Harnack Inequality and Heat Kernel Estimates on Manifolds with Curvature Unbounded Below ⋆

Marc Arnaudon [a], Anton Thalmaier [a], Feng-Yu Wang [b,∗]

[a] *Département de Mathématiques, Université de Poitiers, Téléport 2 – BP 30179, 86962 Futuroscope Chasseneuil, France*

[b] *School of Mathematical Sciences, Beijing Normal University, Beijing 100875, China*

Abstract

Using the coupling by parallel translation, along with Girsanov's theorem, a new version of a dimension-free Harnack inequality is established for diffusion semigroups on Riemannian manifolds with Ricci curvature bounded below by $-c\,(1+\rho_o^2)$, where $c > 0$ is a constant and ρ_o is the Riemannian distance function to a fixed point o on the manifold. As an application, in the symmetric case, a Li-Yau type heat kernel bound is presented for such semigroups.

Key words: Harnack inequality, heat kernel, diffusion semigroup, curvature
2000 *AMS subject Classification:* 58J65, 60H30, 47D07

1 Introduction

The difficulties of extending the elliptic Harnack inequality to the parabolic situation are well studied; see the classical work of Moser [19,20], as well as [9,13,14]. In particular, it is in general not possible to compare, for instance on compact sets, different values of a heat semigroup $P_t f$ (for f non-negative) by a constant only depending on t. There are several ways to deal with this deficiency: typically the parabolic Harnack inequality is formulated by introducing

⋆ Supported in part by NNSFC(10121101) and RFDP(20040027009).
∗ Corresponding author.
Email addresses: marc.arnaudon@math.univ-poitiers.fr (Marc Arnaudon), anton.thalmaier@math.univ-poitiers.fr (Anton Thalmaier), wangfy@bnu.edu.cn (Feng-Yu Wang).

a shift in time; another possibility is by seeking for Hölder type inequalities with an exponent strictly bigger than 1. For a discussion of the difficulties to obtain Harnack inequalities by probabilistic methods from Bismut type formulas, see for instance [23].

In 1997, a dimension-free Harnack inequality (with an exponent bigger than 1) was established in [26] for diffusion semigroups with generators having curvature bounded from below. This inequality has been applied and further developed to the study of functional inequalities (see [1,22,27,29]), heat kernel estimates (see [4,10]), higher order eigenvalues (see [28,30,11]), transportation cost inequalities (see [5]), and short time behavior of transition probabilities (see [2,3,16]). Due to the potential of applications, it would be useful to establish inequalities of this type also for diffusions with curvature unbounded below. On the other hand, since the formulation of the inequality in [26] is equivalent to an underlying lower curvature bound (see [31]), the formulation of the resulting inequality will be slightly different in the present paper.

Let M be a connected complete Riemannian manifold of dimension d, either with convex boundary ∂M or without boundary. Let $o \in M$ be a fixed point, ρ be the Riemannian distance function, and $\rho_o(x) := \rho(o,x)$, $x \in M$. Consider the (reflecting) diffusion semigroup P_t on M generated by $L := \Delta + Z$ for some C^1-vector field Z. We assume that the corresponding (reflecting) diffusion process is non-explosive. We shall prove the dimension-free Harnack inequality for P_t under the following condition.

Assumption 1 *There exists a constant $c > 0$ such that for all $x \in M$,*

$$\begin{aligned} &\mathrm{Ric}_x := \inf\{\mathrm{Ric}(X,X): \ X \in T_xM, \ |X| = 1\} \geq -c\,(1 + \rho_o(x)^2), \\ &h_Z(x) := \sup\{\langle \nabla_X Z, X\rangle: \ X \in T_xM, \ |X| = 1\} \leq c\,(1 + \rho_o(x)), \\ &\langle Z, \nabla \rho_o\rangle(x) \leq c\,(1 + \rho_o(x)). \end{aligned} \tag{1.1}$$

In this case we have no longer a gradient estimate like $|\nabla P_t f| \leq C_t P_t |\nabla f|$, which has been crucial for deriving the original dimension-free Harnack inequality (cf. the proof of [26, Lemma 2.2]). Under our condition it is possible to prove a weaker type estimate such as $|\nabla P_t f|^p \leq C_t P_t |\nabla f|^p$ for $p > 1$, but this is not enough to imply the desired Harnack inequality by following the original proof. Hence, in this paper we develop a new argument in terms of coupling by parallel translation and Girsanov's theorem.

The main idea is as follows. Given two points $x_0 \neq y_0$ on M, let (x_t, y_t) be the coupling by parallel translation of the L-diffusion process starting from (x_0, y_0). To force the two marginal processes to meet before a given time T, we make a Girsanov transformation of y_t, denoted by $\tilde{y}_t$, which is equal to x_t at $t = T$ and is generated by L under a weighted probability $\mathbb{Q} := R\mathbb{P}$ with a density R induced by the Girsanov transform. Then, for any bounded measurable function on M, one has

$$\begin{aligned}|P_T f|^\alpha(y_0) &= \Big|\mathbb{E}_{\mathbb{Q}}[f(\tilde y_T)]\Big|^\alpha = \Big|\mathbb{E}[Rf(x_T)]\Big|^\alpha \\ &\le P_T|f|^\alpha(x_0)\,(\mathbb{E}[R^{\alpha/(\alpha-1)}])^{\alpha-1}, \quad \alpha > 1.\end{aligned}$$

To derive a Harnack inequality, it suffices therefore to prove that $\mathbb{E}[R^p] < \infty$ for $p > 1$ and to estimate this quantity. We will be able to realize this idea under Assumption 1 (cf. Section 2 and Section 3 below for a complete proof).

Theorem 2 *Suppose that Assumption* 1 *holds. For any* $\varepsilon \in \,]0,1]$ *there exits a constant* $c(\varepsilon) > 0$ *such that*

$$\begin{aligned}|P_t f|^\alpha(y) \le P_t|f|^\alpha(x)\,\exp\bigg\{&\frac{\alpha(\varepsilon\alpha+1)\rho(x,y)^2}{2(2-\varepsilon)(\alpha-1)t} \\ &+ \frac{c(\varepsilon)\alpha^2(\alpha+1)^2}{(\alpha-1)^3}\left(1+\rho(x,y)^2\right)\rho(x,y)^2 \\ &+ \frac{\alpha-1}{2}\left(1+\rho_o(x)^2\right)\bigg\}\end{aligned}$$

holds for all $\alpha > 1$, $t > 0$, $x, y \in M$ *and any bounded measurable function* f *on* M, *where* $\rho(x,y)$ *is the Riemannian distance from* x *to* y *and* $\rho_o(x) = \rho(o,x)$.

As an application of the above Harnack inequality, we present a heat kernel estimate as in [10]. Assume that $Z = \nabla V$ for some C^2-function V on M, such that P_t is symmetric w.r.t. the measure $\mu(\mathrm{d}x) := \mathrm{e}^{V(x)}\,\mathrm{d}x$, where $\mathrm{d}x$ is the Riemannian volume measure. Let $p_t(x,y)$ be the transition density of P_t w.r.t. μ; that is,

$$P_t f(x) = \int_M p_t(x,y) f(y)\mu(\mathrm{d}y), \quad x \in M,\ t > 0,\ f \in C_b(M).$$

Corollary 3 *Suppose that Assumption* 1 *holds and let* $Z = \nabla V$. *For any* $\delta > 2$ *there exists a constant* $c(\delta) > 0$ *such that for any* $t > 0$,

$$p_t(x,y) \le \frac{\exp\left\{-\frac{\rho(x,y)^2}{2\delta t} + c(\delta)\left(1 + t + t^2 + \rho_o(x)^2 + \rho_o(y)^2\right)\right\}}{\sqrt{\mu(B(x,\sqrt{2t}))\mu(B(y,\sqrt{2t}))}}, \quad x, y \in M,$$

where $B(x,r)$ *is the geodesic ball centered at* x *in* M *with radius* r.

2 Proof of Theorem 2 without cut-locus

To explain our argument in a simple way, we assume in this section that the cut-locus is empty; that is, $\mathrm{Cut}(M) := \{(x,y) \in M \times M : x \in \mathrm{cut}(y)\} = \varnothing$.

In the next section, we then treat the technical details for $\mathrm{Cut}(M) \neq \varnothing$. Moreover, if ∂M is convex, we may assume that M is a regular domain in a Riemannian manifold such that the minimal geodesic linking any two points in M is contained in M, see [32, Proposition 2.1.5]. Thus, according to the proof of [25, Lemma 2.1], the reflection of the two marginal processes at the boundary makes them move together faster. Hence, without loss of generality, we may and will assume that $\partial M = \varnothing$. Finally, in the sequel we assume that f is a nonnegative measurable bounded function on M.

We now recall the construction of coupling by parallel translation. Let B_t be a d-dimensional Brownian motion. Then the L-diffusion process starting at $x_0 \in M$ can be constructed by solving the following SDE:

$$\mathrm{d}x_t = \sqrt{2}\,\Phi_t \circ \mathrm{d}B_t + Z(x_t)\,\mathrm{d}t, \quad x_0 \in M, \tag{2.1}$$

where Φ_t denotes the horizontal lift of x_t; that is

$$\mathrm{d}\Phi_t = H_{\Phi_t} \circ \mathrm{d}x_t, \quad \Phi_0 \in \mathrm{O}_{x_0}(M),$$

in terms of the horizontal lift operator $H : \pi^* TM \to T\mathrm{O}(M)$.

For given points $x \neq y$, let $e(x,y)\colon [0, \rho(x,y)] \to M$ be the unique minimal geodesic from x to y and let $P_{x,y} : T_xM \to T_yM$ be the parallel translation along the geodesic $e(x,y)$. In particular $P_{x,x} = I$, the identity operator. Consider the Itô equation

$$\mathrm{d}^{\text{Itô}} y_t = \sqrt{2}\, P_{x_t,y_t} \Phi_t\, \mathrm{d}B_t + Z(y_t)\,\mathrm{d}t, \quad y_0 \in M, \tag{2.2}$$

where in coordinates the Itô differential is given by

$$(\mathrm{d}^{\text{Itô}} y_t)^k = \mathrm{d}y_t^k + \frac{1}{2}\sum_{i,j} \Gamma_{ij}^k(y_t)\mathrm{d}[y_t^i, y_t^j],$$

see Emery [8]. Recall that (2.2) is equivalent to the system of equations

$$\begin{aligned} d\Psi_t &= H_{\Psi_t} \circ dy_t, \quad \Psi_0 \in \mathrm{O}_{y_0}(M), \\ \mathrm{d}y_t &= \sqrt{2}\,\Psi_t \circ \mathrm{d}B_t' + Z(y_t)\,\mathrm{d}t, \quad y_0 \in M, \\ \mathrm{d}B_t' &= \Psi_t^{-1} P_{x_t,y_t} \Phi_t\, \mathrm{d}B_t, \end{aligned}$$

where the last equation is an Itô equation in $\mathbb{R}^d$ and Ψ_t is the horizontal lift of y_t. See [17, (2.1)] for an analogous construction (with the mirror reflection operator). Since $P_{x,y}$ is smooth, y_t is a well-defined L-diffusion process starting at y_0. We call the pair (x_t, y_t) the coupling by parallel translation of the L-diffusion process.

To calculate the distance process $\rho(x_t, y_t)$, let $M_{x,y}\colon T_xM \to T_yM$ be the mirror reflection operator along the geodesic $e(x,y)$; that is, $M_{x,y}X := P_{x,y}X$

if $X \perp \dot{e}$, while $M_{x,y}X := -P_{x,y}X$ if $X \parallel \dot{e}$ at the point x. Let $\{u^i\}_{i=0}^{d-1}$ be an orthonormal basis in $\mathbb{R}^d$ such that $\Phi_t u^0 = \dot{e}$ at x_t. Define $v^i := (\Psi_t^{-1} P_{x_t,y_t} \Phi_t) u^i$, $i = 0, \dots, d-1$. Since $\langle \Phi_t u^i, \dot{e} \rangle (x_t) = 0$ for all $i \neq 0$, we have

$$v^0 = -(\Psi_t M_{x_t,y_t} \Phi_t) u^0, \quad v^i = (\Psi_t M_{x_t,y_t} \Phi_t) u^i, \quad i \neq 0.$$

Then [17, Theorem 2 and (2.5)] implies

$$\mathrm{d}\rho(x_t, y_t) \leq I_Z(x_t, y_t)\,\mathrm{d}t, \quad t \leq \tau, \tag{2.3}$$

where $\tau := \inf\{t \geq 0 : x_t = y_t\}$ is the coupling time and

$$\begin{aligned} I_Z(x,y) = \sum_{i=1}^{d-1} \int_0^{\rho(x,y)} \Big(|\nabla_{\dot{e}(x,y)} J_i|^2 - \langle R(\dot{e}(x,y), J_i)\dot{e}(x,y), J_i \rangle \Big)_s \,\mathrm{d}s \\ + Z\rho(\cdot, y)(x) + Z\rho(x, \cdot)(y). \end{aligned}$$

Here R denotes the Riemann curvature tensor, $\dot{e}(x,y)$ the tangent vector of the geodesic $e(x,y)$, and $\{J_i\}_{i=1}^{d-1}$ are Jacobi fields along $e(x,y)$ which, together with $\dot{e}(x,y)$, constitute an orthonormal basis of the tangent space at x and y:

$$J_i(\rho(x,y)) = P(x,y)\, J_i(0), \quad i = 1, \dots, d-1.$$

To calculate $I_Z(x_t, y_t)$ we may take $(\Phi_t(u^i))$ at x_t and $(\Psi_t(v^i))$ at y_t. Let

$$\begin{aligned} K(x,y) &:= \sup_{z \in e(x,y)} (-\mathrm{Ric}_z)^+, \\ \delta(x,y) &:= \sup\{\langle \nabla_X Z, X \rangle_z : z \in e(x,y),\ X \in T_z M,\ |X| = 1\}. \end{aligned}$$

We have

$$Z\rho(\cdot, y)(x) + Z\rho(x, \cdot)(y) = \int_0^{\rho(x,y)} \langle \nabla_{\dot{e}(x,y)} Z, \dot{e}(x,y) \rangle_s \,\mathrm{d}s \leq \delta(x,y)\rho(x,y).$$

Thus, by [32, Theorem 2.1.4] (see also [7] and [6]), we obtain

$$\begin{aligned} I_Z(x,y) \leq 2\sqrt{K(x,y)(d-1)}\ \tanh\left(\frac{\rho(x,y)}{2} \sqrt{K(x,y)/(d-1)} \right) \\ + \delta(x,y)\rho(x,y). \end{aligned} \tag{2.4}$$

To construct a coupling such that the coupling time is less than a given $T > 0$, let us consider the equation

$$\begin{aligned} \mathrm{d}^{\text{Itô}} \tilde{y}_t = \sqrt{2}\, P_{x_t, \tilde{y}_t} \Phi_t \,\mathrm{d}B_t + Z(\tilde{y}_t)\,\mathrm{d}t \\ - \left(I_Z(x_t, \tilde{y}_t) + \frac{\rho(x_0, y_0)}{T} \right) n(\tilde{y}_t, x_t)\,\mathrm{d}t, \quad \tilde{y}_0 = y_0, \end{aligned} \tag{2.5}$$

where $n(y,x) := \dot{e}(y,x)|_y = \nabla\rho(x, \cdot)(y) \in T_y M$ for $x \neq y$. Since $n(x,y)$ is smooth outside the diagonal $D := \{(x,x) : x \in M\}$, the solution $\tilde{y}_t$ exists and is unique up to the coupling time $\tilde{\tau} := \inf\{t \geq 0 : x_t = \tilde{y}_t\}$. We let $\tilde{y}_t = x_t$ for $t \geq \tilde{\tau}$. As in (2.3) we have

$$\mathrm{d}\rho(x_t, \tilde{y}_t) \leq -\frac{\rho(x_0, y_0)}{T}\,\mathrm{d}t, \quad t \leq \tilde{\tau},$$

so that $\tilde{\tau} \leq T$. Let

$$\begin{aligned} N_t &:= \frac{1}{\sqrt{2}} \int_0^{t\wedge\tilde{\tau}} \Big\langle P_{x_s,\tilde{y}_s}\Phi_s \mathrm{d}B_s, \Big(I_Z(x_s,\tilde{y}_s) + \frac{\rho(x_0,y_0)}{T}\Big) n(\tilde{y}_s, x_s)\Big\rangle, \\ R_t &:= \exp\Big(N_t - \frac{1}{2}[N]_t\Big). \end{aligned} \tag{2.6}$$

By Girsanov's theorem, $\{\tilde{y}_t\}$ is an L-diffusion under the weighted probability measure $\mathbb{Q} := R_T\,\mathbb{P}$. Therefore,

$$\begin{aligned} P_T f(y) &= \mathbb{E}_{\mathbb{Q}}[f(\tilde{y}_T)] = \mathbb{E}[R_T f(x_T)] \\ &\leq (\mathbb{E}[f^\alpha(x_T)])^{1/\alpha}\, (\mathbb{E}R_T^\beta)^{1/\beta}, \quad \alpha^{-1} + \beta^{-1} = 1. \end{aligned} \tag{2.7}$$

By (2.4) and (2.6) we have

$$\begin{aligned} [N]_T &\leq \frac{1}{2}\int_0^T \Big(I_Z(x_t,\tilde{y}_t) + \frac{\rho(x_0,y_0)}{T}\Big)^2 \mathrm{d}t \\ &\leq \frac{1}{2}\int_0^T \Big(2\sqrt{(d-1)K(x_t,\tilde{y}_t)} + \delta(x_t,\tilde{y}_t)\rho(x_t,\tilde{y}_t) + \frac{\rho(x_0,y_0)}{T}\Big)^2 \mathrm{d}t. \end{aligned}$$

Exploiting the conditions (1.1) and the fact that $\rho(x_t,\tilde{y}_t) \leq \rho(x_0,y_0)$, we obtain, given $\varepsilon \in\,]0,1]$,

$$[N]_T \leq \int_0^T \Big\{c_1(1+\rho(x_0,y_0)^2)(1+\rho_o(x_t)^2) + \frac{\rho(x_0,y_0)^2}{(2-\varepsilon)T^2}\Big\}\,\mathrm{d}t \tag{2.8}$$

for some constant $c_1 = c_1(\varepsilon) \geq 1$. Next, by (1.1) and the Laplacian comparison theorem, we get

$$\begin{aligned} \mathrm{d}\rho_o(x_t) &\leq \sqrt{2}\,\mathrm{d}b_t \\ &\quad + \sqrt{c\,(1+\rho_o(x_t)^2)/(d-1)}\coth\Big[\rho_o(x_t)\sqrt{c\,(1+\rho_o(x_t)^2)/(d-1)}\Big]\,\mathrm{d}t \\ &\quad + c\,(1+\rho_o(x_t))\,\mathrm{d}t \\ &=: \sqrt{2}\,\mathrm{d}b_t + U_t\,\mathrm{d}t, \end{aligned}$$

where b_t is a one-dimensional Brownian motion. In particular, this gives

$$\begin{aligned} \mathrm{d}\rho_o(x_t)^2 &\leq 2\sqrt{2}\,\rho_o(x_t)\,\mathrm{d}b_t + [2U_t\rho_o(x_t) + 2]\,\mathrm{d}t \\ &\leq 2\sqrt{2}\,\rho_o(x_t)\,\mathrm{d}b_t + c_2(1+\rho_o(x_t)^2)\,\mathrm{d}t \end{aligned}$$

for some constant $c_2 > 0$. Thus, for any $\gamma, \delta > 0$, we have

$$\begin{aligned} &\mathrm{d}\, \mathrm{e}^{\gamma(1+\rho_o(x_t)^2)\,\mathrm{e}^{-\delta t}} \\ &\quad \le \mathrm{d}M_t - \gamma\,\mathrm{e}^{-\delta t}\Big((\delta - c_2)(1+\rho_o(x_t)^2) - 4\gamma\,\mathrm{e}^{-\delta t}\,\rho_o(x_t)^2\Big)\,\mathrm{e}^{\gamma(1+\rho_o(x_t)^2)\,\mathrm{e}^{-\delta t}}\,\mathrm{d}t \end{aligned}$$

for some local martingale M_t. Letting $\delta := c_2 + 4\gamma$ we arrive at

$$\mathbb{E}\left[\exp\{\gamma(1+\rho_o(x_t)^2)\,\mathrm{e}^{-\delta t}\}\right] \le \exp\{\gamma(1+\rho_o(x_0)^2)\}.$$

Therefore, there exists a constant $\delta_0 \in \,]0,1]$ such that

$$\mathbb{E}\left[\exp\{\delta_0(1+\rho_o(x_t)^2)\}\right] \le \exp\{1+\rho_o(x_0)^2\}, \quad t \le 1. \tag{2.9}$$

Let $p := 1+\varepsilon$, $q := (1+\varepsilon)/\varepsilon$. The proof of Theorem 2 is completed in two more steps.

I. Assume that $T \le T_0 := 2\delta_0/[c_1\beta(\beta p - 1)q(1+\rho(x_0,y_0)^2)]$. We have

$$\gamma := \beta(\beta p - 1)qc_1(1+\rho(x_0,y_0)^2)\,T/2 \le \delta_0.$$

Then, by (2.8) and (2.9), we obtain

$$\begin{aligned} &\mathbb{E}\Big[\exp\{\beta(\beta p-1)q\,[N]_T/2\}\Big] \\ &\quad \le \exp\left\{\frac{\beta(\beta p-1)q\rho(x_0,y_0)^2}{2(2-\varepsilon)T}\right\}\mathbb{E}\left[\exp\left\{\frac{\gamma}{T}\int_0^T (1+\rho_o(x_t)^2)\,\mathrm{d}t\right\}\right] \\ &\quad \le \exp\left\{\frac{\beta(\beta p-1)q\rho(x_0,y_0)^2}{2(2-\varepsilon)T}\right\}\frac{1}{T}\int_0^T \mathbb{E}\left[\exp\left\{\delta_0(1+\rho_o(x_t)^2)\right\}\right]\mathrm{d}t \\ &\quad \le \exp\left\{\frac{\beta(\beta p-1)q\rho(x_0,y_0)^2}{2(2-\varepsilon)T} + 1 + \rho_o(x_0)^2\right\}. \end{aligned}$$

Combining this with (2.7) and the fact that

$$\begin{aligned} \mathbb{E}[R_T^\beta] &= \mathbb{E}\left[\exp\left\{\beta N_T - p\beta^2\,[N]_T/2\right\}\exp\{\beta(\beta p-1)\,[N]_T/2\}\right] \\ &\le \Big(\mathbb{E}\exp\left\{p\beta\,N_T - p^2\beta^2\,[N]_T/2\right\}\Big)^{1/p}\Big(\mathbb{E}\exp\{\beta(\beta p-1)q\,[N]_T/2\}\Big)^{1/q} \\ &= \Big(\mathbb{E}\exp\{\beta(\beta p-1)q\,[N]_T/2\}\Big)^{1/q}, \end{aligned}$$

we conclude that

$$\begin{aligned} (P_Tf)^\alpha(y_0) \le P_Tf^\alpha(x_0)\,\exp\bigg\{&\frac{\alpha(p\beta-1)\rho(x_0,y_0)^2}{2(2-\varepsilon)T} \\ &+ \frac{\alpha}{q\beta}(1+\rho_o(x_0)^2)\bigg\}. \end{aligned} \tag{2.10}$$

II. If $T > T_0$, then by (2.10) and Jensen's inequality,

$$\begin{aligned}
(P_T f)^\alpha(y_0) &\le (P_{T_0}(P_{T-T_0} f))^\alpha(y_0) \\
&\le P_{T_0}(P_{T-T_0} f)^\alpha(x_0)\ \exp\left\{\frac{\alpha(p\beta-1)\rho(x_0,y_0)^2}{2(2-\varepsilon)T_0} + \frac{\alpha}{q\beta}(1+\rho_o(x_0)^2)\right\} \\
&\le P_T f^\alpha(x_0)\ \exp\Big\{c_2\alpha\beta(2\beta-1)^2\,(1+\rho(x_0,y_0)^2)\rho(x_0,y_0)^2 \\
&\qquad\qquad + \frac{\alpha}{2\beta}\,(1+\rho_o(x_0)^2)\Big\}
\end{aligned}$$

for some $c_2 = c_2(\varepsilon)$. In conclusion, combining this with (2.10), the desired inequality for some $c(\varepsilon) > 0$ is obtained.

3 Proof of Theorem 2 with cut-locus

As already explained in Section 2, we may assume that $\partial M = \varnothing$. When $\mathrm{Cut}(M) \neq \varnothing$, the idea (originally due to Cranston [7]) is to construct the coupling by parallel translation outside of $\mathrm{Cut}(M)$ and to let the two marginal processes move independently on the cut-locus. This idea has been realized in [32] by an approximation argument. Since the present situation is different due to the Girsanov transformation, we reformulate the procedure here in detail for our purpose of achieving a coupling time smaller than a given $T > 0$.

By Itô's formula it is easy to see that, when $\mathrm{Cut}(M) = \varnothing$, the generator of the coupling $(x_t, \tilde{y}_t)$ is (cf. the proof of [15, Theorem 6.5.1])

$$\begin{aligned}
L(x) + L(y) + 2\sum_{i,j=1}^{d} \langle P_{x,y} X_i(x), Y_j(y)\rangle X_i(x) Y_j(y) \\
- \Big(I_Z(x,y) + \rho(x_0,y_0)/T\Big)\, n(y,x),
\end{aligned}$$

where $L(x)$ and $L(y)$ denote the operator L acting on the first and the second components respectively, and $\{X_i\}$ and $\{Y_i\}$ are local frames normal at x and y, respectively. Note that this operator is independent of the choices of the local frames. Thus, in the general situation, we intend to construct a process generated by

$$\begin{aligned}
\tilde{L}(x,y) := L(x) + L(y) + 2\,1_{\complement \mathrm{Cut}(M)}(x,y) \sum_{i,j=1}^{d} \langle P_{x,y}\, X_i(x), Y_j(y)\rangle X_i(x) Y_j(y) \\
- \Big((1_{\complement \mathrm{Cut}(M)} I_Z)(x,y) + \rho(x_0,y_0)/T\Big)\, n(y,x), \qquad (3.1)
\end{aligned}$$

where $n(y,x)$ is set to be zero on $\mathrm{Cut}(M) \cup D$. To this end, we adopt an approximation argument as in [32, §2.1].

For any $n \geq 1$ and $\varepsilon \in]0,1[$, let $h_{n,\varepsilon} \in C^\infty(M \times M)$ such that

$$0 \leq h_{n,\varepsilon} \leq 1-\varepsilon, \quad h_{n,\varepsilon}|\complement\mathrm{Cut}(M)_n = 1-\varepsilon, \quad h_{n,\varepsilon}|\mathrm{Cut}(M)_{2n} = 0,$$

where

$$\mathrm{Cut}(M)_n := \{(x,y) : \rho_{M\times M}((x,y), \mathrm{Cut}(M)) \leq 1/n\}, \quad n \geq 1$$

and $\rho_{M\times M}$ is the Riemannian distance on $M \times M$. Consider the operator

$$\begin{aligned}\tilde{L}_{n,\varepsilon}(x,y) := L(x) + L(y) + 2h_{n,\varepsilon}(x,y) \sum_{i,j=1}^{d} \langle P_{x,y}X_i(x), Y_j(y)\rangle X_i(x)Y_j(y) \\ - \Big(I_Z(x,y) + (1-h_{n,\varepsilon}(x,y))J(x,y) + \rho(x_0,y_0)/T\Big) n(y,x),\end{aligned}$$

with $J \in C^\infty((M\times M)\setminus D)$ such that

$$L\rho(\cdot,y)(x) + L\rho(x,\cdot)(y) \leq J(x,y), \quad x \neq y.$$

Since $\tilde{L}_{n,\varepsilon}$ is a uniformly elliptic second order differential operator with smooth diffusion coefficients and the drift is smooth outside of D, it generates a unique diffusion process $(x_t, y_t^{n,\varepsilon})$ up to the coupling time $\tau_{n,\varepsilon} := \inf\{t \geq 0 : x_t = y_t^{n,\varepsilon}\}$ (cf. [24, Theorem 6.4.3]), which can be constructed by solving (2.1) and the Itô SDE

$$\begin{aligned}\mathrm{d}^{\text{Itô}} y_t^{n,\varepsilon} &= \sqrt{2h_{n,\varepsilon}(x_t, y_t^{n,\varepsilon})}\, P_{x_t, y_t^{n,\varepsilon}} \Phi_t \,\mathrm{d}B_t \\ &+ \sqrt{2(1-h_{n,\varepsilon}(x_t, y_t^{n,\varepsilon}))}\, \Psi_t^{n,\varepsilon}\,\mathrm{d}B_t' + Z(y_t^{n,\varepsilon})\,\mathrm{d}t \\ &- \Big(I_Z(x_t, y_t^{n,\varepsilon}) + (1-h_{n,\varepsilon}(x_t, y_t^{n,\varepsilon}))J(x_t, y_t^{n,\varepsilon}) + \rho(x_0,y_0)/T\Big) n(y_t^{n,\varepsilon}, x_t)\,\mathrm{d}t\end{aligned}$$

with initial condition $y_0^{n,\varepsilon} = y_0$, where Φ_t and $\Psi_t^{n,\varepsilon}$ are the horizontal lifts of x_t and $y_t^{n,\varepsilon}$ respectively, and B_t and B_t' are two independent d-dimensional Brownian motions. Indeed, the last equation may be solved first neglecting the drift term involving $n(y_t^{n,\varepsilon}, x_t)$ and then by applying Girsanov's theorem. We set $y_t^{n,\varepsilon} = x_t$ for $t \geq \tau_{n,\varepsilon}$. By the choice of J and using Itô's formula of the radial process as presented in [18], we get

$$\mathrm{d}\rho(x_t, y_t^{n,\varepsilon}) \leq 2\sqrt{1-h_{n,\varepsilon}(x_t, y_t^{n,\varepsilon})}\,\mathrm{d}b_t^{n,\varepsilon} - \frac{\rho(x_0,y_0)}{T}\,\mathrm{d}t, \quad t \leq \tau_{n,\varepsilon}, \qquad (3.2)$$

where $b_t^{n,\varepsilon}$ is a one-dimensional Brownian motion. Therefore, letting $\mathbb{P}_{n,\varepsilon}^{x_0,y_0}$ be the distribution of $(x_t, y_t^{n,\varepsilon})_{t\in[0,T]}$, where here and in the sequel, $(\xi_\cdot, \eta_\cdot) \in C([0,T]; M \times M)$ is the canonical path, we have

$$\lim_{N\to\infty} \sup_{n,\varepsilon} \mathbb{P}_{n,\varepsilon}^{x_0,y_0} \left\{ \sup_{t\in[0,T]} \rho(\xi_t, \eta_t) \geq N \right\} = 0. \qquad (3.3)$$

Since the first marginal distribution of $\mathbb{P}^{x_0,y_0}_{n,\varepsilon}$ is $\mathbb{P}^{x_0}$, the distribution of the L-diffusion process starting at x_0, it follows from (3.3) that

$$\begin{aligned}
&\mathbb{P}^{x_0,y_0}_{n,\varepsilon}\left\{\sup_{s,t\in[0,T]}\rho_{M\times M}\Big((\xi_s,\eta_s),(\xi_t,\eta_t)\Big)\geq N\right\}\\
&\quad\leq \mathbb{P}^{x_0,y_0}_{n,\varepsilon}\left\{\sup_{s,t\in[0,T]}\Big(2\rho(\xi_s,\xi_t)+\rho(\xi_s,\eta_s)+\rho(\xi_t,\eta_t)\Big)\geq N\right\}\\
&\quad\leq \mathbb{P}^{x_0,y_0}_{n,\varepsilon}\left\{\sup_{s,t\in[0,T]}\rho(\xi_s,\xi_t)\geq N/4\right\}+\mathbb{P}^{x_0,y_0}_{n,\varepsilon}\left\{\sup_{t\in[0,T]}\rho(\xi_t,\eta_t)\geq N/4\right\}\\
&\quad= \mathbb{P}^{x_0}\left\{\sup_{s,t\in[0,T]}\rho(\xi_s,\xi_t)\geq N/4\right\}+\mathbb{P}^{x_0,y_0}_{n,\varepsilon}\left\{\sup_{t\in[0,T]}\rho(\xi_t,\eta_t)\geq N/4\right\}\to 0,
\end{aligned}$$

as $N\to\infty$. Thus, by [21, Lemma 4] the family $\{\mathbb{P}^{x_0,y_0}_{n,\varepsilon} : n\geq 1,\ \varepsilon\in\,]0,1[\}$ is tight. We take $n_k\to\infty$ and $\varepsilon_\ell\to 0$ such that $\mathbb{P}^{x_0,y_0}_{n_k,\varepsilon_\ell}$ converges weakly to some $\mathbb{P}^{x_0,y_0}_{\varepsilon_\ell}$ $(\ell\geq 1)$ as $k\to\infty$ while $\mathbb{P}^{x_0,y_0}_{\varepsilon_\ell}$ converges weakly to some $\mathbb{P}^{x_0,y_0}$ as $\ell\to\infty$. It is trivial to see that $\mathbb{P}^{x_0,y_0}$ solves the martingale problem for $\tilde L$ up to the coupling time; that is, for any $f\in C_0^\infty((M\times M)\setminus D)$,

$$f(\xi_t,\eta_t)-\int_0^t \tilde L f(\xi_s,\eta_s)\,\mathrm{d}s,\quad t\leq T,$$

is a $\mathbb{P}^{x_0,y_0}$-martingale w.r.t. the natural filtration up to $\inf\{t\geq 0:\xi_t=\eta_t\}$. Moreover, according to the proof of [32, Theorem 2.1.1] and (3.2), there holds

$$\mathrm{d}\rho(\xi_t,\eta_t)\leq -\frac{\rho(x_0,y_0)}{T}\,\mathrm{d}t,\quad \mathbb{P}^{x_0,y_0}\text{-a.s.}\tag{3.4}$$

Hence, there exist two independent d-dimensional Brownian motions B_t and B_t' on a complete probability space $(\Omega,\mathscr{F},\mathscr{F}_t,\mathbb{P})$, and two processes x_t and $\tilde y_t$ on M such that equation (2.1) and

$$\begin{aligned}
\mathrm{d}^{\text{Itô}}\tilde y_t =&\sqrt{2}\,1_{\complement\mathrm{Cut}(M)}(x_t,\tilde y_t)P_{x_t,y_t}\Phi_t\,\mathrm{d}B_t+\sqrt{2}\,1_{\mathrm{Cut}(M)}(x_t,\tilde y_t)\Psi_t\,\mathrm{d}B_t'+Z(\tilde y_t)\,\mathrm{d}t\\
&-\Big(I_Z(x_t,\tilde y_t)+\rho(x_0,y_0)/T\Big)\,n(\tilde y_t,x_t)\,\mathrm{d}t,\quad t\leq\tilde\tau,
\end{aligned}$$

hold, where Φ_t and Ψ_t are the horizontal lifts of x_t and $\tilde y_t$ respectively, and $\tilde\tau:=\inf\{t\geq 0: x_t=\tilde y_t\}$. Moreover, by (3.4) we have

$$\mathrm{d}\rho(x_t,\tilde y_t)\leq -\frac{\rho(x_0,y_0)}{T}\,\mathrm{d}t,$$

as well as $\tilde\tau\leq T$. Let $y_t=x_t$ for $t\geq\tilde\tau$ and let R_t be defined as in Section 2 with

$$\begin{aligned}
N_t=\frac{1}{\sqrt{2}}\int_0^{t\wedge\tilde\tau}\Big\langle &P_{x_s,\tilde y_s}\Big(1_{\complement\mathrm{Cut}(M)}(x_s,\tilde y_s)\Phi_s\,\mathrm{d}B_s+1_{\mathrm{Cut}(M)}(x_s,\tilde y_s)\Psi_s\,\mathrm{d}B_s'\Big),\\
&\Big(I_Z(x_s,\tilde y_s)+\frac{\rho(x_0,y_0)}{T}\Big)n(\tilde y_s,x_s)\Big\rangle.
\end{aligned}$$

We conclude that $\tilde{y}_t$ is generated by L under the probability $\mathbb{Q} := R_T\mathbb{P}$. The remainder of the proof is analogous to the case where $\mathrm{Cut}(M) = \varnothing$.

4 Proof of Corollary 3

Given $t > 0$, let $T > 0$, $p \in]1,2[$ and $q := p/2(p-1)$ be such that $qt < T$. Applying Theorem 2 with $\alpha := 2/p$ and $\varepsilon = 1$, we obtain, for any bounded nonnegative measurable function f,

$$\begin{aligned} I &:= \mu(B(x,\sqrt{2t}))\, \mathrm{e}^{-c_1\left(1+t+t^2+\rho_o(x)^2\right)-t/(T-qt)}(P_t f(x))^2 \\ &\le \int_{B(x,\sqrt{2t})} (P_t f^\alpha(y))^p \exp\Big\{ -c_1(1+t+t^2+\rho_o(x)^2) - \frac{t}{T-qt} \\ &\quad + \frac{\alpha(\alpha+1)p}{\alpha-1} + \frac{2p\, c(1)\alpha^2(\alpha+1)^2(1+2t)t}{(\alpha-1)^3} + \frac{p}{2}(\alpha-1)(1+\rho_o(y)^2)\Big\}\mu(\mathrm{d}y). \end{aligned}$$

Since on $B(x,\sqrt{2t})$ one has $\rho_o(y)^2 \le 2\rho_o(x)^2 + 4t$, there exists a constant $c_1 = c_1(p) > 0$ such that

$$I \le \int_{B(x,\sqrt{2t})} (P_t f^\alpha(y))^p \exp\left\{-\frac{1}{2}\frac{\rho(x,y)^2}{(T-qt)}\right\}\mu(\mathrm{d}y).$$

Combining this with [10, (2.9)] we arrive at

$$I \le \int_M f^2(y) \exp\left\{-\frac{\rho(x,y)^2}{2T}\right\}\mu(\mathrm{d}y). \tag{4.1}$$

Taking $f(y) := (n \wedge p_t(x,y)) \exp\left\{n \wedge \frac{\rho(x,y)^2}{2T}\right\}$, $y \in M$, we obtain from (4.1) that

$$\begin{aligned} &\int_M (n \wedge p_t(x,y))^2 \exp\left\{n \wedge \frac{\rho(x,y)^2}{2T}\right\}\mu(\mathrm{d}y) \\ &\qquad \le \frac{\exp\{c_1(p)(1+t+t^2+\rho_o(x)^2) + t/(T-qt)\}}{\mu(B(x,\sqrt{2t}))}. \end{aligned}$$

For any $\delta > 2$, letting $T := \delta t/2$ and $q := 1/2 + \delta/4$, we obtain

$$\begin{aligned} E_\delta(x,t) &:= \int_M p_t(x,y)^2 \exp\left\{\frac{\rho(x,y)^2}{\delta t}\right\}\mu(\mathrm{d}y) \\ &\le \frac{\exp\{c(\delta)(1+t+t^2+\rho_o(x)^2)\}}{\mu(B(x,\sqrt{2t}))} \end{aligned}$$

for some $c(\delta) > 0$. Therefore, by [12, (3.4)] we have

$$p_t(x,y) \le \exp\left\{\frac{-\rho(x,y)^2}{2\delta t}\right\}\sqrt{E_\delta(x,t)E_\delta(y,t)}$$
$$\le \frac{\exp\left\{c(\delta)\Big(1+t+t^2+\rho_o(x)^2+\rho_o(y)^2\Big)-\rho(x,y)^2/(2\delta t)\right\}}{\sqrt{\mu(B(x,\sqrt{2t})\mu(B(y,\sqrt{2t}))}}.$$

References

[1] S. Aida, *Uniform positivity improving property, Sobolev inequalities, and spectral gaps*, J. Funct. Anal. **158** (1998), no. 1, 152–185.

[2] S. Aida and H. Kawabi, *Short time asymptotics of a certain infinite dimensional diffusion process*, Stochastic analysis and related topics, VII (Kusadasi, 1998), Progr. Probab., vol. 48, Birkhäuser Boston, Boston, MA, 2001, pp. 77–124.

[3] S. Aida and T. Zhang, *On the small time asymptotics of diffusion processes on path groups*, Potential Anal. **16** (2002), no. 1, 67–78.

[4] D. Bakry, M. Ledoux and Z. Qian, *Logarithmic Sobolev inequalities, Poincaré inequalities and heat kernel bounds,* unpublished manuscript (1997).

[5] S. G. Bobkov, I. Gentil and M. Ledoux, *Hypercontractivity of Hamilton-Jacobi equations*, J. Math. Pures Appl. (9) **80** (2001), no. 7, 669–696.

[6] Mu–Fa Chen and Feng–Yu Wang, *Application of coupling method to the first eigenvalue on manifold*, Sci. China Ser. A **37** (1994), no. 1, 1–14.

[7] M. Cranston, *Gradient estimates on manifolds using coupling*, J. Funct. Anal. **99** (1991), no. 1, 110–124.

[8] M. Emery, *Stochastic calculus in manifolds*, Springer-Verlag, Berlin, 1989. With an appendix by P.-A. Meyer.

[9] E. B. Fabes and D. W. Stroock, *A new proof of Moser's parabolic Harnack inequality using the old ideas of Nash*, Arch. Rational Mech. Anal. **96** (1986), no. 4, 327–338.

[10] Fu–Zhou Gong and Feng–Yu Wang, *Heat kernel estimates with application to compactness of manifolds*, Q. J. Math. **52** (2001), no. 2, 171–180.

[11] ———, *On Gromov's theorem and L^2-Hodge decomposition*, Int. J. Math. Math. Sci. (2004), no. 1-4, 25–44.

[12] A. Grigor'yan, *Gaussian upper bounds for the heat kernel on arbitrary manifolds*, J. Differential Geom. **45** (1997), no. 1, 33–52.

[13] A. Grigor'yan and L. Saloff-Coste, *Stability results for Harnack inequalities*, Ann. Inst. Fourier (Grenoble) **55** (2005), no. 3, 825–890.

[14] W. Hebisch and L. Saloff-Coste, *On the relation between elliptic and parabolic Harnack inequalities*, Ann. Inst. Fourier (Grenoble) **51** (2001), no. 5, 1437–1481.

[15] Elton P. Hsu, *Stochastic analysis on manifolds*, Graduate Studies in Mathematics, vol. 38, American Mathematical Society, Providence, RI, 2002.

[16] H. Kawabi, *The parabolic Harnack inequality for the time dependent Ginzburg-Landau type SPDE and its application*, Potential Anal. **22** (2005), no. 1, 61–84.

[17] W. S. Kendall, *Nonnegative Ricci curvature and the Brownian coupling property*, Stochastics **19** (1986), no. 1-2, 111–129.

[18] ———, *The radial part of Brownian motion on a manifold: a semimartingale property*, Ann. Probab. **15** (1987), no. 4, 1491–1500.

[19] J. Moser, *On Harnack's theorem for elliptic differential equations*, Comm. Pure Appl. Math. **14** (1961), 577–591.

[20] ———, *A Harnack inequality for parabolic differential equations*, Comm. Pure Appl. Math. **17** (1964), 101–134. Correction: Comm. Pure Appl. Math. **20** (1967), 231–236.

[21] M.-K. Renesse, *Intrinsic coupling on Riemannian manifolds and polyhedra,* Electronic J. Probab. **9** (2004), 411–435.

[22] M. Röckner and Feng–Yu Wang, *Supercontractivity and ultracontractivity for (non-symmetric) diffusion semigroups on manifolds*, Forum Math. **15** (2003), no. 6, 893–921.

[23] D. W. Stroock, *An introduction to the analysis of paths on a Riemannian manifold*, Mathematical Surveys and Monographs, vol. 74, American Mathematical Society, Providence, RI, 2000.

[24] D. W. Stroock and S. R. S. Varadhan, *Multidimensional diffusion processes*, Grundlehren der Mathematischen Wissenschaften, vol. 233, Springer-Verlag, Berlin, 1979.

[25] Feng–Yu Wang, *Application of coupling methods to the Neumann eigenvalue problem*, Probab. Theory Related Fields **98** (1994), no. 3, 299–306.

[26] ———, *Logarithmic Sobolev inequalities on noncompact Riemannian manifolds*, Probab. Theory Related Fields **109** (1997), no. 3, 417–424.

[27] ———, *Harnack inequalities for log-Sobolev functions and estimates of log-Sobolev constants*, Ann. Probab. **27** (1999), no. 2, 653–663.

[28] ———, *Functional inequalities, semigroup properties and spectrum estimates*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. **3** (2000), no. 2, 263–295.

[29] ———, *Logarithmic Sobolev inequalities: conditions and counterexamples*, J. Operator Theory **46** (2001), no. 1, 183–197.

[30] ———, *Functional inequalities and spectrum estimates: the infinite measure case*, J. Funct. Anal. **194** (2002), no. 2, 288–310.

[31] ———, *Equivalence of dimension-free Harnack inequality and curvature condition*, Integral Equations Operator Theory **48** (2004), no. 4, 547–552.

[32] ———, *Functional Inequalities, Markov Properties, and Spectral Theory,* Science Press, Beijing/New York, 2004.