Probab. Theory Relat. Fields 108, 87 – 101 (1997)

Probability Theory and Related Fields

On estimation of the logarithmic Sobolev constant and gradient estimates of heat semigroups

Feng-Yu Wang★

Department of Mathematics, Beijing Normal University, Beijing 100875, P.R. China (wangfy@maths.warwick.ac.uk)

Received: 19 September 1995 / In revised form 11 April 1996

Summary. This paper presents some explicit lower bound estimates of logarithmic Sobolev constant for diffusion processes on a compact Riemannian manifold with negative Ricci curvature. Let $\mathrm{Ric} \geqq -K$ for some $K > 0$ and d, D be respectively the dimension and the diameter of the manifold. If the boundary of the manifold is either empty or convex, then the logarithmic Sobolev constant for Brownian motion is not less than

$$\max\left\{ \left(\frac{d}{d+2}\right)^d \frac{1}{2(d+1)D^2} \exp[-1-(3d+2)D^2K], \right.$$
$$\left. \left(\frac{d-1}{d+1}\right)^d K \exp[-4D\sqrt{dK}] \right\} .$$

Next, the gradient estimates of heat semigroups (including the Neumann heat semigroup and the Dirichlet one) are studied by using coupling method together with a derivative formula modified from [11]. The resulting estimates recover or improve those given in [7, 21] for harmonic functions.

Mathematics Subject Classification (1991): 35S15, 60J60

1 Introduction

Let M be a complete connected Riemannian manifold with dimension d and boundary ∂M which may be empty. Let $L = \Delta + Z$ for some C^1-vector field Z. For the study of logarithmic Sobolev constant (LSC), we assume that M is compact and $Z = \nabla h$ for some $h \in C^2(M)$ with $\mu(dx) = e^h dx$ being

★ Research supported in part by NSFC and the Foundation of Institution of Higher Education for Doctoral Program

a probability measure. We call the logarithmic Sobolev inequality holds with respect to μ (or the L-diffusion process), if there exists a constant $\alpha > 0$ such that

$$\int u^2 \log u^2 d\mu - \int u^2 \, d\mu \log \int u^2 \, d\mu \leqq \frac{2}{\alpha} \int |\nabla u|^2 \, d\mu \tag{1.1}$$

holds for all $u \in C^1(M) \cap L^2(\mu)$. The so called LSC is the largest possible α, denoted by $\alpha(h)$.

A lot of papers have studied the lower bound estimate of $\alpha(h)$, especially for $\partial M = \emptyset$ (see [12] for detailed references). Some previous estimates are sharp for positive Ricci curvature case, especially, the constant is known when $M = S^d$ and $L = \Delta$ (see [1, 9]). But, to our knowledge, all the known explicit estimates become ineffective if the lower bound of Ricci curvature is very negative. On the other hand, it is well known that $\alpha(h) > 0$ for compact M. The first aim of the paper is to present some explicit lower bound estimates of $\alpha(h)$ for the negative curvature case.

For the first look, it seems hard to get an estimate which is meaningful for any lower bound of Ricci curvature. The reason is that the traditional Bakry–Emery's argument arises trouble if the lower bound of Ricci curvature is very negative (cf. [9]). Fortunately, some recent progress enables us to derive such type of estimate. First, from [13, 22] we have an explicit estimate of LSC for Ornstein–Uhlenbeck process on M (see Lemma 3.1 below). Next, by using the comparison argument of LSC with different potentials (see [9, 6]), we obtain a lower bound of $\alpha(h)$ depending on $c(t)$ which is the constant of the following Harnack inequality:

$$p_t(x, y) \leqq c(t) \, p_t(x, z), \quad x, y, z \in M \ ,$$

where $p_t(x, y)$ is the (Neumann) heat kernel of L, i.e., the transition probability density of the L-diffusion process (with reflecting boundary if $\partial M \neq \emptyset$). Finally, from the Li–Yau's type Harnack inequality for solutions to the heat equation of L [19], we obtain some explicit estimates of $c(t)$ which then provides the desired lower bounds of $\alpha(h)$.

Another purpose of the paper is to study the gradient estimates of heat semigroup P_t for L. To this end, we first recall a derivative formula given by Elworthy–Li [11]. Let $\partial M = \emptyset$ and rewrite L by $\sum_{i=1}^m X_i^2 + A$ for some smooth vector fields X_i $(i \leqq m)$ and C^1-vector field A. Let x_t and W_s solve the following stochastic differential equations:

$$dx_t = \sqrt{2} \sum_{i=1}^m X_i(x_t) \circ db_t^i + A(x_t) \, dt, \quad x_0 = x \ ,$$

$$\frac{\partial W_t(v)}{\partial t} = -\mathrm{Ric}(W_t(v), \cdot)^\# + \nabla_{W_t(v)} Z, \quad W_0(v) = v \in T_x M \ ,$$

where $\mathrm{Ric}(W_t(v), \cdot)^\# \in T_{x_t} M$ is defined as $\langle \mathrm{Ric}(W_t(v), \cdot)^\#, X \rangle = \mathrm{Ric}(W_t(v), X)$ for $X \in T_{x_t} M$. The solution $W_t : TM \to TM$ is called the Ricci flows when $Z = 0$ (see [10, 11]). If $\mathrm{Ric}(\cdot, \cdot) - \langle \nabla \cdot Z, \cdot \rangle$ is bounded from below, then

$$\langle \nabla P_t u, v \rangle = \frac{1}{t\sqrt{2}} E u(x_t) \int_0^t \langle W_s, X(x_s) \, db_s \rangle, \quad v \in T_x M \ , \tag{1.2}$$

holds for all $u \in C_b^1(M)$. Since $|W_t|$ is bounded for each t, from (1.2) we can estimate the gradient of $P_t u$. On the other hand, however, the right hand side of (1.2) depends on the choice of X_i which comes from a certain embedding map of M into $\mathbb{R}^m$ for some $m \geqq d$. In Sect. 4, we prove a more natural version of (1.2) suggested in [10; Remark 1] which depends only on the geometry of M and leads to the exact Bismut's formula [2] for $\nabla \log p_t(\cdot, y)$ (refer to [10, 11]). Moreover, the formula also holds for the Neumann heat semigroup whenever $\partial M \neq \emptyset$.

Next, we know from [7, 21] that the coupling method is powerful in the study of the gradient estimate for harmonic functions. As a continuation, we use this method to study the gradient estimate of heat semigroups. The resulting estimates recover those given in [21] and especially, the gradient estimate for the Dirichlet semigroup presented in Sect. 5 leads to an explicit gradient estimate of harmonic functions on a local domain. This can be consider as an improvement of the previous one given in [7].

2 Harnack type inequality

Suppose that ∂M is either empty or convex. Let $u(x,t) \geqq 0$ solve the heat equation of L:

$$u_t(x,t) = Lu(x,t), \qquad Vu|_{\partial M \times (0,\infty)} = 0 \quad \text{if } \partial M \neq \emptyset\,, \tag{2.1}$$

where $u_t = \partial u / \partial t$ and V is the inward normal vector field of ∂M.

For the case $Z = 0$, Li–Yau [16] studied the heat kernel by estimating $|\nabla u|/u$. The resulting estimate then is improved by Davies [8] as follows: let $\mathrm{Ric} \geqq -K$ for some $K \geqq 0$, then

$$\frac{|\nabla u|^2}{u^2} - \frac{\alpha u_t}{u} \leqq \frac{d\alpha^2}{2t}\left(1 + \frac{Kt}{2(\alpha-1)}\right), \quad \alpha > 1 \tag{2.2}$$

which implies the following parabolic Harnack type inequality

$$u(x,t) \leqq u(y,t+s)\left(\frac{t+s}{t}\right)^{d\alpha/2} \exp\left[\frac{\alpha\rho(x,y)^2}{4s} + \frac{d\alpha K s}{4(\alpha-1)}\right], \quad t,s>0,\ \alpha>1\,, \tag{2.3}$$

where ρ is the Riemannian distance.

For the present operator $L = \Delta + Z$, define

$$R_Z = \max\{0, -\inf\{\mathrm{Ric}(v,v) - \langle \nabla_v Z, v\rangle - \langle Z, v\rangle^2 \colon v \in TM,\ |v| = 1\}\}\,.$$

The proof of [19, Theorem 7] yields (see [18] for further discussion)

$$u(x,t) \leqq u(y,t+s)\left(\frac{t+s}{t}\right)^{(d+1)\alpha/2} \exp\left[\frac{\alpha\rho(x,y)^2}{4s} + \frac{\alpha(d+1)R_Z s}{4(\alpha-1)}\right] \tag{2.4}$$

for t, $s > 0$ and $\alpha > 1$.

Now, we go to estimate the heat kernel $p_t(x, y)$ characterized as the fundamental solution to (2.1): for $u \in C_0^2(M)$, $u(x,t) := \int p_t(x,y)u(y)\mu(dy)$ solves (2.1) for $Z = \nabla h$, $\mu(dy) = e^{h(y)}dy$. The following result is a direct consequence of (2.3) and (2.4).

Proposition 2.1 *If $D := \sup_{x,y \in M} \rho(x,y) < \infty$, choose h such that $\mu(M) = 1$. We have*

$$p_t(x,y) \leqq \inf_{s>0, \alpha>1} \left(\frac{t+s}{t}\right)^{(d+1)\alpha/2} \exp\left[\frac{\alpha D^2}{4s} + \frac{\alpha(d+1)R_{\nabla h}s}{4(\alpha-1)}\right], \tag{2.5}$$

$$p_t(x,y) \geqq \sup_{s\in(0,t), \alpha>1} \left(\frac{t-s}{t}\right)^{(d+1)\alpha/2} \exp\left[-\frac{\alpha D^2}{4s} - \frac{\alpha(d+1)R_{\nabla h}s}{4(\alpha-1)}\right]. \tag{2.6}$$

Next, let $u(x,t) = P_t u(x)$ for some nonnegative $u \in C(M)$, we have

$$u(x,t) \leqq u(y,t) \inf_{s\in(0,t), \alpha>1} \left(\frac{t+s}{t-s}\right)^{(d+1)\alpha/2} \exp\left[\frac{\alpha D^2}{2s} + \frac{\alpha(d+1)R_{\nabla h}s}{2(\alpha-1)}s\right]. \tag{2.7}$$

Finally, if $h = 0$, the number "$d+1$" in (2.5)–(2.7) can be replaced by "d".

Proof. We simply denote the desired upper bound of $p_t(x,y)$ by $c(t)$. For fixed y and large $n \in \mathbb{N}$, take $\alpha(n) > n^{-1}$ such that $\mu(B(y,\alpha(n))) \leqq (1+n^{-1})\mu(B(y,n^{-1}))$. Choose $u_n \in C^\infty(M)$ such that

$$0 \leqq u_n \leqq 1, \qquad u_n|_{B(y,n^{-1})} \equiv 1, \quad u_n|_{B(y,\alpha(n))^c} \equiv 0\,.$$

Let $u_n(x,t) = P_t u_n(x)/\mu(u_n)$, then $\mu(u_n(\cdot,t)) = 1$ and hence there exists x_t such that $u_n(x_t,t) \leqq 1$. By (2.4) we obtain

$$\mu(B(y,\alpha(n)))^{-1}\mu(B(y,n^{-1})) \inf_{B(y,n^{-1})} p_t(x,\cdot) \leqq u_n(x,t) \leqq c(t)\,,$$

the desired upper bound then follows by letting $n \to \infty$. Similarly, we prove the lower bound estimate. Similarly, the claimed estimates for the case $h = 0$ follows from (2.3). □

3 Estimation of the logarithmic Sobolev constant

The main purpose of this section is to present some explicit estimates of $\alpha(h)$ for $K(\nabla h) < 0$, where

$$K(Z) = -\inf\{\mathrm{Ric}(v,v) - \langle \nabla_v Z, v\rangle : v \in TM,\ |v| = 1\}$$

for C^1-vector field Z. The key idea is to compare $\alpha(h)$ with the LSC for the absolute distribution of the L-diffusion process.

For the case $\partial M = \emptyset$, an explicit lower bound estimate is presented in [22] for the logarithmic Sobolev constant with respect to $\delta_x P_t$, the distribution at time t of the L-diffusion process with initial point x. See also [13] for $h = 0$. Here, we claim that the same result holds for M with convex boundary.

Lemma 3.1 *Let M be a complete Riemannian manifold with convex boundary whenever $\partial M \neq \emptyset$. If $K(\nabla h) < \infty$, we have*

$$P_t(u^2 \log u^2) - (P_t u^2)\log(P_t u^2) \leqq 2\frac{e^{2K(\nabla h)t} - 1}{K(\nabla h)} P_t|\nabla u|^2 \tag{3.1}$$

for all $u \in C^1(M)$ with $P_t u^2 < \infty$. Here and in what follows, when $K(\nabla h) = 0$, we take the coefficient of the right-hand side to be the limit as $K(\nabla h) \to 0$.

Proof. (a) We first recall briefly the coupling by parallel displacement. Let $H : TM \to TO(M)$ be the horizontal lift induced by the Riemannian connection. Consider the stochastic differential equations:

$$d\Phi_t = H_{\Phi_t}\Phi_t \circ dM_t$$

$$dM_t = \sqrt{2}\, dB_t + \Phi_t^{-1} Z(x_t)\, dt + \Phi_t^{-1} V(x_t)\, dL_t, \quad x_t = \pi\Phi_t ,$$

where B_t is a Brownian motion on $\mathbb{R}^d$, π is the natural projection of $O(M)$ onto M and L_t is an increasing process called the local time of x_t on ∂M. Then x_t is the reflecting L-diffusion process on M with $x_0 = \pi\Phi_0$.

Next, for given $y_0 \in M$; we construct another reflecting L-diffusion process y_t as follows:

$$d\Psi_t = H_{\Psi_t}\Psi_t \circ dN_t$$

$$dN_t = \sqrt{2}\, d\bar{B}_t + \Psi_t^{-1} Z(y_t)\, dt + \Psi_t^{-1} V(y_t)\, d\bar{L}_t ,$$

$$d\bar{B}_t = \Psi_t^{-1} P_{x_t, y_t} \Phi_t\, dB_t, \quad y_t = \pi\Phi_t ,$$

where $\bar{L}_t$ is the local time of y_t on the boundary, and $P_{x,y} : T_x M \to T_y M$ is the parallel displacement along the unique shortest geodesic from x to y whenever $y \notin C(x)$. As for the case $y_t \in C(x_t)$, we use Cranston's trick [7] so that y_t is constructed for ever. We call (x_t, y_t) the coupling by displacement.

(b) Since the boundary is convex, we have [20, 22] (see [7, 14] for original arguments)

$$d\rho(x_t, y_t) \leqq K(\nabla h)\rho(x_t, y_t)\, dt ,$$

where $\rho(x, y)$ is the Riemannian distance between x and y. Then

$$\rho(x_t, y_t) \leqq \rho(x, y) \exp[K(\nabla h)t], \quad t \geqq 0 . \tag{3.2}$$

Hence, for $u \in C_0^1(M)$ we have

$$\frac{|P_t u(x) - P_t u(y)|}{\rho(x, y)} \leqq \exp[K(\nabla h)t] E^{x,y} \frac{|u(x_t) - u(y_t)|}{\rho(x_t, y_t)} .$$

By letting $y \to x$ (so $y_t \to x_t$), we obtain

$$|\nabla P_t u| \leqq P_t|\nabla u| \exp[K(\nabla h)t] .$$

Now, the remainder of the proof follows from Bakry's argument (see [13, 22]). □

Corollary 3.2 *Under the assumption of Lemma* 3.1. *If* $K(\nabla h) < 0$, *we have* $\alpha(h) \geqq -K(\nabla h)$.

Proof. Note that when $K(\nabla h) < 0$, the L-diffusion process is ergodic. Then the corollary follows from Lemma 3.1 by letting $t \to \infty$. □

For $\partial M = \emptyset$, Corollary 3.2 is a simple consequence of Bakry–Emery criterion. But, the estimate may fail if ∂M is not convex. Actually, due to a famous example by Calabi (see [3, p. 342]), for any $\varepsilon > 0$, there exists a regular domain $\Omega \subset M$ such that the first Neumann eigenvalue of L on Ω is less than ε.

Note that $P_t u(x) = \int p_t(x, y) u(y) e^{h(y)}\, dy$, Lemma 3.1 yields

$$\alpha(h + \log p_t(x, \cdot\,)) \geqq \frac{K(\nabla h)}{e^{2K(\nabla h)t} - 1}, \quad x \in M,\ t > 0\,. \tag{3.3}$$

From a comparison argument between logarithmic Sobolev constants with different potentials (see [6] or [9]), it follows that

$$\alpha(h) \geqq \sup_{t>0} \left\{ \frac{K(\nabla h)}{e^{2K(\nabla h)t} - 1} \inf_{y,z \in M} \frac{p_t(x, y)}{p_t(x,z)} \right\}, \quad x \in M\,. \tag{3.4}$$

By combining this with Proposition 2.1, we obtain the following result.

Theorem 3.3 *Suppose that* M *is a compact connected Riemannian manifold with convex boundary whenever* $\partial M \neq \emptyset$. *We have*

$$\alpha(h) \geqq \sup_{t>s>0,\alpha>1} \frac{K(\nabla h)}{e^{2K(\nabla h)t} - 1} \left(\frac{t-s}{t+s}\right)^{(d+1)\alpha/2} \exp\left[-\frac{\alpha D^2}{2s} - \frac{\alpha(d+1)R_{\nabla h} s}{2(\alpha-1)}\right]\,.$$

When $h = 0$, *the number* "$d+1$" *can be replaced by* "d". *Especially, take* $\alpha = 2,\ s = D^2,\ t = (d+1)D^2$, *we obtain*

$$\alpha(0) \geqq \left(\frac{d}{d+2}\right)^d \frac{K}{e^{2K(d+1)D^2} - 1} \exp[-1 - D^2 d K^+]\,.$$

Corollary 3.4 *Under the assumption of Theorem* 3.3. *If* $K \geqq 0$, *then*

$$\alpha(h) \geqq \left(\frac{d}{d+2}\right)^d \frac{1}{2(d+1)D^2} \exp[-1 - \delta(h) - (3d+2)D^2 K]\,.$$

Proof. The corollary follows from Theorem 3.3 together with the facts $\alpha(h) \geqq \exp[-\delta(h)]\alpha(0)$ and $e^\lambda - 1 \leqq \lambda e^\lambda$ for $\lambda \geqq 0$. □

Remark. Suppose that $K > 0$ and $h = 0$, by taking $\alpha = 2,\ s = D/\sqrt{dK}$ and $t = D\sqrt{d}/\sqrt{K}$, Theorem 3.3 yields

$$\alpha(0) \geqq \left(\frac{d-1}{d+1}\right)^d K \exp[-4D\sqrt{dK}]\,. \tag{3.5}$$

As $K \to \infty$, this lower bound decays with the same order as that of the first eigenvalue given in [5, 20].

4 Gradient estimates of heat semigroups

We begin this section with a new version of (1.2) which is also valid for manifold with boundary.

Given $v_0 \in T_{x_0}M$, let $y_0^l = \exp[lv_0]$ and (x_t, y_t^l) be the coupling by parallel displacement. Define $v_t \in T_{x_t}M$ by

$$\langle v_t, \nabla u(x_t)\rangle = \lim_{l\to 0} \frac{u(y_t^l) - u(x_t)}{l} .$$

It is proved in [13] that v_t is just the Ricci flows when $\partial M = \emptyset$ and $Z = 0$. Then the following result leads to the exact Bismut's formula given in [2, Theorem 2.71] for heat kernel (refer to [10, p. 68]).

Theorem 4.1 *For $u \in C_b^1(M)$, we have*

$$\langle \nabla P_t u(x_0), v_0\rangle = \frac{1}{t\sqrt{2}} Eu(x_t) \int_0^t \langle v_s, \Phi_s\, dB_s\rangle, \quad v_0 \in T_{x_0}M$$

provided $\int_0^t \langle v_s, \Phi_s\, dB_s\rangle$ is a martingale.

Proof. The proof is similar to that of [11, Theorem 2.1]. It follows from Itô's formula that

$$dP_{t-s}u(x_s) = \langle \nabla P_{t-s}u(x_s), \sqrt{2}\Phi_s\, dB_s\rangle .$$

By integrating over s from 0 to t, we obtain

$$u(x_t) = P_t u(x_0) + \int_0^t \langle \nabla P_{t-s}u(x_s), \sqrt{2}\Phi_s\, dB_s\rangle . \tag{4.1}$$

Hence

$$\begin{aligned}\frac{1}{\sqrt{2}} Eu(x_t) \int_0^t \langle v_s, \Phi_s\, dB_s\rangle &= E\int_0^t \langle \nabla P_{t-s}u(x_s), v_s\rangle\, ds \\ &= E\int_0^t \lim_{l\to 0} \frac{P_{t-s}u(y_s^l) - P_{t-s}u(x_s)}{l} \\ &= \int_0^t \lim_{l\to 0} \frac{P_t u(y_0^l) - P_t u(x_0)}{l} = t\langle \nabla P_t u(x_0), v_0\rangle . \qquad \square\end{aligned}$$

Corollary 4.2 *Suppose that ∂M is either convex or empty. If $K(Z) < \infty$, we have*

$$\begin{aligned}|\nabla P_t u(x)| \leqq{}& \frac{((2n-1)!!)^{1/2n}}{2t} \sqrt{\frac{\exp[2K(Z)t] - 1}{K(Z)}} \\ &\times (P_t u^{2n/(2n-1)}(x))^{(2n-1)/2n}, \quad n \in \mathbb{N} .\end{aligned}$$

Proof. Let $R_t = \sqrt{2}\int_0^t \langle v_s, \Phi_s\, dB_s\rangle$. Note that (3.2) implies $|v_s| \leqq \exp[K(Z)s]$, then R_t is a martingale and Theorem 4.1 yields

$$|\nabla P_t u(x)| \leqq \frac{1}{2t}(P_t u^{2n/(2n-1)}(x))^{(2n-1)/2n}(ER_t^{2n})^{1/2n} . \tag{4.2}$$

By Itô's formula we obtain

$$dR_t^{2n} \leqq 2nR_t^{2n-1}\,dR_t + 2n(2n-1)R_t^{2(n-1)}\exp[2K(Z)t]\,dt\;.$$

Hence

$$ER_t^{2n} \leqq 2n(2n-1)\int_0^t ER_s^{2(n-1)}\;\exp[2K(Z)s]\,ds\;.$$

Now, the corollary follows from this and (4.2) by inducing in the number n. □

Next, we go to study the gradient estimate by using coupling. The original idea of the study is due to [7]. Let (x_t, y_t) be a coupling of the L-diffusion process with reflecting boundary whenever $\partial M \neq \emptyset$, and let $T = \{t \geqq 0\colon x_t = y_t\}$ be the coupling time. We have

$$\begin{aligned}\frac{|P_t u(x) - P_t u(y)|}{\rho(x,y)} &= \frac{|P_s P_{t-s}u(x) - P_s P_{t-s}u(y)|}{\rho(x,y)}\\ &\leqq E^{x,y}\frac{|P_{t-s}u(x_s) - P_{t-s}u(y_s)|}{\rho(x,y)}\\ &\leqq \delta(P_{t-s}u)\frac{P^{x,y}(T>s)}{\rho(x,y)},\quad t>s>0\;.\end{aligned}$$

Hence

$$|\nabla P_t u(x)| \leqq \delta(P_{t-s}u)\overline{\lim_{y\to x}}\,\frac{P^{x,y}(T>s)}{\rho(x,y)},\quad t>s>0\;. \tag{4.3}$$

Now, the next step is to estimate the distribution of the coupling time.

Let (x_t, y_t) be the coupling by reflection [7, 14], if ∂M is either convex or empty, we have (see [5, 20])

$$d\rho(x_t, y_t) \leqq 2\sqrt{2}\,db_t + \gamma(\rho(x_t, y_t))\,dt\;, \tag{4.4}$$

where b_t is an one-dimensional Brownian motion and

$$\begin{aligned}\gamma(r) = \min\Big\{&K(Z)r, 2\sqrt{K^+(d-1)}\tanh\left[\frac{r}{2}\sqrt{K^+/(d-1)}\right]\\ &-2\sqrt{K^-(d-1)}\tan\left[\frac{r}{2}\sqrt{K^-/(d-1)}\right] + a(r)\Big\}\end{aligned}$$

with $a(r) \in C(\mathbb{R}_+)$ so that

$$a(r) \geqq \sup_{\rho(x,y)=r}\,(Z\rho(\,\cdot\,,y)(x) + Z\rho(x,\,\cdot\,)(y)),\quad r>0\;.$$

Set $D = \sup_{x,y\in M}\rho(x,y)$ and define

$$C(r) = \exp\left[\frac{1}{4}\int_0^r \gamma(u)\,du\right],\qquad f(r) = \int_0^r \frac{1}{C(u)}\,du,\qquad g(r) = \int_0^r C(u)\,du,$$

$$F_N(r) = \int_0^r \frac{1}{C(s)}\,ds\int_s^N C(u)\,du,\quad N, r\in[0,D]\;.$$

Lemma 4.3 *Suppose that ∂M is either convex or empty. For the coupling by reflection, we have*

$$P^{x,y}(T>t) \leqq \inf_{N\in[\rho(x,y),D]} \left\{ \inf_{\lambda\in[0,4F_N(N)^{-1})} \frac{\lambda F_N(\rho(x,y))}{4-\lambda F_N(N)}(e^{\lambda t}-1)^{-1} + \frac{f(\rho(x,y))}{f(N)} \right\}.$$

Especially, if $F_D(D)<\infty$ (it is the case when $D<\infty$), we have

$$P^{x,y}(T>t) \leqq \inf_{\lambda\in[0,4F_D(D)^{-1})} \frac{\lambda F_D(\rho(x,y))}{4-\lambda F_D(D)}(e^{\lambda t}-1)^{-1}, \quad t>0 .$$

Proof. (a) For given $N \in [\rho(x,y), D]$, (4.4) yields

$$dF_N(\rho(x_t,y_t)) \leqq 2\sqrt{2}F_N'(\rho(x_t,y_t))\,db_t - 4\,dt .$$

Take $G_\lambda(t,r)=e^{\lambda t}F_N(r)$, $\lambda\in[0,4F_N(N)^{-1})$. Let $S_N=\inf\{t\geqq 0\colon \rho(x_t,y_t)\geqq N\}$. We have

$$dG_\lambda(t,\rho(x_t,y_t)) \leqq dM_t + (\lambda e^{\lambda t}F_N(\rho(x_t,y_t)) - 4e^{\lambda t})\,dt$$

for some martingale M_t. Then

$$\begin{aligned} 0 &\leqq E^{x,y}G_\lambda(t\wedge T\wedge S_N, \rho(x_{t\wedge T\wedge S_N}, y_{t\wedge T\wedge S_N})) \\ &= E^{x,y}\int_0^{t\wedge T\wedge S_N} dG_\lambda(s,\rho(x_s,y_s)) + G_\lambda(0,\rho(x,y)) \\ &\leqq \lambda^{-1}(\lambda F_N(N)-4)E^{x,y}(e^{\lambda(t\wedge T\wedge S_N)}-1) + F_N(\rho(x,y)) . \end{aligned}$$

By letting $t\to\infty$ we obtain

$$E^{x,y}(e^{\lambda(T\wedge S_N)}-1) \leqq \frac{\lambda F_N(\rho(x,y))}{4-\lambda F_N(N)} .$$

Hence

$$P^{x,y}(T\wedge S_N>t) \leqq \inf_{\lambda\in[0,4F_N(N)^{-1})} \frac{\lambda F_N(\rho(x,y))}{4-\lambda F_N(N)}(e^{\lambda t}-1)^{-1}, \quad t>0 . \quad (4.5)$$

(b) Note that (4.4) yields $df(x_t,y_t) \leqq 2\sqrt{2}f'(\rho(x_t,y_t))\,db_t$, we have

$$f(\rho(x,y)) \geqq E^{x,y}f(\rho(x_{t\wedge T\wedge S_N}, y_{t\wedge T\wedge S_N})) \geqq f(N)P^{x,y}(t\wedge T \geqq S_N) .$$

Then $P^{x,y}(T\geqq S_N) \leqq f(\rho(x,y))/f(N)$. By combining this with (4.5), we obtain

$$\begin{aligned} P^{x,y}(T>t) &= P^{x,y}(T>t,\ S_N>t) + P^{x,y}(T>t,\ S_N\leqq t) \\ &\leqq P^{x,y}(T\wedge S_N>t) + P^{x,y}(T\geqq S_N) \\ &\leqq \inf_{\lambda\in[0,4F_N(N)^{-1})} \frac{\lambda F_N(\rho(x,y))}{4-\lambda F_N(N)}(e^{\lambda t}-1)^{-1} + \frac{f(\rho(x,y))}{f(N)} . \end{aligned}$$

Finally, if $D < \infty$, the second estimate follows from (a) by replacing $t \wedge T \wedge S_N$ with $t \wedge T$. Next, if $D = \infty$, we have $S_N \to \infty$ as $N \to \infty$ for the non-explosion of the process, the second estimate then follows from (4.5) by letting $N \to \infty$. □

Remark. The argument of (a) was used by Y.Z. Wang to study the exponential convergence in total variation norm for diffusions on compact manifolds. Let $\delta_x P_t$ be the distribution at time t of the L-diffusion process with initial point x, then

$$\|\delta_x P_t - \delta_y P_t\|_{\mathrm{var}} \leqq 2P^{x,y}(T > t)\,.$$

So Lemma 4.3 provides a rate for the process to converge in total variation norm. This improves the main results of [17] in which this topic was studied for Brownian motion on a convex polyhedron of spheres and torus.

By combining (4.3) with Lemma 4.3, we obtain the following result.

Theorem 4.4 *Suppose that ∂M is either convex or empty. For nonnegative $u \in C_b^1(M)$, we have*

$$\frac{\|\nabla P_t u\|_\infty}{\|P_{t-s} u\|_\infty} \leqq \inf_{N\in(0,D)} \left\{ \inf_{\lambda\in[0,4F_N(N)^{-1})} \frac{\lambda g(N)}{4-\lambda F_N(N)}(e^{\lambda s}-1)^{-1} + \frac{1}{f(N)} \right\}$$

for all $t > s > 0$. If in addition $F_D(D) < \infty$, then

$$\|\nabla P_{t-s} u\|_\infty \leqq \|P_{t-s} u\|_\infty \inf_{\lambda\in[0,4F_D(D)^{-1})} \frac{\lambda g(D)}{4-\lambda F_D(D)}(e^{\lambda s}-1)^{-1}, \quad t > s > 0\,.$$

Remark. Suppose that $u \geqq 0$ is L-harmonic, i.e., $Lu = 0$. Then $P_t u = u$ for all t. By letting first $t \uparrow \infty$ then $N \uparrow \infty$, Theorem 4.4 yields

$$\|\nabla u\|_\infty \leqq \|u\|_\infty / f(\infty)$$

which is exactly the main estimate of [21] and hence improves the corresponding one given in [7].

Corollary 4.5 *Suppose that M is compact with convex boundary whenever $\partial M \neq \emptyset$. We have*

$$|\nabla P_t u(x)| \leqq c(t) P_t u(x), \quad u \in C^1(M),\ t > 0$$

with

$$c(t) = \inf_{\delta\in(0,1),\lambda\in[0,4F_D(D)^{-1})} \frac{\lambda g(D)(1-\delta)^{-(d+1)}}{(4-\lambda F_D(D))(e^{\lambda\delta t}-1)} \exp\left[\frac{D^2}{2\delta t} + \frac{(d+1)R_Z \delta t}{2}\right].$$

Proof. The corollary follows from (2.4) and Theorem 4.4 by taking $\alpha = 2$ and $s = \delta t$. □

Remark. Under the assumption of Corollary 4.5, we have

$$|\nabla p_t(\,\cdot\,, y)(x)| \leqq c(t)\, p_t(x, y), \quad x, y \in M\ .$$

Actually, for given y, let $u_s(x) = p_s(x, y)$, $s \in (0, t)$. Then $p_t(x, y) = P_{t-s} u_s(x)$ and the above estimate follows from Corollary 4.5 by letting $s \to 0$.

Before ending this section, we consider the exponential convergence for the gradient of heat semigroup. Theorem 4.4 shows that, when $F_D(D) < \infty$, $\|\nabla P_t u\|_\infty$ goes to zero exponentially fast as $t \to \infty$. But the condition $F_\infty(\infty) < \infty$ is usually too strong for noncompact manifolds. We present here another sufficient condition.

Theorem 4.6 *Under the assumption of Theorem* 4.4. *If* $D = \infty$ *and* $\limsup_{r\to\infty} \gamma(r)/r < 0$, *then there exists* $c, \delta > 0$ *such that*

$$\|\nabla P_t u\|_\infty \leqq \|u\|_\infty c e^{-\delta t}$$

holds for all nonnegative $u \in C_b^1(M)$ *and large* t.

Proof. (a) By taking $t = s = 1$, $N = 1$ and $\lambda = 0$, Theorem 4.4 yields

$$\|P_1 u\|_\infty \leqq \|u\|_\infty (g(1)/4 + f(1)^{-1})\ . \tag{4.6}$$

Next, let (x_t, y_t) be the coupling by reflection. For $t > 1$ we have

$$\frac{|P_t u(x) - P_t u(y)|}{\rho(x, y)} \leqq \frac{E^{x,y}|P_1 u(x_{t-1}) - P_1 u(y_{t-1})|}{\rho(x, y)} \leqq \|\nabla P_1 u\|_\infty \frac{E^{x,y}\rho(x_{t-1}, y_{t-1})}{\rho(x, y)}\ .$$

Hence

$$|\nabla P_t u(x)| \leqq \|u\|_\infty \left(\frac{g(1)}{4} + \frac{1}{f(1)}\right) \limsup_{y\to x} \frac{E^{x,y}\rho(x_{t-1}, y_{t-1})}{\rho(x, y)}\ . \tag{4.7}$$

(b) Since $\limsup_{r\to\infty} \gamma(r)/r < 0$, there exist $r_0, \delta_0 > 0$ such that $\gamma(r) \leqq -\delta_0 r$ for $r \geqq r_0$. Define

$$\delta_1 = \exp\left[-\frac{1}{4}\int_0^{r_0} [\delta_0 s + \gamma(s)]^+\, ds\right],$$

$$G(r) = \int_0^r \exp\left[-\frac{1}{4}\int_0^s [\delta_0 t + \gamma(t)]^+\, dt\right] ds, \quad r > 0\ .$$

Then

$$\lim_{r\to 0} G(r)/r = 1 \quad \text{and} \quad r \geqq G(r) \geqq \delta_1^{-1} r\ . \tag{4.8}$$

Next, by (4.4) we obtain

$$dG(\rho(x_t, y_t)) \leqq dM_t - \delta_0\delta_1 \rho(x_t, y_t) \leqq dM_t - \delta_0\delta_1 G(\rho(x_t, y_t))$$

for some martingale M_t. Then

$$E^{x,y}\rho(x_t, y_t) \leqq \delta_1^{-1} E^{x,y} G(\rho(x_t, y_t)) \leqq \delta_1^{-1} G(\rho(x, y)) e^{-\delta_0\delta_1 t}, \quad t \geqq 0\ .$$

The proof is completed by combining this with (4.7) and (4.8). □

5 Gradient estimates of Dirichlet heat semigroup

Suppose that $\partial M \neq \emptyset$ and $|Z| \leqq b$. Let x_t be the L-diffusion process and $\tau = \inf\{t \geqq 0\colon x_t \in \partial M\}$ be the exit time. The Dirichlet heat semigroup is defined as

$$P_t^D u(x) = E^x u(x_{t\wedge\tau}), \quad t \geqq 0,\ x \in M,\ u \in C(M)\,.$$

We also use the coupling method developed in [7]. Let (x_t, y_t) be the coupling by reflection with coupling time T, denote by τ_x and τ_y respectively the exit times of the two marginal processes. For $u \geqq 0$, we have

$$|P_t^D u(x) - P_t^D(y)| \leqq E^{x,y}|u(x_{t\wedge\tau_x}) - u(y_{t\wedge\tau_y})| \leqq \|u\|_\infty P^{x,y}(T > t\wedge\tau_x\wedge\tau_y)\,.$$

Hence

$$|\nabla P_t^D u(x)| \leqq \|u\|_\infty \limsup_{y\to x} P^{x,y}(T > t\wedge\tau_x\wedge\tau_y)/\rho(x,y)\,. \tag{5.1}$$

Next, let $\delta_x = \mathrm{dist}(x, \partial M)$. For any $\delta \in (0, \delta_x]$, define

$$\tau_x^\delta = \inf\{t \geqq 0\colon \rho(x, x_t) \geqq \delta\}, \qquad \tau_y^\delta = \inf\{t \geqq 0\colon \rho(x, y_t) \geqq \delta\}$$

and let $\tau^\delta = T\wedge\tau_x^\delta\wedge\tau_y^\delta$. Then

$$\begin{aligned} P^{x,y}(T > t\wedge\tau_x\wedge\tau_y) &\leqq P^{x,y}(T > t\wedge\tau_x^\delta\wedge\tau_y^\delta) \\ &\leqq P^{x,y}(\tau^\delta > t) + P^{x,y}(T > \tau_x^\delta\wedge\tau_y^\delta) \\ &\leqq P^{x,y}(\tau^\delta > t) + P^{x,y}(\rho(x, x_{\tau^\delta}) \geqq \delta) \\ &\quad + P^{x,y}(\rho(x, y_{\tau^\delta}) \geqq \delta)\,. \end{aligned} \tag{5.2}$$

Note that (4.4) holds up to $T\wedge\tau_x\wedge\tau_y$ (see [7]), by replacing $t\wedge T\wedge S_N$ with $t\wedge\tau^\delta$, the proof of Lemma 4.3 gives

$$P^{x,y}(\tau^\delta > t) \leqq \inf_{\lambda\in[0,4F_{2\delta}(2\delta)^{-1})} \frac{\lambda F_{2\delta}(\rho(x,y))}{4 - \lambda F_{2\delta}(2\delta)}(e^{\lambda t} - 1)^{-1}, \quad t > 0\,. \tag{5.3}$$

Finally, define

$$j(r) = \begin{cases} \sin(r\sqrt{-K/(d-1)}) & \text{if } K < 0\,, \\ r & \text{if } K = 0\,, \\ \sinh(r\sqrt{K/(d-1)}) & \text{if } K > 0\,. \end{cases}$$

$$J(r) = \int_0^r j(s)^{d-1} e^{-bs}\, ds \int_0^s j(u)^{d-1} e^{bu}\, du, \quad r \in [0, D]\,.$$

We have the following result.

Lemma 5.1 *For* $\delta \in (0, \delta_x)$ *and* $\rho(x, y) < \delta$, *we have*

$$P^{x,y}(\rho(x, x_{\tau^\delta}) \geqq \delta) \leqq \frac{F_{2\delta}(\rho(x, y))}{4J(\delta)} ,$$

$$P^{x,y}(\rho(x, y_{\tau^\delta}) \geqq \delta) \leqq \frac{F_{2\delta}(\rho(x, y)) + 4J(\rho(x, y))}{4J(\delta)} .$$

Proof. By Laplacian comparison theorem we have

$$\Delta\rho(x, \cdot)(y) \leqq (d-1)j'(\rho(x, y))/j(\rho(x, y)) ,$$

then (see [15])

$$d\rho(x, x_t) \leqq \sqrt{2}\, db_t + [(d-1)j'(\rho(x, x_t))/j(\rho(x, x_t)) + b]\, dt ,$$

where b_t is an one-dimensional Brownian motion. By Itô's formula we obtain

$$dJ(\rho(x, x_t)) \leqq \sqrt{2}J'(\rho(x, x_t))\, db_t + dt$$

which implies $E^{x,y}J(\rho(x, x_{\tau^\delta})) \leqq E^{x,y}\tau^\delta$.

Next, by replacing $T \wedge S_N$ with τ^δ and taking $\lambda = 0$ in the proof of Lemma 4.3, we obtain $E^{x,y}\tau^\delta \leqq \frac{1}{4}F_{2\delta}(\rho(x, y))$. Therefore

$$P^{x,y}(\rho(x, x_{\tau^\delta}) \geqq \delta) \leqq \frac{E^{x,y}J(\rho(x, x_{\tau^\delta}))}{J(\delta)} \leqq \frac{F_{2\delta}(\rho(x, y))}{4J(\delta)} .$$

Finally, the proof of the second inequality is similar, the only difference is that

$$E^{x,y}J(\rho(x, y_{\tau^\delta})) \leqq E^{x,y}\tau^\delta + J(\rho(x, y)) . \qquad \square$$

By (5.1), (5.2), (5.3) and Lemma 5.1, we obtain the following result immediately.

Theorem 5.2 *Let* $u \geqq 0$, *then*

$$|\nabla P_t^D u(x)| \leqq \|u\|_\infty \inf_{\delta \leqq \delta_x} \left\{ \inf_{\lambda \in [0, 4F_{2\delta}(2\delta)^{-1})} \frac{\lambda g(2\delta)}{4 - \lambda F_{2\delta}(2\delta)} (e^{\lambda t} - 1)^{-1} + \frac{g(2\delta)}{2J(\delta)} \right\} .$$

Proof. Simply note that $\lim_{r \to 0} J(r)/r = 0$. $\square$

Corollary 5.3 *If* $u \geqq 0$ *and* $Lu = 0$, *then*

$$|\nabla u(x)| \leqq \|u\|_\infty \inf_{\delta \leqq \delta_x} \frac{g(2\delta)}{2J(\delta)} \leqq 2de \max\{\sqrt{K^+(d-1)} + 2b, \delta_x^{-1}\}\|u\|_\infty .$$

Proof. Note that $P_t^D u = u$, then the first estimate follows from Theorem 5.2 by letting $t \to \infty$. Next, since $\mathrm{Ric} \geqq -K^+$, we assume that $K \geqq 0$.

Note that $\gamma(r) \leqq 2[\sqrt{K(d-1)}+b]$, then

$$g(2\delta) \leqq \int_0^{2\delta} \exp\left[\frac{r}{2}(\sqrt{K(d-1)}+b)\right] dr \leqq 2\delta \exp[\delta(\sqrt{K(d-1)}+b)] .$$

On the other hand,

$$J(\delta) = \int_0^{\delta} e^{br}\{\sinh[r\sqrt{K/(d-1)}]\}^{1-d} \int_0^{r} e^{bs}\{\sinh[s\sqrt{K/(d-1)}]\}^{d-1} ds$$

$$\geqq e^{-b\delta}\int_0^{\delta} r^{1-d}\int_0^{r} s^{d-1} ds = \frac{\delta^2}{2d}e^{-b\delta} .$$

Then the proof is completed by taking $\delta = \delta_x \wedge (\sqrt{K(d-1)}+2b)^{-1}$ in the first inequality. □

Recall that for $K \geqq 0$, it is proved in [7] that there exists $c(K,d,b)$ such that

$$|\nabla u| \leqq c(K,d,b)(1+\delta_x^{-1})\|u\|_\infty$$

holds for all $u \geqq 0$ with $Lu = 0$. Hence, Corollary 5.3 can be considered as an improvement of this result.

Acknowledgments. The author thanks Prof. M.F. Chen for useful discussion. The comments and suggestions of D. Bakry, D. Elworthy, L. Gross, E. Hsu, W. Kendall and A. Thalmaier, who participated the 1995 Warwick workshop on Differential Geometry and Stochastic Analysis, are gratefully acknowledged. Thanks are also given to a referee for helpful comments on the first version of the paper.

References

1. Bakry, D., Emery, M.: Diffusions hypercontractive. In: Séminair de Probabilités XIX (Lect. Notes Math., Vol. 1123, pp. 175–206) Berlin: Springer 1985
2. Bismut, J.M.: The Witten complex and the degenerate Morse inequalities. J. Differential Geom. **23**, 207–240 (1986)
3. Chavel, I.: Eigenvalues in Riemannian geometry. New York: Academic Press, 1984
4. Chen, M.F., Li, S.F.: Coupling methods for multi-dimensional diffusion process. Ann. Probab. **17**, 151–177 (1989)
5. Chen, M.F., Wang, F.Y.: Application of coupling method to the first eigenvalue on manifolds. Sci. Sin. (English Edition) **37**, 1–14 (1994)
6. Chen, M.F., Wang, F.Y.: Estimates of logarithmic Sobolev constant – An improvement of Bakry–Emery criterion. J. Funct. Anal. (to appear)
7. Cranston, M.: Gradient estimates on manifolds using coupling. J. Funct. Anal. **99**, 110–124 (1991)
8. Davies, E.B.: Heat kernels and spectral theory, New York: Cambridge University Press 1989
9. Deuschel, J.D., Stroock, D.W.: Hypercontractivity and spectral gap of symmetric diffusions with applications to the stochastic Ising models. J. Funct. Anal. **92**, 30–48 (1990)
10. Elworthy, K.D.: Stochastic flows on Riemannian manifolds. In: Pinsky, M.A., Wihstutz, V. (eds.) Diffusion processes and related problems in analysis, Vol. II, pp. 37–72. Basel: Birkhäser 1992

11. Elworthy, K.D., Li, X.M.: Formulae for the derivatives of heat semigroups. J. Funct. Anal. **125**, 252–286 (1994)
12. Gross, L.: Logarithmic Sobolev inequalities and contractivity properties of semigroups. In: Dirichlet forms (Lect. Notes Math., Vol. 1563, pp. 64–82) Berlin: Springer 1993
13. Hsu, E.: Logarithmic Sobolev inequalities on path spaces (preprint 1994)
14. Kendall, W.S.: Nonnegative Ricci curvature and the Brownian coupling property. Stochastics **19**, 111–129 (1986)
15. Kendall, W.S.: The radial part of Brownian motion on a manifold: a semimartingale property. Ann. Probab. **15**, 1491–1500 (1987)
16. Li, P., Yau, S.T.: On the parabolic kernel of the Schrödinger operator. Acta Math. **156**, 153–201 (1986)
17. Matthews, P.: Mixing rates for Brownian motion in a convex polyhedron. J. Appl. Probab. **27**, 259–286 (1990)
18. Qian, Z.: Gradient estimates and heat kernel estimates. Proc. Roy. Soc. Edinburgh A **125**, 975–990 (1995)
19. Setti, A.G.: Gaussian estimates for the heat kernel of the weighted Laplacian and fractal measures. Canad. J. Math. **44**, 1061–1078 (1992)
20. Wang, F.Y.: Application of coupling method to the Neumann eigenvalue problem. Probab. Theory Relat. Fields **98**, 299–306 (1994)
21. Wang, F.Y.: Gradient estimates for generalized harmonic functions on manifolds. Chinene Sci. Bull. (English edition) **39**, 1849–1852 (1994)
22. Wang, F.Y.: Logarithmic Sobolev inequalities for diffusion processes with application to path space (preprint 1995)