Bernoulli **17**(4), 2011, 1136–1158
DOI: 10.3150/10-BEJ308

Coupling for Ornstein–Uhlenbeck processes with jumps

FENG-YU WANG

School of Mathematical Sciences, Beijing Normal University, Beijing 100875, China and Department of Mathematics, Swansea University, Singleton Park, SA2 8PP, UK
E-mail: wangfy@bnu.edu.cn; F.Y.Wang@swansea.ac.uk

Consider the linear stochastic differential equation (SDE) on $\mathbb{R}^n$:

$$\mathrm{d}X_t = AX_t\,\mathrm{d}t + B\,\mathrm{d}L_t,$$

where A is a real $n \times n$ matrix, B is a real $n \times d$ real matrix and L_t is a Lévy process with Lévy measure ν on $\mathbb{R}^d$. Assume that $\nu(\mathrm{d}z) \ge \rho_0(z)\,\mathrm{d}z$ for some $\rho_0 \ge 0$. If $A \le 0$, $\mathrm{Rank}(B) = n$ and $\int_{\{|z-z_0|\le\varepsilon\}} \rho_0(z)^{-1}\,\mathrm{d}z < \infty$ holds for some $z_0 \in \mathbb{R}^d$ and some $\varepsilon > 0$, then the associated Markov transition probability $P_t(x, \mathrm{d}y)$ satisfies

$$\|P_t(x,\cdot) - P_t(y,\cdot)\|_{\mathrm{var}} \le \frac{C(1+|x-y|)}{\sqrt{t}}, \qquad x, y \in \mathbb{R}^d, t > 0,$$

for some constant $C > 0$, which is sharp for large t and implies that the process has successful couplings. The Harnack inequality, ultracontractivity and the strong Feller property are also investigated for the (conditional) transition semigroup.

Keywords: coupling; Harnack inequality; Lévy process; quasi-invariance; strong Feller

1. Introduction

Lévy processes are fundamental models of Markov processes, from which more general diffusion-jump-type Markov processes can be constructed by solving stochastic differential equations or martingale problems. It is well known that a Lévy process can be decomposed into two independent parts, that is, the Brownian (or Gaussian) part and the jump part. Comparing with the analysis on the Brownian motion, that on the pure jump part is far from complete. For instance, except for stable-like processes that can be treated as subordinations of diffusion processes [24] (see also [5,13] for heat kernel upper bounds for α-stable processes with drifts), little is known concerning regularities of the transition probabilities of O–U-type jump processes. Most existing regularity results for O–U (or generalized Mehler) semigroups were derived by using the Gaussian part as the leading term (cf. [10,21,22] and references within). In contrast, besides known results on the transition density for Lévy processes (see [11,14,26] and references therein), the strong Feller property was recently proved by Priola and Zabczyk [19] for O–U jump processes by considering a Hörmander condition and Lévy measures. The main purpose of this paper is to investigate more regular properties on O–U semigroups in the same spirit, so that our results work well for the pure jump case as emphasized in the Abstract.

1350-7265

Recall that a Lévy measure ν on $\mathbb{R}^d$ is such that $\nu(\{0\}) = 0$ and (see [2])

$$\int_{\mathbb{R}^d} (|z|^2 \wedge 1)\nu(\mathrm{d}z) < \infty.$$

Let $b \in \mathbb{R}^d$ and Q be a non-negatively definite $d \times d$ matrix. The underlying Lévy process L_t is the Markov process on $\mathbb{R}^d$ generated by

$$\mathcal{L} f := \langle b, \nabla f\rangle + \mathrm{Tr}(Q\nabla^2 f) + \int_{\mathbb{R}^d} \big\{ f(z + \cdot) - f - \langle \nabla f, z\rangle 1_{\{|z|\le 1\}} \big\} \nu(\mathrm{d}z),$$

which is well defined for $f \in C_b^2(\mathbb{R}^d)$.

Now, let A be a real $n \times n$ matrix and B be a real $n \times d$ matrix. We shall investigate the solution to the following linear stochastic differential equation

$$\mathrm{d}X_t^x = (AX_t^x)\,\mathrm{d}t + B\,\mathrm{d}L_t, \qquad X_0^x = x \in \mathbb{R}^n. \tag{1.1}$$

We shall investigate the following properties of the solution:

(A) The coupling property.
(B) The Harnack inequality and ultracontractivity.
(C) The strong Feller property.

The coupling method is a powerful tool in the study of Markov processes, and the coupling property that we are going to study is closely related to long-time behaviors, Liouville-type properties and the 0–1 law of tail-/shift-invariant events. Recall that a couple (X_t, Y_t) is called a coupling of the Markov process associated with a given transition probability if both X_t and Y_t are Markov processes associated with the same transition probability (possibly with different initial distributions). In this case, X_t and Y_t are called the marginal processes of the coupling. A coupling (X_t, Y_t) is called successful if the coupling time

$$\mathbf{T} := \inf\{t \ge 0 : X_t = Y_t\} < \infty, \qquad \text{a.s.}$$

A Markov process is said to have a coupling property (or to have successful couplings) if, for any initial distributions μ_1 and μ_2, there exists a successful coupling with marginal processes starting from μ_1 and μ_2, respectively. A slightly weaker notion is the shift-coupling property: for any two initial distributions there exists a coupling (X_t, Y_t) with marginal processes starting from them respectively, such that "$X_{\mathbf{T}_1} = Y_{\mathbf{T}_2}$" holds for some finite stopping times $\mathbf{T}_1, \mathbf{T}_2$ (see [1]). In general, the coupling property is stronger than the shift-coupling property, but they are equivalent if the Markov semigroup satisfies a weak parabolic Harnack inequality (see [9]).

Consider a strong Markov process with transition semigroup P_t. For any (not necessarily successful) coupling with initial distributions μ_1 and μ_2, one has (see [6,15])

$$\|\mu_1 P_t - \mu_2 P_t\|_{\mathrm{var}} \le 2\mathbb{P}(\mathbf{T} > t), \qquad t \ge 0, \tag{1.2}$$

where $\|\cdot\|_{\mathrm{var}}$ is the total variational norm. This follows by setting $X_t = Y_t$ for $t \ge \mathbf{T}$ due to the strong Markov property. Moreover, for any coupling (X_t, Y_t) with initial distributions δ_x and δ_y, a bounded harmonic function f (i.e. $P_t f = f$ for $t \geq 0$) satisfies

$$|f(x) - f(y)| \leq \inf_{t>0} \mathbb{E}|f(X_t) - f(Y_t)| \leq 2\|f\|_\infty \mathbb{P}(\mathbf{T} = \infty).$$

Consequently, if a strong Markov process has a coupling property, then its bounded harmonic functions have to be constant, that is, the Liouville property holds for bounded harmonic functions. In general, the coupling property of a strong Markov process on $\mathbb{R}^n$ with semigroup P_t is equivalent to each of the following statements (see [8], Section 4, and [15], Chapters 3 and 5):

(i) For any $\mu_1, \mu_2 \in \mathscr{P}(\mathbb{R}^n)$, $\lim_{t\to\infty} \|\mu_1 P_t - \mu_2 P_t\|_{\mathrm{var}} = 0$.
(ii) All bounded time–space harmonic functions are constant, that is, a bounded measurable function u on $[0, \infty) \times \mathbb{R}^n$ has to be constant if $u(t, \cdot) = P_s u(t+s, \cdot)$ holds for all $s, t \geq 0$.
(iii) The tail σ-algebra of the process is trivial, that is, $\mathbb{P}(X \in A) = 0$ or 1 holds for any initial distribution and any $A \in \bigcap_{t>0} \sigma((\mathbb{R}^n)^{[0,\infty)} \ni w \mapsto w_s : s \geq t)$.

Correspondingly, each of the following statements are equivalent to the shift-coupling property (see [25], Section 4, or [1]):

(iv) For any $\mu_1, \mu_2 \in \mathscr{P}(\mathbb{R}^n)$, $\lim_{t\to\infty} \|\frac{1}{t}\int_0^t (\mu_1 - \mu_2) P_s \,\mathrm{d}s\|_{\mathrm{var}} = 0$.
(v) All bounded harmonic functions are constant.
(vi) The shift-invariant σ-algebra of the process is trivial, that is, $\mathbb{P}(X \in B) = 0$ or 1 for any initial distribution and any shift-invariant measurable set $B \subset (\mathbb{R}^n)^{[0,\infty)}$.

In Section 3, we shall present explicit conditions on A, B and the Lévy measure ν such that the coupling property holds (see Theorem 3.1).

Next, we aim to establish the following Harnack inequality for P_t initiated in [27] for diffusion semigroups:

$$(P_t f(x))^\alpha \leq (P_t f^\alpha(y)) H_\alpha(t, x, y), \qquad t > 0, x, y \in \mathbb{R}^n, \alpha > 1,$$

for positive measurable functions f, where H_α is a positive function on $(0, \infty) \times (\mathbb{R}^n)^2$.

When $\nu(\mathbb{R}^d) < \infty$, with a positive probability the process does not jump before a fixed time $t > 0$, so that this inequality could not hold for the pure jump case. This is the main reason why all existing results in this direction only work for the case with a non-degenerate Gaussian part (cf. [21,22]). To work out the Harnack inequality also for the pure jump case, we shall be restricted on the event that the process jumps before time t. More precisely, let τ_1 be the first jump time of the Lévy process induced by an absolutely continuous part of ν. If $\mathrm{Rank}(B) = n$, then the Harnack inequality and ultracontractivity are investigated in Section 4 for the following modified sub-Markov operator P_t^1 (see Theorem 4.1):

$$P_t^1 f(x) := \mathbb{E}\{f(X_t^x) 1_{\{t \geq \tau_1\}}\}.$$

Finally, we look at the strong Feller property of P_t. By the same reasoning that leads to the invalidity of the Harnack inequality, when ν is finite the pure jump semigroup cannot be strong Feller. Therefore, in [19] the authors only considered the case that ν is infinite. More precisely, if ν has an infinite absolutely continuous part and if there exists $m \geq 1$ such that the rank condition

$$\mathrm{Rank}(B, AB, \ldots, A^{m-1}B) = n$$

holds, then [19], Theorem 1.1 and Proposition 2.1, imply the strong Feller property of P_t. We shall extend this result by allowing the absolutely continuous part of ν to be finite. In this case the number m in the rank condition will refer to the strong Feller property of the semigroup conditioned by the event that the mth jump happens before time $t > 0$ (see Theorem 5.2).

It might be interesting to indicate that for jump processes the strong Feller property is incomparable with the coupling property. Indeed, the latter is a long-time property but the former is somehow a short-time property. For the strong Feller property, we need the process to be able to visit any area before any fixed time, for which the jump measure has to be infinite as mentioned above. However, the situation in the diffusion case is very different: Whenever the diffusion process is able to visit any area for a long time, it will be able to do so before any fixed time.

The remainder of the paper is organized as follows. To study the coupling property and the Harnack inequality, we shall first investigate in Section 2 the quasi-invariance of random shifts for compound Poisson processes, which in particular leads to a conditional Girsanov theorem. Then we will study the properties included in (A), (B) and (C) in Sections 3, 4 and 5, respectively.

2. Quasi-invariance and the Girsanov theorem

Throughout of this section, we assume $\lambda := \nu(\mathbb{R}^d) \in (0, \infty)$ and let $L := \{L_t\}_{t\ge 0}$ be the compound Poisson process with Lévy measure ν and $L_0 = 0$. Let Λ be the distribution of L, which is a probability measure on the path space

$$W = \left\{ \sum_{i=1}^{\infty} x_i 1_{[t_i,\infty)} : i \in \mathbb{N}, x_i \in \mathbb{R}^d \setminus \{0\}, 0 \le t_i \uparrow \infty \text{ as } i \uparrow \infty \right\}$$

equipped with the σ-algebra induced by $\{w \mapsto w_t : t \ge 0\}$. Let $\Delta w_t = w_t - w_{t-}$ for $t > 0$.

For any $T > 0$, let Λ_T be the distribution of $L_{[0,T]} := \{L_t\}_{t\in[0,T]}$, and let τ and ξ be random variables with distributions $\frac{1}{T} 1_{[0,T]}(t)\,\mathrm{d}t$ on $[0, T]$ and $\frac{1}{\lambda}\nu$ on $\mathbb{R}^d$, respectively, such that ξ, τ and L are independent. It is shown in [28] that the distribution of $L_{[0,T]} + \xi 1_{[\tau,T]}$ is $\frac{1}{\lambda T} n_T(w)\Lambda_T(\mathrm{d}w)$, where

$$n_T(w) := \#\{t \in [0, T] : w_t \ne w_{t-}\}, \qquad w \in W. \tag{2.1}$$

We shall extend this result to more general random variables ξ and τ. To this end, write

$$L_t = \sum_{i=1}^{N_t} \xi_i, \qquad t \ge 0, \tag{2.2}$$

where N_t is the Poisson process on $\mathbb{Z}_+$ with rate λ and $\{\xi_i\}_{i\ge 1}$ are i.i.d. random variables on $\mathbb{R}^d$, which are independent of $N := \{N_t\}_{t\ge 0}$ and have common distribution $\frac{1}{\lambda}\nu$.

Theorem 2.1. *Let (ξ, τ) be a random variable on $\mathbb{R}^d \times [0, \infty)$. Then the distribution of $L + \xi 1_{[\tau,\infty)}$ is absolutely continuous with respect to Λ if and only if the joint distribution of (L, ξ, τ)*

has the form

$$\varepsilon\Lambda(\mathrm{d}w)\delta_0(\mathrm{d}z)\Theta(w,\mathrm{d}t)+g(w,z,t)\Lambda(\mathrm{d}w)\nu(\mathrm{d}z)\,\mathrm{d}t,$$

where $\varepsilon\in[0,1]$ *is a constant*, g *is a non-negative measurable function on* $W\times\mathbb{R}^d\times[0,\infty)$ *and* $\Theta(w,\mathrm{d}t)$ *is a transition probability from* W *to* $[0,\infty)$. *In this case*, *the distribution of* $L+\xi 1_{[\tau,\infty)}$ *is formulated as*

$$\bigg\{\varepsilon+\sum_{\Delta w_t\neq 0} g\big(w-\Delta w_t 1_{[t,\infty)},\Delta w_t,t\big)\bigg\}\Lambda(\mathrm{d}w).$$

According to Theorem 2.1, the random shift $L\mapsto L+\xi 1_{[\tau,\infty)}$ is quasi-invariant if and only if the conditional distribution of (L,ξ,τ) given $\{\xi\neq 0\}$ is absolutely continuous w.r.t. the product measure $\Lambda\times\nu\times\mathrm{d}t$. Since, when $\xi=0$, the random shift does not help the coupling, below we will only choose non-zero ξ.

To prove this result, we shall make use of the Mecke formula for the Poisson measure. Let E be a Polish space with Borel σ-field $\mathscr{F}$, and let σ be a locally finite measure on E. Then π_σ, the Poisson measure with intensity σ, is a probability measure on the configuration space

$$\Gamma:=\left\{\sum_{i=1}^n \delta_{x_i}: n\in\mathbb{Z}_+\cup\{\infty\}, x_i\in E\right\}$$

fixed by the Laplace transform

$$\int_\Gamma \mathrm{e}^{\gamma(f)}\pi_\sigma(\mathrm{d}\gamma)=\mathrm{e}^{\sigma(\mathrm{e}^f-1)},\qquad f\in C_0(E).$$

Note that the corresponding σ-field on Γ is induced by $\{\gamma\mapsto\gamma(f): f\in C_0(E)\}$. The Mecke formula [17] (see also [20]) says that for any non-negative measurable function F on $\Gamma\times E$,

$$\int_{\Gamma\times E} F(\gamma+\delta_z,z)\pi_\sigma(\mathrm{d}\gamma)\sigma(\mathrm{d}z)=\int_\Gamma \pi_\sigma(\mathrm{d}\gamma)\int_E F(\gamma,z)\gamma(\mathrm{d}z). \tag{2.3}$$

By considering the Poisson measure with intensity $\nu\times\mathrm{d}t$ on $\mathbb{R}^d\times[0,\infty)$, we will be able to prove Theorem 2.1 by using the following result.

Theorem 2.2. *Let* $A\subset E$ *be measurable*, X *be a random variable on* Γ *with distribution* π_σ *and* η *be a random variable on* E. *Then the measure* $\mathbb{P}(X+\delta_\eta\in\cdot,\eta\in A)$ *is absolutely continuous with respect to* π_σ *if and only if the measure* $\mathbb{P}((X,\eta)\in\cdot,\eta\in A)$ *is absolutely continuous with respect to* $\pi_\sigma\times\sigma$.

Proof. Let $D_{X,\eta}$ be the distribution of (X,η).

(a) The sufficiency. Assume that $\mathbb{P}((X,\eta)\in\cdot,\eta\in A)=g(\gamma,z)\pi_\sigma(\mathrm{d}\gamma)\sigma(\mathrm{d}z)$ for some non-negative measurable function g on $\Gamma\times E$. For any bounded measurable function f on Γ, by the Mecke formula (2.3) for

$$F(\gamma,z):=f(\gamma)g(\gamma-\delta_z,z)1_{\{\gamma\geq\delta_z\}}1_A(z),$$

we have

$$\begin{aligned}\mathbb{E}\{1_A(\eta)f(X+\delta_\eta)\} &= \int_{\Gamma\times A} f(\gamma+\delta_z)g(\gamma,z)\pi_\sigma(\mathrm{d}\gamma)\sigma(\mathrm{d}z)\\ &= \int_\Gamma f(\gamma)\left\{\int_A g(\gamma-\delta_z,z)\gamma(\mathrm{d}z)\right\}\pi_\sigma(\mathrm{d}\gamma).\end{aligned}$$

So, $\mathbb{P}(X+\delta_\eta\in\cdot,\eta\in A)$ is absolutely continuous with respect to π_σ with density function $\gamma\mapsto\int_A g(\gamma-\delta_z,z)\gamma(\mathrm{d}z)$.

(b) The necessity. Assume that $\mathbb{P}(X+\delta_\eta\in\cdot,\eta\in A)$ is absolutely continuous with respect to π_σ. For any measurable set $N\subset\Gamma\times E$ with $(\pi_\sigma\times\sigma)(N)=0$, we intend to prove

$$\mathbb{P}\big((X,\eta)\in N,\eta\in A\big)=D_{X,\eta}(N)=0. \tag{2.4}$$

Let

$$A_N=\{\gamma+\delta_z:(\gamma,z)\in N,z\in A\}\subset\Gamma,\qquad F(\gamma,z)=1_{N\cap(\Gamma\times A)}(\gamma-\delta_z,z)$$
$$\text{for } (\gamma,z)\in\Gamma\times E.$$

If $\gamma\in A_N$, then there exists $z_0\in A$ such that $(\gamma-\delta_{z_0},z_0)\in N$. This means that $\gamma\ge\delta_{z_0}$ and

$$\int_E F(\gamma,z)\gamma(\mathrm{d}z)\ge h(\gamma,z_0)=1.$$

Therefore,

$$\int_E F(\gamma,z)\gamma(\mathrm{d}z)\ge 1_{A_N}(\gamma),\qquad \gamma\in\Gamma.$$

Combining this with (2.3) and noting that $(\pi_\sigma\times\sigma)(N)=0$, we obtain

$$\begin{aligned}\pi_\sigma(A_N) &\le \int_\Gamma\pi_\sigma(\mathrm{d}\gamma)\int_E F(\gamma,z)\gamma(\mathrm{d}z)=\int_{\Gamma\times E}F(\gamma+\delta_z,z)\pi_\sigma(\mathrm{d}\gamma)\sigma(\mathrm{d}z)\\ &\le\int_{\Gamma\times E}1_N(\gamma,z)\pi_\sigma(\mathrm{d}\gamma)\sigma(\mathrm{d}z)=(\pi_\sigma\times\sigma)(N)=0.\end{aligned}$$

Since $\mathbb{P}(X+\delta_\eta\in\cdot,\eta\in A)$ is absolutely continuous with respect to π_σ, this implies that

$$\mathbb{P}\big((X,\eta)\in N,\eta\in A\big)\le\mathbb{P}(X+\delta_\eta\in A_N,\eta\in A)=0.$$

Thus, (2.4) holds. □

Proof of Theorem 2.1. (1) The sufficiency. Let π_σ be the Poisson measure with intensity $\sigma:=\nu(\mathrm{d}z)\times\mathrm{d}t$. Since $\sigma(\{0\})=0$ and the Lebesgue measure $\mathrm{d}t$ is infinite on $[0,\infty)$ without atom, π_σ is supported on

$$\Gamma_0:=\left\{\sum_{i=1}^\infty\delta_{(x_i,t_i)}:i\in\mathbb{N},x_i\in\mathbb{R}^d\setminus\{0\},0\le t_i\uparrow\infty\text{ as }i\uparrow\infty\right\}.$$

Let

$$\psi : W \to \Gamma_0; \qquad \sum_{i=1}^{\infty} x_i 1_{[t_i,\infty)} \mapsto \sum_{i=1}^{\infty} \delta_{(x_i,t_i)}.$$

We have (see [4], page 12)

$$\pi_\sigma = \Lambda \circ \psi^{-1}, \qquad \Lambda = \pi_\sigma \circ \psi. \tag{2.5}$$

By (2.3), for any non-negative measurable function h on $\Gamma_0 \times \mathbb{R}^d \times [0,\infty)$, we have

$$\int_{\Gamma_0} \pi_\sigma(\mathrm{d}\gamma) \int_{\mathbb{R}^d \times [0,\infty)} h(\gamma, x, t)\gamma(\mathrm{d}x, \mathrm{d}t) = \int_{\Gamma_0 \times \mathbb{R}^d \times [0,\infty)} h\big(\gamma + \delta_{(x,t)}, x, t\big)\pi_\sigma(\mathrm{d}\gamma)\nu(\mathrm{d}x)\,\mathrm{d}t.$$

Combining this with (2.5) we conclude that

$$\int_W \sum_{\Delta w_t \neq 0} H(w, \Delta w_t, t)\Lambda(\mathrm{d}w) = \int_{W \times \mathbb{R}^d \times [0,\infty)} H\big(w + x 1_{[t,\infty)}, x, t\big)\Lambda(\mathrm{d}w)\nu(\mathrm{d}x)\,\mathrm{d}t$$

holds for any non-negative measurable function H on $W \times \mathbb{R}^d \times [0,\infty)$. Therefore, for any non-negative measurable function F on W, we have

$$\begin{aligned}
&\mathbb{E}F\big(L + \xi 1_{[\tau,\infty)}\big) \\
&\quad = \mathbb{E}\big\{F(L)1_{\{\xi=0\}}\big\} + \int_{W \times \mathbb{R}^d \times [0,\infty)} F\big(w + x 1_{[t,\infty)}\big) g(w, x, t)\Lambda(\mathrm{d}w)\nu(\mathrm{d}x)\,\mathrm{d}t \\
&\quad = \int_W F(w)\Big\{\varepsilon + \sum_{\Delta w_t \neq 0} g\big(w - \Delta w_t 1_{[t,\infty)}, \Delta w_t, t\big)\Big\}\Lambda(\mathrm{d}w).
\end{aligned}$$

This completes the proof of the sufficiency.

(2) The necessity. Let the distribution of $L + \xi 1_{[\tau,\infty)}$ be absolutely continuous with respect to Λ. Let $\varepsilon = \mathbb{P}(\xi = 0)$ and let $\Theta(w, \mathrm{d}t)$ be the regular conditional distribution of τ given L and $\xi = 0$. Then for any non-negative measurable function f on $W \times \mathbb{R}^d \times [0,\infty)$,

$$\mathbb{E}f(L, \xi, \tau) = \varepsilon \int_{\Gamma \times [0,\infty)} f(w, 0, t)\Lambda(\mathrm{d}w)\Theta(w, \mathrm{d}t) + \mathbb{E}\big\{f(L, \xi, \tau)1_{\{\xi \neq 0\}}\big\}.$$

So, to prove that the distribution of (L, ξ, τ) has the required form, it suffices to show that for any $\Lambda \times \nu \times \mathrm{d}t$-null set N, we have

$$\mathbb{P}\big((L, \xi, \tau) \in N, \xi \neq 0\big) = 0. \tag{2.6}$$

To this end, we shall make use of Theorem 2.2. Let $E = \mathbb{R}^d \times [0,\infty)$ and $X = \psi(L)$. We have

$$\psi\big(L + \xi 1_{[\tau,\infty)}\big) = X + \delta_{(\xi,\tau)} \qquad \text{for } \xi \neq 0.$$

Let

$$\tilde{N} = \{(\psi(w), z, t) : (w, z, t) \in N, z \neq 0\}.$$

By (2.5) we have

$$(\pi_\sigma \times \nu \times \mathrm{d}t)(\tilde{N}) \leq (\Lambda \times \nu \times \mathrm{d}t)(N) = 0. \tag{2.7}$$

Now, since the distribution of $L + \xi 1_{[\tau,\infty)}$ is absolutely continuous with respect to Λ, due to (2.5) so is $\mathbb{P}(X + \delta_{(\xi,\tau)} \in \cdot,\ \xi \neq 0)$ with respect to π_σ. Hence, according to Theorem 2.2, $\mathbb{P}((X, \xi, \tau) \in \cdot,\ \xi \neq 0)$ is absolutely continuous with respect to $\pi_\sigma \times \nu \times \mathrm{d}t$. Combining this with (2.5) and (2.7), we arrive at

$$\mathbb{P}\big((L, \xi, \tau) \in N, \xi \neq 0\big) = \mathbb{P}\big((X, \xi, \tau) \in \tilde{N}, \xi \neq 0\big) = 0.$$

Therefore, (2.6) holds. □

In the situation of Theorem 2.1, let

$$U(w) = \varepsilon + \sum_{\Delta w_t \neq 0} g\big(w - \Delta w_t 1_{[t,\infty)}, \Delta w_t, t\big), \qquad w \in W. \tag{2.8}$$

As a direct consequence of Theorem 2.1, the following result says that the distribution of $L + \xi 1_{[\tau,\infty)}$ under probability

$$\frac{1_{\{U>0\}}}{\Lambda(U>0)U}\big(L + \xi 1_{[\tau,\infty)}\big)\mathbb{P}$$

coincides with that of L under probability $\frac{1_{\{U(L)>0\}}}{\Lambda(U>0)}\mathbb{P}$. This can be regarded as a conditional Girsanov theorem.

Corollary 2.3. *In the situation of Theorem 2.1 let U be in (2.8). Then for any non-negative measurable function F on W,*

$$\mathbb{E}\big\{\big(F 1_{\{U>0\}}\big)(L)\big\} = \mathbb{E}\bigg\{\frac{F 1_{\{U>0\}}}{U}\big(L + \xi 1_{[\tau,\infty)}\big)\bigg\}.$$

3. The coupling property

Recall that for the Brownian motion the equality in (1.2) is reached by the coupling by reflection covered by Lindvall and Rogers in [16]. More precisely, let $P_t^B(x, \mathrm{d}y)$ be the transition probability of the Brownian motion on $\mathbb{R}^d$ and let $\mathbf{T}_{x,y}$ be the coupling time of the coupling by reflection for initial distributions δ_x and δ_y. One has (see [7], Section 5)

$$\begin{aligned}\frac{1}{2}\|P_t^B(x,\cdot) - P_t^B(y,\cdot)\|_{\mathrm{var}} &= \mathbb{P}(\mathbf{T}_{x,y} > t) \\ &= \frac{\sqrt{2}}{\sqrt{\pi}}\int_0^{|x-y|/(2\sqrt{t})} \mathrm{e}^{-u^2/2}\,\mathrm{d}u \leq \frac{\sqrt{2}|x-y|}{\sqrt{t}}, \qquad t > 0. \end{aligned} \tag{3.1}$$

Our first result aims to provide an analogous estimate for Lévy jump processes, which in particular implies the coupling property of the process according to the equivalent statement (i). Intuitively, to ensure the coupling property for a Lévy jump process, the Lévy measure should have a non-discrete support to make the process active enough. In this paper, we shall assume that $\mathrm{Rank}(B) = n$ and ν has a non-trivial absolutely continuous part.

Theorem 3.1. *Let $P_t(x, \mathrm{d}y)$ be the transition probability for the solution to (1.1). Let $\mathrm{Rank}(B) = n$ and $\langle Ax, x\rangle \le 0$ hold for $x \in \mathbb{R}^n$. If $\nu \ge \rho_0(z)\,\mathrm{d}z$ such that*

$$\int_{\{|z-z_0|\le\varepsilon\}} \rho_0(z)^{-1}\,\mathrm{d}z < \infty$$

holds for some $z_0 \in \mathbb{R}^d$ and some $\varepsilon > 0$, then

$$\|P_t(x,\cdot) - P_t(y,\cdot)\|_{\mathrm{var}} \le \frac{C(1+|x-y|)}{\sqrt{t}}, \qquad x, y \in \mathbb{R}^n, t > 0, \tag{3.2}$$

holds for some constant $C > 0$, and hence, the coupling property and assertions (i)–(vi) *hold.*

Remark 3.1. (1) According to [18], Theorem 3.5(ii), if A has an eigenvalue with a positive real part, then, under an assumption on large jumps, the coupling property fails. In this sense the assumption $A \le 0$ is somehow reasonable for the coupling property. On the other hand, by [18], Theorem 3.8, in the diffusion case, all bounded harmonic functions could be constant (i.e., the shift-coupling property holds) provided all eigenvalues of A have non-positive real parts. It would be interesting to extend this result to the jump case.

(2) The condition $\int_{\{|z-z_0|\le\varepsilon\}} \rho_0(z)^{-1}\,\mathrm{d}z < \infty$ follows from $\inf_{|z-z_0|\le\varepsilon} \rho_0(z) > 0$, which corresponds to the uniformly elliptic condition in the diffusion setting. Similarly to (3.1) in the Brownian motion case, (3.2) is sharp for large $t > 0$ in the pure jump case. To see this, let $n = d = B = 1$, $A = Q = 0$ and let ν be a probability measure such that

$$\int_{\mathbb{R}} z\nu(\mathrm{d}z) = 0, \qquad \int_{\mathbb{R}} z^2\nu(\mathrm{d}z) = 1, \qquad \int_{\mathbb{R}} |z|^3\nu(\mathrm{d}z) < \infty.$$

Then the corresponding Lévy process reduces to the compound Poisson process up to a constant drift b_0:

$$X_t = \sum_{i=1}^{N_t} \xi_i + b_0 t,$$

where N_t is the Poisson process on $\mathbb{Z}_+$ with rate 1 and $\{\xi_i\}_{i\ge1}$ are i.i.d. and independent of N_t with common distribution ν. By the Berry–Esseen inequality (see [23]),

$$\sup_{r\in\mathbb{R}} \bigl|\mathbb{P}\bigl(X_t < r\sqrt{t} + b_0 t\bigr) - \Phi(r)\bigr| \le \frac{c_0}{\sqrt{t}}, \qquad t > 0,$$

holds for some constant $c_0 > 0$, where Φ is the standard Gaussian distribution function. Therefore,

$$\begin{aligned}\|P_t(x,\cdot)-P_t(0,\cdot)\|_{\mathrm{var}} &\ge 2\sup_{r\in\mathbb{R}}\big|\mathbb{P}\big(X_t<r\sqrt{t}+b_0t\big)-\mathbb{P}\big(X_t<r\sqrt{t}+b_0t-x\big)\big| \\ &\ge 2\sup_{r\in\mathbb{R}}\big|\Phi(r)-\Phi\big(r-x/\sqrt{t}\big)\big|-\frac{4c_0}{\sqrt{t}}\ge\frac{c_1|x|-4c_0}{\sqrt{t}}, \qquad t\ge x^2,\end{aligned}$$

holds for some constant $c_1 > 0$.

It is well known that the solution to (1.1) can be formulated as

$$X_t^x = \mathrm{e}^{At}x + \int_0^t \mathrm{e}^{A(t-s)}B\,\mathrm{d}L_s, \qquad x\in\mathbb{R}^n, t\ge 0. \tag{3.3}$$

To make use of Theorem 2.1, we shall split L_t into two independent parts:

$$L_t = L_t^1 + L_t^0,$$

where $L^0 := \{L_t^0\}_{t\ge0}$ is the compound Poisson process with Lévy measure $\nu_0(\mathrm{d}z) := \rho_0(z)\,\mathrm{d}z$, and $L^1 := \{L_t^1\}_{t\ge0}$ is the Lévy process with Lévy measure $\nu-\nu_0$ generated by $\mathscr{L}-\mathscr{L}_0$ for

$$\mathscr{L}_0 f := \int_{\mathbb{R}^d}\big(f(\cdot+z)-f(z)\big)\nu_0(\mathrm{d}z).$$

So, (3.3) reduces to

$$X_t^x = \mathrm{e}^{At}x + \int_0^t \mathrm{e}^{A(t-s)}B\,\mathrm{d}L_s^1 + \int_0^t \mathrm{e}^{A(t-s)}B\,\mathrm{d}L_s^0, \qquad x\in\mathbb{R}^n, t\ge 0. \tag{3.4}$$

Moreover, let

$$L_t^0 = \sum_{i=1}^{N_t}\xi_i, \tag{3.5}$$

where $N := \{N_t\}_{t\ge0}$ is the Poisson process on $\mathbb{Z}_+$ with rate $\lambda_0 := \nu_0(\mathbb{R}^d)$, and $\{\xi_i\}_{i\ge1}$ are i.i.d. real random variables with common distribution ν_0/λ_0 such that N, $\{\xi_i\}_{i\ge1}$ and L^1 are independent.

To prove Theorem 3.1, we introduce the following fundamental lemma.

Lemma 3.2. *Let $\lambda_0\in(0,\infty)$, and let $\{\eta_i\}_{i\ge1}$ be a sequence of square-integrable real random variables that are conditional independent given N such that $\mathbb{E}(\eta_i|N)=1$ and $\mathbb{E}(\eta_i^2|N)\le\sigma$ hold for some constant $\sigma\in(0,\infty)$ and all $i\ge1$. Then*

$$E\left(1-\frac{1}{\lambda_0 T}\sum_{i=1}^{N_T}\eta_i\right)^2 \le \frac{\sigma}{\lambda_0 T}.$$

Proof. Since $\mathbb{E}(\eta_i\eta_j|N) = 1$ for $i \neq j$ and $\mathbb{E}(\eta_i^2|N) \leq \sigma$ for $i \geq 1$, we have

$$\begin{aligned}\mathbb{E}\left(1 - \frac{1}{\lambda_0 T}\sum_{i=1}^{N_T}\eta_i\right)^2 &= \frac{1}{(\lambda_0 T)^2}\mathbb{E}\left(\sum_{i=1}^{N_T}\eta_i\right)^2 - \frac{2}{\lambda_0 T}\mathbb{E}\sum_{i=1}^{N_T}\eta_i + 1\\ &= \frac{1}{(\lambda_0 T)^2}\mathbb{E}\left\{\sum_{i,j=1}^{N_T}\mathbb{E}(\eta_i\eta_j|N)\right\} - \frac{2}{\lambda_0 T}\mathbb{E}\left\{\sum_{i=1}^{N_T}\mathbb{E}(\eta_i|N)\right\} + 1\\ &\leq \frac{1}{(\lambda_0 T)^2}\sum_{n=1}^{\infty}\frac{(n^2 - n + \sigma n)(\lambda_0 T)^n e^{-\lambda_0 T}}{n!} - 1 = \frac{\sigma}{\lambda_0 T}.\end{aligned}$$

□

Proof of Theorem 3.1. We simply denote $\mathbf{B}_r = \{z : |z - z_0| \leq r\}$ for $r > 0$. Using $\rho_0 \wedge 1$ to replace ρ_0, we may and do assume that $\rho_0 \leq 1$. In this case $\nu_0(dz) := \rho_0(z)\, dz$ is finite. For $T > 0$, let τ be a random variable on $[0, \infty)$ with distribution $\frac{1}{T}1_{[0,T]}(t)\, dt$ and ξ on $\mathbb{R}^n$ with distribution

$$\frac{1_{\mathbf{B}_{\varepsilon/2}}(z)\nu_0(dz)}{\nu_0(\mathbf{B}_{\varepsilon/2})},$$

such that L^0, L^1, ξ, τ are independent. Let $\Lambda(dw)$ be the distribution of L^0. It is easy to see that the distribution of (L^0, ξ, τ) is

$$\frac{1_{\mathbf{B}_{\varepsilon/2}}(z)1_{[0,T]}(t)}{T\nu_0(\mathbf{B}_{\varepsilon/2})}\Lambda(dw)\nu_0(dz)\, dt.$$

By Theorem 2.1 and (3.5), for any $z \in \mathbb{R}^d$ we have

$$\begin{aligned}&\mathbb{E}f\left(e^{AT}z + \int_0^T e^{A(T-t)}B\, d\big(L^0 + \xi 1_{[\tau,\infty)}\big)_t\right)\\ &\quad = \mathbb{E}\left\{\frac{f(e^{AT}z + \int_0^T e^{A(T-t)}B\, dL_t^0)}{T\nu_0(\mathbf{B}_{\varepsilon/2})}\sum_{t\leq T}1_{\mathbf{B}_{\varepsilon/2}\setminus\{0\}}(\Delta L_t^0)\right\}\\ &\quad = \mathbb{E}\left\{\frac{f(e^{AT}z + \int_0^T e^{A(T-t)}B\, dL_t^0)}{T\nu_0(\mathbf{B}_{\varepsilon/2})}\sum_{i=1}^{N_T}1_{\mathbf{B}_{\varepsilon/2}}(\xi_i)\right\}.\end{aligned}$$

Letting $\pi_{x,T}$ be the distribution of $x + \int_0^T e^{-At}B\, dL_t^1$ and combining this with (3.4) and the independence of L^0 and L^1, we obtain

$$\begin{aligned}&\mathbb{E}f\big(X_T^x + e^{A(T-\tau)}B\xi\big)\\ &\quad = \mathbb{E}f\left(e^{AT}\left\{x + \int_0^T e^{-At}B\, dL_t^1\right\} + \int_0^T e^{A(T-t)}B\, d\big(L^0 + \xi 1_{[\tau,\infty)}\big)_t\right)\end{aligned}$$

$$= \int_{\mathbb{R}^d} \left\{ \mathbb{E} f\left(\mathrm{e}^{AT} z + \int_0^T \mathrm{e}^{A(T-t)} B \, \mathrm{d}\big(L^0 + \xi 1_{[\tau,\infty)}\big)_t \right) \right\} \pi_{x,T}(\mathrm{d}z)$$

$$= \int_{\mathbb{R}^d} \mathbb{E} \left\{ \frac{f(\mathrm{e}^{AT} z + \int_0^T \mathrm{e}^{A(T-t)} B \, \mathrm{d}L_t^0)}{T \nu_0(\mathbf{B}_{\varepsilon/2})} \sum_{i=1}^{N_T} 1_{\mathbf{B}_{\varepsilon/2}}(\xi_i) \right\} \pi_{x,T}(\mathrm{d}z)$$

$$= \mathbb{E} \left\{ \frac{f(X_T^x)}{T \nu_0(\mathbf{B}_{\varepsilon/2})} \sum_{i=1}^{N_T} 1_{\mathbf{B}_{\varepsilon/2}}(\xi_i) \right\}. \tag{3.6}$$

Next, since $\operatorname{Rank}(B) = n$, we have $d \geq n$ and up to a permutation of coordinates in $\mathbb{R}^d$, we may and do assume that $B = (B_1, B_2)$ for some invertible $n \times n$ matrix B_1 and some $n \times (d-n)$ matrix B_2. If, in particular, $n = d$, then $B_1 = B$. Moreover, for simplicity we write

$$\mathbb{R}^n = \mathbb{R}^n \times \{\bar{0}\} \subset \mathbb{R}^d,$$

where $\bar{0}$ is the original in $\mathbb{R}^{d-n}$. In other words, for any $x \in \mathbb{R}^n$, we set $x = (x, \bar{0}) \in \mathbb{R}^d$. Since $\langle Ax, x \rangle \leq 0$ for $x \in \mathbb{R}^n$, if

$$\|B_1^{-1}\| \cdot |x - y| \leq \frac{\varepsilon}{2},$$

then

$$|B_1^{-1} \mathrm{e}^{\tau A}(x - y)| \leq \|B_1^{-1}\| \cdot |\mathrm{e}^{\tau A}(x - y)| \leq \|B_1^{-1}\| \cdot |x - y| \leq \frac{\varepsilon}{2}.$$

So the distribution of $(L^0, \xi + B_1^{-1} \mathrm{e}^{A\tau}(x - y), \tau)$ is

$$\frac{1_{[0,T]}(t) 1_{\mathbf{B}_{\varepsilon/2} + B_1^{-1} \mathrm{e}^{At}(x-y)}(z)}{T \nu_0(\mathbf{B}_{\varepsilon/2})} \Lambda(\mathrm{d}w) \nu_0\big(\mathrm{d}z - B_1^{-1} \mathrm{e}^{At}(x - y)\big) \, \mathrm{d}t$$

$$= \frac{1_{[0,T]}(t) 1_{\mathbf{B}_{\varepsilon/2} + B_1^{-1} \mathrm{e}^{At}(x-y)}(z) \rho_0(z - B_1^{-1} \mathrm{e}^{At}(x - y))}{T \nu_0(\mathbf{B}_{\varepsilon/2}) \rho_0(z)} \Lambda(\mathrm{d}w) \nu_0(\mathrm{d}z) \, \mathrm{d}t.$$

Similarly to (3.6), due to Theorem 2.1, (3.5) and the independence of L^0 and L^1, we have

$$\mathbb{E} f\big(X_T^x + \mathrm{e}^{A(T-\tau)} B \xi\big)$$

$$= \mathbb{E} f\left(\mathrm{e}^{AT} y + \int_0^T \mathrm{e}^{A(T-t)} B \, \mathrm{d}\big(L^1 + L^0 + \{\xi + B_1^{-1} \mathrm{e}^{A\tau}(x - y)\} 1_{[\tau,\infty)}\big)_t \right)$$

$$= \mathbb{E} \left\{ \frac{f(X_T^y)}{T \nu_0(\mathbf{B}_{\varepsilon/2})} \sum_{t \leq T} 1_{(\mathbf{B}_{\varepsilon/2} + B_1^{-1} \mathrm{e}^{At}(x-y)) \setminus \{0\}}(\Delta L_t^0) \frac{\rho_0(\Delta L_t^0 - B_1^{-1} \mathrm{e}^{At}(x - y))}{\rho_0(\Delta L_t^0)} \right\}$$

$$= \mathbb{E} \left\{ \frac{f(X_T^y)}{T \nu_0(\mathbf{B}_{\varepsilon/2})} \sum_{i=1}^{N_T} 1_{\mathbf{B}_{\varepsilon/2} + B_1^{-1} \mathrm{e}^{A\tau_i}(x-y)}(\xi_i) \frac{\rho_0(\xi_i - B_1^{-1} \mathrm{e}^{A\tau_i}(x - y))}{\rho_0(\xi_i)} \right\},$$

where τ_i is the ith jump time of N_t for $i \geq 1$. Combining this with (3.6), we arrive at

$$\begin{aligned}
&|P_T f(x) - P_T f(y)| \\
&\leq \mathbb{E}\Bigg| f(X_T^y)\Bigg(1 - \frac{1}{T\nu_0(\mathbf{B}_{\varepsilon/2})}\sum_{i=1}^{N_T} 1_{\mathbf{B}_{\varepsilon/2}+B_1^{-1}e^{A\tau_i}(x-y)}(\xi_i)\frac{\rho_0(\xi_i - B_1^{-1}e^{A\tau_i}(x-y))}{\rho_0(\xi_i)}\Bigg) \\
&\quad + f(X_T^x)\Bigg(\frac{1}{T\nu_0(\mathbf{B}_{\varepsilon/2})}\sum_{i=1}^{N_T} 1_{\mathbf{B}_{\varepsilon/2}}(\xi_i) - 1\Bigg)\Bigg|, \qquad |x-y| \leq \frac{\varepsilon}{2\|B_1^{-1}\|}.
\end{aligned} \tag{3.7}$$

To apply Lemma 3.2, let

$$\eta_i = \frac{\lambda_0 1_{\mathbf{B}_{\varepsilon/2}}(\xi_i)}{\nu_0(\mathbf{B}_{\varepsilon/2})}, \qquad \tilde{\eta}_i = \frac{\lambda_0\rho_0(\xi_i - B_1^{-1}e^{A\tau_i}(x-y))}{\nu_0(\mathbf{B}_{\varepsilon/2})\rho_0(\xi_i)} 1_{\mathbf{B}_{\varepsilon/2}+B_1^{-1}e^{A\tau_i}(x-y)}(\xi_i), i \geq 1.$$

Then $\{\eta_i\}_{i\geq 1}$ are i.i.d. and independent of N with

$$\begin{aligned}
\mathbb{E}\eta_i &= \frac{1}{\lambda_0}\int_{\mathbf{B}_{\varepsilon/2}} \frac{\lambda_0}{\nu_0(\mathbf{B}_{\varepsilon/2})}\nu_0(\mathrm{d}z) = 1, \\
\mathbb{E}\eta_i^2 &= \frac{1}{\lambda_0}\int_{\mathbf{B}_{\varepsilon/2}} \frac{\lambda_0^2}{\nu_0(\mathbf{B}_{\varepsilon/2})^2}\nu_0(\mathrm{d}z) = \frac{\lambda_0}{\nu_0(\mathbf{B}_{\varepsilon/2})} < \infty,
\end{aligned}$$

while $\{\tilde{\eta}_i\}_{i\geq 1}$ are conditional independent given N such that

$$\begin{aligned}
\mathbb{E}(\tilde{\eta}_i|N) &= \frac{1}{\lambda_0}\int_{\mathbf{B}_{\varepsilon/2}+B_1^{-1}e^{A\tau_i}(x-y)} \frac{\lambda_0\rho_0(z - B_1^{-1}e^{A\tau_i}(x-y))}{\nu_0(\mathbf{B}_{\varepsilon/2})\rho_0(z)}\nu_0(\mathrm{d}z) \\
&= \frac{1}{\lambda_0}\int_{\mathbf{B}_{\varepsilon/2}+B_1^{-1}e^{A\tau_i}(x-y)} \frac{\lambda_0\rho_0(z - B_1^{-1}e^{A\tau_i}(x-y))}{\nu_0(\mathbf{B}_{\varepsilon/2})}\,\mathrm{d}z = 1,
\end{aligned}$$

and since $\rho_0 \leq 1$ and $|B_1^{-1}e^{A\tau_i}(x-y)| \leq \frac{\varepsilon}{2}$,

$$\begin{aligned}
\mathbb{E}(\tilde{\eta}_i^2|N) &= \frac{1}{\lambda_0}\int_{\mathbf{B}_{\varepsilon/2}+B_1^{-1}e^{A\tau_i}(x-y)} \frac{\lambda_0^2\rho_0(z - B_1^{-1}e^{A\tau_i}(x-y))^2}{\nu_0(\mathbf{B}_{\varepsilon/2})^2\rho_0(z)^2}\nu_0(\mathrm{d}z) \\
&\leq \lambda_0\int_{\mathbf{B}_{\varepsilon/2}+B_1^{-1}e^{A\tau_i}(x-y)} \frac{\mathrm{d}z}{\nu_0(\mathbf{B}_{\varepsilon/2})^2\rho_0(z)} \\
&\leq \frac{\lambda_0}{\nu_0(\mathbf{B}_{\varepsilon/2})^2}\int_{\mathbf{B}_{\varepsilon}} \frac{\mathrm{d}z}{\rho_0(z)} < \infty.
\end{aligned}$$

Therefore, by (3.7) and Lemma 3.2,

$$\|P_T(x,\cdot) - P_T(y,\cdot)\|_{\mathrm{var}} \leq \frac{c}{\sqrt{T}}, \qquad T > 0, |x-y| \leq \frac{\varepsilon}{2\|B_1^{-1}\|},$$

holds for some constant $c > 0$. This implies (3.2) for some constant $C > 0$ since for $|x - y| > \frac{\varepsilon}{2\|B_1^{-1}\|}$ and $m_{x,y} := \inf\{i \in \mathbb{N} : i \geq 2\|B_1^{-1}\| \cdot |x - y|/\varepsilon\}$, we have

$$\begin{aligned}
&\|P_t(x, \cdot) - P_t(y, \cdot)\|_{\mathrm{var}} \\
&\quad \leq \sum_{i=1}^{m_{x,y}} \left\| P_t\left(x + \frac{i(y-x)}{m_{x,y}}, \cdot\right) - P_t\left(x + \frac{(i-1)(y-x)}{m_{x,y}}, \cdot\right) \right\|_{\mathrm{var}}.
\end{aligned}$$

Finally, it is easy to see that (3.2) implies the statement (i) and hence, the coupling property of the process. □

To conclude this section, we present a result on the equivalence of the coupling property and the shift-coupling property by using a criterion in [9].

Proposition 3.3. *Let $\nu(\mathbb{R}^d) < \infty$ and $A = 0$. If either $b = \int_{\{|z|\leq 1\}} z\nu(\mathrm{d}z)$ or* Rank$(B) = n$ *and Q is non-degenerate, then the coupling property is equivalent to the shift-coupling property.*

Proof. Let $\lambda := \nu(\mathbb{R}^d) < \infty$ and $A = 0$. Let $L_t = L_t^1 + L_t^0$ as before for L_t^0 being the compound Poisson process specified in (3.5) for ν in place of ν_0. If $Q = 0$ and $b = \int_{\{|z|\leq 1\}} z\nu(\mathrm{d}z)$, then $X_t^x = x + BL_t^0$. So, for any non-negative measurable function f on $\mathbb{R}^n$, and any $t, s > 0$, we have

$$P_{t+s} f(x) = \mathbb{E} f(BL_{t+s}^0 + x) \geq \mathbb{E}\big[f(BL_t^0 + x)1_{\{N_{t+s} - N_t = 0\}}\big] = \mathrm{e}^{-\lambda s} P_t f(x). \tag{3.8}$$

Therefore, by [9], Theorem 5, the coupling property is equivalent to the shift-coupling property.

Next, let $A = 0$, Q be non-degenerate and Rank$(B) = n$. Let P_t^J and P_t^D be the semigroups of BL_t^0 and BL_t^1, respectively. Then it is easy to see that the generator of P_t^D is an elliptic operator with constant coefficients and hence, satisfies the Bakry–Emery curvature-dimension condition. Therefore, according to [3], there exists a constant $k \geq n$ such that

$$P_t^D f \leq \left(\frac{t+s}{s}\right)^{k/2} P_{t+s}^D f, \qquad t, s > 0, \tag{3.9}$$

holds for non-negative measurable function f. Since $A = 0$ implies that the diffusion part and the jump part are independent, we have $P_t = P_t^D P_t^J$, where P_t^J is the semigroup associated to BL_t^0, which satisfies (3.8). Therefore,

$$P_t f \leq \left(\frac{t+s}{s}\right)^{k/2} \mathrm{e}^{\lambda s} P_{t+s} f, \qquad f \geq 0, s, t > 0.$$

This implies the equivalence of the coupling property and the shift-coupling property according to [9], Theorem 5. □

The condition $b = \int_{\{|z|\le 1\}} z\nu(\mathrm{d}z)$ is used to ensure the desired inequality (3.8). If this condition does not hold, there exists $b_0 \neq 0$ such that $X_t^x = x + BL_t^0 + b_0 t$, so that instead of (3.8) one has

$$P_{t+s} f(x) \ge \mathbb{E}\big[f(BL_t^0 + x + b_0 s)1_{\{N_{t+s}-N_t=0\}}\big] = \mathrm{e}^{-\lambda s} P_t f(x + b_0 s),$$

which is not enough to apply [9], Theorem 5.

4. Harnack inequality and ultracontractivity

Let $\nu \ge \nu_0 := \rho_0(z)\,\mathrm{d}z > 0$ for some $\rho_0 > 0$ with $\lambda_0 := \nu_0(\mathbb{R}^d) \in (0,\infty)$. As in Section 3, let $L_t = L_t^1 + L_t^0$ such that L^0 and L^1 are independent, where L^0 is the compound Poisson process with Lévy measure ν_0. Let τ_1 be the first jump time of L_t^0. We shall establish the Harnack inequality for

$$P_t^1 f(x) := \mathbb{E}\big\{f(X_t^x)1_{\{\tau_1\le t\}}\big\}, \qquad t \ge 0, x \in \mathbb{R}^n, f \in \mathscr{B}_b(\mathbb{R}^n). \tag{4.1}$$

Theorem 4.1. *Let $\nu \ge \nu_0 := \rho_0(z)\,\mathrm{d}z$ with $\lambda_0 := \nu_0(\mathbb{R}^d) \in (0,\infty)$, and let P_t^1 be defined above. Let* $\mathrm{Rank}(B) = n$. *There exists a constant $c = c(B) > 0$ such that if*

$$V_p(r) := \frac{1}{\lambda_0} \sup_{|z'|\le r} \int_{\mathbb{R}^d} \frac{\rho_0(z-z')^{p/(p-1)}}{\rho_0(z)^{1/(p-1)}}\,\mathrm{d}z < \infty, \qquad r \ge 0,$$

holds for some $p > 1$, then for any positive measurable function f on $\mathbb{R}^n$,

$$(P_t^1 f(x))^p \le (P_t^1 f^p(y))\big\{(1-\mathrm{e}^{-\lambda_0 t})V_p\big(c\mathrm{e}^{\|A\|t}|x-y|\big)\big\}^{p-1}, \qquad x, y \in \mathbb{R}^d, t > 0,$$

holds. Consequently,

$$\|P_t^1\|_{p\to\infty} \le (1-\mathrm{e}^{-\lambda_0 t})\mathrm{e}^{\|A\|t/p}\left\{\int_{\mathbb{R}^d} \frac{\mathrm{d}x}{V_p(c\mathrm{e}^{\|A\|t}|x|)^{p-1}}\right\}^{-1/p} < \infty,$$

where $\|\cdot\|_{p\to q}$ is the operator norm from $L^p(\mathbb{R}^n;\mathrm{d}x)$ to $L^q(\mathbb{R}^n;\mathrm{d}x)$ for any $p, q \ge 1$.

Proof. Let L^0, L^1, ξ, τ be independent such that the distributions of ξ and τ are ν_0/λ_0 and $\frac{1}{T}1_{[0,T]}(t)\,\mathrm{d}t$, respectively. As in the proof of Theorem 3.1, let $B = (B_1, B_2)$ such that B_1 is invertible. Since the distribution of (L^0, ξ, τ) is

$$\frac{1_{[0,T]}(t)}{\lambda_0 T}\Lambda(\mathrm{d}w)\nu_0(\mathrm{d}z)\,\mathrm{d}t,$$

Corollary 2.3 holds for

$$U(w) = \frac{1}{\lambda_0 T} n_T(w),$$

where n_T is defined by (2.1). Since $\tau \le T$ and $\xi \ne 0$, which are independent of L^0 and L^1, we have

$$U\big(L^0 + \xi 1_{[\tau,\infty)}\big) = \frac{1}{\lambda_0 T} n_T\big(L^0 + \xi 1_{[\tau,\infty)}\big) > 0.$$

Therefore, by Corollary 2.3 and noting that $\tau_1 \le T$ a.s. for the process $L^0 + \xi 1_{[\tau,\infty)}$,

$$\begin{aligned}
P_T^1 f(x) &= \mathbb{E}\big[f(X_T^x) 1_{\{\tau_1 \le T\}}\big] \\
&= \mathbb{E}\left\{\frac{\lambda_0 T f(\mathrm{e}^{AT}x + \int_0^T \mathrm{e}^{A(T-t)} B\, \mathrm{d}(L^1 + L^0 + \xi 1_{[\tau,\infty)})_t)}{n_T(L^0 + \xi 1_{[\tau,\infty)})}\right\} \\
&= \mathbb{E}\left\{\frac{\lambda_0 T f(\mathrm{e}^{AT}y + \int_0^T \mathrm{e}^{A(T-t)} B\, \mathrm{d}(L^1 + L^0 + \{\xi + B_1^{-1}\mathrm{e}^{A\tau}(x-y)\} 1_{[\tau,\infty)})_t)}{n_T(L^0 + \{\xi + B_1^{-1}\mathrm{e}^{A\tau}(x-y)\} 1_{[\tau,\infty)})}\right\} \\
&= \mathbb{E}\left\{\frac{f(X_T^y) 1_{\{\tau_1 \le T\}}}{N_T} \sum_{i=1}^{N_T} \frac{\rho_0(\xi_i - B_1^{-1}\mathrm{e}^{A\tau_i}(x-y))}{\rho_0(\xi_i)}\right\}, \qquad (4.2)
\end{aligned}$$

where $c = \|B_1^{-1}\|$. By the Hölder inequality, we obtain

$$\begin{aligned}
(P_T^1 f(x))^p &\le P_T^1 f^p(y) \left\{\mathbb{E}\left(\frac{1_{\{N_T \ge 1\}}}{N_T} \sum_{i=1}^{N_T} \frac{\rho_0(\xi_i - B_1^{-1}\mathrm{e}^{A\tau_i}(x-y))}{\rho_0(\xi_i)}\right)^{p/(p-1)}\right\}^{p-1} \\
&\le P_T^1 f^p(y) \left\{\sum_{n=1}^{\infty} \frac{(\lambda_0 T)^n \mathrm{e}^{-\lambda_0 T}}{n(n!)} \sum_{i=1}^{n} \sup_{|z'| \le c\mathrm{e}^{\|A\|T}|x-y|} \mathbb{E}\left(\frac{\rho_0(\xi_i - z')}{\rho_0(\xi_i)}\right)^{p/(p-1)}\right\}^{p-1} \\
&= P_T^1 f^p(y) \big\{(1 - \mathrm{e}^{-\lambda_0 T}) V_p\big(c\mathrm{e}^{\|A\|T}|x-y|\big)\big\}^{p-1}.
\end{aligned}$$

This implies the desired Harnack inequality.

Next, since there exists a probability μ_T on $\mathbb{R}^n$ such that

$$P_T f^p(x) =: \mathbb{E} f^p(X_T^x) = \int_{\mathbb{R}^n} f^p(\mathrm{e}^{AT}x + y) \mu_T(\mathrm{d}y),$$

if $\int_{\mathbb{R}^n} f^p(x)\,\mathrm{d}x \le 1$, then

$$\int_{\mathbb{R}^n} P_T^1 f^p(x)\,\mathrm{d}x \le \int_{\mathbb{R}^n} P_T f^p(x)\,\mathrm{d}x = \int_{\mathbb{R}^n} \mu_T(\mathrm{d}y) \int_{\mathbb{R}^n} f^p(\mathrm{e}^{TA}x + y)\,\mathrm{d}x \le \mathrm{e}^{\|A\|T}.$$

Therefore, by the Harnack inequality, for any non-negative f with $\int_{\mathbb{R}^d} f^p(z)\,\mathrm{d}z \le 1$,

$$\begin{aligned}
&(P_T^1 f(x))^p \int_{\mathbb{R}^d} \frac{\mathrm{d}y}{(V_p(c\mathrm{e}^{\|A\|T}|x-y|))^{p-1}} \\
&\quad \le (1 - \mathrm{e}^{-\lambda_0 T})^{p-1} \int_{\mathbb{R}^d} P_T^1 f^p(y)\,\mathrm{d}y \le (1 - \mathrm{e}^{-\lambda_0 T})^p \mathrm{e}^{\|A\|T}.
\end{aligned}$$

This implies the desired upper bound of $\|P_T^1\|_{p\to\infty}$. □

It is easy to see that $V_p < \infty$ holds for many concrete choices of ρ_0, including $\rho_0(z) := c_1 e^{-c_2|z|^r}$ for some constants $c_1, c_2, r > 0$ and $\rho_0(z) := c(1+|z|)^{-r}$ for some $r > d$ and $c > 0$.

Finally, when ν has a large enough absolutely continuous part, we may derive the ultracontractivity by comparing with the α-stable process.

Theorem 4.2. *Assume that $n = d$ and $B = I$. Let $\alpha \in (0,2)$. If*

$$\nu(\mathrm{d}z) \ge \frac{c}{|z|^{\alpha+d}} 1_{\{|z|<r\}}\, \mathrm{d}z$$

holds for some constants $c, r > 0$, then

$$\|P_t\|_{1\to\infty} \le \frac{c'}{(1\wedge t)^{d/\alpha}}, \qquad t > 0,$$

holds for some constant $c' > 0$.

Proof. (a) We first observe that if $r = \infty$, that is,

$$\nu(\mathrm{d}z) \ge \frac{c}{|z|^{\alpha+d}}\, \mathrm{d}z, \tag{4.3}$$

then

$$\|P_t\|_{1\to\infty} \le \frac{c'}{t^{\mathrm{d}/\alpha}}, \qquad t \in (0,1],$$

holds. When $A = 0$ and $\nu(\mathrm{d}z) \ge \frac{c}{|z|^{\alpha+d}}\, \mathrm{d}z$ this is well known according to the heat kernel upper bound of the α-stable process. In general, let η be the symbol of the Lévy process L with characteristics (b, Q, ν). Let μ_t be the probability measure on $\mathbb{R}^d$ with Fourier transform

$$\hat{\mu}_t(z) = \exp\left[-\int_0^t \eta(\mathrm{e}^{sA^*}z)\, \mathrm{d}s\right], \qquad z \in \mathbb{R}^d.$$

We have

$$P_t f(x) = \int_{\mathbb{R}^d} f(\mathrm{e}^{tA}x + y)\mu_t(\mathrm{d}y).$$

Let $c_1 > 0$ be such that

$$c\int_0^t |\mathrm{e}^{sA^*}z|^\alpha\, \mathrm{d}s \ge c_1 t|z|^\alpha, \qquad t \in [0,1].$$

According to (4.3) there are two probability measures μ_t^1 and μ_t^2 on $\mathbb{R}^d$ such that $\mu_t = \mu_t^1 * \mu_t^2$ and the Fourier transform of μ_t^1 is

$$\hat{\mu}_t^1(z) = \exp[-c_1 t|z|^\alpha].$$

Combining this with the known heat kernel bound of the α-stable process, we can find a constant $c' > 0$ such that for any $f \geq 0$,

$$\begin{aligned} P_t f(x) &= \int_{\mathbb{R}^d} \mu_t^1(\mathrm{d}z) \int_{\mathbb{R}^d} f(\mathrm{e}^{tA}x + y + z)\mu_t^2(\mathrm{d}y) \\ &\leq \frac{c'}{t^{d/\alpha}} \int_{\mathbb{R}^d} f(z)\,\mathrm{d}z, \qquad x \in \mathbb{R}^d, t \in (0, 1]. \end{aligned}$$

This implies the desired estimate.

(b) Let $r \in (0, \infty)$. To apply (a), let L^0 be the compound Poisson process independent of L with Lévy measure

$$\nu_0(\mathrm{d}z) := \frac{c}{(|z| \vee r)^{d+\alpha}}\,\mathrm{d}z.$$

Then $\bar{L} := L + L^0$ is a Lévy process with Lévy measure

$$\bar{\nu}(\mathrm{d}z) = \nu(\mathrm{d}z) + \nu_0(\mathrm{d}z) \geq \frac{c}{|z|^{\alpha+d}}\,\mathrm{d}z.$$

Let $\bar{P}_t$ be the semigroup associated with the equation

$$\mathrm{d}\bar{X}_t = A\bar{X}_t\,\mathrm{d}t + \mathrm{d}\bar{L}_t.$$

By (a)

$$\|\bar{P}_t\|_{1\to\infty} \leq \frac{c'}{t^{d/\alpha}}, \qquad t \in (0, 1], \tag{4.4}$$

holds for some constant $c' > 0$. Let τ_1 be the first jump time of L^0. We have

$$\begin{aligned} \bar{P}_t f(x) &:= \mathbb{E} f\left(\mathrm{e}^{At}x + \int_0^t \mathrm{e}^{A(t-s)}\,\mathrm{d}L_s + \int_0^t \mathrm{e}^{A(t-s)}\,\mathrm{d}L_s^0\right) \\ &\geq \mathbb{E}\left\{1_{\{\tau_1 > t\}} f\left(\mathrm{e}^{At}x + \int_0^t \mathrm{e}^{A(t-s)}\,\mathrm{d}L_s\right)\right\} \\ &= \mathrm{e}^{-\lambda_0 t} P_t f(x), \qquad f \geq 0, \end{aligned}$$

where $\lambda_0 := \nu_0(\mathbb{R}^d) < \infty$. Combining this with (4.4) we complete the proof. □

5. Strong Feller property

As in Sections 3 and 4, let $\nu \geq \nu_0 := \rho_0(z)\,\mathrm{d}z$ for some non-negative measurable function ρ_0 on $\mathbb{R}^d$ such that $\lambda_0 := \nu_0(\mathbb{R}^d) > 0$. Let $L_t = L_t^1 + L_t^0$ for independent L^1 and L^0 such that L_t^0 is the compound Poisson process with Lévy measure ν_0. For any $i \geq 1$, let τ_i be the ith jump time of L_t^0. If $\lambda_0 = \infty$, we set $\tau_i = 0$ for all $i \geq 1$ by convention. We shall prove the strong Feller property for the operator P_t^m defined by

$$P_t^m f(x) = \mathbb{E}\big\{f(X_t^x)1_{\{\tau_m \leq t \wedge (\tau_1 + t_m)\}}\big\}, \tag{5.1}$$

where $m \geq 1$ and

$$t_m := \sup\{t \geq 0 : \mathrm{Rank}(\mathrm{e}^{s_1 A}B, \dots, \mathrm{e}^{s_m A}B) = n, \forall 0 \leq s_1 < \cdots < s_m \leq t\}.$$

According to the following lemma, we have $t_m > 0$ provided the rank condition

$$\mathrm{Rank}(B, AB, \dots, A^{m-1}B) = n \tag{5.2}$$

holds. This extends [19], Lemma 2.2, by allowing $m \neq n$.

Lemma 5.1. *If (5.2) holds for some $m \geq 1$, then $t_m > 0$. Consequently, for $0 \leq s_1 < \cdots < s_m \leq t_m$ and*

$$\psi_{s_1,\dots,s_m}(z_1, \dots, z_m) := \sum_{i=1}^m \mathrm{e}^{s_i A}Bz_i, \qquad z_1, \dots, z_m \in \mathbb{R}^d,$$

$\gamma \circ \psi_{s_1,\dots,s_m}^{-1}$ is an absolutely continuous probability measure on $\mathbb{R}^n$ provided so is γ on $\mathbb{R}^{md}$.

Proof. By [19], Lemma 2.3, it suffices to prove the first assertion. For $0 \leq s_1 < \cdots < s_m$, let

$$F_{i,0}^{(0)} = \mathrm{e}^{s_i A}, \qquad 1 \leq i \leq m,$$

$$F_{i,k}^{(k)} = \frac{F_{i,k-1}^{(k-1)} - F_{k,k-1}^{(k-1)}}{s_i - s_k}, \qquad 1 \leq k \leq m-1, k+1 \leq i \leq m.$$

Since

$$\frac{\mathrm{d}^i}{\mathrm{d}s^i}\mathrm{e}^{sA}\bigg|_{s=0} = A^i, \qquad i \geq 0,$$

for any $1 \leq i \leq m$, $F_{i,i-1}^{(i-1)}$ approximates $A^{(i-1)}$ as $s_m \downarrow 0$. Therefore, there exist real matrices $U_1, \dots, U_m$ depending on $(s_1, \dots, s_m)$ such that

$$\lim_{s_m \to 0} \|U_i\| = 0, \qquad 1 \leq i \leq m,$$

and

$$F_{i,i-1}^{(i-1)} = A^{i-1} + U_i, \qquad 1 \leq i \leq m.$$

Since $\{F_{i,i-1}^{(i-1)} : 1 \leq i \leq m\}$ are linear combinations of $\{\mathrm{e}^{s_i A} : 1 \leq i \leq m\}$, we have

$$\mathrm{Rank}(\mathrm{e}^{s_1 A}B, \mathrm{e}^{s_2 A}B, \dots, \mathrm{e}^{s_m A}B) \geq \mathrm{Rank}(B + U_1 B, AB + U_2 B, \dots, A^{m-1}B + U_m B). \tag{5.3}$$

Since $(B, AB, \dots, A^{m-1}B)$ has full rank n, and since $U_i B \to 0$ as $s_m \to 0$, there exists $t > 0$ such that if $0 \le s_1 < \dots < s_m \le t$, then

$$\mathrm{Rank}(B + U_1 B, AB + U_2 B, \dots, A^{m-1}B + U_m B) = n.$$

Combining this with (5.3) we complete the proof. □

Theorem 5.2. *If $t_m > 0$, then P_t^m is strong Feller for $t > 0$. Consequently, if (5.2) holds for some $m \ge 1$, then $t_{m\wedge n} > 0$ such that $P_t^{n\wedge m}$ is strong Feller for $t > 0$.*

Proof. According to Lemma 5.1 and the fact that (5.2) with $m \ge n$ is equivalent to the condition with $m = n$ (cf. [29]), it suffices to prove the first assertion. We shall complete the proof in four easy steps.

(a) We first observe that P_t^m is strong Feller if

$$P_t^m(0, \mathrm{d}x) := \mathbb{P}(X_t^0 \in \mathrm{d}x, t \ge \tau_m, t_m \ge \tau_m - \tau_1)$$

is absolutely continuous. Indeed, let $P_t^m(0, \mathrm{d}x) = g(x)\,\mathrm{d}x$. Then

$$P_t^m f(x) = \int_{\mathbb{R}^n} f(\mathrm{e}^{At}x + y) g(y)\,\mathrm{d}y.$$

Therefore, P_t^m is strong Feller according to [12], Lemma 11.

(b) Next, we claim that it suffices to prove the result for $\lambda_0 < \infty$. If $\lambda_0 = \infty$, then for any $l \ge 1$ let $\nu_l = (\rho_0 \wedge l)(z)\,\mathrm{d}z$ and $\lambda_l = \nu_l(\mathbb{R}^d)$. Let $\tau_i(l)$ be the ith jump time for the corresponding compound Poisson process with Lévy measure ν_l. If the assertion holds for finite λ_0, then we may use ν_l to replace ν_0 so that

$$\mathbb{P}\big(X_t^0 \in \mathrm{d}x, t \ge \tau_m(l), t_m \ge \tau_m(l) - \tau_1(l)\big)$$

is absolutely continuous. Therefore, for any measurable set $D \subset \mathbb{R}^n$ with volume $|D| = 0$,

$$P_t(0, D) \le \mathbb{P}\big(X_t^0 \in D, t \ge \tau_m(l), t_m \ge \tau_m(l) - \tau_1(l)\big) + \mathbb{P}\big(\tau_m(l) \ge t \wedge t_m\big) = \mathrm{e}^{-\lambda_l(t\wedge t_m)/m}.$$

Since $\lambda_l \uparrow \lambda_0 = \infty$ as $l \uparrow \infty$, we see that $P_t(0, \cdot)$ is absolutely continuous.

(c) We aim to show that it suffices to prove for the case that $L_t = L_t^0$, that is, $\nu = \nu_0$ and the Lévy process is the compound Poisson process with Lévy measure ν_0. Indeed, since

$$X_t^0 = \int_0^t \mathrm{e}^{(t-s)A} B\,\mathrm{d}L_s^1 + \int_0^t \mathrm{e}^{(t-s)A} B\,\mathrm{d}L_s^0,$$

where L^1 and L^0 are independent, $P_t^m(0, \mathrm{d}x)$ is absolutely continuous provided so is

$$\mathbb{P}\left(\int_0^t \mathrm{e}^{(t-s)A} B\,\mathrm{d}L_s^0 \in \mathrm{d}x, t \ge \tau_m, t_m \ge \tau_m - \tau_1\right).$$

(d) Now, assume that $\nu = \nu_0$ with $\lambda_0 \in (0, \infty)$ and $L_t = L_t^0$. Let $\pi(\mathrm{d}s_1, \dots, \mathrm{d}s_m)$ be the distribution of $(\tau_1, \dots, \tau_m)$, and let

$$K = \{(s_1, \dots, s_m) : s_m - s_1 \le t_m, 0 < s_1 < \cdots s_m \le t\}.$$

Since by (3.4) and (3.5) with $L_t^1 = 0$

$$X_t^0 = \int_0^t \mathrm{e}^{(t-s)A} B \,\mathrm{d}L_s^0 = \mathrm{e}^{(t-\tau_m)A} \sum_{i=1}^m \mathrm{e}^{(\tau_m - \tau_i)A} B\xi_i + \int_{\tau_m}^t \mathrm{e}^{(t-s)A} B \,\mathrm{d}L_s^0$$

provided $\tau_m \le t$, for any non-negative measurable function f on $\mathbb{R}^n$, we have

$$P_t^m f(0) = \int_K \mathbb{E} f\left(\mathrm{e}^{(t-s_m)A} \sum_{i=1}^m \mathrm{e}^{(s_m - s_i)A} B\xi_i + \int_{s_m}^t \mathrm{e}^{(t-s)A} B \,\mathrm{d}L_s^0\right) \pi(\mathrm{d}s_1, \dots, \mathrm{d}s_m), \qquad (5.4)$$

where $\{\xi_i\}$ are i.i.d. random variables with distribution ν_0/λ_0 independent of $(L_s^0)_{s \ge s_m}$. Since $\mathrm{e}^{(t-s_m)A}$ is invertible and $s_m - s_i < t_m$, by the definition of t_m the mapping

$$(z_1, \dots, z_m) \mapsto \mathrm{e}^{(t-s_m)A} \sum_{i=1}^m \mathrm{e}^{(s_m - s_i)A} B z_i$$

is onto, so that the distribution of the random variable

$$\mathrm{e}^{(t-s_m)A} \sum_{i=1}^m \mathrm{e}^{(s_m - s_i)A} B\xi_i$$

is absolutely continuous (see [19], Lemma 2.3). By (5.4) and the independence of this random variable and

$$\int_{s_m}^t \mathrm{e}^{(t-s)A} B \,\mathrm{d}L_s^0,$$

we conclude that $P_t^m(0, \mathrm{d}x)$ is absolutely continuous. □

Remark 5.1. In concrete examples we may have $t_m = \infty$ so that P_t^m reduces to

$$P_t^m f(x) := \mathbb{E}\big\{f(X_t^x) 1_{\{\tau_m \le t\}}\big\},$$

which refers to the conditional distribution of X_t^x in the event that L_t^0 jumps at least m times before t. For instance, as in [19], formula (1.3), let $n = 2$, $d = 1$ and

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

We have $A^2 = I$ and $AB = \binom{0}{1}$. So,

$$e^{sA}B = \left(\sum_{n=0}^{\infty} \frac{s^{2n}}{(2n)!}\right)\begin{pmatrix}0\\1\end{pmatrix} + \left(\sum_{n=0}^{\infty} \frac{s^{2n+1}}{(2n+1)!}\right)\begin{pmatrix}1\\0\end{pmatrix} = \cosh(s)\begin{pmatrix}0\\1\end{pmatrix} + \sinh(s)\begin{pmatrix}1\\0\end{pmatrix}$$

holds for all $s \geq 0$. Since $\sinh(s_2 - s_1) > 0$ for $s_2 - s_1 > 0$, and since $e^{s_1 A}$ is invertible, we have

$$\mathrm{Rank}(e^{s_1 A}B, e^{s_2 A}B) = \mathrm{Rank}\bigl(B, e^{(s_2 - s_1)A}B\bigr) = 2 = n.$$

Therefore, $t_2 = \infty$.

Acknowledgements

The author would like to thank Enrico Priola and a referee for helpful comments on the first version of the paper and for introducing him to the very interesting paper [19]. This paper supported in part by WIMCS and SRFDP.

References

[1] Aldous, D. and Thorisson, H. (1993). Shift-coupling. *Stochastic Proc. Appl.* **44** 1–14. MR1198659

[2] Applebaum, D. (2004). *Lévy Processes and Stochastic Calculus*. Cambridge Univ. Press. MR2072890

[3] Bakry, D. and Qian, Z. (1999). Harnack inequalities on a manifold with positive or negative Ricci curvature. *Rev. Mat. Iberoamericana* **15** 143–179. MR1681640

[4] Bertoin, J. (1996). *Lévy Processes*. Cambridge Univ. Press. MR1406564

[5] Bogdan, K. and Jakubowski, T. (2007). Estimate of heat kernel of fractional Laplacian perturbed by gradient operators. *Comm. Math. Phys.* **271** 179–198. MR2283957

[6] Chen, M.-F. (1992). *From Markov Chains to Non-Equilibrium Particle Systems*. Singapore: World Scientific. MR2091955

[7] Chen, M.-F. and Li, S.-F. (1989). Coupling methods for multidimensional diffusion processes. *Ann. Probab.* **17** 151–177. MR0972776

[8] Cranston, M. and Greven, A. (1995). Coupling and harmonic functions in the case of continuous time Markov processes. *Stochastic Process Appl.* **60** 261–286. MR1376804

[9] Cranston, M. and Wang, F.-Y. (2000). A condition for the equivalence of coupling and shift-coupling. *Ann. Probab.* **28** 1666–1679. MR1813838

[10] Da Prato, G. and Zabczyk, J. (1992). *Stochastic Equations in Infinite Dimensions*. Cambridge Univ. Press. MR1207136

[11] Fisz, M. and Varadarajan, V.S. (1962/63). A condition for absolute continuity of infinitely divisible distribution functions. *Z. Wahrsch. Verw. Gebiete* **1** 335–339. MR0149521

[12] Hawkes, J. (1979). Potential theory of Lévy processes. *Proc. London Math. Soc.* **38** 335–352. MR0531166

[13] Jakubowski, T. and Szczypkowski, K. (2010). Time-dependent gradient perturbations of fractional Laplacian. *J. Evol. Equ.* **10** 319–339. MR2643799

[14] Knopova, V. and Schilling, R.L. (2010). A note on the existence of transition probability densities for Lévy processes. Available at arXiv:1003.1419.

[15] Lindvall, T. (1992). *Lectures on the Coupling Methods*. New York: Wiley. MR1180522
[16] Lindvall, T. and Rogers, C. (1986). Coupling of multidimensional diffusions by reflection. *Ann. Probab.* **14** 860–872. MR0841588
[17] Mecke, J. (1967). Stationaire Zufällige Maße auf lokalkompakten abelschen Gruppen. *Z. Wahrsch. Verw. Gebiete* **9** 36–58. MR0228027
[18] Priola, E. and Zabczyk, J. (2004). Liouville theorems for non-local operators. *J. Funct. Anal.* **216** 455–490. MR2095690
[19] Priola, E. and Zabczyk, J. (2009). Densities for Ornstein–Uhlenbeck processes with jumps. *Bull. Lond. Math. Soc.* **41** 41–50. MR2481987
[20] Röckner, M. (1998). Stochastic analysis on configuration spaces: Basic ideas and recent results. In *New Directions in Dirichlet Forms. AMS/IP Stud. Math.* **8** 157–231. Providence, RI: Amer. Math. Soc. MR1652281
[21] Röckner, M., Ouyang, S.-X. and Wang, F.-Y. (2009). Harnack inequalities and applications for Ornstein–Uhlenbeck semigroups with jump. Available at arXiv:0908.2889.
[22] Röckner, M. and Wang, F.-Y. (2003). Harnack and functional inequalities for generalized Mehler semigroups. *J. Funct. Anal.* **203** 237–261. MR1996872
[23] Shevtsova, I.G. (2006). A refinement of the upper estimate of the absolute constant in the Berry–Esseen inequality. *Theory Probab. Appl.* **51** 622–626. MR2325552
[24] Song, R. and Vondraček, Z. (2009). Potential theory of subordinate Brownian motion. In *Potential Analysis of Stable Processes and its Extensions. Lecture Notes in Math.* **1980** 87–176. Berlin: Springer.
[25] Thorisson, H. (1994). Shift-coupling in continuous time. *Probab. Theory Related Fields* **99** 477–483. MR1288066
[26] Tucker, H.G. (1965). On a necessary and sufficient condition that an infinitely divisible distribution be absolutely continuous. *Trans. Amer. Math. Soc* **118** 316–330. MR0182061
[27] Wang, F.-Y. (1997). Logarithmic Sobolev inequalities on noncompact Riemannian manifolds. *Probab. Theory Related Fields* **109** 417–424. MR1481127
[28] Wang, F.-Y. and Yuan, C. (2010). Poincaré inequality on the path space of Poisson point processes. *J. Theory Probab.* **23** 824–833.
[29] Zabczyk, J. (1992). *Mathematical Control Theory: An Introduction*. Boston, MA: Birkhäuser. MR1193920

Received March 2010 and revised June 2010