The Annals of Probability
2000, Vol. 28, No. 4, 1666–1679

A CONDITION FOR THE EQUIVALENCE OF COUPLING AND SHIFT COUPLING

BY M. CRANSTON[1] AND FENG-YU WANG[2]

University of Rochester and Beijing Normal University

It is proved in this paper that a weak parabolic Harnack inequality for a Markov semigroup implies the existence of a coupling and a shift coupling for the corresponding process with equal chances of success. This implies equality of the tail and invariant σ-fields for the diffusion as well as equality of the class of bounded parabolic functions and the class of bounded harmonic functions.

1. Introduction. Coupling methods have been widely used in the study of Markov processes (see [12] and references therein.) In this paper we will compare coupling and shift coupling of a Markov process. The tail σ-field is naturally associated to coupling and the invariant σ-field is naturally associated to shift coupling. To describe this association, let $(E, \mathscr{B})$ be a Polish space and $(P_t, t \in I)$ a conservative Markov semigroup on $C(E)$, where I is either $\mathbb{Z}_+$ (discrete time) or $\mathbb{R}_+$ (continuous time). Let $(\Omega, \mathscr{F}_t, \theta_t)$ be the canonical path space on E (cadlag paths) with the usual σ-fields $\mathscr{F}_t$ and shift operator θ_t. Then $\mathscr{J} = \bigcap_t \theta_t^{-1}\mathscr{F}_\infty$ and $\mathscr{I} = \{\Lambda \in \mathscr{F}_\infty : \theta_t^{-1}\Lambda = \Lambda, t \in I\}$ are the tail and invariant σ-fields respectively. For any two initial distributions μ and ν in $\mathscr{P}(E)$, the space of probability measures on $(E, \mathscr{B})$, let X and Y be two copies of the Markov process on E with semigroup P_t and initial distributions μ and ν, respectively.

We shall say (X, Y) is a shift coupling if there are stopping times S and T for X and Y, respectively, such that $X_S = Y_T$ a.s. on the set $\{S < \infty\} \cap \{T < \infty\}$. We shall call (X, Y) a coupling if $S = T$ a.s. For clarity we shall use T' for S (or T) in couplings. By $\mathbb{P}_{\mu,\nu}$ (or $\mathbb{P}$ for simplicity) we denote the distribution law of a coupling (or a shift coupling) with initial distribution pair (μ, ν), and let $\mathbb{P}_\mu$ be the distribution law of the Markov process with initial distribution μ. A shift coupling (respectively, a coupling) is called successful if $S \vee T < \infty$ (respectively, $T' < \infty$) a.s. We say a shift coupling is maximally successful if $\mathbb{P}(S < \infty) + \mathbb{P}(T < \infty)$ is as large as possible and that a coupling is maximally successful if $\mathbb{P}(T' < \infty)$ is as large as possible.

Existence of maximally successful couplings and shift couplings have been established in various contexts in [9], [1], [5] and [19]. We will say coupling

Received April 1999; revised March 2000.

[1]Supported by a grant from the NSA.

[2]Supported by the Alexander von Humboldt Foundation NNSFC (19631060) and Fok Ying-Tung Educational Foundation.

AMS 1991 *subject classifications*. 60J60, 60J45.

Key words and phrases. Coupling, shift coupling, parabolic inequality, tail σ-field, invariant σ-field.

and shift coupling are equivalent when $P(S<\infty)+P(T<\infty)=2P(T'<\infty)$ for maximally successful coupling and shift coupling.

Given a measure ν on $(E,\mathscr{B})$ with sub σ-field $\mathscr{A}\subseteq\mathscr{B}$, $\nu\big|_{\mathscr{A}}$ denotes the restriction of ν to $\mathscr{A}$, $\|\nu\|$ denotes the total variation norm of ν and denote $\nu\mathrm{V}=\int_0^\infty \nu P_t dt$. Let us recall some relevant results which will be used in our present study.

THEOREM 1 [12], [19]. *The following statements are equivalent*:

(a) *A successful coupling exists for any initial distributions* $\mu,\nu\in\mathscr{P}(E)$;
(b) $\|(\mu-\nu)P_t\|\to 0$ *as* $t\to\infty$ *for any* $\mu,\nu\in\mathscr{P}(E)$;
(c) $\mathbb{P}_\mu$ *is trivial on* $\mathscr{J}$ *for any* $\mu\in\mathscr{P}(E)$;
(d) $\mathbb{P}_\mu=\mathbb{P}_\nu$ *on* $\mathscr{J}$ *for any* $\mu,\nu\in\mathscr{P}(E)$.

THEOREM 2 [1], [19]. *The following statements are equivalent:*

(e) *A successful shift coupling exists for any* $\mu,\nu\in\mathscr{P}(E)$;
(f) $\frac{1}{t}\left\|\int_0^t(\mu-\nu)P_s ds\right\|\to 0$ *as* $t\to\infty$ *for any* $\mu,\nu\in\mathscr{P}(E)$;
(g) $\mathbb{P}_\mu$ *is trivial on* $\mathscr{I}$ *for any* $\mu\in\mathscr{P}(E)$;
(h) $\mathbb{P}_\mu=\mathbb{P}_\nu$ *on* $\mathscr{I}$ *for any* $\mu,\nu\in\mathscr{P}(E)$.

We let $\mathscr{D}^1$ denote the set of parabolic functions $h(t,x)$ (i.e., $P_s h(t+s,x)=h(t,x)$ for any $t>0$, $x\in E$) for which $\sup_{t>0,x\in E}|h(t,x)|\le 1$ and $\mathscr{H}^1$ the set of harmonic functions ($P_t h(x)=h(x)$ for any $t>0$, $x\in E$ for which $\sup_x|h(x)|\le 1$. Similarly, $\mathscr{D}_b$ and $\mathscr{H}_b$ will denote the sets of bounded parabolic and harmonic functions respectively. The following result expresses the relation between $\mathscr{D}^1$ and coupling and that between $\mathscr{H}^1$ and shift-coupling.

THEOREM 3 [5]. *Assume that* $I=\mathbb{R}_+$ *and* $\mu\mathrm{V},\nu\mathrm{V}$ *are* σ-*finite.*

(a) *There exists a coupling* (X,Y) *with coupling time* T' *such that*

$$2\mathbb{P}(T'=\infty)=\sup_{h\in\mathscr{D}^1}\langle h,(\mu-\nu)\otimes\delta_0\rangle=\lim_{t\to\infty}\|(\nu-\mu)P_t\|\,.$$

(b) *There exists a shift coupling* (X,Y) *with coupling epochs* (S,T) *such that*

$$\mathbb{P}(S=\infty)+\mathbb{P}(T=\infty)=\sup_{h\in\mathscr{H}^1}\langle h,\mu-\nu\rangle=\lim_{t\to\infty}\left\|\frac{1}{t}\int_0^t(\mu-\nu)P_s ds\right\|\,.$$

We also mention the following results which can be found in [8]. For the coupling in Theorem 3,

$$\left\|\mathbb{P}_\mu(X\in\cdot)\big|_{\mathscr{J}}-\mathbb{P}_\nu(Y\in\cdot)\big|_{\mathscr{J}}\right\|=\sup\left\{\int_\Omega Z(d\mathbb{P}_\mu-d\mathbb{P}_\nu);Z\in\mathscr{J},|Z|\le 1\ \text{ a.s.}\right\}$$
$$=\sup_{h\in\mathscr{D}^1}\langle h,(\mu-\nu)\otimes\delta_0\rangle\,.$$

For the shift coupling in Theorem 3,

$$\|\mathbb{P}_\mu(X \in \cdot)|_{\mathscr{I}} - \mathbb{P}_\nu(Y \in \cdot)|_{\mathscr{I}}\| = \sup\left\{\int_\Omega Z(d\mathbb{P}_\mu - d\mathbb{P}_\nu); Z \in \mathscr{I}, |Z| \le 1 \quad \text{a.s.}\right\}$$
$$= \sup_{h\in\mathscr{H}^1} \langle h, \mu - \nu\rangle\ .$$

A consequence of the above is:

COROLLARY 1. *Under the assumption of Theorem* 3.

(a) *There is a coupling* (X, Y) *with coupling epoch* T' *such that*

$$\|\mathbb{P}_\mu(X \in \cdot)|_{\mathscr{J}} - \mathbb{P}_\nu(Y \in \cdot)_{\mathscr{J}}\| = 2\mathbb{P}_\mu(T' = \infty) = \sup_{h\in\mathscr{D}^1}\langle h, (\mu - \nu) \otimes \delta_0\rangle.$$

(b) *There is a shift coupling* (X, Y) *with coupling epochs* (S, T) *such that*

$$\|\mathbb{P}_\mu(X \in \cdot)|_{\mathscr{I}} - \mathbb{P}_v(Y \in \cdot)|_{\mathscr{I}}\| = P(S = \infty) + P(T = \infty) = \sup_{h\in\mathscr{H}^1} \langle h, \mu - \nu\rangle\ .$$

The application of the parabolic Harnack inequality to the equivalence of coupling and shift coupling is enabled by the Derrienic 0-2 Law. Define for $x \in E$, $h > 0$,

$$\alpha(x, h) = \lim_{t\to\infty} \|\delta_x P_{t+h} - \delta_x P_t\|$$

and

$$\alpha(h) = \sup_{x\in E} \alpha(x, h)\ .$$

THEOREM 4 (DERRIENIC 0-2 LAW [8]). *For any* $h > 0$, $\alpha(h) = 0$ *or* 2. *Moreover,* $\alpha(h) = 0$ *for some* $h > 0$ *if and only if* $\mathscr{I} = \mathscr{J}$ *(equivalently,* $\mathscr{D}_b = \mathscr{H}_b$*),* $\mathbb{P}_\mu$ *a.s. for every Borel probability measure* μ *on* E.

The plan of the paper is as follows: In Section 2 we will show how a parabolic Harnack inequality can be used to apply Derrienic's 0-2 law. In Section 3 we offer examples which satisfy this parabolic Harnack inequality. In Section 4 we show how gradient estimates can be used to apply Derrienic. In Section 5 we show by means of an example that parabolic Harnack is not a necessary condition for the equivalence of coupling and shift coupling.

2. Equivalence of coupling and shift coupling. In terms of Theorems 1, 2, 3 and 4, the key point in proving the equivalence of coupling and shift coupling is to show that $\alpha(h) = 0$ for some $h > 0$. To this end, we shall use the following parabolic Harnack-type inequality which will be studied in the next section: there exist $t, h \in I\setminus\{0\}$ and a nonnegative increasing $\Phi \in C[0, 1]$ such that $\Phi(0) < 1$ and

$$P_t f \le \Phi(P_{t+h} f), \qquad 0 \le f \le 1. \tag{2.1}$$

Note that Φ may depend on t and h.

THEOREM 5. *Assume that* (2.1) *holds for some* $t, h \in I \setminus \{0\}$ *and increasing* $\Phi \in C[0,1]$ *with* $\Phi(0) < 1$. *Then* $\alpha(h) = 0$. *Consequently, under* (2.1) *a successful coupling exists for any pair of initial distributions if and only if so does a successful shift coupling.*

PROOF. By the contractivity of P_t, we have

$$\alpha(x,h) \le \|\delta_x P_t - \delta_x P_{t+h}\|, \qquad x \in E. \tag{2.2}$$

By (2.1), for any $0 \le f \le 1$ we have

$$P_{t+h} f(x) \ge \Phi^{-1}(P_t f(x)) := \inf\{r \ge 0 : \Phi(r) \ge P_t f(x)\}. \tag{2.3}$$

It then follows from (2.2) and (2.3) that

$$\begin{aligned}\alpha(h) = \sup_x \alpha(x,h) &= 2 \sup_{0 \le f \le 1, x \in E} \{P_t f(x) - P_{t+h} f(x)\}\\ &\le 2 \sup_{s \in [0,1]} \left\{s - \Phi^{-1}(s)\right\} < 2\end{aligned}$$

since $\Phi(0) < 1$ and $\Phi^{-1}(r) > 0$ for $r > \Phi(0)$ by our assumption on Φ. This implies that $\alpha(h) = 0$ by Theorem 4. The proof is now complete by Theorems 1, 2 and 4. □

REMARK. Condition (2.1) is in some sense sharp for $\alpha(h) = 0$. Actually, by the definition of $\alpha(h)$ and the proof of Theorem 2, $\alpha(h) = 0$ if and only if for any x, there exists $t_x \in I$ such that

$$P_{t_x} f(x) \le P_{t_x + h} f(x) + \tfrac{1}{2}, \qquad 0 \le f \le 1. \tag{2.4}$$

If t_x is independent of x, then (2.1) holds for $\Phi(s) = s + \frac{1}{2}$.

COROLLARY 2. *Assume that* $I = \mathbb{R}_+$ *and* μV *is* σ*-finite for any* $\mu \in \mathscr{P}(E)$. *If* (2.1) *holds, then* $\mathscr{I} = \mathscr{J}$, $\mathscr{D}_b = \mathscr{H}_b$,

$$\lim_{t\to\infty} \frac{1}{t} \left\| \int_0^t (\nu - \mu) P_t \, dt \right\| = \lim_{t\to\infty} \|(\nu - \mu) P_t\|, \tag{2.5}$$

and for any $\mu, \nu \in \mathscr{P}(E)$, *there exist a shift coupling* (X, Y, S, T) *and a coupling* (X', Y', T') *with initial distribution* (μ, ν) *such that*

$$\begin{aligned} 2\mathbb{P}(T' = \infty) &= \max\nolimits_{h \in \mathscr{D}^1} \langle (\nu - \mu) \otimes \delta_0, h\rangle,\\ \mathbb{P}(S = \infty) + \mathbb{P}(T = \infty) &= \max\nolimits_{h \in \mathscr{H}^1} \langle \nu - \mu, h\rangle. \end{aligned} \tag{2.6}$$

PROOF. By Theorem 5, $\alpha(h) = 0$ for any h, thus $\mathscr{I} = \mathscr{J}$ and $\mathscr{D}_b = \mathscr{H}_b$ according to [8]. Consequently, (2.5) and (2.6) follows from Theorem 3 and Corollary 1. □

Before ending this section, we present a result on the weak parabolic Harnack inequality (2.1) under perturbations.

THEOREM 6. *Let $E = M$ be a complete Riemannian manifold, and $P_t^{(i)}$ the diffusion semigroups generated by $L_i = \Delta + Z_i$, where $Z_i (i = 1, 2)$ are two C^1-vector fields. If $c := \|Z_1 - Z_2\|_\infty < \infty$, then $P_t^{(1)}$ satisfies* (2.1) *for some t, h and Φ_1 with $\Phi_1(0) = 0$ if and only if $P_t^{(2)}$ also satisfies* (2.1) *for the same t, h and some Φ_2 satisfying $\Phi_2(0) = 0$.*

PROOF. The key point of the proof is the following comparison between semigroups:

$$|P_t^{(i)} f|^r \le \exp\left[\frac{c^2 rt}{4(r-1)}\right] P_t^{(j)} |f|^r, \quad r > 1,\ i, j = 1, 2. \tag{2.7}$$

The proof of this inequality is similar to that of Lemma 2.1 in [23]. For positive $f \in C_0^\infty(M)$, let $\phi(s) = P_s^{(i)}(P_{t-s}^{(j)} f)^r$, $s \in [0, t]$. It is easy to see that

$$\begin{aligned} \phi'(s) &\ge \frac{r}{P_s^{(i)}(P_{t-s}^{(j)} f)^r} P_s^{(i)}(P_{t-s}^{(j)} f)^r \left\{ (r-1)\frac{|\nabla P_{t-s}^{(j)} f|^2}{(P_{t-s}^{(j)} f)^2} - c\frac{|\nabla P_{t-s}^{(j)} f|}{P_{t-s}^{(j)} f} \right\} \\ &\ge -\frac{rc^2}{4(r-1)}. \end{aligned} \tag{2.8}$$

This implies (2.7) immediately.

Now, assume that $P_t^{(1)}$ satisfies (2.1) for t, s and Φ_1. By (2.7) with $r = 2$, for any f with $0 \le f \le 1$ we have

$$\begin{aligned} [P_t^{(2)} f]^2 &\le \exp[c^2 t/2][P_t^{(1)} f] \le \exp[c^2 t/2]\Phi_1(P_{t+s}^{(1)} f) \\ &\le \exp[c^2 t/2]\Phi_1\left(\exp[c^2 t/4]\sqrt{P_{t+s}^{(j)} f}\right). \end{aligned}$$

Therefore, (2.1) also holds for $P_t^{(2)}$ with

$$\Phi_2(r) = \exp[c^2 t/4]\sqrt{\Phi_1\big(\exp[c^2 t/4] r^{1/2}\big)}. \qquad \square$$

3. Weak parabolic Harnack inequalities. The aim of this section is to present sufficient conditions for the weak parabolic Harnack inequality (2.1). Let us first recall some results on parabolic Harnack inequalities which are much stronger than (2.1).

Let $E = M$ be a connected, complete Riemannian manifold and let P_t be the diffusion semigroup generated by $L = \Delta + Z$ for some C^1-vector field Z. A well known parabolic Harnack inequality was proved by Li and Yau[14] for $Z = 0$ under the assumption that the Ricci curvature is bounded from below. This inequality has been up to now extended and refined by several papers, for instance, Setti[17] (for the case $Z = \nabla V$ with Ric $+\exp[V]\mathrm{Hess}_{\exp[-V]}$ bounded from below), Qian [15] (for bounded $|Z|$ with Ric $-\langle \nabla_\cdot Z, \cdot\rangle$ bounded from below), Yau [24] (for general non-self-adjoint operators) and Bakry and Qian [2] (under a curvature-dimension condition). We state here the result in [2] which recovers and improves those in [14], [15] and [17].

Given a second order operator L, let $\Gamma(f, g) = Lfg - fLg - gLf$ and

$$\Gamma_2(f, g) = \tfrac{1}{2}\{L\Gamma(f, g) - \Gamma(f, Lg) - \Gamma(Lf, g)\}.$$

Define a distance ρ on E by setting $\rho(x, y) = \sup\{\psi(x) - \psi(y) : \Gamma(\psi, \psi) \le 1\}$. In the case $E = M$, a Riemannian manifold and $L = \Delta + Z$, ρ is the usual Riemannian distance.

If there exist $K \ge 0$ and $n \in (0, \infty)$ such that

$$\Gamma_2(f, f) \ge -K\Gamma(f, f) + \frac{1}{n}(Lf)^2, \quad f \in C_0^\infty(M). \tag{3.1}$$

Then

$$\begin{aligned} P_t f(x) \le P_{t+s} f(y)\Big[\frac{t+s}{t}\Big]^{n/2} \exp\Big[& \frac{\big(\rho + \sqrt{nKs}\big)^2}{4s} \\ & + \frac{\sqrt{nK}}{2} \min\Big\{(\sqrt{2}-1)\rho, \frac{\sqrt{nK}}{2}s\Big\}\Big], \end{aligned} \tag{3.2}$$

Taking $y = x$ in (3.2) so that $\rho = \rho(x, x) = 0$, we see (2.1) holds with $\Phi(r) = r(t + h/t)^{n/2} e^{nK/4}$. Note that when $n = \infty$, (3.1) becomes the usual curvature condition: Ric $-\langle \nabla_{\cdot} Z, \cdot\rangle \ge -K$.

Next, we introduce a result by Saloff-Coste [16] using the doubling property, $D(R)$, and local Poincaré inequality, $P(R)$, described below. Let $Z = \nabla V$ and denote $d\mu = \exp[V]dx$, where dx is the Riemannian volume element on M. Let $\mu(x, r)$ be the measure of μ on $B(x, r)$, the geodesic ball with center x and radius r. Given $R > 0$, the conditions $D(R)$ and $P(R)$ read as follows.

$D(R)$: there exists $C_1 > 0$ such that for all $r \in (0, R)$ and all $x \in M$, $\mu(x, 2r) \le C_1 \mu(x, r)$.

$P(R)$: there exists $C_2 > 0$ such that for all $r \in (0, R)$ and all $x \in M$, if $f_r = (1/\mu(x, r)) \int_{B(x,r)} f d\mu$, then

$$\int_{B(x,r)} |\psi - \psi_r|^2 d\mu \le C_2 r^2 \int_{B(x,2r)} |\nabla\psi|^2 d\mu, \quad \psi \in C^1(B(x, 2r)) .$$

The properties $D(R)$ and $P(R)$ are particularly nice as they are essentially geometric properties and are equivalent to a parabolic Harnack inequality which we will discuss below. Also, these properties make sense in settings more general than the class of Riemannian manifolds which also we will discuss later on. Finally the properties $P(R)$ and $D(R)$ are preserved by quasi-isometries and isometries at infinity (see [16] and [6]). On the other hand, however, $D(R)$ fails to be true if the measure μ is finite and decays faster than polynomial. For instance, $\mu(dx) = \exp[-c|x|^\varepsilon]dx$ in $\mathbb{R}^d$ for some $\varepsilon, c > 0$.

According to [16], if

$$P(R) \text{ and } D(R) \tag{3.3}$$

hold, then there is a $C > 0$ such that given $x \in M$, $0 < r < R$ and any positive f, one has

$$P_t f(x_1) \le P_{t+s} f(x_2) \exp\left\{C\left[\frac{\rho^2(x_1, x_2)}{s} + \left(\frac{1}{t} + \frac{1}{r}\right) s\right]\right\} \tag{3.4}$$

for $t, s > 0$ and $x_1, x_2 \in B(x, \frac{1}{2}r)$. Once again, with $x_2 = x_1$, (2.1) holds under $D(R)$ and $P(R)$ with $\Phi(r) = r \exp\left\{C\left(\frac{1}{t} + \frac{1}{R}\right) h\right\}$.

REMARKS. We would like to examine various examples where $P(R)$ and $D(R)$ and therefore (3.4) hold. Consequently, Theorem 5 holds for all of the examples below. Our discussion follows [16].

1. Assume (M, g) is a Riemannian manifold for which $P(R)$ and $D(R)$ hold for $Z = 0$ and some $R > 0$, and $\tilde{g}$ is another metric on M for which $\lambda g \le \tilde{g} \le \Lambda g$ for two positive constants λ and Λ. Then $(M, \tilde{g})$ also satisfies $D(R)$ and $P(R)$. We note that $D(R)$ follows from the Bishop comparison theorem when (M, g) is a manifold satisfying $\mathrm{Ric} \ge -Kg$. Furthermore, Buser has proven an L^1 version of the Poincaré inequality (which is stronger than Poincaré): with $f_B = 1/\mu(B) \int_B f d\mu$,

$$\int_B |f - f_B| d\mu \le \exp\left[C_n(1 + \sqrt{\kappa}\, r) r\right] \int_B |\nabla f| d\mu$$

under the same Ricci curvature assumption, $\mathrm{Ric} \ge -Kg$. Thus, If $(M, \tilde{g})$ is conformally equivalent to (M, g) and $\mathrm{Ric} \ge -Kg$, then (3.4) holds for $Z = 0$.
2. By the results of [6], $P(R)$ and $D(R)$ are preserved by rough isometries. This allows one to deduce $\mathscr{H}_b = \mathscr{D}_b$ for a discretization of a manifold which satisfies $P(R)$ and $D(R)$ or for a manifold isometric at infinity to one which satisfies $P(R)$ and $D(R)$. See [6] for details.
3. If G is a unimodular Lie group and μ Haar measure on G, let $X_1, \ldots, X_k$ be left invariant vector fields on G and set $L = -\sum_{i=1}^k X_i^2$. A distance can be defined by setting

$$\rho(x, y) = \sup\{\psi(x) - \psi(y) : \psi \in C_0^\infty(G),\ \Gamma(\psi, \psi) \le 1\}$$

where Γ is the carré du champ above. Under this distance, metric balls on G either satisfy

$$\forall r > 1, \quad cr^N \le \mu(r) \le Cr^N$$

or

$$\forall r > 1, \quad c \exp[Cr] \le \mu(r) \le C \exp[Cr]\,.$$

These are results of [10]. In either case, $D(R)$ is satisfied. Varopoulos [20] showed $P(R)$ is satisfied for these operators. Thus the parabolic Harnack inequality holds for these operators.

4. Subelliptic operators on $\mathbb{R}^d$ satisfy $P(R)$ and $D(R)$ under certain conditions. These are operators of the form

$$Lf = -m(x)^{-1} \sum_{i,j=1}^{n} \partial_i (m(x) a_{ij}(x) \partial_j f)$$

with $0 < m(x) < \infty$, $(a_{ij}(x))$ positive semi-definite and m and a smooth. L is self-adjoint on $L^2(\mathbb{R}^d, \mu)$ with $\mu(dx) = m(x)dx$, $\Gamma(f, f) = \sum_{i,j=1}^n a_{ij}(x) \partial_i f \partial_j f$. L is locally subelliptic if for each bounded domain $U \subset \mathbb{R}^d$

$$\|(I - \Delta)^{2\varepsilon} f\|_2 \le C(\|Lf\|_2 + \|f\|_2), \quad f \in C_0^\infty(U) \tag{3.5}$$

for some $C, \varepsilon > 0$. Suppose (3.5) holds with $U_3 = \{x : \|x\| \le 3\}$ (Euclidean ball of radius 3) and assume m, m^{-1}, a_{ij} and their derivatives of any order are bounded by a constant A in U_3. Then if $B(x, r) = \{y : \rho(x, y) \le r\}$, there exist constants D and P such that

$$\mu(B(x, 2r)) \le D\mu(B(x, r))\,, \quad x \in U_1,\ 0 < r < 1,$$

$$\int_{B(x,n)} |f - f_B| d\mu \le P \int_B \sqrt{\Gamma(f, f)}\, d\mu, \quad x \in U_1, 0 < r < 1\,.$$

Thus, if the subellipticity and boundedness condition hold on $U_3(x)$ for any $x \in \mathbb{R}^d$ with the same constant A, then $D(1)$ and $P(1)$ hold.

We now present some new criteria for our weak parabolic Harnack inequality (2.1).

THEOREM 7. *Let $E = M$ be a complete, connected Riemannian manifold, and let P_t be generated by $L = \Delta + Z$ for some C^1-vector field Z. If*

$$\begin{aligned} &(3.1) \textit{ holds for some } K \ge 0 \textit{ and } n = \infty, \textit{ and there exist} \\ &r > 0, h > 0 \textit{ such that } p_r := \inf\nolimits_{x \in M} P_h 1_{B(x,r)}(x) > 0. \end{aligned} \tag{3.6}$$

Then, for any $t > 0$ and the above $h > 0$, (2.1) holds for

$$\Phi(s) = \exp\left[\frac{Kr^2}{2(1 - \exp[-2Kt])}\right] \sqrt{p_r s}.$$

PROOF. By Lemma 2.1 in [23], for $0 \le f \le 1$ we have if x_h is the Markov process at time h,

$$\left[P_t f(x)\right]^2 \exp\left[-\frac{K\rho(x, x_h)^2}{1 - \exp[-2Kt]}\right] \le P_t f^2(x_h) \le P_t f(x_h).$$

This proves the theorem since it implies that

$$\left[P_t f(x)\right]^2 p_r \exp\left[-\frac{Kr^2}{1 - \exp[-2Kt]}\right] \le P_{t+h} f(x), \quad x \in M. \qquad \square$$

We now go back to the general case and use a sort of Lipschtz continuity of semigroups. Let γ be a nonnegative symmetric function on $E \times E$, define $B_\gamma(x, r) = \{y \in E : \gamma(x, y) \le r\}$. Assume that there exist $c > 0$ and $t > 0$ such that

$$|P_t f(x) - P_t f(y)| \le \gamma(x, y)\|f\|_\infty, \quad x, y \in E. \tag{3.7}$$

THEOREM 8. *If there exist $t, h \in I \setminus \{0\}$ and $r \in (0, 1)$ such that (3.7) holds, and*

$$p_r := \inf_{x \in E} P_h 1_{B_\gamma(x,r)}(x) > 0, \tag{3.8}$$

then (2.1) holds for the above $t, h > 0$ and $\Phi(s) = p_r^{-1}s + r$.

PROOF. By (3.6) we have

$$P_t f(x_h) \ge \left[P_t f(x) - \rho(x, x_h)\right]^+, \quad 0 \le f \le 1,$$

where x_h denotes the corresponding Markov process starting from x. Then, by taking expectation, $P_{t+h} f(x) \ge p_r[P_t f(x) - r]$. The proof is complete. □

To conclude this section, we present the following consequence of Theorem 6 and results mentioned and proved above.

COROLLARY 3. *There exist $t, h \in I \setminus \{0\}$ and increasing $\Phi \in C[0, 1]$ with $\Phi(0) < 1$ such that (2.1) holds. Hence the assertions in Theorem 5 and Corollary 2 are true, provided (3.7) and (3.8) hold. They also hold provided P_t is generated by $L = \Delta + Z$ on M and one of the following holds: for some Z' with $\|Z - Z'\|_\infty < \infty$ in place of Z: (3.1), (3.3)* (for the case that $Z' = \nabla V$) *and (3.6).*

4. Further results for diffusion processes using gradient estimates. We first consider diffusions on $\mathbb{R}^d$. Let

$$L = \frac{1}{2}\sum_{i,j=1}^n a_{ij}(x)\partial_i\partial_j + \sum_{i=1}^n b_i(x)\partial_i$$

where $\partial_i = \partial/\partial x_i$. Assume that there exist $\lambda > 0, C \ge 0$ such that

$$\sum_{i,j} a_{ij}(x)\xi_i\xi_j \ge \lambda|\xi|^2, \qquad \xi, x \in \mathbb{R}^d, \tag{4.1}$$

$$\|\nabla_v a(x)\|^2 + 2\langle\nabla_v b(x), v\rangle \le C, \qquad v, x \in \mathbb{R}^d, |v| \le 1. \tag{4.2}$$

Here $\langle , \rangle$ and $\|.\|$ denote, respectively, the Eucleadien inner product and Hermite-Schmidt norm, and $\nabla_v f(x) := \sum_i v_i\partial_i f(x) = \langle\nabla f, v\rangle$.

Next, consider the SDEs on $\mathbb{R}^d$:

$$dx_t = a(x_t)dB_t + b(x_t)dt,$$

$$dv_t = [\nabla_{v_t} a(x_t)]dB_t + [\nabla_{v_t} b(x_t)]dt,$$

where B_t is the d-dimensional Brownian motion. Then x_t is the L-diffusion process and v_t its derivative flow. By Itô's formula,

$$d|v_t|^2 = 2\langle [\nabla_{v_t} a(x_t)]dB_t, v_t\rangle + \{2\langle \nabla_{v_t} b(x_t), v_t\rangle + \|\nabla_{v_t} a(x_t)\|^2\}dt.$$

Then by (4.2),

$$E|v_t|^2 \le |v_0|^2 \exp[Ct], \quad t \ge 0. \tag{4.3}$$

This then implies

$$\begin{aligned} E\left\{\int_0^t \langle a(x_s)^{-1} v_s, dB_s\rangle\right\}^2 &= \int_0^t E|a(x_s)^{-1} v_s|^2 ds \\ &\le \frac{1}{\lambda^2}\int_0^t \exp[Cs]ds = \frac{\exp[Ct]-1}{c\lambda^2}. \end{aligned}$$

Hence $\int_0^t \langle a(x_s)^{-1} v_s, dB_s\rangle$ is a martingale (square integrable). By Theorem 5 in [7], and furthermore, according to an observation in [18], for any $g \in C^1[0,t]$ with $g_0 = 0$, $g_t = 1$, one has

$$\langle \nabla P_t f(x_0), v_0\rangle = Ef(x_t)\int_0^t g_s' \langle a(x_s)^{-1} v_s, dB_s\rangle.$$

Taking $g_s = 1 - \exp[-Cs]/(1-\exp[-Ct])$, the above formula yields that

$$\begin{aligned} |\nabla P_t f(x)|^2 &\le P_t f^2(x)\int_0^t (g_s')^2 E|a(x_s)^{-1} v_s|^2 ds \\ &\le \frac{C^2 P_t f^2(x)}{\lambda^2}\int_0^t \frac{\exp[-Cs]ds}{(1-\exp[Ct])^2} = \frac{CP_t f^2(x)}{\lambda^2(1-\exp[-Ct])}. \end{aligned} \tag{4.4}$$

Now, taking $\gamma(x,y) = C|x-y|/(\lambda^2(1-\exp[-Ct]))$ in Theorem 8, we obtain the following result.

THEOREM 9. *Assume that* (4.1) *and* (4.2) *hold. If there exist* $r < \lambda/\sqrt{C}$ *and* $h > 0$ *such that*

$$p_r := \inf_x \mathbb{P}_x(|x_s - x| \le r) > 0, \tag{4.5}$$

then (2.1) *holds for some* $t > 0$, *the above* h *and some increasing* $\Phi \in C[0,1]$ *with* $\Phi(0) < 1$. *Consequently, the assertions in Theorem* 5 *and Corollary* 2 *hold.*

The gradient estimates of P_t can also be obtained by using coupling, see for instance [22], in which the mirror coupling is used to estimate the gradient of diffusion semigroups. This coupling has been developed and exploited in [13], [11], [4], [3], [21] and many other places.

THEOREM 10. *Let P_t be generated by $L = \Delta + Z$ on a complete, connected Riemannian manifold M. Suppose that the Ricci curvature is bounded from below and for any $r > 0$, there exists $h > 0$ such that*

$$\sup_{\rho(x,y)=r,\, y\notin \operatorname{cut}(x)} \big[Z\rho(\cdot, y)(x) + Z\rho(x, \cdot)(y)\big] < \infty,$$

$$\inf_{x\in M} \mathbb{P}_x(x_h \in B(x, r)) > 0,$$

where $\rho(x, y)$ denotes the distance between x and y, and $\operatorname{cut}(x)$ *denotes the cut locus of x. Then for any $t > 0$, there exists $h > 0$ and increasing $\Phi \in C[0, 1]$ with $\Phi(0) < 1$ such that* (2.1) *holds, and therefore, the assertions in Theorem* 5 *and Corollary* 2 *are true.*

PROOF. Simply note that by Theorem 4.4 in [22], our assumption implies the gradient estimate $\|\nabla P_t f\|_\infty \le C(t)\|F\|_\infty$ for some $C(t) > 0$. $\square$

5. Supplementary example. We now analyze the following reflecting diffusion process on $[1, \infty)$

$$r_t = r + b_t + \int_0^t r_s^a ds + \ell_t(r), \tag{5.1}$$

where b_t is one-dimensional Brownian motion, $\ell_t(r)$ is local time of r at 1, $r \ge 1$, and $0 \le a \le 1$ is fixed. This provides an example, when $0 \le a \le \frac{1}{3}$, of a diffusion for which the condition $D(R)$ clearly fails yet $\mathscr{I} = \mathscr{J}$. Since $D(R)$ and $P(R)$ together are equivalent to a parabolic Harnack inequality, this example shows a parabolic Harnack inequality is not a necessary condition for the conclusion $\mathscr{I} = \mathscr{J}$. This diffusion is (essentially) the radial part of Brownian motion on a model manifold with $\operatorname{Ric}(\partial r, \partial r) \simeq -r^{2a}$ as $r \to \infty$. Obviously since $r_t \to \infty$ a.s., given two starting points $1 \le r < \rho$ we can always shift couple with $T = \inf\{t > 0 : r_t = \rho\}$, $S = 0$. Thus $\mathscr{J}$ is always trivial for $0 \le a \le 1$.

THEOREM 11. *If r_t satisfies* (5.1) *and $0 \le a \le \frac{1}{3}$, then $\mathscr{J} = \mathscr{I}$ is trivial. If $\frac{1}{3} < a \le 1$, then $\mathscr{I} \subsetneq \mathscr{J}$.*

PROOF. By Itô's formula,

$$r_t^{1-a} = r^{1-a} + (1-a)\int_0^t r_u^{-a} db_u + (1-a)\int_0^t r_u^{-a} d\ell_u(r) + \frac{a(a-1)}{2}\int_0^t r_u^{-1-a} du + (1-a)t\,.$$

Since $r_u \ge 1$ for all $u \ge 0$ and $\varepsilon > 0$, there is a finite $T(\omega, \varepsilon)$ such that for $t \ge T(\omega, \varepsilon)$

$$-t^{\frac{1}{2}+\varepsilon} \le (1-a)\int_0^t r_u^{-a} db_u \le t^{\frac{1}{2}+\varepsilon}\,.$$

Thus, for $t \geq T(\omega, \varepsilon)$ both

$$r_t^{1-a} \geq r^{1-a} - t^{\frac{1}{2}+\varepsilon} + (1-a-\delta)t$$

and

$$r_t^{1-a} \leq r^{1-a} + t^{\frac{1}{2}+\varepsilon} + (1-a)\ell_\infty + (1-a)t$$

for some $\delta \in (0, 1-a)$. Note that $\ell_\infty < \infty$ by the transience of r_t.

Thus, there exists finite $T(\omega, \ell_\infty(r), \eta, \varepsilon, \delta)$ such that for all $t \geq T(\omega, \ell_\infty(r), \eta, \varepsilon, \delta)$,

$$r^{1-a} + (1-a-\eta)t \leq r_t^{1-a} \leq r^{1-a} + (1-a+\eta)t\ . \tag{5.2}$$

This inequality implies that the quadratic variation of the martingale term $(1-a)\int_0^t r_u^{-a} db_u$ diverges for $0 \leq a \leq \frac{1}{3}$ and converges for $\frac{1}{3} < a \leq 1$ since it is $(1-a)^2 \int_0^t r_u^{-2a} du$ and $r_u^{-2a} \simeq t^{-2a/(1-a)}$ for large t, and $-2a/(1-a) \geq -1$ for $0 \leq a \leq \frac{1}{3}$, $-2a/(1-a) < -1$ for $\frac{1}{3} < a \leq 1$. The inequality (5.2) also implies $\int_0^\infty r_u^{-1-a} du$ is convergent a.s. for all $a \in [0,1]$. Consider now the case $0 \leq a \leq \frac{1}{3}$ and take $\rho > r \geq 1$. Put $\rho_t = \rho - b_t + \int_0^t \rho_s^a ds + \ell_t(\rho)$ where b_t is the Brownian motion used to drive r_t and $\ell_t(\rho)$ is local time of ρ at 1 up to time t. Then

$$\begin{aligned}\rho_t^{1-a} - r_t^{1-a} &= \rho^{1-a} - r^{1-a} - (1-a)\int_0^t (\rho_s^{-a} + r_s^{-a}) db_s \\ &\quad + (1-a)\int_0^t (\rho_u^{-a} d\ell_u(\rho) - r_u^{-a} d\ell_u(r)) \\ &\quad + \frac{a(a-1)}{2}\int_0^t (\rho_s^{-1-a} - r_s^{-1-a}) ds\ .\end{aligned}$$

But

$$\int_0^\infty [\rho_t^{-a} d\ell_t(\rho) + r_t^{-a} d\ell_t(r)] \leq \ell_\infty(\rho) + \ell_\infty(r) \in \mathbb{R}$$

by transience of r_t and ρ_t, and $\int_0^\infty (\rho_s^{-1-a} - r_s^{-1-a}) ds$ is also a.s. finite. Since $E(\int_0^t (\rho_s^{-a} + r_s^{-a}) db_s)^2 \geq E(\int_0^t (\rho_s^{-2a} + r_s^{-2a}) ds)$ tends to infinity a.s., as the martingale $\int_0^t (\rho_s^{-a} + r_s^{-a}) db_s$ does not converge a.s. and so must be unbounded. Thus, $T = \inf\{t > 0 : \rho_t = r_t\} < \infty$ a.s. and $\mathscr{J}$ is trivial so $\mathscr{I} = \mathscr{J}$ a.s.

For $\frac{1}{3} < a < 1$ we use Derrienic's 0-2 law (Theorem 4). For $a > \frac{1}{3}$, recall that the martingale $\int_0^t r_s^{-a} db_s$ converges a.s. since $\mathbb{P}_r\big(\int_0^\infty r_s^{-2a} ds < \infty\big) = 1$. Put

$$g(t,s) = r^{1-a} + (1-a)t + \tfrac{1}{2}(1-a)s\ .$$

Then for any $t > 0$,

$$\begin{aligned}&\mathbb{P}_r\big(r_t \leq g(t,s)^{\frac{1}{1-a}}\big) \\ &\quad = \mathbb{P}_r\big(r_t^{1-a} \leq g(t,s)\big)\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{P}_r \left\{ r_t^{1-a} - r^{1-a} - (1-a)t \le \frac{(1-a)s}{2} \right\} \\
&\ge \mathbb{P}_r \left\{ (1-a)\int_0^t r_u^{-a}\, db_u + \frac{(a-1)a}{2}\int_0^t r_u^{-1-a}\, du \right. \\
&\qquad\qquad\qquad \left. + (1-a)\ell_t(r) \le \frac{(1-a)s}{2} \right\} \\
&\ge \frac{3}{4},
\end{aligned}
$$

provided s is large enough since the integrals on the $\ell.h.s.$ inside the probability all converge by (5.2).

Similarly,

$$
\begin{aligned}
&\mathbb{P}_r \left(r_{t+s} \ge g(t,s)^{\frac{1}{1-a}} \right) \\
&\quad = \mathbb{P}_r \big(r_{t+s}^{1-a} \ge g(t,s) \big) \\
&\quad \ge \mathbb{P}_r \left\{ (1-a)\int_0^{t+s} r_u^{-a}\, db_u + (1-a)s + \frac{a(a-1)}{2}\int_0^{t+s} r_u^{-1-a}\, du \ge \frac{(1-a)s}{2} \right\} \\
&\quad = \mathbb{P}_r \left\{ (1-a)\int_0^{t+s} r_u^{-a}\, db_u + \frac{(a-1)a}{2}\int_0^{t+s} r_u^{-1-a}\, du \ge -\frac{(1-a)s}{2} \right\} \\
&\quad \ge \frac{3}{4},
\end{aligned}
$$

for s sufficiently large since all integrals converge as $t+s \to \infty$. Thus, since for any A and probability measures ν and μ, $\|\nu - \mu\| \ge 2[\nu(A) - \mu(A)]$, taking $A = (g(t,s)^{\frac{1}{1-a}}, \infty)$ we have, for large enough s,

$$
\begin{aligned}
\|P_{t+s}\delta_r - P_t\delta_r\| &\ge 2[P_{t+s}\delta_r(A) - P_t\delta_r(A)] \\
&= 2\big[\mathbb{P}_r\big(r_{t+s} \ge g(t,s)^{\frac{1}{1-a}}\big) - \mathbb{P}_r\big(r_t \ge g(t,s)^{\frac{1}{1-a}}\big)\big] \\
&\ge \tfrac{3}{2} - \tfrac{1}{2} = 1.
\end{aligned}
$$

Thus, by Theorem 4, $\sup_x \alpha(x,s) = 2$ and $\mathscr{I} \subsetneq \mathscr{J}$. $\square$

REFERENCES

[1] ALDOUS, D. and THORISSON, H. (1993). Shift-coupling. *Stochastic Process. Appl.* **44** 1–14.

[2] BAKRY, D. and QIAN, Z. (1997). On Harnack estimates for positive solutions of the heat equation on a complete manifold. *C. R. Acad. Sci. Paris Sér. I* **324** 1037–1042.

[3] CHEN, M. F. and WANG, F.Y. (1997). General formula for lower bound of the first eigenvalue on Riemannian manifolds. *Sci. Sin. U.* **40** 384–394.

[4] CRANSTON, M. (1991). Gradient estimates on manifolds using coupling. *J. Funct. Anal.* **99** 110–124.

[5] CRANSTON, M. and GREVEN, A. (1995). Coupling and harmonic functions in the case of continuous time Markov processes. *Stochastic Process. Appl.* **60** 261–286.

[6] COULHON, T. and SALOFF-COSTE, L. (1995). Variétés riemanniennes isométrique à l'infini. *Rev. Mat. Iberoamericana* **2** 687–726.

[7] ELWORTHY, D. and LI, X.M. (1994). Formulae for the derivatives of heat semigroups. *J. Funct. Anal.* **125** 252–286.

[8] DERRIENNIC, Y. (1976). Lois "zéro ou deux" pour les processus de Markov. *Ann. Inst. Henri Poincaré Sec. B* **12** 111–129.

[9] GREVEN, A. (1987). Couplings of Markov chains by randomized stopping times, Part I. *Probab. Theory Related Fields* **75** 195–212.

[10] GUIVARC'H, Y. (1973). Croissance polynômiale et période des fonctions harmoniques. *Bull. Soc. Math. France* **101** 149–152.

[11] KENDALL. W. (1986). Nonnegative Ricci curvature and the Brownian coupling property. *Stochastics* **19** 111–129.

[12] LINDVALL, T. (1992). *Lectures On The Coupling Method.* Wiley, New York.

[13] LINDVALL, T. and ROGERS, C. (1986). Coupling of multidimensional diffusions by reflection. *Ann. Probab.* **14** 860–872.

[14] LI, P. and YAU, S.T. (1986). On the parabolic kernel of the Schrödinger operator. *Acta. Math.* **156** 153–201.

[15] QIAN, Z. M. (1995). Gradient estimates and heat kernel estimates. *Proc. Roy. Soc. Edinburgh A* **125** 975–990.

[16] SALOFF-COSTE, L. (1995). Parabolic Harnack inequality for divergence form second order differential operators. *Potential Anal.* **4** 429–467.

[17] SETTI, A. (1992). Gaussian estimates for the heat kernel of the weighted Laplacian and fractal measures. *Canad. J. Math.* **44** 1061–1078.

[18] THALMAIER, A. (1997). On the differentiation of heat semigroups and Poisson integrals. *Stochastics Stochastics Rep.* **61** 297–321.

[19] THORISSON, H. (1994). Shift-coupling in continuous time. *Probab. Theory Related Fields* **99** 477–483.

[20] VAROPOULOS, N. (1987). Fonctions harmoniques sur les groupes de Lie. *C. R. Acad. Sci. Paris Ser. I Math.* **309** 519–521.

[21] WANG, F.Y. (1994). Application of coupling method to the Neumann eigenvalue problem. *Probab. Theory Related Fields* **98** 299–306.

[22] WANG, F. Y. (1997). On estimation of the logarithmic Sobolev constant and gradient estimates of heat semigroups. *Probab. Theory Related Fields* **108** 87–101.

[23] WANG, F. Y. (1997). Logarithmic Sobolev inequalities on noncompact Riemannian manifolds. *Probab. Theory Related Fields* **109** 417–424.

[24] YAU, S. T. (1995). Harnack inequality for non-self-adjoint evolution equations. *Math. Res. Lett.* **2** 387–399.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ROCHESTER
ROCHESTER, NEW YORK 14627
E-MAIL: cran@math.rochester.edu

DEPARTMENT OF MATHEMATICS
BEIJING NORMAL UNIVERSITY
BEIJING 100875
P.R. CHINA
E-MAIL: wangfy@bnu.edu.cn
fwang@mathematik.uni-bielefeld.de