

[39] Daniel W. Stroock, *An estimate on the Hessian of the heat kernel*, Itô's stochastic calculus and probability theory, Springer, Tokyo, 1996, pp. 355–371. MR 1439536

[40] Anton Thalmaier and Feng-Yu Wang, *Derivative estimates of semigroups and Riesz transforms on vector bundles*, Potential Anal. **20** (2004), no. 2, 105–123. MR 2032944

[41] Caitlin Wang, *The Calderón-Zygmund inequality on a compact Riemannian manifold*, Pacific J. Math. **217** (2004), no. 1, 181–200. MR 2105773

[42] Feng-Yu Wang, *Analysis for diffusion processes on Riemannian manifolds*, Advanced Series on Statistical Science & Applied Probability, vol. 18, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2014. MR 3154951