Some inequalities on Riemannian manifolds linking Entropy, Fisher information, Stein discrepancy and Wasserstein distance

Li-Juan Cheng[1], Anton Thalmaier[2], Feng-Yu Wang[3,4]

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

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

[3]Center for Applied Mathematics, Tianjin University,
Tianjin 300072, People's Republic of China
[4]Department of Mathematics, Swansea University, Bay Campus,
Swansea SA1 8EN, United Kingdom
wangfy@tju.edu.cn

August 29, 2022

Abstract

For a complete connected Riemannian manifold M let $V \in C^2(M)$ be such that $\mu(\mathrm{d}x) = \mathrm{e}^{-V(x)}\,\mathrm{vol}(\mathrm{d}x)$ is a probability measure on M. Taking μ as reference measure, we derive inequalities for probability measures on M linking relative entropy, Fisher information, Stein discrepancy and Wasserstein distance. These inequalities strengthen in particular the famous log-Sobolev and transportation-cost inequality and extend the so-called Entropy/Stein-discrepancy/Information (HSI) inequality established by Ledoux, Nourdin and Peccati (2015) for the standard Gaussian measure on Euclidean space to the setting of Riemannian manifolds.

1 Introduction

Let $\gamma(\mathrm{d}x) = (2\pi)^{-n/2}\mathrm{e}^{-|x|^2/2}\,\mathrm{d}x$ be the standard Gaussian measure on $\mathbb{R}^n$ and let $\mathscr{P}(\mathbb{R}^n)$ denote the set of probability measures on $\mathbb{R}^n$. The classical log-Sobolev inequality [9] indicates that

$$H(\nu\,|\,\gamma) \leq \frac{1}{2} I(\nu\,|\,\gamma), \quad \nu \in \mathscr{P}(\mathbb{R}^n), \tag{1.1}$$

2010 *Mathematics Subject Classification*. Primary: 60E15; Secondary: 35K08, 46E35, 42B35.
Key words and phrases. Relative entropy, Fisher information, Stein discrepancy, Wasserstein distance.

and the transportation-cost inequality [18] states that

$$\mathbb{W}_2(\nu,\gamma)^2 \le 2H(\nu\,|\,\gamma), \quad \nu \in \mathscr{P}(\mathbb{R}^n), \tag{1.2}$$

where for $\nu,\mu \in \mathscr{P}(\mathbb{R}^n)$ we have

1. the relative entropy of ν with respect to μ,

$$H(\nu\,|\,\mu) := \begin{cases} \int_{\mathbb{R}^n} h\log h\,\mathrm{d}\mu, & \text{if } \nu(\mathrm{d}x) = h(x)\mu(\mathrm{d}x),\\ \infty, & \text{otherwise,}\end{cases} \tag{1.3}$$

2. the Fisher information of ν with respect to μ

$$I(\nu\,|\,\mu) := \begin{cases} \int_{\mathbb{R}^n} \frac{|\nabla h|^2}{h}\,\mathrm{d}\mu, & \text{if } \nu(\mathrm{d}x) = h(x)\mu(\mathrm{d}x),\ \sqrt{h}\in W^{1,2}(\mu),\\ \infty, & \text{otherwise,}\end{cases} \tag{1.4}$$

3. the L^2-Wasserstein distance $\mathbb{W}_2$ of μ and ν, i.e.

$$\mathbb{W}_2(\mu,\nu) := \inf_{\pi\in\mathscr{C}(\mu,\nu)} \left(\int_{\mathbb{R}^n\times\mathbb{R}^n} |x-y|^2\,\pi(\mathrm{d}x,\mathrm{d}y)\right)^{1/2} \tag{1.5}$$

with $\mathscr{C}(\mu,\nu)$ being the set of all couplings of μ and ν.

Inspired by [14], Ledoux, Nourdin and Peccati [10] established some new type of inequalities improving (1.1) and (1.2) by adopting the Stein discrepancy $S(\nu\,|\,\gamma)$ of ν with respect to γ as further ingredient. This quantity is defined as

$$S(\nu\,|\,\gamma) := \inf_{\tau\in\mathbb{S}_\nu}\left(\int_{\mathbb{R}^n} |\tau-\mathrm{id}|_{\mathrm{HS}}^2\,\mathrm{d}\nu\right)^{1/2} \tag{1.6}$$

where id is the $n\times n$-identity matrix and $\mathbb{S}_\nu$ the set of measurable maps $\tau \in L^1_{\mathrm{loc}}(\mathbb{R}^n\to\mathbb{R}^n\otimes\mathbb{R}^n;\nu)$ such that

$$\int_{\mathbb{R}^n} x\cdot\nabla\varphi\,\mathrm{d}\nu = \int_{\mathbb{R}^n}\langle\tau,\mathrm{Hess}_\varphi\rangle_{\mathrm{HS}}\,\mathrm{d}\nu, \quad \varphi\in C_0^\infty(\mathbb{R}^n).$$

A map $\tau\in\mathbb{S}_\nu$ is called a Stein kernel of ν. In general, the set $\mathbb{S}_\nu$ may contain infinitely many maps; for instance, for the Gaussian measure γ,

$$\left\{x\mapsto(1+r\mathrm{e}^{|x|^2/2})\,\mathrm{id}\colon\, r\in\mathbb{R}\right\}\subset\mathbb{S}_\gamma.$$

Recall that however the Gaussian measure γ is characterized as the only probability distribution on $\mathbb{R}^n$ satisfying

$$\int_{\mathbb{R}^n} x\cdot\nabla\varphi\,\mathrm{d}\gamma = \int_{\mathbb{R}^n}\Delta\varphi\,\mathrm{d}\gamma, \quad \varphi\in C_0^\infty(\mathbb{R}^n).$$

Hence for $\nu\in\mathscr{P}(\mathbb{R}^n)$ it holds that $\mathrm{id}\in\mathbb{S}_\nu$ if and only if $\nu=\gamma$.

This equivalence shows that the Stein discrepancy $S(\nu\,|\,\gamma)$ with respect to the Gaussian distribution γ provides a natural measure for the proximity of ν to γ and allows to quantify how far ν is away from γ. It is a crucial quantity for normal approximations and appears implicitly in many works on Stein's method [17]. The Stein method was initially developed to quantify the rate of convergence in the Central Limit Theorem [16], which was developed to probability distributions on Riemannian manifolds [21]. For Gamma approximations the Stein discrepancy represents the bound one customarily obtains when applying Stein's method to measure distance to the one-dimensional Gamma distribution, see [3, 5, 10, 13].

Recall that the relative entropy $H(\nu\,|\,\gamma)$ is another measure of the proximity between ν and γ (note that $H(\nu\,|\,\gamma) \geq 0$ and $H(\nu\,|\,\gamma) = 0$ if and only if $\nu = \gamma$) which is moreover stronger than the total variation distance, $2\mathrm{TV}(\nu,\gamma)^2 \leq H(\nu\,|\,\gamma)$, see [22, 10].

Considering the Stein discrepancy $S(\nu\,|\,\gamma)$ as a new ingredient, according to [10, Theorem 2.2], one has the following HSI inequality which strengthens (1.1):

$$H(\nu\,|\,\gamma) \leq \frac{1}{2} S^2(\nu\,|\,\gamma) \log\Big(1 + \frac{I(\nu\,|\,\gamma)}{S^2(\nu\,|\,\gamma)}\Big), \quad \nu \in \mathscr{P}(\mathbb{R}^n), \tag{1.7}$$

$$\mathbb{W}_2(\nu\,|\,\gamma) \leq S(\nu\,|\,\gamma) \arccos\left(\exp\left(-\frac{H(\nu\,|\,\gamma)}{S^2(\nu\,|\,\gamma)}\right)\right), \quad \nu \in \mathscr{P}(\mathbb{R}^n), \tag{1.8}$$

improves the transportation-cost inequality (1.2). Moreover, [10, Theorem 2.8] gives the existence of a constant $C > 0$ whereas the inequality [10, Theorem 3.2], such that

$$\left(\int |f|^p \,\mathrm{d}\nu\right)^{1/p} \leq C\left(S_p(\nu\,|\,\gamma) + \sqrt{p}\Big(\int |\tau|_{\mathrm{op}}^{p/2} \mathrm{d}\nu\Big)^{1/p}\right), \quad \nu(f) = 0,\ |\nabla f| \leq 1,\ \tau \in \mathbb{S}_\nu, \tag{1.9}$$

where for $p \geq 1$,

$$S_p(\nu\,|\,\gamma) := \inf_{\tau \in \mathbb{S}_\nu} \left(\int_{\mathbb{R}^n} |\tau - \mathrm{id}|_{\mathrm{HS}}^p \,\mathrm{d}\nu\right)^{1/p}. \tag{1.10}$$

In particular $S_2(\nu\,|\,\gamma)$ is the Stein discrepancy as defined above.

These inequalities have been extended in [10] to probability measures $\mu(\mathrm{d}x) := \mathrm{e}^{V(x)}\mathrm{d}x$ on $\mathbb{R}^n$ which are stationary distributions of an elliptic symmetric diffusion process on $\mathbb{R}^n$ satisfying some conditions on the Bakry-Émery Γ_i operators ($i = 1, 2, 3$).

The aim of this paper is to put forward this framework and to investigate inequalities of the type (1.7), (1.8) and (1.9) on general Riemannian manifolds. Our results on Riemannian manifolds include the above inequalities as special cases.

We start with some basic notations. Let M be a complete connected Riemannian manifold equipped with the probability measure

$$\mu(\mathrm{d}x) = \mathrm{e}^{-V(x)}\mathrm{vol}(\mathrm{d}x)$$

for some $V \in C^2(M)$, where $\mathrm{vol}(\mathrm{d}x)$ denotes the Riemannian volume measure. As well known, the diffusion semigroup $P_t = \mathrm{e}^{\frac{1}{2}tL}$ generated by $L := \Delta + \nabla V$ is symmetric on $L^2(\mu)$. Denote by $\mathrm{Ric}_V := \mathrm{Ric} + \mathrm{Hess}_V$ the Bakry-Émery curvature tensor.

Let $H(\nu\,|\,\mu)$, $I(\nu\,|\,\mu)$, $\mathbb{W}_2(\nu,\mu)$ and $S(\nu\,|\,\mu)$ for $\nu \in \mathscr{P}(M)$ be defined as in (1.3), (1.4), (1.5) and (1.6) respectively with (M,μ) replacing $(\mathbb{R}^n,\gamma)$, the Riemannian distance $\rho(x,y)$ replacing $|x-y|$, and $\mathbb{S}_\nu$ being the class of measurable 2-tensors τ which are locally integrable with respect to ν such that

$$\int_M \langle \nabla V, \nabla f \rangle \, d\nu = \int_M \langle \tau, \mathrm{Hess}_f \rangle_{\mathrm{HS}} \, d\nu, \quad f \in C_0^\infty(M).$$

Assume $\mathbb{S}_\nu$ is non-empty, i.e. the Stein kernel of ν exists. When $M = \mathbb{R}^n$, it is ensured by the existence of a spectral gap (see [4]). The existence of the Stein kernel on Riemannian manifolds are currently under development and will be publicized later.

Our results on Riemannian manifolds are presented in the Sections 3, 4 and 5. The estimates take the most concise form in case when the function V satisfies $\mathrm{Hess}_V = K$ for some constant $K > 0$. In this case, for instance, we obtain inequalities of the same form as in the Euclidean case:

$$H(\nu\,|\,\mu) \le \frac{1}{2} S^2(\nu\,|\,\mu) \log\Big(1 + \frac{I(\nu\,|\,\mu)}{K S^2(\nu\,|\,\mu)}\Big),$$
$$\mathbb{W}_2(\nu\,|\,\mu) \le \frac{S(\nu\,|\,\mu)}{K^{1/2}} \arccos\left(\exp\left(-\frac{H(\nu\,|\,\mu)}{S^2(\nu\,|\,\mu)}\right)\right), \quad \nu \in \mathscr{P}(M),$$

and there exists a constant $C > 0$ such that

$$\left(\int |f|^p \, d\nu\right)^{1/p} \le C\left(S_p(\nu\,|\,\gamma) + \sqrt{p}\Big(\int |\tau|_{\mathrm{op}}^{p/2} d\nu\Big)^{1/p}\right), \quad \nu(f) = 0, \ |\nabla f| \le 1, \ \tau \in \mathbb{S}_\nu.$$

The remainder of this paper is organized as follows. In Section 2 we study Hessian estimates for P_t following the lines of [25]. Such estimates which are interesting in themselves, serve as crucial tools for extending (1.7), (1.8) and (1.9) to the general geometric setting in Sections 3, 4 and 5 respectively. We work out some examples in Section 3.1.

2 Hessian estimate of P_t

Let (M, g) be a n-dimensional complete Riemannian manifold. We write $\langle u, v \rangle = g(u, v)$ and $|u| = \sqrt{\langle u, u \rangle}$ for $u, v \in T_xM$ and $x \in M$. Let R, Ric be the Riemann curvature tensor and Ricci curvature tensor respectively. Recall that $R \in \Gamma(T^*M \otimes T^*M \otimes T^*M \otimes TM)$ where

$$R(X, Y, Z) \equiv R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z, \quad X, Y, Z \in \Gamma(TM),$$

and $\mathrm{Ric} \in \Gamma(T^*M \otimes T^*M)$ given as $\mathrm{Ric}(Y, Z) = \mathrm{tr}(X \mapsto R(X, Y)Z)$.

1. For $f, h \in C^2(M)$ and $x \in M$, we consider the Hilbert-Schmidt inner product of the Hessian tensors Hess_f and Hess_h, i.e.

$$\langle \mathrm{Hess}_f, \mathrm{Hess}_h \rangle_{\mathrm{HS}} = \sum_{i,j=1}^n \mathrm{Hess}_f(X_i, X_j) \mathrm{Hess}_h(X_i, X_j),$$

where $(X_i)_{1 \le i \le n}$ denotes an orthonormal base of T_xM. Then the Hilbert-Schmidt norm of Hess_f is given by

$$|\mathrm{Hess}_f|_{\mathrm{HS}}(x) = \sqrt{\langle \mathrm{Hess}_f, \mathrm{Hess}_f \rangle_{\mathrm{HS}}}.$$

2. For a symmetric 2-tensor T and a constant K, we write $T \geq K$ if

$$T(w,w) \geq K|w|^2, \quad w \in T_xM, \ x \in M,$$

and $T \leq K$ if

$$T(w,w) \leq K|w|^2, \quad w \in T_xM, \ x \in M.$$

3. Given a symmetric 2-tensor T, let $T^\sharp \colon TM \to TM$ be defined by

$$\langle T^\sharp(v), w\rangle = T(v,w), \quad v,w \in T_xM, \ x \in M.$$

Then $T^\sharp$ is a symmetric endomorphism, i.e., $\langle T^\sharp(w), v\rangle = \langle T^\sharp(v), w\rangle$ for $v,w \in T_xM$, $x \in M$. Let

$$|T|(x) = \sup\left\{|T^\sharp(w)| \colon w \in T_xM, \ |w| \leq 1\right\}, \quad x \in M.$$

Then, in particular, $|\mathrm{Hess}_f|(x)$ gives the operator norm of the Hessian of a function f at x.

4. Furthermore, denoting by $\mathrm{Bil}(TM)$ the vector bundle of bilinear forms on TM, we consider $\tilde{R} \in \Gamma(T^*M \otimes T^*M \otimes \mathrm{Bil}(TM))$ given by

$$\tilde{R}(v_1, v_2) = \langle R(\cdot, v_1)v_2, \cdot\rangle, \quad v_1, v_2 \in T_xM,$$

and let

$$|\tilde{R}|(x) = \left||\tilde{R}(\cdot,\cdot)|_{\mathrm{HS}}\right|_{\mathrm{HS}}(x) \text{ for } x \in M \ \text{ and } \ \|\tilde{R}\|_\infty = \sup_{x\in M} |\tilde{R}|(x).$$

Note that in explicit terms

$$\|\tilde{R}\|_\infty = \sup_{x\in M}\left(\sum_{k,\ell}\sum_{i,j}\langle R(e_i, v_k)v_\ell, e_j\rangle^2\right)^{1/2}$$

where $(v_k)_{1\leq k\leq n}$ and $(e_i)_{1\leq i\leq n}$ denote orthonormal bases for T_xM.

5. For a general symmetric 2-tensor T, we adopt the notation

$$(RT)(v_1, v_2) := \mathrm{tr}\,\langle R(\cdot, v_1)v_2, T^\sharp(\cdot)\rangle = \sum_{i=1}^n \langle R(e_i, v_1)v_2, T^\sharp(e_i)\rangle,$$

where $v_1, v_2 \in T_xM$, $x \in M$ and $(e_i)_{1\leq i\leq n}$ is an orthonormal base of T_xM. Let

$$|R|(x) = \sup\left\{|(RT)(v_1,v_2)| \colon |v_1| \leq 1, \ |v_2| \leq 1, \ |T| \leq 1\right\} \quad \text{and} \quad \|R\|_\infty = \sup_{x\in M}|R|(x).$$

It is easy to see that $|\tilde{R}|(x) \leq n|R|(x)$. In particular, if $\|R\|_\infty < \infty$ then $\|\tilde{R}\|_\infty < \infty$ as well.

6. In addition, let

$$\mathrm{d}^*R = -\mathrm{tr}\, \nabla_{\bullet} R,$$

i.e.,

$$(\mathrm{d}^*R)(v_1, v_2) = -\mathrm{tr}\, \nabla_{\bullet} R(\cdot, v_1)v_2, \quad v_1, v_2 \in T_xM.$$

Note that

$$\langle (\mathrm{d}^*R)(v_1, v_2), v_3\rangle = \langle (\nabla_{v_3}\mathrm{Ric}^\sharp)(v_1), v_2\rangle - \langle (\nabla_{v_2}\mathrm{Ric}^\sharp)(v_3), v_1\rangle, \quad v_1, v_2, v_3 \in T_xM.$$

7. Finally, for $v, w \in T_xM$, let

$$R(\nabla V)(v, w) := R(\nabla V, v)w.$$

In this section, we get the explicit Hessian estimates of the semigroups, which are derived from the second derivative formula of the semigroup obtained by first identifying appropriate local martingales. Actrally, an martingales approach to derivative formulas was first developed by Elworthy and Li [8], after which an approach based on local martingales was given by Thalmaier [19] and Driver and Thalmaier [7]. Although various formulas for the Hessian have been given, for example, in [1, 8, 11, 25, 20, 21], the Hessian estimate of semigoup are not well calculated with explicit constants depending on the curvature tensor on general Riemannian manifolds.

2.1 Hessian estimates of semigroup: type I

Let us introduce the following type of Hessian estimate of semigroup. When M is Ricci parallel and the generator of the diffusion equals half the Laplacian Δ, such kind of formula is given in [25] which bounds the norm of the Hessian of P_tf from above by $P_t|\nabla f|^2$.

Theorem 2.1 (Hessian estimate: type I)**.** *Assume that* $\mathrm{Ric}_V \geq K$*,* $\|R\|_\infty < \infty$ *and*

$$\beta := \|\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V)\|_\infty < \infty.$$

Let $\alpha_1 := \|R\|_\infty$ *and* $\alpha_2 := \|\tilde{R}\|_\infty$*. Then for* $f \in C_b^2(M)$*,*

$$\begin{aligned}&|\mathrm{Hess}_{P_tf}|\\ &\leq \left(\frac{K - 2\alpha_1}{\mathrm{e}^{(2K-2\alpha_1)t} - \mathrm{e}^{Kt}}\right)^{1/2} \left((P_t|\nabla f|^2)^{1/2} + \left(\frac{\mathrm{e}^{Kt} - 1}{K}\right)^{1/2} \frac{\beta}{K}(P_t|\nabla f|)\right).\end{aligned}$$

Moreover, if $\mathrm{Ric}_V = K$*, then*

$$\begin{aligned}&|\mathrm{Hess}_{P_tf}|_{\mathrm{HS}}\\ &\leq \left(\frac{K - 2\alpha_2}{\mathrm{e}^{(2K-2\alpha_2)t} - \mathrm{e}^{Kt}}\right)^{1/2} \left((P_t|\nabla f|^2)^{1/2} + \left(\frac{\mathrm{e}^{Kt} - 1}{K}\right)^{1/2} \frac{n\beta}{K}(P_t|\nabla f|)\right).\end{aligned} \tag{2.1}$$

To prove Theorem 2.1, we first introduce a probabilistic representation formula for $\mathrm{Hess}_{P_t f}$. For the semigroup P_t generated by $\Delta/2$, a Bismut type Hessian formula has been established in [1], which was then extended to general Schrödinger operators on M [11, 20].

Denote by $\mathrm{Ric}_V^\sharp = \mathrm{Ric}^\sharp + \mathrm{Hess}_V^\sharp$ the Bakry-Émery tensor (written as endomorphism of TM). The damped parallel transport $Q_t\colon T_xM \to T_{X_t}M$ is defined as the solution, along the paths of X_t, to the covariant ordinary differential equation

$$DQ_t = -\frac{1}{2}\mathrm{Ric}_V^\sharp Q_t\,\mathrm{d}t, \quad Q_0 = \mathrm{id},$$

where the covariant differential is given by $/\!/_t^{-1} D = \mathrm{d}\,/\!/_t^{-1}$.

For $w \in T_xM$, we define an operator-valued process $W_t(\cdot, w) : T_xM \to T_{X_t}M$ by

$$\begin{aligned} W_t(\cdot, w) :=& Q_t \int_0^t Q_r^{-1} R(/\!/_r\,\mathrm{d}B_r, Q_r(\cdot))Q_r(w) \\ &- \frac{1}{2} Q_t \int_0^t Q_r^{-1}(\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^* R + R(\nabla V))(Q_r(\cdot), Q_r(w))\,\mathrm{d}r. \end{aligned}$$

Note that $W_t(\cdot, w)$ is the solution to the covariant Itô equation

$$\begin{aligned} DW_t(\cdot, w) =& R(/\!/_t\,\mathrm{d}B_t, Q_t(\cdot))Q_t(w) - \frac{1}{2}\mathrm{Ric}_V^\sharp(W_t(\cdot, w))\,\mathrm{d}t \\ &- \frac{1}{2}(\mathrm{d}^* R + \nabla \mathrm{Ric}_V^\sharp + R(\nabla V))(Q_t(\cdot), Q_t(w))\,\mathrm{d}t, \end{aligned}$$

with initial condition $W_0(\cdot, w) = 0$.

Lemma 2.2. *Let ρ be the Riemannian distance to a fixed point $o \in M$. Assume that*

$$\lim_{\rho\to\infty} \frac{\log\left(\left|\mathrm{d}^* R + \nabla \mathrm{Ric}_V^\sharp + R(\nabla V)\right| + |R|\right)}{\rho^2} = 0,$$

and

$$\mathrm{Ric}_V \geq -h(\rho) \quad \text{for some positive function } h \in C([0,\infty)) \text{ such that } \lim_{r\to\infty} \frac{h(r)}{r^2} = 0.$$

Then

$$\mathrm{Hess}_{P_t f}(v, w) = \mathbb{E}\left[\mathrm{Hess}_f(Q_t(v), Q_t(w)) + \langle \nabla f(X_t), W_t(v, w)\rangle\right].$$

Proof. For fixed $T > 0$, set

$$N_t(v, w) := \mathrm{Hess}_{P_{T-t} f}(Q_t(v), Q_t(w)) + \langle \nabla P_{T-t} f(X_t), W_t(v, w)\rangle.$$

We first show that $N_t(v, w)$ is a local martingale, which has been shown e.g. in [21, Lemma 11.3]. We include it for readers' convenience. We first observe that

$$\mathrm{d}(\Delta - \nabla V)f = \left(\mathrm{tr}\,\nabla^2 - \nabla_{\nabla V}\right)\mathrm{d}f - \mathrm{d}f(\mathrm{Ric}_V^\sharp),$$

$$\nabla\mathrm{d}(\Delta f) = \mathrm{tr}\,\nabla^2(\nabla\mathrm{d}f) - (\nabla\mathrm{d}f)(\mathrm{Ric}^\sharp\odot\mathrm{id} + \mathrm{id}\odot\mathrm{Ric}^\sharp - 2R^{\sharp,\sharp}) - \mathrm{d}f(\mathrm{d}^*R + \nabla\mathrm{Ric}^\sharp),$$
$$\nabla\mathrm{d}(\nabla V(f)) = \nabla_{\nabla V}(\nabla\mathrm{d}f) + (\nabla\mathrm{d}f)(\mathrm{Hess}_V^\sharp\odot\mathrm{id} + \mathrm{id}\odot\mathrm{Hess}_V^\sharp) + \mathrm{d}f(\nabla\mathrm{Hess}_V^\sharp + R(\nabla V)),$$

where $\odot$ denotes the symmetric tensor product. Thus, for the Itô differential of $N_t(v,w)$, we obtain

$$\begin{aligned}
\mathrm{d}N_t(v,w) &= (\nabla_{//_t\,\mathrm{d}B_t}\mathrm{Hess}_{P_{T-t}f})(Q_t(v),Q_t(w)) + \mathrm{Hess}_{P_{T-t}f}\left(\frac{D}{\mathrm{d}t}Q_t(v),Q_t(w)\right)\mathrm{d}t\\
&\quad + \mathrm{Hess}_{P_{T-t}f}\left(Q_t(v),\frac{D}{\mathrm{d}t}Q_t(w)\right)\mathrm{d}t + \partial_t(\mathrm{Hess}_{P_{T-t}f})(Q_t(v),Q_t(w))\,\mathrm{d}t\\
&\quad + \frac{1}{2}\mathrm{tr}(\nabla^2 - \nabla_{\nabla V})(\mathrm{Hess}_{P_{T-t}f})(Q_t(v),Q_t(w))\,\mathrm{d}t + (\nabla_{//_t\,\mathrm{d}B_t}\mathrm{d}P_{T-t}f)(W_t(v,w))\\
&\quad + (\mathrm{d}P_{T-t}f)(DW_t(v,w)) + \langle D(\mathrm{d}P_{T-t}f), DW_t(v,w)\rangle + \partial_t(\mathrm{d}P_{T-t}f)(W_t(v,w))\,\mathrm{d}t\\
&\quad + \frac{1}{2}\mathrm{tr}(\nabla^2 - \nabla_{\nabla V})(\mathrm{d}P_{T-t}f)(W_t(v,w))\,\mathrm{d}t\\
&\overset{\mathrm{m}}{=} -\frac{1}{2}\mathrm{Hess}_{P_{T-t}f}\left(\mathrm{Ric}_V^\sharp(Q_t(v)),Q_t(w)\right)\mathrm{d}t - \frac{1}{2}\mathrm{Hess}_{P_{T-t}f}\left(Q_t(v),\mathrm{Ric}_V^\sharp(Q_t(w))\right)\mathrm{d}t\\
&\quad - \frac{1}{2}(\nabla\mathrm{d}(\Delta-\nabla V)P_{T-t}f)(Q_t(v),Q_t(w))\,\mathrm{d}t + \frac{1}{2}\left(\mathrm{tr}\nabla^2 - \nabla_{\nabla V}\right)(\mathrm{Hess}_{P_{T-t}f})(Q_t(v),Q_t(w))\,\mathrm{d}t\\
&\quad - \frac{1}{2}(\mathrm{d}P_{T-t}f)(\mathrm{d}^*R + \nabla\mathrm{Ric}_V^\sharp + R(\nabla V))(Q_t(v),Q_t(w))\,\mathrm{d}t\\
&\quad - \frac{1}{2}(\mathrm{d}P_{T-t}f)(\mathrm{Ric}_V^\sharp(W_t(v,w)))\,\mathrm{d}t + \mathrm{tr}\left\{\mathrm{Hess}_{P_{T-t}f}(\cdot,R(\cdot,Q_t(v))Q_t(w))\right\}\mathrm{d}t\\
&\quad - \frac{1}{2}(\mathrm{d}(\Delta-\nabla V)P_{T-t}f)(W_t(v,w))\,\mathrm{d}t + \frac{1}{2}\left(\mathrm{tr}\nabla^2 - \nabla_{\nabla V}\right)(\mathrm{d}P_{T-t}f)(W_t(v,w))\,\mathrm{d}t\\
&= 0,
\end{aligned}$$

where $\overset{\mathrm{m}}{=}$ denotes equality modulo differentials of local martingales, so that N_t is a local martingale. Assume that

$$\lim_{\rho\to\infty}\frac{\log\left(|\mathrm{d}^*R + \nabla\mathrm{Ric}_V^\sharp + R(\nabla V)| + |R|\right)}{\rho^2} = 0,$$

and

$$\mathrm{Ric}_V \geq -h(\rho) \quad \text{for some positive } h\in C([0,\infty)) \text{ with } \lim_{r\to\infty}\frac{h(r)}{r^2} = 0.$$

Then by [25, Proposition 3.1], for $t>0$ we have

$$\mathbb{E}\left[\sup_{s\in[0,t]}|Q_s|^2\right] < \infty \quad \text{and} \quad \mathbb{E}\left[\sup_{s\in[0,t]}|W_s|^2\right] < \infty.$$

In addition, $|\nabla P_{T-t}|(x)$ and $|\mathrm{Hess}_{P_{T-t}}|(x)$ are easy to bound by local Bismut type formulae [1, 20]. Under our curvature assumptions these local bounds then provide global bounds uniformly in $(t,x)\in[0,T-\varepsilon]\times M$ for every small $\varepsilon>0$. Thus the local martingale N_t is a true martingale on the time interval $[0,T-\varepsilon]$. By taking expectations, we first obtain $\mathbb{E}[N_0] = \mathbb{E}[N_{T-\varepsilon}]$ and then

$$\mathrm{Hess}_{P_Tf}(v,w) = \mathbb{E}\left[\mathrm{Hess}_f(Q_T(v),Q_T(w)) + \langle\nabla f(X_T), W_T(v,w)\rangle\right]$$

by passing to the limit as $\varepsilon\downarrow 0$. □

According to the definition of W_t, we have

$$\begin{aligned}\mathbb{E}\langle\nabla f(X_t), W_t(v,w)\rangle =&\mathbb{E}\Big\langle\nabla f(X_t), Q_t\int_0^t Q_r^{-1}R(/\!/_r\,\mathrm{d}B_r, Q_r(v))Q_r(w)\Big\rangle\\ &-\frac{1}{2}\mathbb{E}\Big\langle\nabla f(X_t), Q_t\int_0^t Q_r^{-1}(\nabla\mathrm{Ric}_V^\sharp+\mathrm{d}^*R+R(\nabla V))(Q_r(v),Q_r(w))\,\mathrm{d}r\Big\rangle.\end{aligned}$$

To deal with the first term on the right hand side, we observe that

Lemma 2.3. *Keeping the assumptions of Lemma 2.6, we have*

$$\mathbb{E}\left[\Big\langle\nabla f(X_t), Q_t\int_0^t Q_r^{-1}R(/\!/_r\,\mathrm{d}B_r, Q_r(v))Q_r(w)\Big\rangle\right]=\mathbb{E}\left[\int_0^t(R\mathrm{Hess}_{P_{t-s}f})(Q_s(v),Q_s(w)\,\mathrm{d}s\right].$$

Proof. Let

$$H_s(v,w)=\Big\langle\nabla P_{t-s}f(X_s), Q_s\int_0^s Q_r^{-1}R(/\!/_r\,\mathrm{d}B_r, Q_r(v))Q_r(w)\Big\rangle.$$

It is easy to see that

$$\begin{aligned}\mathrm{d}(H_s(v,w))=&\Big\langle\nabla_{/\!/_s\mathrm{d}B_s}(\nabla P_{t-s}f)(X_s), Q_s\int_0^s Q_r^{-1}R(/\!/_r\,\mathrm{d}B_r, Q_r(v))Q_r(w)\Big\rangle\\ &+\Big\langle\mathrm{Ric}_V^\sharp(\nabla P_{t-s}f)(X_s), Q_s\int_0^s Q_r^{-1}R(/\!/_r\,\mathrm{d}B_r, Q_r(v))Q_r(w)\Big\rangle\,\mathrm{d}s\\ &-\Big\langle(\nabla P_{t-s}f)(X_s), \mathrm{Ric}_V^\sharp\Big(Q_s\int_0^s Q_r^{-1}R(/\!/_r\,\mathrm{d}B_r, Q_r(v))Q_r(w)\Big)\Big\rangle\,\mathrm{d}s\\ &+\Big\langle(\nabla P_{t-s}f)(X_s), R(/\!/_s\,\mathrm{d}B_s, Q_s(v))Q_s(w)\Big\rangle\\ &+\mathrm{tr}\,\langle\nabla_\bullet(\nabla P_{t-s}f), R(\cdot, Q_s(v))Q_s(w)\rangle\,\mathrm{d}s\\ \overset{\mathrm{m}}{=}&\ \mathrm{tr}\Big(\mathrm{Hess}_{P_{t-s}f}(\cdot, R(\cdot, Q_s(v))Q_s(w))\Big)\,\mathrm{d}s\end{aligned}$$

which implies

$$\mathbb{E}\left[\Big\langle\nabla f(X_t), Q_t\int_0^t Q_s^{-1}R(/\!/_s\,\mathrm{d}B_s, Q_s(v))Q_s(w)\Big\rangle\right]=\mathbb{E}\left[\int_0^t\mathrm{tr}\Big(\mathrm{Hess}_{P_{t-s}f}(\cdot, R(\cdot, Q_s(v))Q_s(w))\Big)\,\mathrm{d}s\right].$$

□

With these two lemmas we are now in position to prove Theorem 2.1.

Proof of Theorem 2.1. We begin with the following observation obtained by combining the formulas in Lemmas 2.6 and 2.3:

$$\begin{aligned}\mathrm{Hess}_{P_tf}(v,w)=&\,\mathbb{E}\Big[\mathrm{Hess}_f(Q_t(v),Q_t(w))\Big]+\mathbb{E}\left[\langle\nabla f(X_t), W_t(v,w)\rangle\right]\\ =&\,\mathbb{E}\Big[\mathrm{Hess}_f(Q_t(v),Q_t(w))\Big]+\mathbb{E}\left[\Big\langle\nabla f(X_t), Q_t\int_0^t Q_r^{-1}R(/\!/_r\mathrm{d}B_r, Q_r(v))Q_r(w)\Big\rangle\right]\\ &-\frac{1}{2}\mathbb{E}\left[\Big\langle\nabla f(X_t), Q_t\int_0^t Q_r^{-1}(\nabla\mathrm{Ric}_V^\sharp+\mathrm{d}^*R+R(\nabla V))(Q_r(v),Q_r(w))\,\mathrm{d}r\Big\rangle\right]\end{aligned}$$

$$\begin{aligned}&= \mathbb{E}\Big[\mathrm{Hess}_f(Q_t(v),Q_t(w))\Big] + \mathbb{E}\Big[\int_0^t \mathrm{tr}\,\Big(\mathrm{Hess}_{P_{t-s}f}(\cdot, R(\cdot, Q_s(v))Q_s(w))\Big)\,\mathrm{d}s\Big]\\ &\quad - \frac{1}{2}\mathbb{E}\Big[\Big\langle \nabla f(X_t), Q_t\int_0^t Q_r^{-1}(\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V))(Q_r(v),Q_r(w))\,\mathrm{d}r\Big\rangle\Big].\end{aligned}$$

Noting that $|Q_tQ_r^{-1}| \le \mathrm{e}^{-K(t-r)/2}$, $|Q_r| \le \mathrm{e}^{-Kr/2}$, and

$$\mathrm{tr}\,\Big(\mathrm{Hess}_{P_{t-s}f}(\cdot, R(\cdot, Q_s(v))Q_s(w))\Big) \le \mathrm{e}^{-Ks}\,|\mathrm{Hess}_{P_{t-s}f}|(X_s)\,\|R\|_\infty,$$

where $(e_i)_{1\le i\le n}$ is an orthonormal base of T_xM, we derive

$$\begin{aligned}|\mathrm{Hess}_{P_tf}| &\le \mathrm{e}^{-Kt}\,P_t|\mathrm{Hess}_f| + \|R\|_\infty\int_0^t \mathrm{e}^{-Ks}P_s|\mathrm{Hess}_{P_{t-s}f}|\,\mathrm{d}s + \frac{\beta}{2}\Big(\int_0^t \mathrm{e}^{-K(t+r)/2}\mathrm{d}r\Big)P_t|\nabla f|\\ &= \mathrm{e}^{-Kt}\,P_t|\mathrm{Hess}_f| + \frac{\beta(\mathrm{e}^{-Kt/2}-\mathrm{e}^{-Kt})}{K}P_t|\nabla f| + \|R\|_\infty\int_0^t \mathrm{e}^{-Ks}P_s|\mathrm{Hess}_{P_{t-s}f}|\,\mathrm{d}s,\quad t\ge 0.\end{aligned}$$

Now let

$$\phi(r) := \mathrm{e}^{-K(t-r)}P_{t-r}|\mathrm{Hess}_{P_rf}|,\quad r\in[0,t].$$

Applying the above estimate for P_rf instead of P_tf, and noting that $\frac{\mathrm{e}^{Kr/2}-1}{K}$ is increasing in r, we obtain

$$\begin{aligned}\phi(r) &\le \phi(0) + \beta\mathrm{e}^{-Kt}\frac{\mathrm{e}^{Kr/2}-1}{K}P_{t-r}(P_r|\nabla f|) + \|R\|_\infty\int_0^r \phi(r-s)\,\mathrm{d}s\\ &\le \phi(0) + \frac{\beta(\mathrm{e}^{-Kt/2}-\mathrm{e}^{-Kt})}{K}P_t|\nabla f| + \|R\|_\infty\int_0^r\phi(s)\,\mathrm{d}s,\quad r\in[0,t].\end{aligned}$$

By Gronwall's lemma, this implies

$$\begin{aligned}|\mathrm{Hess}_{P_tf}| = \phi(t) &\le \Big\{\phi(0) + \frac{\beta(\mathrm{e}^{-Kt/2}-\mathrm{e}^{-Kt})}{K}P_t|\nabla f|\Big\}\,\mathrm{e}^{\|R\|_\infty t}\\ &= \mathrm{e}^{(\|R\|_\infty-K)t}P_t|\mathrm{Hess}_f| + \frac{\beta\,\mathrm{e}^{\|R\|_\infty t}(\mathrm{e}^{-Kt/2}-\mathrm{e}^{-Kt})}{K}P_t|\nabla f|.\end{aligned}\tag{2.2}$$

On the other hand, by Itô's formula we have

$$\begin{aligned}\mathrm{d}|\nabla P_{t-s}f|^2(X_s) &= \frac{1}{2}\Big(L|\nabla P_{t-s}f|^2(X_s) - \langle \nabla P_{t-s}f, \nabla LP_{t-s}f\rangle(X_s)\Big)\,\mathrm{d}s\\ &\quad + \langle \nabla|\nabla P_{t-s}f|^2(X_s), //_s\mathrm{d}B_s\rangle,\qquad s\in[0,t].\end{aligned}$$

Using the Bochner-Weitzenböck formula and the assumption $\mathrm{Ric}_V \ge K$, we obtain

$$\begin{aligned}&\mathrm{d}|\nabla P_{t-s}f|^2(X_s)\\ &\ge \Big(\mathrm{Ric}_V(\nabla P_{t-s}f, \nabla P_{t-s}f) + |\mathrm{Hess}_{P_{t-s}f}|_{\mathrm{HS}}^2\Big)(X_s)\,\mathrm{d}s + \langle \nabla|\nabla P_{t-s}f|^2(X_s), //_s\mathrm{d}B_s\rangle\\ &\ge K|\nabla P_{t-s}f|^2(X_s)\,\mathrm{d}s + |\mathrm{Hess}_{P_{t-s}f}|_{\mathrm{HS}}^2(X_s)\,\mathrm{d}s + \langle \nabla|\nabla P_{t-s}f|^2(X_s), //_s\mathrm{d}B_s\rangle.\end{aligned}$$

From this, we conclude that

$$P_t|\nabla f|^2 - \mathrm{e}^{Kt}|\nabla P_t f|^2 \geq \int_0^t \mathrm{e}^{K(t-s)} P_s|\mathrm{Hess}_{P_{t-s}f}|_{\mathrm{HS}}^2\,\mathrm{d}s.$$

By the inequalities of Jensen and Schwartz, this yields

$$\begin{aligned}\mathrm{e}^{-Kt/2}(P_t|\nabla f|^2)^{1/2} &\geq \left(\int_0^t \mathrm{e}^{-2\|R\|_\infty s}\mathrm{e}^{(2\|R\|_\infty - K)s}(P_s|\mathrm{Hess}_{P_{t-s}f}|_{\mathrm{HS}})^2\,\mathrm{d}s\right)^{1/2}\\ &\geq \left(\frac{K-2\|R\|_\infty}{\mathrm{e}^{(K-2\|R\|_\infty)t}-1}\right)^{1/2}\int_0^t \mathrm{e}^{-\|R\|_\infty s} P_s|\mathrm{Hess}_{P_{t-s}f}|_{\mathrm{HS}}\,\mathrm{d}s.\end{aligned}$$

Combining this with (2.2) for $(P_s, P_{t-s}f)$ instead of (P_t, f), and noting that $|\nabla P_{t-s}f| \leq \mathrm{e}^{-K(t-s)/2}P_{t-s}|\nabla f|$, we arrive at

$$\begin{aligned}&\mathrm{e}^{-Kt/2}\left(\frac{\mathrm{e}^{(K-2\|R\|_\infty)t}-1}{K-2\|R\|_\infty}\right)^{1/2}(P_t|\nabla f|^2)^{1/2}\\ &\geq \int_0^t \mathrm{e}^{-\|R\|_\infty s}P_s|\mathrm{Hess}_{P_{t-s}f}|_{\mathrm{HS}}\mathrm{d}s\\ &\geq \int_0^t \mathrm{e}^{-\|R\|_\infty s}\left(\mathrm{e}^{(K-\|R\|_\infty)s}|\mathrm{Hess}_{P_t f}| - \frac{\beta\,\mathrm{e}^{Ks}(\mathrm{e}^{-Ks/2}-\mathrm{e}^{-Ks})}{K}P_s|\nabla P_{t-s}f|\right)\mathrm{d}s\\ &\geq \frac{\mathrm{e}^{(K-2\|R\|_\infty)t}-1}{K-2\|R\|_\infty}|\mathrm{Hess}_{P_t f}| - \beta(P_t|\nabla f|)\,\mathrm{e}^{-Kt/2}\int_0^t \mathrm{e}^{(K-\|R\|_\infty)s}\frac{1-\mathrm{e}^{-Ks/2}}{K}\,\mathrm{d}s\\ &\geq \frac{\mathrm{e}^{(K-2\|R\|_\infty)t}-1}{K-2\|R\|_\infty}|\mathrm{Hess}_{P_t f}| - \frac{\beta}{K}\,\mathrm{e}^{-Kt/2}\left(\frac{\mathrm{e}^{(K-2\|R\|_\infty)t}-1}{K-2\|R\|_\infty}\right)^{1/2}\left(\frac{\mathrm{e}^{Kt}-1}{K}\right)^{1/2}(P_t|\nabla f|).\end{aligned}$$

This completes the proof of the first inequality.

For the second case, when $\mathrm{Ric}_V = K$ we realize that $Q_t(v) = \mathrm{e}^{-Kt/2}/\!/_t v$ for $v \in T_xM$, and that for all $f \in C_b^2(M)$ and $v, w \in T_xM$ such that $|v| = |w| = 1$,

$$\begin{aligned}&\mathrm{Hess}_{P_t f}(v,w) \qquad (2.3)\\ &= \mathbb{E}\left[\mathrm{Hess}_f(Q_t(v), Q_t(w))\right] + \mathbb{E}\int_0^t (R\mathrm{Hess}_{P_s f})(Q_{t-s}v, Q_{t-s}w)\,\mathrm{d}s\\ &\quad - \frac{1}{2}\mathbb{E}\left[\left\langle \nabla f(X_t), Q_t\int_0^t Q_r^{-1}(\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V))(Q_r(v), Q_r(w))\,\mathrm{d}r\right\rangle\right]\\ &= \mathrm{e}^{-tK}\mathbb{E}\left[\mathrm{Hess}_f(/\!/_t v, /\!/_t w)(X_t)\right] + \frac{\beta(\mathrm{e}^{-Kt/2}-\mathrm{e}^{-Kt})}{K}P_t|\nabla f|\\ &\quad + \int_0^t \mathbb{E}\left[\mathrm{e}^{-K(t-s)}(R\mathrm{Hess}_{P_s f})(/\!/_{t-s}v, /\!/_{t-s}w)\right]\mathrm{d}s\\ &\leq \mathrm{e}^{-tK}\mathbb{E}\left[\mathrm{Hess}_f(/\!/_t v, /\!/_t w)(X_t)\right] + \frac{\beta(\mathrm{e}^{-Kt/2}-\mathrm{e}^{-Kt})}{K}P_t|\nabla f|\\ &\quad + \int_0^t \mathbb{E}\left[\mathrm{e}^{-K(t-s)}\,\mathrm{tr}\langle \mathrm{Hess}_{P_s f}(\cdot), R(\cdot, /\!/_{t-s}v)/\!/_{t-s}w\rangle\right]\mathrm{d}s\end{aligned}$$

$$\leq e^{-tK}\mathbb{E}\left[\mathrm{Hess}_f(/\!/_t v, /\!/_t w)(X_t)\right] + \frac{\beta(e^{-Kt/2}-e^{-Kt})}{K}P_t|\nabla f| + \int_0^t e^{-K(t-s)}\mathbb{E}\left[\left|\mathrm{Hess}_{P_s f}\right|_{\mathrm{HS}}(X_{t-s})\,|\tilde{R}(/\!/_{t-s}v, /\!/_{t-s}w)|_{\mathrm{HS}}(X_{t-s})\right]\,\mathrm{d}s. \tag{2.4}$$

This gives us

$$\begin{aligned}
&|\mathrm{Hess}_{P_t f}|_{\mathrm{HS}}\\
&\leq \mathbb{E}\left[e^{-Kt}|\mathrm{Hess}_f|_{\mathrm{HS}}(X_t)\right] + \frac{n\beta(e^{-Kt/2}-e^{-Kt})}{K}P_t|\nabla f| +\\
&+\sqrt{\sum_{i,j}\left(\mathbb{E}\left[\int_0^t e^{-K(t-s)}\left(\left|\mathrm{Hess}_{P_s f}\right|_{\mathrm{HS}}(X_{t-s})\,|\tilde{R}(/\!/_{t-s}e_i, /\!/_{t-s}e_j)|_{\mathrm{HS}}(X_{t-s})\right)\mathrm{d}s\right]\right)^2}\\
&\leq \mathbb{E}\left[e^{-Kt}|\mathrm{Hess}_f|_{\mathrm{HS}}(X_t)\right] + \frac{n\beta(e^{-Kt/2}-e^{-Kt})}{K}P_t|\nabla f| +\\
&+\sqrt{\sum_{i,j}\mathbb{E}\left[\int_0^t e^{-K(t-s)}\left|\mathrm{Hess}_{P_s f}\right|_{\mathrm{HS}}(X_{t-s})\,\mathrm{d}s\right]\mathbb{E}\left[\int_0^t e^{-K(t-s)}\left(\left|\mathrm{Hess}_{P_s f}\right|_{\mathrm{HS}}|\tilde{R}(/\!/_{t-s}e_i, /\!/_{t-s}e_j)|^2_{\mathrm{HS}}\right)(X_{t-s})\,\mathrm{d}s\right]}\\
&\leq \mathbb{E}\left[e^{-Kt}|\mathrm{Hess}_f|_{\mathrm{HS}}(X_t)\right] + \frac{n\beta(e^{-Kt/2}-e^{-Kt})}{K}P_t|\nabla f| + \|\tilde{R}\|_\infty\,\mathbb{E}\left[\int_0^t e^{-K(t-s)}\left|\mathrm{Hess}_{P_s f}\right|_{\mathrm{HS}}(X_{t-s})\,\mathrm{d}s\right].
\end{aligned}$$

The remaining steps are similar to the first part of the proof; we skip the details. □

Important examples in the sequel will be Ricci parallel manifolds which is the class of Riemannian manifolds where Ricci curvature is constant under parallel transport, that is $\nabla\mathrm{Ric} = 0$ for the Levi-Civita connection ∇. Recall that an Einstein manifold is Ricci parallel but in general the inverse is not true.

Recently F.-Y. Wang [25] used functional inequalities for the semigroup to identify constant curvature manifolds, Einstein manifolds, and Ricci parallel manifolds. Here we list the results for the Hessian estimate of P_t generated by the operator $\frac{1}{2}L$ when M is a Ricci parallel manifold and ∇V is a Killing field on (M, g). Here, a vector field X on a Riemannian manifold (M, g) is called a *Killing field* if the local flows generated by X act by isometries i.e., for $Y, Z \in TM$,

$$\nabla^2_{Y,Z}(X) = -R(X, Y)Z.$$

We conclude that $\|\mathrm{d}^*R + \nabla\mathrm{Ric}^\sharp_V + R(\nabla V)\|_\infty = 0$ if ∇V is a Killing field on a Ricci parallel manifold (M, g).

Corollary 2.4. *Assume that M is a Ricci parallel manifold, ∇V is a Killing field and $\|R\|_\infty < \infty$. Then for any constant $K \in \mathbb{R}$,*

(i) *if* $\mathrm{Ric}_V \geq K > 0$, *then for any $f \in C_b^2(M)$ and $t \geq 0$,*

$$|\mathrm{Hess}_{P_t f}|^2_{\mathrm{HS}} \leq \frac{n(2\|R\|_\infty - K)}{e^{Kt} - e^{2(K-\|R\|_\infty)t}}P_t|\nabla f|^2;$$

(ii) *if* $\mathrm{Ric}_V = K > 0$*, then for any* $f \in C_b^2(M)$ *and* $t \geq 0$*,*

$$|\mathrm{Hess}_{P_t f}|_{\mathrm{HS}}^2 \leq \frac{2\|\tilde{R}\|_\infty - K}{\mathrm{e}^{Kt} - \mathrm{e}^{2(K-\|\tilde{R}\|_\infty)t}} P_t |\nabla f|^2.$$

Proof. These items are direct consequences of Theorem 2.1. The second assertion can also be proved by an argument as in [25, Theorem 4.1] with some straightforward modifications. □

2.2 Hessian estimate of semigroup: type II

We now introduce the following another type of Hessian estimate of semigroup.

Theorem 2.5 (Hessian estimate: type II)**.** *Assume that* $\mathrm{Ric}_V \geq K > 0$*,* $\alpha_1 := \|R\|_\infty < \infty$ *(or* $\alpha_2 := \|\tilde{R}\|_\infty < \infty$ *) and*

$$\beta := \|\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^* R + R(\nabla V)\|_\infty < \infty.$$

Then for $f \in C_b^2(M)$*,*

$$|\mathrm{Hess}_{P_t f}| \leq \left(\frac{\mathrm{e}^{-\frac{K}{2}t}}{\sqrt{\int_0^t \mathrm{e}^{Kr}\,\mathrm{d}r}} + \frac{\alpha_1 \mathrm{e}^{-\frac{K}{2}t}}{\sqrt{K}} \right) (P_t |\nabla f|^2)^{1/2} + \frac{\beta \mathrm{e}^{-\frac{K}{2}t}}{K} (P_t |\nabla f|).$$

Moreover, if $\mathrm{Ric}_V = K$*, then for* $f \in C_b^2(M)$*,*

$$|\mathrm{Hess}_{P_t f}|_{\mathrm{HS}} \leq \left(\frac{\mathrm{e}^{-\frac{K}{2}t}}{\sqrt{\int_0^t \mathrm{e}^{Kr}\,\mathrm{d}r}} + \frac{\alpha_2 \mathrm{e}^{-\frac{K}{2}t}}{\sqrt{K}} \right) (P_t |\nabla f|^2)^{1/2} + \frac{n\beta \mathrm{e}^{-\frac{K}{2}t}}{K} (P_t |\nabla f|).$$

To prove this theorem, we need the following Hessian and gradient formula of semigroup which is similar to [21, Theorem 11.6] with a difference by introducing $W^k(\cdot,\cdot) : TM \times TM \to M$.

Lemma 2.6. *Let* ρ *be the Riemannian distance to a fixed point* $o \in M$*. Assume that*

$$\lim_{\rho \to \infty} \frac{\log\left(\left| \mathrm{d}^* R + \nabla \mathrm{Ric}_V^\sharp + R(\nabla V) \right| + |R| \right)}{\rho^2} = 0,$$

and

$$\mathrm{Ric}_V \geq -h(\rho) \quad \text{for some positive function } h \in C([0,\infty)) \text{ such that } \lim_{r\to\infty} \frac{h(r)}{r^2} = 0.$$

Then for $k \in C^1([0,t])$ *with* $k(0) = 1$ *and* $k(t) = 0$*,*

$$\mathrm{Hess}_{P_t f}(v, w) = \mathbb{E}^x \left[-\nabla f(Q_t(v)) \int_0^t \langle Q_s(\dot{k}(s)w),\ /\!/_s \mathrm{d}B_s \rangle + \langle \nabla f(X_t), W_t^k(v,w) \rangle \right],$$

for $v, w \in T_x M$*, where*

$$\begin{aligned} W_t^k(\cdot, w) :=& Q_t \int_0^t Q_r^{-1} R(/\!/_r \,\mathrm{d}B_r, Q_r(\cdot)) Q_r(k(r)w) \\ &- \frac{1}{2} Q_t \int_0^t Q_r^{-1} (\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^* R + R(\nabla V))(Q_r(\cdot), Q_r(k(r)w))\,\mathrm{d}r. \end{aligned}$$

Proof. Fixed $T > 0$, set

$$N_t(v,w) := \mathrm{Hess}_{P_{T-t}f}(Q_t(v), Q_t(w)) + \langle \nabla P_{T-t}f(X_t), W_t(v,w)\rangle.$$

Furthermore, define

$$N_t^k(v,w) = \mathrm{Hess}_{P_{T-t}f}(Q_t(v), Q_t(k(t)w)) + (\mathrm{d}P_{T-t}f)(W_t^k(v,w)).$$

According to the definition of $W_t^k(v,w)$, resp. $W_t(v,w)$, and in view of the fact that $N_t(v,w)$ is a local martingale, it is easy to see that

$$N_t^k(v,w) - \int_0^t (\mathrm{Hess}_{P_{T-s}f})(Q_s(v), Q_s(\dot{k}(s)w))\,\mathrm{d}s. \tag{2.5}$$

From the formula

$$\mathrm{d}P_{T-t}f(Q_t(v)) = \mathrm{d}P_T f(v) + \int_0^t (\mathrm{Hess}_{P_{T-s}f})(//_s\mathrm{d}B_s, Q_s(v)),$$

it follows that

$$\int_0^t (\mathrm{Hess}_{P_{T-s}f})(Q_s(v), Q_s(\dot{k}(s)w))\,\mathrm{d}s - \mathrm{d}P_{T-t}f(Q_t(v))\int_0^t \langle Q_s(\dot{k}(s)w), //_s\mathrm{d}B_s\rangle \tag{2.6}$$

is also a local martingale. Concerning the last term in (2.6), we note that

$$M_t := \mathrm{Hess}_{P_{T-t}f}(Q_t(v), Q_t(k(t)w)) + (\mathrm{d}P_{T-t}f)(W_t^k(v,w)) - \mathrm{d}P_{T-t}f(Q_t(v))\int_0^t \langle Q_s(\dot{k}(s)w), //_s\mathrm{d}B_s\rangle$$

is a local martingale as well. As explained in the proof of Theorem 2.1, the local martingale M_t is a true martingale on the time interval $[0, T-\varepsilon]$. By taking expectations, we first obtain $\mathbb{E}[M_0] = \mathbb{E}[M_{T-\varepsilon}]$ and then

$$\mathrm{Hess}_{P_T f}(v,w) = \mathbb{E}\left[-\mathrm{d}f(Q_T(v))\int_0^T \langle Q_s(\dot{k}(s)w), //_s\mathrm{d}B_s\rangle + \mathrm{d}f(W_T^k(v,w))\right]$$

by passing to the limit as $\varepsilon \downarrow 0$. □

Proof of Theorem 2.5. As $\alpha_1 := \|R\|_\infty < \infty$, $\mathrm{Ric}_Z \geq K$ for some constants K and

$$\beta := \|\mathbf{d}^* R + \nabla \mathrm{Ric}_Z^\sharp - R(Z)\|_\infty < \infty,$$

then for all $t > 0$,

$$\begin{aligned}&\mathbb{E}\left[\mathrm{d}f(Q_t(v))\int_0^t \langle Q_s(\dot{k}(s)w), //_s\mathrm{d}B_s\rangle\right]\\ &\quad\leq \mathrm{e}^{-\frac{K}{2}t}(P_t|\nabla f|^2)^{1/2}\left(\int_0^t \mathrm{e}^{-Ks}\dot{k}(s)^2\,\mathrm{d}s\right)^{1/2},\end{aligned}$$

$$
\begin{aligned}
&\mathbb{E}\left[\mathrm{d}f\left(Q_t\int_0^t Q_r^{-1}R(//_r\,\mathrm{d}B_r, Q_r(k(r)w))Q_r(v)\right)\right]\\
&\leq \alpha_1 \mathrm{e}^{-\frac{Kt}{2}}(P_t|\nabla f|^2)^{1/2}\left(\int_0^t \mathrm{e}^{-Ks}k(s)^2\,\mathrm{d}s\right)^{1/2},
\end{aligned}
$$

and

$$
\begin{aligned}
&\frac{1}{2}\mathbb{E}\left[\mathrm{d}f\left(Q_t\int_0^t Q_r^{-1}(\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V))(Q_r(k(r)w), Q_r(v))\,\mathrm{d}r\right)\right]\\
&\leq \frac{\beta}{2}\mathrm{e}^{-\frac{Kt}{2}}\left(\int_0^t \mathrm{e}^{-\frac{Ks}{2}}k(s)\,\mathrm{d}s\right)(P_t|\nabla f|).
\end{aligned}
$$

Keeping the assumptions of Lemma 2.6, we have

$$
\begin{aligned}
|\mathrm{Hess}_{P_tf}| \leq &\mathrm{e}^{-\frac{K}{2}t}(P_t|\nabla f|^2)^{1/2}\left[\left(\int_0^t \mathrm{e}^{-Ks}\dot{k}(s)^2\,\mathrm{d}s\right)^{1/2} + \alpha_1\left(\int_0^t \mathrm{e}^{-Ks}k(s)^2\,\mathrm{d}s\right)^{1/2}\right]\\
&+\frac{\beta}{2}\mathrm{e}^{-\frac{K}{2}t}(P_t|\nabla f|)\left(\int_0^t \mathrm{e}^{-\frac{K}{2}s}k(s)\,\mathrm{d}s\right).
\end{aligned}
$$

Choose the function

$$
k(s) := \frac{\int_0^s \mathrm{e}^{Kr}\,\mathrm{d}r}{\int_0^t \mathrm{e}^{Kr}\,\mathrm{d}r}.
$$

Then we obtain

$$
|\mathrm{Hess}_{P_tf}| \leq \left(1+\frac{\alpha_1}{\sqrt{K}}\sqrt{\int_0^t \mathrm{e}^{Kr}\,\mathrm{d}r}\right)\frac{\mathrm{e}^{-\frac{K}{2}t}}{\sqrt{\int_0^t \mathrm{e}^{Kr}\,\mathrm{d}r}}(P_t|\nabla f|^2)^{1/2} + \frac{\beta}{K}\mathrm{e}^{-\frac{K}{2}t}(P_t|\nabla f|).
$$

□

Corollary 2.7. *Assume that* $\alpha_1 := \|R\|_\infty < \infty$ *(or* $\alpha_2 := \|R\|_\infty < \infty$*) and* $\beta := \|\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V)\|_\infty < \infty$*. Then for any constant* $K \in \mathbb{R}$*,*

(i) *if* $\mathrm{Ric}_V \geq K > 0$*, then for any* $f \in C_b^2(M)$ *and* $t > 0$*,*

$$
|\mathrm{Hess}_{P_tf}|_{\mathrm{HS}}^2 \leq n\left(1+(\frac{\alpha_1}{\sqrt{K}}+\frac{\beta}{K})\sqrt{\int_0^t \mathrm{e}^{Kr}\,\mathrm{d}r}\right)^2\frac{\mathrm{e}^{-Kt}}{\int_0^t \mathrm{e}^{Kr}\,\mathrm{d}r}P_t|\nabla f|^2;
$$

(ii) *if* $\mathrm{Ric}_V = K > 0$*, then for any* $f \in C_b^2(M)$ *and* $t > 0$*,*

$$
|\mathrm{Hess}_{P_tf}|_{\mathrm{HS}}^2 \leq \left(1+(\frac{\alpha_2}{\sqrt{K}}+\frac{\beta n}{K})\sqrt{\int_0^t \mathrm{e}^{Kr}\,\mathrm{d}r}\right)^2\frac{\mathrm{e}^{-Kt}}{\int_0^t \mathrm{e}^{Kr}\,\mathrm{d}r}P_t|\nabla f|^2.
$$

Proof. These items are direct consequences of Theorem 2.1. The second assertion can also be proved by an argument as in [25, Theorem 4.1] with some straightforward modifications. □

In Theorem 2.5, $|\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V)|$ is uniformly bounded on the whole space. We will relax this condition by viewing $|\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V)|(x)$ as a space dependent function which satisfies some conditions. Let

$$\beta(x) = |\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V)|(x); \tag{2.7}$$

$$K_V(x) := \inf\{\mathrm{Ric}_V(v,v)(x) : \ v \in T_xM\}. \tag{2.8}$$

Theorem 2.8. *Assume that there exist* $K > 0$, $p > 1$ *and* $\delta > 0$ *such that* $K_V(x) - \frac{2(p-1)}{p}(\delta\beta(x))^{\frac{p}{p-1}} - K \geq 0$ *for all* $x \in M$. *Let* $\alpha_1 := \|R\|_\infty < \infty$. *Then for* $f \in C_b^2(M)$,

$$|\mathrm{Hess}_{P_tf}| \leq \left(1 + \frac{\alpha_1}{\sqrt{K}}\sqrt{\int_0^t \mathrm{e}^{Kr}\,\mathrm{d}r}\right)\frac{\mathrm{e}^{-\frac{K}{2}t}}{\sqrt{\int_0^t \mathrm{e}^{Kr}\,\mathrm{d}r}}(P_t|\nabla f|^2)^{1/2} + \frac{1}{\delta 2^{(p-1)/p}(pK)^{1/p}}\mathrm{e}^{-\frac{K}{2}t}P_t|\nabla f|.$$

Proof. It is easy to see from the condition that $K_V(x) \geq K > 0$, i.e. $\mathrm{Ric}_V \geq K > 0$. Following the same steps of the proof of Theorem 2.5, it suffices to estimate

$$\left|Q_t\int_0^t Q_r^{-1}(\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V))(Q_r(w), Q_r(k(r)w))\,\mathrm{d}r\right|.$$

For $p > 1$, by Itô' s formula,

$$\begin{aligned}
&\mathrm{d}\left|Q_t\int_0^t Q_r^{-1}(\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V))(Q_r(w), Q_r(k(r)w))\,\mathrm{d}r\right|^p\\
&= -\frac{p}{2}\left|Q_t\int_0^t Q_r^{-1}(\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V))(Q_r(w), Q_r(k(r)w))\,\mathrm{d}r\right|^{p-2}\\
&\quad\times \mathrm{Ric}_V\Big(Q_t\int_0^t Q_r^{-1}(\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V))(Q_r(w), Q_r(k(r)w))\,\mathrm{d}r,\\
&\qquad\quad Q_t\int_0^t Q_r^{-1}(\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V))(Q_r(w), Q_r(k(r)w))\,\mathrm{d}r\Big)\,\mathrm{d}t\\
&\quad + p\left|Q_t\int_0^t Q_r^{-1}(\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V))(Q_r(w), Q_r(k(r)w))\,\mathrm{d}r\right|^{p-2}\\
&\quad\times\Big\langle(\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V))(Q_t(w), Q_t(k(t)w)),\\
&\qquad\quad Q_t\int_0^t Q_r^{-1}(\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V))(Q_r(w), Q_r(k(r)w))\,\mathrm{d}r\Big\rangle\,\mathrm{d}t\\
&\leq -\frac{p}{2}K_V(X_t)\left|Q_t\int_0^t Q_r^{-1}(\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V))(Q_r(w), Q_r(k(r)w))\,\mathrm{d}r\right|^p\,\mathrm{d}t\\
&\quad + p\beta(X_t)|Q_t|^2k(t)\left|Q_t\int_0^t Q_r^{-1}(\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V))(Q_r(w), Q_r(k(r)w))\,\mathrm{d}r\right|^{p-1}\,\mathrm{d}t.
\end{aligned}$$

Using Young's inequality, we further obtain

$$
\begin{aligned}
&\mathrm{d}\left|Q_t\int_0^t Q_r^{-1}(\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V))(Q_r(w), Q_r(k(r)w))\,\mathrm{d}r\right|^p\\
&\le \left[(p-1)(\delta\beta(X_t))^{\frac{p}{p-1}} - \frac{p}{2}K(X_t)\right]\left|Q_t\int_0^t Q_r^{-1}(\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V))(Q_r(w), Q_r(k(r)w))\,\mathrm{d}r\right|^p \mathrm{d}t\\
&\qquad + \frac{1}{\delta^p}|Q_t(w)|^{2p}\,\mathrm{d}t\\
&\le -\frac{p}{2}K\left|Q_t\int_0^t Q_r^{-1}(\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V))(Q_r(w), Q_r(k(r)w))\,\mathrm{d}r\right|^p \mathrm{d}t + \frac{1}{\delta^p}|Q_t(w)|^{2p}\,\mathrm{d}t,
\end{aligned}
$$

which further implies

$$
\begin{aligned}
&\left|Q_{t\wedge\tau_D}\int_0^{t\wedge\tau_D} Q_r^{-1}(\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V))(Q_r(w), Q_r(k(r)w))\,\mathrm{d}r\right|\\
&\le \frac{1}{\delta}\mathrm{e}^{-\frac{1}{2}K(t\wedge\tau_D)}\left(\int_0^{t\wedge\tau_D}\mathrm{e}^{\frac{p}{2}Ks}|Q_s(w)|^{2p}\,\mathrm{d}s\right)^{1/p} \le \frac{1}{\delta}\left(\frac{2}{pK}\right)^{1/p}\mathrm{e}^{-\frac{1}{2}K(t\wedge\tau_D)},
\end{aligned}
$$

where τ_D is the first exit time of the compact set $D \subset M$. Letting D increase to M yields

$$
\left|Q_t\int_0^t Q_r^{-1}(\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V))(Q_r(w), Q_r(k(r)w))\,\mathrm{d}r\right| \le \frac{1}{\delta}\left(\frac{2}{pK}\right)^{1/p}\mathrm{e}^{-\frac{1}{2}Kt}.
$$

□

3 The HSI inequality

We first recall the formula relating relative entropy and Fisher information. From now on, we always assume that ν is a distribution which is absolutely continuous with respect to μ such that $h := \mathrm{d}\nu/\mathrm{d}\mu \in C_b^2(M)$.

Proposition 3.1. *Assume that*

$$\mathrm{Ric}_V := \mathrm{Ric} - \mathrm{Hess}_V \ge K$$

for some positive constant K. Recall that $\mathrm{d}\nu^t = P_t h\,\mathrm{d}\mu$ for $t > 0$. Then

(i) *(Integrated de Bruijn's formula)*

$$H(\nu\,|\,\mu) = \mathrm{Ent}_\mu(h) = \frac{1}{2}\int_0^\infty I_\mu(P_t h)\,\mathrm{d}t;$$

(ii) *(Exponential decay of Fisher information) for every $t \ge 0$,*

$$I_\mu(P_t h) = I(\nu^t\,|\,\mu) \le \mathrm{e}^{-Kt} I(\nu\,|\,\mu) = \mathrm{e}^{-Kt} I_\mu(h).$$

The HSI inequality connects the entropy H, the Stein discrepancy S and the Fisher information I. We first give a bound for the Fisher information by Stein's discrepancy S. More precisely, we have the following result.

Theorem 3.2. *Let ν be a distribution satisfying $\mathrm{d}\nu = h\,\mathrm{d}\mu$. Assume that $\alpha_1 := \|R\|_\infty < \infty$ (or $\alpha_2 := \|\tilde{R}\|_\infty < \infty$) and*

$$\beta := \|\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V)\|_\infty < \infty.$$

(i) *If $\mathrm{Ric}_V \geq K$, then for $t > 0$ and $f \in C_b^2(M)$,*

$$I_\mu(P_t h) \leq \Psi(t)\, S^2(\nu\,|\,\mu), \quad t > 0, \tag{3.1}$$

where $\Psi(t) = \min\{\Psi_1(t), \Psi_2(t)\}$ and

$$\Psi_1(t) := \frac{Kn}{\mathrm{e}^{2Kt} - \mathrm{e}^{Kt}} \left(1 + (\frac{\alpha_1}{\sqrt{K}} + \frac{\beta}{K}) \left(\frac{\mathrm{e}^{Kt} - 1}{K}\right)^{1/2}\right)^2;$$

$$\Psi_2(t) := \frac{(K - 2\alpha_1)n}{\mathrm{e}^{(2K-2\alpha_1)t} - \mathrm{e}^{Kt}} \left(1 + \frac{\beta}{K} \left(\frac{\mathrm{e}^{Kt} - 1}{K}\right)^{1/2}\right)^2.$$

(ii) *If $\mathrm{Ric}_V = K > 0$, then Ψ in (3.1) also can be chosen as*

$$\min\left\{\tilde{\Psi}_1(t), \tilde{\Psi}_2(t)\right\}, \quad \textit{where} \tag{3.2}$$

$$\tilde{\Psi}_1(t) := \frac{K}{\mathrm{e}^{2Kt} - \mathrm{e}^{Kt}} \left(1 + (\frac{\alpha_2}{\sqrt{K}} + \frac{\beta n}{K}) \left(\frac{\mathrm{e}^{Kt} - 1}{K}\right)^{1/2}\right)^2;$$

$$\tilde{\Psi}_2(t) := \frac{K - 2\alpha_2}{\mathrm{e}^{(2K-2\alpha_2)t} - \mathrm{e}^{Kt}} \left(1 + \frac{\beta n}{K} \left(\frac{\mathrm{e}^{Kt} - 1}{K}\right)^{1/2}\right)^2.$$

Proof. By Theorem 2.1, if $\mathrm{Ric}_V \geq K$, $\|R\|_\infty < \infty$, and $\beta < \infty$, then

$$|\mathrm{Hess}_{P_t f}|_{\mathrm{HS}}^2 \leq \Psi(t)(P_t|\nabla f|^2). \tag{3.3}$$

Let $g_t = \log P_t h$. By the symmetry of $(P_t)_{t\geq 0}$ in $L^2(\mu)$,

$$I_\mu(P_t h) = -\int (Lg_t)P_t h\,\mathrm{d}\mu = -\int (LP_t g_t)h\,\mathrm{d}\mu = -\int LP_t g_t\,\mathrm{d}\nu.$$

Hence, according to the definition of a Stein kernel, we have

$$\begin{aligned} I_\mu(P_t h) &= -\int \langle \mathrm{id}, \mathrm{Hess}_{P_t g_t}\rangle_{\mathrm{HS}}\,\mathrm{d}\nu - \int \langle \nabla V, \nabla P_t g_t\rangle\,\mathrm{d}\nu \\ &= \int \langle \tau_\nu - \mathrm{id}, \mathrm{Hess}_{P_t g_t}\rangle_{\mathrm{HS}}\,\mathrm{d}\nu \end{aligned}$$

and hence by the Cauchy-Schwartz inequality,

$$\begin{aligned} I_\mu(P_t h) &= \int \langle \tau_\nu - \mathrm{id}, \mathrm{Hess}_{P_t g_t}\rangle_{\mathrm{HS}}\,\mathrm{d}\nu \\ &\leq \left(\int |\tau_\nu - \mathrm{id}|_{\mathrm{HS}}^2\,\mathrm{d}\nu\right)^{1/2} \left(\int |\mathrm{Hess}_{P_t g_t}|_{\mathrm{HS}}^2\,\mathrm{d}\nu\right)^{1/2} \end{aligned}$$

$$\leq \left(\int |\tau_\nu - \mathrm{id}|_{\mathrm{HS}}^2 \,\mathrm{d}\nu\right)^{1/2} \left(\Psi(t) \int P_t |\nabla g_t|^2 \,\mathrm{d}\nu\right)^{1/2},$$

here we use (3.3) by taking the function $g_t = \log P_t h$ inside. Since

$$\int P_t |\nabla g_t|^2 \,\mathrm{d}\nu = \int P_t |\nabla g_t|^2 h \,\mathrm{d}\mu = \int |\nabla g_t|^2 P_t h \,\mathrm{d}\mu$$
$$= \int \frac{|\nabla P_t h|^2}{P_t h} \,\mathrm{d}\mu = I_\mu(P_t h),$$

it then follows that

$$I_\mu(P_t h) \leq \Psi(t) \int |\tau_\nu - \mathrm{id}|_{\mathrm{HS}}^2 \,\mathrm{d}\nu.$$

Taking the infimum over all Stein kernels of ν, we finish the proof of (i). The second item can be proved following the same steps as above by replacing the upper bound in (3.3) by that in (2.1). □

Corollary 3.3. *Assume that $\beta = 0$ and $\|R\|_\infty < \infty$. Let ν be a distribution satisfying $\mathrm{d}\nu = h\,\mathrm{d}\mu$ and $h \in C_b^2(M)$. Then,*

(i) *if $\mathrm{Ric}_V \geq K$, then for $t > 0$,*

$$I_\mu(P_t h) \leq \frac{n(2\|R\|_\infty - K)}{\mathrm{e}^{Kt} - \mathrm{e}^{2(K - \|R\|_\infty)t}} S(\nu\,|\,\mu)^2; \tag{3.4}$$

(ii) *if $\mathrm{Ric}_V = K$, then for $t > 0$,*

$$I_\mu(P_t h) \leq \frac{2\|\tilde{R}\|_\infty - K}{\mathrm{e}^{Kt} - \mathrm{e}^{2(K - \|\tilde{R}\|_\infty)t}} S(\nu\,|\,\mu)^2.$$

Remark 3.4. When M is a Ricci parallel manifold, ∇V is a Killing field, we have $\beta = 0$ and in this case, we observe that when $\mathrm{Ric}_V = K > 0$, both inequalities can be used to bound $I_\mu(P_t h)$. It is easy to see that when $K < 2(K - \|R\|_\infty)$, the first inequality may give a smaller upper bound as the main decay rate is $\mathrm{e}^{-2(K - \|R\|_\infty)t}$ which is faster than e^{-Kt}. When $K < 2(K - \|\tilde{R}\|_\infty)$ and if $\|\tilde{R}\|_\infty$ is small, then the second inequality is likely to give the sharper upper bound as the upper bound in (3.5) has an additional n.

When $|\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^* R + R(\nabla V)|$ is not uniformly bounded, we also have the following result.

Theorem 3.5. *Let ν be a distribution satisfying $\mathrm{d}\nu = h\,\mathrm{d}\mu$ and $h \in C_b^2(M)$. Assume that there exists $K > 0$, $p > 1$ and $\delta > 0$ such that $K_V(x) - \frac{2(p-1)}{p}(\delta\beta(x))^{\frac{p}{p-1}} \geq K$ for all $x \in M$, where K_V and β are defined as in (2.7) and (2.8). Moreover, assume that $\alpha_1 := \|R\|_\infty < \infty$. Then for $f \in C_b^2(M)$,*

$$I_\mu(P_t h) \leq n\left(1 + \Big(\frac{\alpha_1}{\sqrt{K}} + \frac{1}{\delta 2^{(p-1)/p}(pK)^{1/p}}\Big)\sqrt{\int_0^t \mathrm{e}^{Kr}\,\mathrm{d}r}\right)^2 \frac{\mathrm{e}^{-Kt}}{\int_0^t \mathrm{e}^{Kr}\,\mathrm{d}r} S(\nu\,|\,\mu)^2. \tag{3.5}$$

Using Theorem 3.2, we have the following inequality connecting the entropy H, S and I.

Theorem 3.6 (HSI inequality)**.** *Let ν be a distribution satisfying $\mathrm{d}\nu = h\,\mathrm{d}\mu$. Assume that*

$$\|\mathrm{Hess}_{P_t f}\|_{\mathrm{HS}}^2 \le \Psi(t) P_t |\nabla f|^2,$$

for some function $\Psi \in C([0,\infty))$. then

$$H(\nu\,|\,\mu) \le \frac{1}{2}\inf_{u>0}\left\{ I(\nu\,|\,\mu)\int_0^u \mathrm{e}^{-Kt}\,\mathrm{d}t + S(\nu\,|\,\mu)^2 \int_u^\infty \Psi(t)\,\mathrm{d}t \right\}.$$

Proof. By Proposition 3.1 (i), we have

$$H(\nu\,|\,\mu) = \frac{1}{2}\int_0^\infty I_\mu(P_t h)\,\mathrm{d}t.$$

Combining this with the following facts:

$$I_\mu(P_t h) \le \mathrm{e}^{-Kt} I(\nu\,|\,\mu),$$

and

$$I_\mu(P_t h) \le \Psi(t) S^2(\nu\,|\,\mu),$$

we obtain

$$H(\nu\,|\,\mu) \le \frac{1}{2}\inf_{u>0}\left\{ I(\nu\,|\,\mu)\int_0^u \mathrm{e}^{-Kt}\,\mathrm{d}t + S(\nu\,|\,\mu)^2 \int_u^\infty \Psi(t)\,\mathrm{d}t \right\}.$$

□

Remark 3.7. If β is a constant, i.e. $\|\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^* R + R(\nabla V)\|_\infty$ is bounded, then according to Remark 2.2, when $\beta = 0$, we have

$$H(\nu\,|\,\mu) \le \frac{1}{2}\inf_{u>0}\left\{ I(\nu\,|\,\mu)\int_0^u \mathrm{e}^{-Kt}\,\mathrm{d}t + nS(\nu\,|\,\mu)^2 \int_u^\infty \left(\frac{K - 2\|R\|_\infty}{\mathrm{e}^{(2K-2\|R\|_\infty)t} - \mathrm{e}^{Kt}} \right)\mathrm{d}t \right\}.$$

The following term

$$\frac{K - 2\|R\|_\infty}{\mathrm{e}^{(2K-2\|R\|_\infty)t} - \mathrm{e}^{Kt}}$$

at least has the decay rate e^{-Kt}. If $\beta \neq 0$, then the decay rate

$$n\left(\frac{K - 2\|R\|_\infty}{\mathrm{e}^{(2K-2\|R\|_\infty)t} - \mathrm{e}^{Kt}} \right)\left(1 + \frac{\beta}{K}\left(\frac{\mathrm{e}^{Kt} - 1}{K} \right)^{1/2} \right)^2$$

won't be faster than e^{-Kt}. For this case, if we use Corollary 2.7, then the decay rate

$$n\left(1 + \Big(\frac{\|R\|_\infty}{\sqrt{K}} + \frac{\beta}{K} \Big)\left(\frac{\mathrm{e}^{Kt} - 1}{K} \right)^{1/2} \right)^2 \frac{\mathrm{e}^{-Kt}}{\int_0^t \mathrm{e}^{Kr}\,\mathrm{d}r}$$

has the same rate as e^{-Kt}. From this point of view, when $\beta \neq 0$, it is better to choose the estimate from Corollary 2.7 to establish the HSI inequality.

To make the upper bounds in Theorem 3.6 more explicitly, we continue our discussion by assuming $\|\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^* R + R(\nabla V)\|$ is bounded with $\beta = 0$ and $\beta \neq 0$ separately, and is not bounded but satisfies some specific condition as in Corollary 3.5.

3.1 Case I: $\beta = 0$.

We first introduce the main result in this subsection.

Theorem 3.8. *Assume that* $\|R\|_\infty < \infty$ *and* $\beta = 0$.

(i) *If* $\mathrm{Ric}_V \geq K > 0$ *and* $\alpha := K - 2\|R\|_\infty > 0$*, then*

$$H(\nu\,|\,\mu) \leq \frac{I(\nu\,|\,\mu)}{2K}\left(1 - \left(\frac{I(\nu\,|\,\mu)}{I(\nu\,|\,\mu) + \alpha n S^2(\nu\,|\,\mu)}\right)^{K/\alpha}\right) + \frac{nS^2(\nu\,|\,\mu)}{2}\int_0^{\frac{I(\nu\,|\,\mu)}{I(\nu\,|\,\mu)+\alpha n S^2(\nu\,|\,\mu)}} \frac{r^{K/\alpha}}{1-r}\,\mathrm{d}r. \tag{3.6}$$

(i') *If* $\mathrm{Ric}_V \geq K > 0$ *and* $\alpha = K - 2\|R\|_\infty = 0$*, then*

$$H(\nu\,|\,\mu) \leq \frac{I(\nu\,|\,\mu)}{2K}\left(1 - \mathrm{e}^{-nKS^2(\nu\,|\,\mu)/I(\nu\,|\,\mu)}\right) + \frac{nS^2(\nu\,|\,\mu)}{2}\,\mathrm{li}(\mathrm{e}^{-nKS^2(\nu\,|\,\mu)/I(\nu\,|\,\mu)}),$$

where $\mathrm{li}(x) = \int_0^x \frac{1}{\ln t}\,\mathrm{d}t$ *is the logarithmic integral function.*

(ii) *If* $\mathrm{Ric}_V = K > 0$ *and* $\tilde{\alpha} := K - 2\|\tilde{R}\|_\infty > 0$*, then*

$$H(\nu\,|\,\mu) \leq \frac{I(\nu\,|\,\mu)}{2K}\left(1 - \left(\frac{I(\nu\,|\,\mu)}{I(\nu\,|\,\mu) + \tilde{\alpha} S^2(\nu\,|\,\mu)}\right)^{K/\tilde{\alpha}}\right) + \frac{S^2(\nu\,|\,\mu)}{2}\int_0^{\frac{I(\nu\,|\,\mu)}{I(\nu\,|\,\mu)+\tilde{\alpha} S^2(\nu\,|\,\mu)}} \frac{r^{K/\tilde{\alpha}}}{1-r}\,\mathrm{d}r. \tag{3.7}$$

Moreover, if $\mathrm{Hess}_V = K$*, then*

$$H(\nu\,|\,\mu) \leq \frac{1}{2}S^2(\nu\,|\,\mu)\log\left(1 + \frac{I(\nu\,|\,\mu)}{KS^2(\nu\,|\,\mu)}\right).$$

(ii') *If* $\mathrm{Ric}_V \geq K > 0$ *and* $\tilde{\alpha} = 0$*, then*

$$H(\nu\,|\,\mu) \leq \frac{I(\nu\,|\,\mu)}{2K}\left(1 - \mathrm{e}^{-KS^2(\nu\,|\,\mu)/I(\nu\,|\,\mu)}\right) + \frac{S^2(\nu\,|\,\mu)}{2}\,\mathrm{li}(\mathrm{e}^{-KS^2(\nu\,|\,\mu)/I(\nu\,|\,\mu)}),$$

where $\mathrm{li}(x) = \int_0^x \frac{1}{\ln t}\,\mathrm{d}t$ *is again the logarithmic integral function.*

Proof. We only need to prove the first two estimates (i) and (i'); then (ii) and (ii') are obtained through replacing $nS^2(\nu\,|\,\mu)$ by $S^2(\nu\,|\,\mu)$, and $\|R\|_\infty$ by $\|\tilde{R}\|_\infty$, respectively. We write $I = I(\nu\,|\,\mu)$ and $S = S(\nu\,|\,\mu)$ for simplicity. By Theorem 3.6 (i) and (ii), we have

$$\begin{aligned} H(\nu\,|\,\mu) &\leq \frac{1}{2}\inf_{u>0}\left\{I(\nu\,|\,\mu)\int_0^u \mathrm{e}^{-Kt}\,\mathrm{d}t + nS(\nu\,|\,\mu)^2\int_u^\infty \frac{\alpha}{\mathrm{e}^{Kt}(\mathrm{e}^{\alpha t}-1)}\,\mathrm{d}t\right\} \\ &= \frac{1}{2}\inf_{u>0}\left\{\frac{I(\nu\,|\,\mu)(1-\mathrm{e}^{-Ku})}{K} + nS(\nu\,|\,\mu)^2\int_0^{\mathrm{e}^{-\alpha u}} \frac{r^{K/\alpha}}{1-r}\,\mathrm{d}r\right\}. \end{aligned}$$

It is easy to see that inf is reached for $e^{\alpha u} = (\alpha n S^2 + I)/I$ so that

$$H(\nu\,|\,\mu) \le \frac{I}{2K}\left(1-\left(\frac{I}{I+\alpha n S^2}\right)^{K/\alpha}\right)+\frac{nS^2}{2}\int_0^{\frac{I}{I+\alpha n S^2}} \frac{r^{K/\alpha}}{1-r}\,\mathrm{d}r. \tag{3.8}$$

We thus obtain (i). The case $\alpha = 0$ can be dealt as limiting result of (3.8) when α tends to 0, i.e.,

$$\begin{aligned}
&\lim_{\alpha\to 0}\left\{\frac{I}{2K}\left(1-\left(\frac{I}{I+\alpha n S^2}\right)^{K/\alpha}\right)+\frac{nS^2}{2}\int_0^{\frac{I}{I+\alpha n S^2}} \frac{r^{K/\alpha}}{1-r}\,\mathrm{d}r\right\}\\
&\quad= \frac{I}{2K}\left(1-\mathrm{e}^{-Kn\frac{S^2}{I}}\right)+\frac{nS^2}{2}\lim_{\alpha\to 0}\int_0^{\frac{I}{I+\alpha n S^2}} \frac{r^{K/\alpha}}{1-r}\,\mathrm{d}r\\
&\quad= \frac{I}{2K}\left(1-\mathrm{e}^{-nK\frac{S^2}{I}}\right)+\frac{nS^2}{2}\lim_{\alpha\to 0}\int_{\frac{nS^2}{I+\alpha n S^2}}^{1/\alpha} \frac{(1-\alpha t)^{K/\alpha}}{t}\,\mathrm{d}t\\
&\quad= \frac{I}{2K}\left(1-\mathrm{e}^{-nK\frac{S^2}{I}}\right)+\frac{nS^2}{2}\int_{\frac{nS^2}{I}}^{\infty} \frac{\mathrm{e}^{-Kt}}{t}\,\mathrm{d}t\\
&\quad= \frac{I}{2K}\left(1-\mathrm{e}^{-nK\frac{S^2}{I}}\right)+\frac{nS^2}{2}\,\mathrm{li}(\mathrm{e}^{-K\frac{nS^2}{I}})
\end{aligned}$$

which proves (i').

If $\mathrm{Hess}_V = K$, by Obata's Rigidity Theorem (see [26, Theorem 6.3]), then M is isometric to $\mathbb{R}^n$, which implies $\mathrm{Ric}_V = K$, $\alpha_n = K$ and $\beta = 0$. Thus by (3.7),

$$\begin{aligned}
H(\nu\,|\,\mu) &\le \frac{I(\nu\,|\,\mu)}{2K}\left(1-\left(\frac{I(\nu\,|\,\mu)}{I(\nu\,|\,\mu)+KS^2(\nu\,|\,\mu)}\right)\right)+\frac{S^2(\nu\,|\,\mu)}{2}\int_0^{\frac{I(\nu|\mu)}{I(\nu|\mu)+KS^2(\nu|\mu)}} \frac{r}{1-r}\,\mathrm{d}r\\
&= \frac{S^2(\nu\,|\,\mu)I(\nu\,|\,\mu)}{2(I(\nu\,|\,\mu)+KS^2(\nu\,|\,\mu))}+\frac{S^2(\nu\,|\,\mu)}{2}\int_0^{\frac{I(\nu|\mu)}{I(\nu|\mu)+KS^2(\nu|\mu)}}\left(\frac{1}{1-r}-1\right)\mathrm{d}r\\
&= \frac{1}{2}S^2(\nu\,|\,\mu)\log\left(1+\frac{I(\nu\,|\,\mu)}{KS^2(\nu\,|\,\mu)}\right),
\end{aligned}$$

which covers the result in [10, Theorem 2.2] for the case $M = \mathbb{R}^n$. □

Remark 3.9. In the case $\mathrm{Ric}_V = K > 0$ and $\tilde\alpha > 0$ (which implies $\alpha > 0$), both inequalities (4.3) and (3.7) hold. Hence one may choose the one which provides the sharper estimate.

The case that $\beta = 0$ and α or $\tilde\alpha$ is less than 0, can be dealt as follows.

Theorem 3.10. *Assume that $\beta = 0$ and $\|R\|_\infty < \infty$.*

(i) *If $\mathrm{Ric}_V \ge K > 0$ and $\alpha := K - 2\|R\|_\infty < 0$, then*

$$H(\nu\,|\,\mu) \le \frac{nS^2(\nu\,|\,\mu)\max\{-\alpha,K\}}{2K}\,\Theta\left(\frac{I(\nu\,|\,\mu)}{nS^2(\nu\,|\,\mu)\max\{-\alpha,K\}}\right), \tag{3.9}$$

where

$$\Theta(r) = \begin{cases} 1+\log r, & r\ge 1;\\ r, & 0<r<1.\end{cases}$$

(ii) *If* $\mathrm{Ric}_V = K > 0$ *and* $\tilde{\alpha} := K - 2\|\tilde{R}\|_\infty < 0$*, then*

$$H(\nu\,|\,\mu) \leq \frac{S^2(\nu\,|\,\mu)\max\{-\tilde{\alpha}, K\}}{2K}\,\Theta\left(\frac{I(\nu\,|\,\mu)}{S^2(\nu\,|\,\mu)\max\{-\tilde{\alpha}, K\}}\right).$$

Proof. As $\alpha := K - 2\|R\|_\infty < 0$, we have

$$1 - \mathrm{e}^{-Ku} \leq \max\{1, -K/\alpha\}(1 - \mathrm{e}^{\alpha u}),$$

and then

$$\frac{-\alpha}{\mathrm{e}^{Ku} - \mathrm{e}^{(K+\alpha)u}} \leq \max\{-\alpha, K\}\,\frac{1}{\mathrm{e}^{Ku} - 1},$$

which implies

$$\begin{aligned} H(\nu\,|\,\mu) &\leq I(\nu\,|\,\mu)\frac{1 - \mathrm{e}^{-Ku}}{2K} + \frac{n}{2}S^2(\nu\,|\,\mu)\int_u^\infty \max\{-\alpha, K\}\,\frac{1}{\mathrm{e}^{Kt} - 1}\,\mathrm{d}t \\ &= I(\nu\,|\,\mu)\frac{1 - \mathrm{e}^{-Ku}}{2K} - \frac{n}{2K}S^2(\nu\,|\,\mu)\max\{-\alpha, K\}\ln(1 - \mathrm{e}^{-Ku}). \end{aligned}$$

This further implies

$$\begin{aligned} H(\nu\,|\,\mu) &\leq \frac{1}{2}\inf_u\left\{I(\nu\,|\,\mu)\frac{1 - \mathrm{e}^{-Ku}}{K} - \frac{nS^2(\nu\,|\,\mu)}{K}\max\{-\alpha, K\}\ln(1 - \mathrm{e}^{-Ku})\right\} \\ &= \frac{nS^2(\nu\,|\,\mu)\max\{-\alpha, K\}}{2K}\,\Theta\left(\frac{I(\nu\,|\,\mu)}{nS^2(\nu\,|\,\mu)\max\{-\alpha, K\}}\right). \end{aligned}$$

□

3.2 Case II : $\beta \neq 0$.

We first introduce the the main theorem in this subsection, which also provides general way to establish the HSI inequality.

Theorem 3.11. *Assume that* $\alpha_1 := \|R\|_\infty < \infty$, $\beta := \|\nabla\mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V)\|_\infty < \infty$. *Let* $\mathrm{d}\nu = h\,\mathrm{d}\mu$ *for* $h \in C_0^\infty(M)$.

(i) *If* $\mathrm{Ric}_V \geq K$*, then*

$$H(\nu\,|\,\mu) \leq \frac{n\,(1+\varepsilon)\,S^2(\nu\,|\,\mu)}{2\varepsilon}\left[c_0 + \Theta\left(\frac{\varepsilon I(\nu\,|\,\mu)}{n\,(1+\varepsilon)\,KS^2(\nu\,|\,\mu)} - c_0\right)\right],$$

for any $\varepsilon > 0$*, where*

$$c_0 = \frac{\varepsilon(\alpha_1\sqrt{K} + \beta)^2}{K^3} - 1.$$

Moreover, if $\alpha_1 = 0$ *and* $\beta = 0$*, then*

$$H(\nu\,|\,\mu) \leq \frac{n}{2}S^2(\nu\,|\,\mu)\ln\left(1 + \frac{I}{nKS^2(\nu\,|\,\mu)}\right).$$

(ii) *If* $\mathrm{Ric}_V = K$*, then*

$$H(\nu\,|\,\mu) \le \frac{(1+\varepsilon)\,S^2(\nu\,|\,\mu)}{2\varepsilon}\left[\tilde{c}_0 + \Theta\left(\frac{\varepsilon I(\nu\,|\,\mu)}{(1+\varepsilon)\,KS^2(\nu\,|\,\mu)} - \tilde{c}_0\right)\right],$$

for any $\varepsilon > 0$*, where*

$$\tilde{c}_0 = \frac{\varepsilon(\alpha_2\sqrt{K} + n\beta)^2}{K^3} - 1.$$

Moreover, if $\alpha_2 = \beta = 0$*, then*

$$H(\nu\,|\,\mu) \le \frac{1}{2} S^2(\nu\,|\,\mu) \ln\left(1 + \frac{I}{KS^2(\nu\,|\,\mu)}\right).$$

Proof. We only need to prove the first estimate. Denote $I = I(\nu\,|\,\mu)$ and $S = S(\nu\,|\,\mu)$ for simplicity. By Theorem 3.2, we have

$$\begin{aligned}
I_\mu(P_t h) &\le n\left(\frac{1}{\sqrt{\int_0^t \mathrm{e}^{Kr}\,\mathrm{d}r}} + \frac{\alpha_1}{\sqrt{K}} + \frac{\beta}{K}\right)^2 \mathrm{e}^{-Kt} S^2(\nu\,|\,\mu)\\
&\le n\left(1 + \frac{1}{\varepsilon}\right) S^2(\nu\,|\,\mu) \frac{\mathrm{e}^{-Kt}}{\int_0^t \mathrm{e}^{Kr}\,\mathrm{d}r} + n(1+\varepsilon)\left(\frac{\alpha_1}{\sqrt{K}} + \frac{\beta}{K}\right)^2 \mathrm{e}^{-Kt} S^2(\nu\,|\,\mu)
\end{aligned}$$

for any $\varepsilon > 0$. Using this inequality, we first need to estimate

$$\begin{aligned}
H(\nu\,|\,\mu) &\le \frac{1}{2}\inf_{u>0}\left\{A\int_0^u \mathrm{e}^{-Kt}\,\mathrm{d}t + B\int_u^\infty \frac{K}{\mathrm{e}^{Kt}(\mathrm{e}^{Kt}-1)}\,\mathrm{d}t + C\int_u^\infty \mathrm{e}^{-Kt}\,\mathrm{d}t\right\}\\
&= \frac{1}{2}\inf_{u>0}\left\{\frac{A(1-\mathrm{e}^{-Ku}) + C\mathrm{e}^{-Ku}}{K} + B\int_0^{\mathrm{e}^{-Ku}} \frac{r}{1-r}\,\mathrm{d}r\right\},
\end{aligned}$$

where

$$\begin{aligned}
A &= I(\nu\,|\,\mu); \quad B = n\left(1 + \frac{1}{\varepsilon}\right) S^2(\nu\,|\,\mu);\\
C &= n(1+\varepsilon)\left(\frac{\alpha_1}{\sqrt{K}} + \frac{\beta}{K}\right)^2 S^2(\nu\,|\,\mu).
\end{aligned}$$

It is easy to see that if $A \le C$, then inf is reached when u goes to ∞; if $A > C$, then inf is reached for $\mathrm{e}^{Ku} = \frac{A-C+BK}{A-C}$ so

$$H(\nu\,|\,\mu) \le \frac{C}{2K} + \frac{B}{2}\ln\left[1 + \frac{A-C}{BK}\right].$$

We then conclude that

$$H(\nu\,|\,\mu) \le \frac{B}{2}\left[c_0 + \Phi\left(\frac{A}{BK} - c_0\right)\right], \tag{3.10}$$

where

$$c_0 = \frac{C - BK}{BK} = \frac{\varepsilon(\alpha_1 \sqrt{K} + \beta)^2}{K^3} - 1.$$

The proof of (ii) is the same by taking B with

$$\left(1 + \frac{1}{\varepsilon}\right) S^2(\nu\,|\,\mu),$$

and C with

$$(1+\varepsilon)\left(\frac{\alpha_2}{\sqrt{K}} + \frac{n\beta}{K}\right)^2 S^2(\nu\,|\,\mu),$$

we omit the details here. □

3.3 Case III: $|\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^* R + R(\nabla V)|$ is not bounded

For the case $|\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^* R + R(\nabla V)|$ may not bounded on whole space M, we have the following result by using Theorem 3.5.

Theorem 3.12. *Assume that there exists $K > 0$, $p > 1$ and $\delta > 0$ such that $K_V(x) - \frac{2(p-1)}{p}(\delta\beta(x))^{\frac{p}{p-1}} - K \geq 0$ for all $x \in M$. Let $\alpha_1 := \|R\|_\infty < \infty$. Then for $f \in C_b^2(M)$,*

$$H(\nu\,|\,\mu) \leq \frac{n^2\,(1+\varepsilon)\,S^2(\nu\,|\,\mu)}{2\varepsilon}\left[\tilde{c}_0 + \Theta\left(\frac{\varepsilon I(\nu\,|\,\mu)}{n^2\,(1+\varepsilon)\,KS^2(\nu\,|\,\mu)} - \tilde{c}_0\right)\right],$$

for any $\varepsilon > 0$, where

$$\tilde{c}_0 = \frac{\varepsilon}{K}\Big(\frac{\alpha_1}{\sqrt{K}} + \frac{1}{\delta 2^{(p-1)/p}(pK)^{1/p}}\Big)^2 - 1.$$

Proof. By Theorem 3.5, taking

$$A = I(\nu\,|\,\mu); \quad B = n\left(1 + \frac{1}{\varepsilon}\right) S^2(\nu\,|\,\mu);$$

$$C = n(1+\varepsilon)\Big(\frac{\alpha_1}{\sqrt{K}} + \frac{1}{\delta 2^{(p-1)/p}(pK)^{1/p}}\Big)^2 S^2(\nu\,|\,\mu)$$

into the inequality (3.10), we complete the proof. □

3.4 Examples

In order to elucidate the conditions in Theorem 3.8 and Theorem 3.10 we consider some examples. For simplicity, we restrict ourselfs to the case when $\beta = 0$. For the case $\beta > 0$, one may work out specific examples by using Theorem 3.8 directly.

Example 3.13. Let $M = \mathbb{R}^n$. Consider the operator $L = \Delta - x \cdot \nabla$. We have $\mathrm{Ric}_V = 1$, $R = 0$ and $\nabla V = x$. Then $\mu(\mathrm{d}x) = (2\pi)^{-n/2} \mathrm{e}^{-|x|^2/2}\,\mathrm{d}x$, and by Theorem 3.8 (ii), we have

$$H(\nu\,|\,\mu) \leq \frac{1}{2} S^2(\nu\,|\,\mu) \log\left(1 + \frac{I(\nu\,|\,\mu)}{S^2(\nu\,|\,\mu)}\right),$$

which covers the result in [10].

Example 3.14. Let $M = \mathbb{R}$. We consider a family of diffusion operator on the line of the type

$$Lf = f'' - u'f'$$

associated with a symmetric invariant probability measure $\mathrm{d}\mu = \mathrm{e}^{-u}\,\mathrm{d}x$, where u is a smooth potential on $\mathbb{R}$. It is easy to see that Ric $= 0$ and $R = 0$. Then

$$\mathrm{Ric}_V = u'', \quad \nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^*R + R(\nabla V) = u'''$$

So if there exists $K > 0$, $p > 1$ and $\delta > 0$ such that $u'' - \frac{2(p-1)}{p}|\delta u'''|^{\frac{p}{p-1}} \geq K > 0$, then

$$\begin{aligned} &H(\nu\,|\,\mu) \\ &\leq \frac{(1+\varepsilon)\,S^2(\nu\,|\,\mu)}{2\varepsilon}\left[\frac{\varepsilon}{\delta^2 2^{2(p-1)/p}(pK)^{2/p}K} - 1 + \Theta\left(\frac{\varepsilon I(\nu\,|\,\mu)}{(1+\varepsilon)\,KS^2(\nu\,|\,\mu)} - \frac{\varepsilon}{\delta^2 2^{2(p-1)/p}(pK)^{2/p}K} + 1\right)\right] \end{aligned}$$

for any $\varepsilon > 0$.

In particular, if $\varepsilon = \delta^2 2^{2(p-1)/p}(pK)^{2/p}K$, then

$$H(\nu\,|\,\mu) \leq \frac{\left(1 + \delta^2 2^{2(p-1)/p}(pK)^{2/p}K\right)S^2(\nu\,|\,\mu)}{\delta^2 2^{1+2(p-1)/p}(pK)^{2/p}K}\Theta\left(\frac{\delta^2 2^{2(p-1)/p}(pK)^{2/p}I(\nu\,|\,\mu)}{(1+\delta^2 2^{2(p-1)/p}(pK)^{2/p}K)\,S^2(\nu\,|\,\mu)}\right).$$

For instance, $u = \frac{1}{2}(x^2 + ax^4)$ for $a > 0$. Then $u'' = 1 + 6ax^2$ and $u''' = 12ax$. It is easy to see that $|u'''|$ is bounded. Let $p = 2$ and $\delta^2 = \frac{1}{24a}$. Then

$$u'' - (\delta u''')^2 \geq 1$$

and

$$H(\nu\,|\,\mu) \leq \frac{1}{2}(1+6a)S^2(\nu\,|\,\mu)\Theta\left(\frac{I(\nu\,|\,\mu)}{(6a+1)\,S^2(\nu\,|\,\mu)}\right).$$

Meanwhile, from [10, Proposition 4.5], it asks the following conditions are all satisfied: there exists $c > 0$ such that

$$\begin{aligned} &u'' \geq c, \\ &u^{(4)} - u'u''' + 2(u'')^2 - 6cu'' \geq 0, \\ &3(u''')^2 \leq 2(u''-c)(u^{(4)} - u'u''' + 2(u'')^2 - 6cu''). \end{aligned}$$

Then,

$$H(\nu\,|\,\mu) \leq \frac{1}{2}S^2(\nu\,|\,\mu)\,\Theta\left(\frac{I(\nu\,|\,\mu)}{cS^2(\nu\,|\,\mu)}\right).$$

So this result depends on properly choosing c and the cost of computation is higher compared with our conditions.

Example 3.15. Let $M = \mathbb{S}^n$. Consider the operator $L = \Delta$ with $V \equiv 0$ and let $\mu(\mathrm{d}x) = \mathrm{vol}(\mathrm{d}x)/\mathrm{vol}(M)$. Then $R_{ijk\ell} = (\delta_{ik}\delta_{j\ell} - \delta_{i\ell}\delta_{jk})$, Ric $= n-1$, $\|\tilde{R}\|_\infty = \sqrt{2n(n-1)}$ and

$$\alpha = K - 2\|\tilde{R}\|_\infty = (n-1) - 2\sqrt{2n(n-1)} < 0.$$

By Theorem 3.10, we have

$$H(\nu\,|\,\mu) \le \frac{(2\sqrt{2n(n-1)} - (n-1))}{2(n-1)} S^2(\nu\,|\,\mu)\,\Theta\left(\frac{I(\nu\,|\,\mu)}{(2\sqrt{2n(n-1)} - (n-1))S^2(\nu\,|\,\mu)}\right).$$

On the other hand, let us introduce the following result if we apply the method from [10]. Let

$$\begin{aligned}
\Gamma_1(f, g) &:= \langle \nabla f, \nabla g\rangle;\\
\Gamma_2(f, g) &:= \mathrm{Ric}_V(\nabla f, \nabla g) + \langle \mathrm{Hess}_f, \mathrm{Hess}_g\rangle_{\mathrm{HS}};\\
\Gamma_3(f, g) &:= \frac{1}{2}\Big[L\Gamma_2(f, g) - \Gamma_2(Lf, g) - \Gamma_2(f, Lg)\Big]
\end{aligned}$$

We use the method from [10, Theorem 4.1] to conclude that

Theorem 3.16. *if there exist positive constants κ, ρ and σ such that*

$$\Gamma_2(f) \ge \rho\Gamma_1(f); \quad \Gamma_3(f) \ge \kappa\Gamma_2(f); \quad \Gamma_2(f) \ge \sigma|\mathrm{Hess}_f|^2_{\mathrm{HS}}.$$

Then

$$H(\nu\,|\,\mu) \le \frac{1}{2\sigma} S^2(\nu\,|\,\mu)\,\Theta\left(\frac{\sigma\max\{\rho,\kappa\} I(\nu\,|\,\mu)}{\rho\kappa S^2(\nu\,|\,\mu)}\right).$$

For general Riemmanian case, we see that the most difficulty for people to use this theorem is to check if there exists $\kappa > 0$ such that $\Gamma_3(f) \ge \kappa\Gamma_2(f)$. For the special case $\mathbb{S}^n$,

$$\begin{aligned}
&\Gamma_2(f) = (n-1)|\nabla f|^2 + |\mathrm{Hess}_f|^2_{\mathrm{HS}} \ge (n-1)|\nabla f|^2;\\
&\Gamma_3(f) = (n-1)[(n-1)|\nabla f|^2 + |\mathrm{Hess}_f|^2_{\mathrm{HS}}] + \frac{1}{2}|\nabla \mathrm{Hess}_f|^2 + 2(n-1)|\mathrm{Hess}_f|^2_{\mathrm{HS}} - 2\langle \mathrm{Hess}_f(R^{\sharp,\sharp}), \mathrm{Hess}_f\rangle\\
&\qquad \ge \min\{(3(n-1) - 2\|\tilde{R}\|_\infty), (n-1)\}\Gamma_2(f) \ge (3(n-1) - 2\|\tilde{R}\|_\infty)\Gamma_2(f);\\
&\Gamma_2(f) \ge |\mathrm{Hess}_f|^2_{\mathrm{HS}}.
\end{aligned}$$

Thus $\rho = (n-1)$, $\sigma = 1$, and $\kappa = \min\{(3(n-1) - 2\|\tilde{R}\|_\infty), (n-1)\}$. If $\kappa = 3(n-1) - 2\sqrt{2n(n-1)} > 0$, i.e. $n \ge 9$, by Theorem 3.16, we have

$$H(\nu\,|\,\mu) \le \frac{1}{2} S^2(\nu\,|\,\mu)\,\Theta\left(\frac{I(\nu\,|\,\mu)}{[3(n-1) - 2\sqrt{2n(n-1)}]S^2(\nu\,|\,\mu)}\right).$$

We first observe that this inequality holds for $n \ge 0$ and when

$$I(\nu\,|\,\mu) \le (3(n-1) - 2\sqrt{2n(n-1)})S^2(\nu\,|\,\mu),$$

this inequality can not become the classical log-Sobolev inequality. But our HSI inequality turely improve the classical log-Sobolev inequality. In particular, for general Riemannian case, if $|R|$ is small such that $K - 2\|R\|_\infty > 0$, the HSI inequality improve the classical HI inequality certainly no matter if $S^2(\nu\,|\,\mu)$ is small enough or not.

Example 3.17. Let G be a n-dimensional Lie group with a bi-invariant metric g. Let $\mathfrak{g}$ be its Lie algebra. Let $L = \Delta - \nabla V$ for $V \in C^2(M)$ such that $\mu(\mathrm{d}x) = \mathrm{e}^{-V(x)}\mathrm{d}x$. Then for $X, Y, Z \in \mathfrak{g}$,

$$\nabla_X Y = \frac{1}{2}[X, Y] \quad \text{and} \quad R(X, Y)Z = \frac{1}{4}[Z, [X, Y]].$$

By the Jacobi identity, we have

$$\begin{aligned}&(\nabla \mathrm{Hess}_V + R(\nabla V))(X, Y)\\&= \nabla_X(\nabla_Y \nabla V) - \nabla_{\nabla_X^Y} \nabla V + R(\nabla V, X)Y\\&= \frac{1}{4}[X, [Y, \nabla V]] + \frac{1}{4}[\nabla V, [X, Y]] + \frac{1}{4}[Y, [\nabla V, X]] = 0.\end{aligned}$$

We conclude that if G is a Ricci parallel Lie group with $\mathrm{Ric}_V \geq K > 0$ and $\|R\|_\infty < \infty$, then the inequalities in Theorem 3.8 (i) and Theorem 3.10 (i) hold. When the condition $\mathrm{Ric}_V = K > 0$ is satisfied, both of the inequalities in Theorems 3.8 and 3.10 (i) and (ii) hold true.

4 The WS inequality and HWSI inequality

Denote by $\mathscr{P}(M)$ the set of probability measures on M. For $\mu_1, \mu_2 \in \mathscr{P}(M)$ the L^2-Wasserstein distance is given by

$$\mathbb{W}_2(\mu_1, \mu_2) := \inf_{\pi \in \mathscr{C}(\mu_1, \mu_2)} \left(\int_{M \times M} \rho(x, y)^2 \, \mathrm{d}\pi(x, y) \right)^{1/2}$$

where ρ denotes the Riemannian distance on M and $\mathscr{C}(\mu_1, \mu_2)$ consists of all couplings of μ_1 and μ_2. The Wasserstein distance has various characterizations and plays an important role in the study of SDEs, partial differential equations, optimal transportation problems, etc. For more background, one may consult [18, 23, 24] and the references therein. The following Theorem describes the relationship between Wasserstein distance and Stein discrepancy.

Theorem 4.1 (WS inequality)**.** *Assume that* $\mathrm{Ric}_V \geq K > 0$, $\alpha_1 := \|R\|_\infty < \infty$ *(or* $\alpha_2 := \|\tilde{R}\|_\infty < \infty$*) and*

$$\beta := \|\nabla \mathrm{Ric}_V^\sharp + \mathrm{d}^* R + R(\nabla V)\|_\infty < \infty.$$

Then for $\nu \in \mathscr{P}(M)$ *satisfying* $\mathrm{d}\nu/\mathrm{d}\mu \in C_b^2(M)$*, we have*

$$\mathbb{W}_2(\nu, \mu) \leq \left(\int_0^\infty \sqrt{\Psi(t)} \, \mathrm{d}t \right) S(\nu \,|\, \mu),$$

where Ψ *is defined as in* (3.5) *(and also as in* (3.2) *when* $\mathrm{Ric} = K > 0$*).*

Proof. Recall that $h = \mathrm{d}\nu/\mathrm{d}\mu \in C_b^2(M)$ and let $\mathrm{d}\nu^t = P_t h \, \mathrm{d}\mu$. Dividing $\mathbb{W}_2(\nu, \nu^t)$ by t and the formula in [15, Lemma 2] or [22, Theorem 24.2(iv)], we obtain

$$\frac{\mathrm{d}^+}{\mathrm{d}t} \mathbb{W}_2(\nu, \nu^t) \leq \left(\int_M \frac{|\nabla P_t h|^2}{P_t h} \, \mathrm{d}\mu \right)^{1/2} = I_\mu(P_t h)^{1/2}, \tag{4.1}$$

where $\frac{\mathrm{d}^+}{\mathrm{d}t}$ stands for the upper right derivative. On the other hand, by Theorem 3.2,

$$I_\mu(P_t h) \le \Psi(t) S(\nu\,|\,\mu)^2.$$

Combining this with (4.1), we obtain

$$\mathbb{W}_2(\nu,\mu) \le \int_0^\infty (I_\mu(P_t h))^{1/2}\,\mathrm{d}t \le S(\nu\,|\,\mu)\int_0^\infty \sqrt{\Psi(t)}\,\mathrm{d}t. \qquad \square$$

Corollary 4.2. *Assume that M is a Ricci parallel manifold, ∇V is a Killing field and $\|R\|_\infty < \infty$. Let $\nu \in \mathscr{P}(M)$ satisfying $\mathrm{d}\nu/\mathrm{d}\mu \in C_b^2(M)$.*

(i) *If* $\mathrm{Ric}_V \ge K > 0$, *then*

$$\mathbb{W}_2(\nu,\mu) \le \left(\int_0^\infty \sqrt{\frac{n(2\|R\|_\infty - K)}{\mathrm{e}^{Kt} - \mathrm{e}^{2(K-\|R\|_\infty)t}}}\,\mathrm{d}t\right) S(\nu\,|\,\mu);$$

(ii) *if* $\mathrm{Ric}_V = K > 0$, *then*

$$\mathbb{W}_2(\nu,\mu) \le \left(\int_0^\infty \sqrt{\frac{2\|\tilde{R}\|_\infty - K}{\mathrm{e}^{Kt} - \mathrm{e}^{2(K-\|\tilde{R}\|_\infty)t}}}\,\mathrm{d}t\right) S(\nu\,|\,\mu).$$

One may compare this inequality with the classical Talagrand-type transportation cost inequality

$$\mathbb{W}_2(\nu,\mu)^2 \le \frac{1}{2K} H(\nu\,|\,\mu). \tag{4.2}$$

We can go further and improve this inequality to the following HWSI inequality by assuming $\beta = 0$.

Theorem 4.3 (HWSI inequality)**.** *Assume that $\|R\|_\infty < \infty$ and $\beta = 0$. If $\mathrm{Ric}_V \ge K > 0$ and $\alpha := K - 2\|R\|_\infty > 0$. Let $\mathrm{d}\nu = h\,\mathrm{d}\mu$. Then*

$$\mathbb{W}_2(\nu,\,\mu) \le \frac{S(\nu\,|\,\mu)}{2K}\int_0^{L^{-1}\left(\frac{2KH(\nu|\mu)}{S^2(\nu|\mu)}\right)} \frac{1}{\sqrt{y}}\left(1 - \left(\frac{y}{y+\alpha n}\right)^{K/\alpha}\right)\mathrm{d}y$$

where

$$L(x) = x + Kn\int_0^x \frac{r^{K/\alpha-1}(r-x)}{(r+\alpha n)^{K/\alpha+1}}\,\mathrm{d}r.$$

Remark 4.4. Since $L(r) \le r$ for $r \ge 0$, this inequality improves the Talagrand quadratic transportation cost inequality (4.2).

Proof of Theorem 4.3. Recall that $\mathrm{d}\nu^t = P_t h\,\mathrm{d}\mu$. Then

$$H(\nu^t\,|\,\mu) = \frac{1}{2}\int_0^\infty I_\mu(P_{s+t}h)\,\mathrm{d}s.$$

Togather with Proposition 3.1 implies

$$\begin{aligned} H(\nu^t\,|\,\mu) &\le \frac{1}{2}\inf_{u>0}\left\{\{I(\nu^t\,|\,\mu)\int_0^u \mathrm{e}^{-Ks}\,\mathrm{d}s + S(\nu\,|\,\mu)^2\int_{u+t}^\infty \Psi(s)\,\mathrm{d}s\right\} \\ &\le \frac{1}{2}\inf_{u>0}\left\{\{I(\nu^t\,|\,\mu)\int_0^u \mathrm{e}^{-Ks}\,\mathrm{d}s + S(\nu\,|\,\mu)^2\int_{u}^\infty \Psi(s)\,\mathrm{d}s\right\} \end{aligned}$$

If $\beta = 0$, $\alpha = K - 2\|R\|_\infty \ge 0$ and $\Psi(s) = \frac{\alpha n}{\mathrm{e}^{Ks}(\mathrm{e}^{\alpha s}-1)}$, then

$$\begin{aligned} H(\nu^t\,|\,\mu) &\le \frac{I(\nu^t\,|\,\mu)}{2K}\left(1 - \left(\frac{I(\nu^t\,|\,\mu)}{I(\nu^t\,|\,\mu) + \alpha n S^2(\nu\,|\,\mu)}\right)^{K/\alpha}\right) \\ &\quad + \frac{nS^2(\nu\,|\,\mu)}{2}\int_0^{\frac{I(\nu^t\,|\,\mu)}{I(\nu^t\,|\,\mu)+\alpha n S^2(\nu\,|\,\mu)}} \frac{r^{K/\alpha}}{1-r}\,\mathrm{d}r \\ &= \frac{S^2(\nu\,|\,\mu)}{2K} L\left(\frac{I(\nu^t\,|\,\mu)}{S^2(\nu\,|\,\mu)}\right), \end{aligned} \tag{4.3}$$

where

$$L(x) = x + Kn\int_0^x \frac{r^{K/\alpha-1}(r-x)}{(r+\alpha n)^{K/\alpha+1}}\,\mathrm{d}r.$$

It is easy to see that

$$L'(x) = 1 - \left(\frac{x}{x+\alpha n}\right)^{K/\alpha} > 0$$

for $x > 0$. Then L^{-1} exists and

$$I(\nu^t\,|\,\mu) \ge S^2(\nu\,|\,\mu) L^{-1}\left(\frac{2KH(\nu^t\,|\,\mu)}{S^2(\nu\,|\,\mu)}\right).$$

Dividing $\mathbb{W}_2(\mu, \nu^t)$ by t and using the above estimate, we have

$$\begin{aligned} \frac{\mathrm{d}^+}{\mathrm{d}t}\mathbb{W}_2(\mu,\nu^t) \le I_\mu(P_t h)^{1/2} &= \frac{\frac{\mathrm{d}}{\mathrm{d}t}H(\nu^t\,|\,\mu)}{\sqrt{I(\nu^t\,|\,\mu)}} \\ &\le \frac{-\frac{\mathrm{d}}{\mathrm{d}t}H(\nu^t\,|\,\mu)}{S(\nu\,|\,\mu)\sqrt{L^{-1}\left(\frac{2KH(\nu^t\,|\,\mu)}{S^2(\nu\,|\,\mu)}\right)}}. \end{aligned}$$

Therefore, integrating both sides from 0 to ∞ yields

$$\mathbb{W}_2(\nu,\mu) \le \int_0^\infty \frac{-\frac{\mathrm{d}}{\mathrm{d}t}H(\nu^t\,|\,\mu)}{S(\nu\,|\,\mu)\sqrt{L^{-1}\left(\frac{2KH(\nu^t\,|\,\mu)}{S^2(\nu\,|\,\mu)}\right)}}$$

$$= S(\nu\,|\,\mu)\int_0^{\frac{H(\nu|\mu)}{S^2(\nu|\mu)}} \frac{\mathrm{d}x}{\sqrt{L^{-1}(2Kx)}}$$
$$= \frac{S(\nu\,|\,\mu)}{2K}\int_0^{L^{-1}\left(\frac{2KH(\nu|\mu)}{S^2(\nu|\mu)}\right)} \frac{1}{\sqrt{y}}\left(1-\left(\frac{y}{y+\alpha n}\right)^{K/\alpha}\right)\mathrm{d}y.$$

□

In particular, if $\mathrm{Hess}_V = K$ for some positive constant K, by Obata's Rigidity Theorem (see [26, Theorem 6.3]), then M is isometric to $\mathbb{R}^n$, then we have

Corollary 4.5. *Assume that* $\mathrm{Hess}_V = K > 0$. *Let* $\mathrm{d}\nu = h\mathrm{d}\mu$. *Then*

$$\mathbb{W}_2(\nu,\mu) \le \frac{S(\nu\,|\,\mu)}{K^{1/2}} \arccos\left(\exp\left(-\frac{H(\nu\,|\,\mu)}{S^2(\nu\,|\,\mu)}\right)\right).$$

Proof. As $\mathrm{Hess}_V = K$, we know that M is isometric to $\mathbb{R}^n$. First, repeating the same steps of proof of Theorem 4.3 by puting $\Psi(t) = \frac{K}{\mathrm{e}^{Kt}(\mathrm{e}^{Kt}-1)}$. By this and (4.1), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{W}_2(\nu,\nu^t) \le \sqrt{I(\nu^t\,|\,\mu)} \le -\frac{\frac{\mathrm{d}}{\mathrm{d}t}H(\nu^t\,|\,\mu)}{\sqrt{K}S(\nu\,|\,\mu)\sqrt{\exp\left(\frac{2H(\nu^t|\mu)}{S^2(\nu|\mu)}\right)-1}}$$
$$= -\frac{\mathrm{d}}{\mathrm{d}t}\left\{\frac{S(\nu\,|\,\mu)}{K^{1/2}}\arccos\left(\exp\left(-\frac{H(\nu^t\,|\,\mu)}{S^2(\nu\,|\,\mu)}\right)\right)\right\}.$$

Consequently,

$$\mathbb{W}_2(\nu,\mu) = \int_0^\infty \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{W}_2(\mu,\nu^t)\,\mathrm{d}t \le \frac{S(\nu\,|\,\mu)}{K^{1/2}}\arccos\left(\exp\left(-\frac{H(\nu\,|\,\mu)}{S^2(\nu\,|\,\mu)}\right)\right).$$ □

5 Moment bounds and Stein discrepancy

In [10], the authors investigate another feature of Stein's discrepancy applied to concentration inequalities on $\mathbb{R}^d$. It is well known that the classical log-Sobolev inequalities on the manifolds is a powerful tool towards the invariant measure. In this section, we extend to discuss the Stein discrepancy to concentration inequality on Riemannian manifold. Let

$$S_p(\nu\,|\,\mu) = \inf\left(\int |\tau_\nu - \mathrm{id}|_{\mathrm{HS}}^p\,\mathrm{d}\nu\right)^{1/p}.$$

It has been explained in [10] that the growth of the Stein discrepancy $S_p(\nu\,|\,\mu)$ in p entails concentration properties of the measure ν in terms of the growth of its moments. The following result shows how to directly transfer information on the Stein kernel to concentration properties on the manifold.

Theorem 5.1 (Moment bounds). *Assume that* $\mathrm{Ric}_V \geq K > 0$*, and*

$$|\mathrm{Hess}_{P_t f}|_{\mathrm{HS}}^2 \leq \Psi(t) P_t |\nabla f|^2$$

where Ψ *satisfies*

$$\int_0^\infty \Psi^{1/2}(r)\,\mathrm{d}r < \infty.$$

There exists a numerical constant $C > 0$ *such that for every* 1*-Lipshitz function* $f\colon M \to \mathbb{R}$ *with* $\int f\,\mathrm{d}\nu = 0$*, and every* $p \geq 2$*,*

$$\left(\int |f|^p\,\mathrm{d}\nu\right)^{1/p} \leq C\left(S_p(\nu\,|\,\mu) + \sqrt{p}\Big(\int |\tau_\nu|_{\mathrm{op}}^{p/2}\mathrm{d}\nu\Big)^{1/p}\right),$$

where the constant C *depends on the constants* K*,* p *and* $\int_0^\infty \Psi^{1/2}(r)\,\mathrm{d}r$*.*

Proof. We only prove the result for p an even integer, the general case follows similarly with some further technicalities. We may also replace the assumption $\int_M f\,\mathrm{d}\nu = 0$ by $\int_M f\,\mathrm{d}\mu = 0$ via a simple use of the triangle inequality. Let $f\colon M \to \mathbb{R}$ be 1-Lipshitz, and assume f to be smooth and bounded. Let $q \geq 1$ be an integer and set

$$\phi(t) = \int_M (P_t f)^{2q}\,\mathrm{d}\nu, \quad t \geq 0.$$

Since $\mu(f) = 0$, it follows that $\phi(\infty) = 0$. Now using the calculation with respect to the semigroup P_t, we have

$$\begin{aligned}
\phi'(t) &= 2q\int_M (P_t f)^{2q-1} L P_t f\,\mathrm{d}\nu \\
&= 2q\int (P_t f)^{2q-1}\Delta P_t f\,\mathrm{d}\nu - \int \langle \tau_\nu, \mathrm{Hess}((P_t f)^{2q})\rangle_{\mathrm{HS}}\,\mathrm{d}\nu \\
&= 2q\int (P_t f)^{2q-1}\langle \mathrm{id} - \tau_\nu, \mathrm{Hess}(P_t f)\rangle_{\mathrm{HS}}\,\mathrm{d}\nu \\
&\quad - 2q(2q-1)\int_M (P_t f)^{2q-2}\langle \tau_\nu, \nabla P_t f \otimes \nabla P_t f\rangle\,\mathrm{d}\nu.
\end{aligned} \tag{5.1}$$

Next, Theorem 2.1 implies

$$\begin{aligned}
\langle \tau_\nu - \mathrm{id}, \mathrm{Hess}(P_t f)\rangle_{\mathrm{HS}} &\leq |\tau_\nu - \mathrm{id}|_{\mathrm{HS}}\,|\mathrm{Hess}(P_t f)||_{\mathrm{HS}} \\
&\leq |\tau_\nu - \mathrm{id}|_{\mathrm{HS}}\left(\Psi(t) P_t |\nabla f|^2\right)^{1/2} \\
&\leq |\tau_\nu - \mathrm{id}|_{\mathrm{HS}}\,\Psi^{1/2}(t).
\end{aligned}$$

Combining these inequalities with (5.1) and observing that,

$$|\nabla P_t f| \leq \mathrm{e}^{-K/2} P_t |\nabla f| \leq \mathrm{e}^{-K/2},$$

we arrive at

$$-\phi'(t) \leq \Psi^{1/2}(t) \int 2q|P_t f|^{2q-1}|\tau_\nu - \mathrm{id}|_{\mathrm{HS}}\,\mathrm{d}\nu + \mathrm{e}^{-Kt}\int 2q(2q-1)(P_t f)^{2q-2}|\tau_\nu|_{\mathrm{op}}\,\mathrm{d}\nu.$$

Therefore, using the Young-Hölder inequality, we obtain

$$-\phi'(t) \leq C(t)\phi(t) + D(t),$$

where

$$D(t) = \Psi^{1/2}(t)\int |\tau_\nu - \mathrm{id}|^{2q}_{\mathrm{HS}}\,\mathrm{d}\nu + \mathrm{e}^{-Kt}\int ((2q-1)|\tau_\nu|_{\mathrm{op}})^q \mathrm{d}\nu$$

and

$$C(t) = \Psi^{1/2}(t)\,(2q)^{2q/(2q-1)} + \mathrm{e}^{-Kt}(2q)^{2q/(2q-2)}.$$

Thus we obtain

$$\phi(t) \leq \int_t^\infty \exp\left(\int_t^s C(r)\,dr\right) D(s)\,\mathrm{d}s$$

and it follows that

$$\phi(0) \leq \frac{1}{(2q)^{2q/(2q-1)}}\exp\left((2q)^{2q/(2q-1)}\int_0^\infty \Psi^{1/2}(s)\,\mathrm{d}s\right)\int |\tau_\nu - \mathrm{id}|^{2q}_{\mathrm{HS}}\,\mathrm{d}\nu + \frac{\mathrm{e}^{(2q)^{2q/(2q-2)}/K}}{(2q)^{2q/(2q-2)}}\int \left((2q-1)|\tau_\nu|_{\mathrm{op}}\right)^q\,\mathrm{d}\nu.$$

Therefore, there exists a constant $C > 0$ such that

$$\int_M |f|^{2q}\,\mathrm{d}\nu \leq C\left(\int |\tau_\nu - \mathrm{id}|^{2q}_{\mathrm{HS}}\,\mathrm{d}\nu + \int (2q|\tau_\nu|_{\mathrm{op}})^q\,\mathrm{d}\nu\right). \qquad \square$$

Remark 5.2. We see that when $\mathrm{Hess}_V = K$, by Obata's Rigidity Theorem (see [26, Theorem 6.3]), then M is isometric to $\mathbb{R}^n$, which implies $\mathrm{Ric}_V = K$, $\alpha_n = K$, $\|R\|_\infty = 0$, then the constant C is independent of the dimension n. In the general case however, as Ψ depends on the dimension, the constant C will not be dimension free.

When $p = 2$, we observe that $|\tau_\nu|_{\mathrm{op}} \leq 1 + |\tau_\nu - \mathrm{id}|_{\mathrm{HS}}$ which implies that

$$\mathrm{Var}_\nu(f) \leq C(1 + S(\nu\,|\,\mu) + S^2(\nu\,|\,\mu)).$$

Thus, the Stein discrepancy $S(\nu\,|\,\mu)$ with respect to the invariant measure gives another control of the spectral properties for log-concave measures, see [12] for the Lipshitz characterization of Poincaré inequalities for measures of this type.

References

[1] 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

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

[3] Louis H.Y. Chen, Larry Goldstein, and Qi-Man Shao, *Normal approximation by Stein's method*, Probability and its Applications (New York), Springer, Heidelberg, 2011. MR 2732624

[4] Thomas A. Courtade, Max Fathi, and Ashwin Pananjady, *Existence of Stein kernels under a spectral gap, and discrepancy bounds*, Ann. Inst. Henri Poincaré Probab. Stat. **55** (2019), no. 2, 777–790. MR 3949953

[5] Christian Döbler and Giovanni Peccati, *The gamma Stein equation and noncentral de Jong theorems*, Bernoulli **24** (2018), no. 4B, 3384–3421. MR 3788176

[6] Bruce K. Driver, *A Cameron-Martin type quasi-invariance theorem for pinned Brownian motion on a compact Riemannian manifold*, Trans. Amer. Math. Soc. **342** (1994), no. 1, 375–395. MR 1154540

[7] Bruce K. Driver and Anton Thalmaier, *Heat equation derivative formulas for vector bundles*, J. Funct. Anal. **183** (2001), no. 1, 42–108. MR 1837533

[8] 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

[9] Leonard Gross, *Logarithmic Sobolev inequalities*, Amer. J. Math. **97** (1975), no. 4, 1061–1083. MR 420249

[10] Michel Ledoux, Ivan Nourdin, and Giovanni Peccati, *Stein's method, logarithmic Sobolev and transport inequalities*, Geom. Funct. Anal. **25** (2015), no. 1, 256–306. MR 3320893

[11] Xue-Mei Li, *Hessian formulas and estimates for parabolic Schrödinger operators*, arXiv:1610.09538 (2016).

[12] Emanuel Milman, *On the role of convexity in isoperimetry, spectral gap and concentration*, Invent. Math. **177** (2009), no. 1, 1–43. MR 2507637

[13] Ivan Nourdin and Giovanni Peccati, *Normal approximations with Malliavin calculus*, Cambridge Tracts in Mathematics, vol. 192, Cambridge University Press, Cambridge, 2012, From Stein's method to universality. MR 2962301

[14] Ivan Nourdin, Giovanni Peccati, and Yvik Swan, *Entropy and the fourth moment phenomenon*, J. Funct. Anal. **266** (2014), no. 5, 3170–3207. MR 3158721

[15] Felix Otto and Cédric Villani, *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality*, J. Funct. Anal. **173** (2000), no. 2, 361–400. MR 1760620

[16] Charles Stein, *A bound for the error in the normal approximation to the distribution of a sum of dependent random variables*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory, 1972, pp. 583–602. MR 0402873

[17] ______, *Approximate computation of expectations*, Institute of Mathematical Statistics Lecture Notes—Monograph Series, vol. 7, Institute of Mathematical Statistics, Hayward, CA, 1986. MR 882007

[18] Michel Talagrand, *Transportation cost for Gaussian and other product measures*, Geom. Funct. Anal. **6** (1996), no. 3, 587–600. MR 1392331

[19] Anton Thalmaier, *On the differentiation of heat semigroups and Poisson integrals*, Stochastics Stochastics Rep. **61** (1997), no. 3-4, 297–321. MR 1488139

[20] James Thompson, *Derivatives of Feynman-Kac semigroups*, J. Theoret. Probab. **32** (2019), no. 2, 950–973. MR 3959634

[21] ______, *Approximation of riemannian measures by stein's method*, Arxiv (2020).

[22] Cédric Villani, *Optimal transport, old and new*, Grundlehren der Mathematischen Wissenschaften, vol. 338, Springer, 2009.

[23] Max-K. von Renesse and Karl-Theodor Sturm, *Transport inequalities, gradient estimates, entropy, and Ricci curvature*, Comm. Pure Appl. Math. **58** (2005), no. 7, 923–940. MR 2142879

[24] 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

[25] ______, *Identifying constant curvature manifolds, Einstein manifolds, and Ricci parallel manifolds*, J. Geom. Anal. **29** (2019), no. 3, 2374–2409. MR 3969430

[26] Guoqiang Wu and Rugang Ye, *A note on Obata's rigidity theorem*, Commun. Math. Stat. **2** (2014), no. 3-4, 231–252. MR 3326231