Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications 121 (2011) 466–478
stochastic processes and their applications
www.elsevier.com/locate/spa

Gradient estimate for Ornstein–Uhlenbeck jump processes ☆

Feng-Yu Wang *

School of Mathematical Sci. and Lab. Math. Com. Sys., Beijing Normal University, Beijing 100875, China
Department of Mathematics, Swansea University, Singleton Park, SA2 8PP, UK

Received 16 June 2010; received in revised form 23 September 2010; accepted 6 December 2010
Available online 13 December 2010

Abstract

By using absolutely continuous lower bounds of the Lévy measure, explicit gradient estimates are derived for the semigroup of the corresponding Lévy process with a linear drift. A derivative formula is presented for the conditional distribution of the process at time t under the condition that the process jumps before t. Finally, by using bounded perturbations of the Lévy measure, the resulting gradient estimates are extended to linear SDEs driven by Lévy-type processes.

MSC: 60J75; 60J45

Keywords: Lévy process; Gradient estimate; Subordination; Compound Poisson process

1. Introduction

It is well-known that a Lévy process can be decomposed into two independent parts, i.e. the diffusion part and the jump part. If the diffusion part is non-degenerate, regularity properties for the semigroup of the Brownian motion can be easily confirmed for the Lévy semigroup. On the other hand, when the Lévy process is a pure jump, existence and regularities of the transition density have been derived by using conditions on the symbol or the Lévy measure (see [10–12] and references within); see also [5,9] for heat kernel upper bounds for α-stable processes with

☆ Supported in part by WIMCS, SRFDP.
* Corresponding address: School of Mathematical Sci. and Lab. Math. Com. Sys., Beijing Normal University, Beijing 100875, China.
E-mail addresses: wangfy@bnu.edu.cn, F.Y.Wang@swansea.ac.uk.

0304-4149/$ - see front matter
doi:10.1016/j.spa.2010.12.002

drifts. As a continuation to the recent work [15], where the coupling property and applications are studied by using absolutely continuous lower bounds of the Lévy measure, this note aims to derive gradient estimates of the Lévy semigroup in the same spirit.

Let L_t be the Lévy process on $\mathbb{R}^d$ with symbol (see e.g. [1])

$$\eta(u) = \mathrm{i}\langle u, b\rangle - \langle Qu, u\rangle + \int_{\mathbb{R}^d} \big(\mathrm{e}^{\mathrm{i}\langle u,z\rangle} - 1 - \mathrm{i}\langle u, z\rangle 1_{\{|z|<1\}}\big)\nu(\mathrm{d}z),$$

where $b \in \mathbb{R}^d$, Q is a non-negatively definite $d \times d$ matrix, and ν is a Lévy measure on $\mathbb{R}^d$. In references the Lévy symbol is also called the characteristic exponent or the Lévy exponent, and in e.g. [8], $-\eta$ rather than η is called the Lévy symbol. It is well known that L_t is a strong Markov process on $\mathbb{R}^d$ generated by

$$\mathscr{L} f := \langle b, \nabla f\rangle + \mathrm{Tr}(Q\nabla^2 f) + \int_{\mathbb{R}^d} \big\{ f(z + \cdot) - f - \langle \nabla f, z\rangle 1_{\{|z|\le 1\}}\big\}\nu(\mathrm{d}z) \tag{1.1}$$

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

Let P_t be the semigroup for the solution of the linear stochastic differential equation

$$\mathrm{d}X_t = AX_t\mathrm{d}t + \mathrm{d}L_t, \tag{1.2}$$

where A is a $d \times d$ matrix. According to [4], we have

$$P_t f(x) = \int_{\mathbb{R}^d} f(\mathrm{e}^{tA}x + y)\mu_t(\mathrm{d}y), \tag{1.3}$$

where μ_t is the probability measure on $\mathbb{R}^d$ with characteristic function

$$\hat{\mu}_t(z) = \exp\left[\int_0^t \eta(\mathrm{e}^{sA^*}z)\mathrm{d}s\right], \quad z \in \mathbb{R}^d. \tag{1.4}$$

Let $\mathscr{B}_b(\mathbb{R}^d)$ be the set of all bounded measurable functions on $\mathbb{R}^d$. We shall estimate $\|\nabla P_t f\|_\infty$, the uniform norm of the gradient $\nabla P_t f$, for $t > 0$ and $f \in \mathscr{B}_b(\mathbb{R}^d)$. When the Lévy measure is finite, with a positive probability the process does not jump before a fixed time $t > 0$. So, in this case, the semigroup is not strong Feller and thus, does not have a finite uniform gradient estimate. Therefore, to derive the uniform gradient estimate, it is essential to assume that ν is infinite. Since ν is always finite outside a neighborhood of 0, the behavior of ν around the origin will be crucial for the study.

We will make use of the following lower bound condition of ν:

$$\nu(\mathrm{d}z) \ge |z|^{-d} S(|z|^{-2}) 1_{\{|z|<r_0\}}\mathrm{d}z, \tag{1.5}$$

where $r_0 \in (0, \infty]$ is a constant and S is a Bernstein function with $S(0) = 0$. Let

$$c_0 = \int_{\{|z|\le \mathrm{e}^{-\|A\|}\}} (1 - \cos z_1)|z|^{-d}\mathrm{d}z,$$
$$\lambda_0 = \int_{\mathbb{R}^d} (r_0 \vee |z|)^{-d} S((r_0 \vee |z|)^{-2})\mathrm{d}z,$$

where z_1 stands for the first coordinate of z, and $\|A\|$ is the operator norm of A. We have $c_0 \in (0, \infty)$. Since $S(r) \le cr$ holds for some constant $c \in (0, \infty)$, we have $\lambda_0 < \infty$. In particular, if $r_0 = \infty$ then $\lambda_0 = 0$. We will estimate $\|\nabla P_t f\|_\infty$ by using the upper bound of A and the function

$$\alpha(t) := \int_0^\infty \frac{1}{\sqrt{r}} e^{-tS(r)} dr, \quad t > 0.$$

Obviously, if $\lim_{r\to\infty} \frac{S(r)}{\log r} = \infty$ then $\alpha(t) < \infty$ for all $t > 0$.

Theorem 1.1. *Let (1.5) hold and let $c_0, \lambda_0, \alpha(t)$ be defined above, let $\theta \in \mathbb{R}$ be such that $A \leq -\theta I$. Then there exists a constant $c_1 \in (0, \infty)$ depending only on d and θ such that*

$$\|\nabla P_t f\|_\infty \leq \|f\|_\infty c_1 e^{\lambda_0(t\wedge 1) - \theta^+ t}\left\{\alpha(c_0(t \wedge 1)) + \frac{(t\wedge 1)S(r_0^{-2})}{r_0}\right\} \tag{1.6}$$

holds for any $t > 0$ and $f \in \mathscr{B}_b(\mathbb{R}^d)$. If moreover $A = 0$, then there exists c_1 depending on d such that

$$\|\nabla P_t f\|_\infty \leq \|f\|_\infty e^{\lambda_0 t}\left\{\frac{1}{\sqrt{2\pi}}\alpha(c_0 t) + \frac{c_1(1 - e^{-t\lambda_0})S(r_0^{-2})}{r_0\lambda_0}\right\} \tag{1.7}$$

holds for any $t > 0$ and $f \in \mathscr{B}_b(\mathbb{R}^d)$, where $\lambda_0 = \frac{1-e^{-t\lambda_0}}{r_0\lambda_0} = 0$ for $r_0 = \infty$.

Now, we consider the gradient estimate for the semigroup associated to the linear SDE driven by a Lévy-type process. Let $\sigma(x, dy)$ be a signed kernel on $\mathbb{R}^d$, i.e. for each $x \in \mathbb{R}^d$, $\sigma(x, \cdot)$ is a signed measure while for each measurable set A, $\sigma(\cdot, A)$ is a measurable function. We call σ bounded if

$$\|\sigma\|_\infty := \sup_{x\in\mathbb{R}^d} |\sigma(x, \cdot)|(\mathbb{R}^d) < \infty.$$

Let $L_t^{+\sigma}$ be the Lévy-type process with jump measure

$$q(x, dz) := \nu(dz - x) + \sigma(x, dz)$$

for a bounded σ. In other words, there exist $b \in \mathbb{R}^d$ and non-negatively definite $d \times d$-matrix Q such that $L_t^{+\sigma}$ is generated by

$$\mathscr{L}^{+\sigma} f(x) = \mathscr{L} f(x) + \int_{\mathbb{R}^d} \{f(z) - f(x)\}\sigma(x, dz) =: \mathscr{L} f(x) + \sigma f(x) \tag{1.8}$$

for $f \in C_b^2(\mathbb{R}^d)$, where $\mathscr{L}$ is in (1.6). Let $P_t^{+\sigma}$ be the semigroup associated to the linear SDE

$$dX_t = AX_t dt + dL_t^{+\sigma}.$$

Combining Theorem 1.1 with a standard perturbation argument, we prove the following result on the gradient estimate of $P_t^{+\sigma}$.

Corollary 1.2. *If (1.5) holds for some S such that $\int_0^1 \alpha(t)dt < \infty$, then there exists a constant $c \in (0, \infty)$ such that*

$$\|\nabla P_t^{+\sigma} f\|_\infty \leq c\{\alpha(c_0(t \wedge 1)) + \|\sigma\|_\infty\}\|f\|_\infty, \quad t > 0, f \in \mathscr{B}_b(\mathbb{R}^d)$$

holds for any bounded σ.

To illustrate our results, we consider below two typical choices of S.

Example 1.3. (1) If $\nu(\mathrm{d}z) \geq c|z|^{-d-\alpha}1_{\{|z|\leq r_0\}}$ for some $c, r_0 > 0$ and $\alpha \in (0, 2)$, then

$$\|\nabla P_t f\|_\infty \leq \frac{c'}{(t \wedge 1)^{1/\alpha}} \mathrm{e}^{-\theta^+ t}\|f\|_\infty, \quad t > 0,\ f \in \mathscr{B}_b(\mathbb{R}^d)$$

holds for some constant $c' \in (0, \infty)$. If $\alpha \in (1, 2)$, then there exists a constant $c \in (0, \infty)$ such that

$$\|\nabla P_t^{+\sigma} f\|_\infty \leq c\|f\|_\infty \left\{\frac{1}{(t \wedge 1)^{1/\alpha}} + \|\sigma\|_\infty\right\}, \quad t > 0,\ f \in \mathscr{B}_b(\mathbb{R}^d)$$

holds for any bounded σ.

(2) If $\nu(\mathrm{d}z) \geq c|z|^{-d} \log^{1+\varepsilon}(1 + |z|^{-2})1_{\{|z|\leq r_0\}}$ for some $c, r_0, \varepsilon > 0$, then

$$\|\nabla P_t f\|_\infty \leq c_1\|f\|_\infty \exp[c_2 t^{-1/\varepsilon} - \theta^+ t], \quad t > 0,\ f \in \mathscr{B}_b(\mathbb{R}^d)$$

holds for some constants $c_1, c_2 \in (0, \infty)$.

Note that for the α-stable process one has (see Corollary 2.2(2) below for a more general result)

$$\sup_{\|f\|_\infty\leq 1} \|\nabla P_t f\|_\infty \geq \frac{c}{t^{1/\alpha}}$$

for some constant $c > 0$. Thus, the upper bound in Example 1.3(1) is sharp.

The main idea of the proof is to compare the process with the S-subordinate semigroup of the Brownian motion. To this end, we shall study in the next section the gradient estimate for subordinate semigroups. We will see that to compare the original semigroup with the subordinate semigroup, the error term is given by the conditional distribution of a compound Poisson process under the condition that the process jumps before time t. Thus, in Section 3 we will study the gradient estimate for the corresponding conditional distribution for compound Poisson processes. In this case, a derivative formula is presented. By combining results derived in Sections 2 and 3, we prove Theorem 1.1 in Section 4. Finally, the proofs of Corollary 1.2 and Example 1.3 are addressed in Section 5.

2. Gradient estimates for subordinate semigroups

This section is a counterpart of the recent work [7] where a dimension-free Harnack inequality is investigated for subordinate semigroups, see e.g. [14] and references within for potential theory and historical remarks on subordinations of the Brownian motion.

Let (E, ρ) be a Polish space. For a function f on E, define

$$|\nabla f|(x) := \limsup_{y\to x} \frac{|f(y) - f(x)|}{\rho(x, y)}, \quad x \in E.$$

Let P_t^0 be a (sub-)Markov semigroup on $\mathscr{B}_b(E)$ such that for some positive function φ on $(0, \infty)$,

$$|\nabla P_t^0 f| \leq \|f\|_\infty\varphi(t), \quad t > 0,\ f \in \mathscr{B}_b(E) \tag{2.1}$$

holds. We intend to estimate the gradient of a subordinate semigroup P_t^S of P_t^0 induced by a Bernstein function S. More precisely, for any $t \geq 0$ let μ_t^S be the probability measure on $[0, \infty)$

with the Laplace transformation

$$\int_0^\infty e^{-\lambda s}\mu_t^S(ds) = e^{-tS(\lambda)}, \quad \lambda \ge 0. \tag{2.2}$$

Then the S-subordination of P_t^0 is given by

$$P_t^S = \int_0^\infty P_s^0 \mu_t^S(ds), \quad t \ge 0. \tag{2.3}$$

The following assertion follows immediately from (2.3) and the dominated convergence theorem.

Theorem 2.1. *If (2.1) holds with $\int_0^\infty \varphi(s)\mu_t^S(ds) < \infty$, then*

$$|\nabla P_t^S f| \le \|f\|_\infty \int_0^\infty \varphi(s)\mu_t^S(ds), \quad f \in \mathscr{B}_b(E).$$

In particular, we have the following explicit gradient estimates by using known results on diffusion semigroups.

Corollary 2.2. (1) *Let E be a complete connected Riemannian manifold and P_t^0 be the diffusion semigroup generated by $\Delta + Z$ for a vector field Z on E such that*

$$Ric - \nabla Z \ge 0$$

holds. Then

$$\|\nabla P_t^S f\|_\infty \le \frac{\|f\|_\infty}{\sqrt{2\pi}} \int_0^\infty \frac{1}{\sqrt{r}} e^{-tS(r)} dr, \quad t > 0, \quad f \in \mathscr{B}_b(E).$$

(2) *Let P_t^0 be generated by Δ on $\mathbb{R}^d$. We have*

$$\sup_{\|f\|_\infty \le 1} \|\nabla P_t^S f\|_\infty \ge \frac{1}{\sqrt{2\pi}} \int_0^\infty \frac{1}{\sqrt{r}} e^{-tS(r)} dr.$$

Proof. (1) It is well-known that the curvature condition implies (cf. [2])

$$P_t^0 f^2 - (P_t^0 f)^2 \ge t|\nabla P_t^0 f|^2.$$

This implies that

$$\|\nabla P_t^0 f\|_\infty \le \frac{1}{\sqrt{t}} \|f\|_\infty.$$

Then the proof of (1) is finished by combining this with Theorem 2.1 and noting that

$$\begin{aligned}\int_0^\infty \frac{\mu_t^S(ds)}{\sqrt{s}} &= \int_0^\infty \frac{1}{\sqrt{2\pi}} \int_0^\infty \frac{1}{\sqrt{r}} e^{-rs} dr \mu_t^S(ds) \\ &= \frac{1}{\sqrt{2\pi}} \int_0^\infty \frac{1}{\sqrt{r}} \int_0^\infty e^{-rs}\mu_t^S(ds)\, dr = \frac{1}{\sqrt{2\pi}} \int_0^\infty r^{-1/2} e^{-tS(r)} dr.\end{aligned}$$

(2) Let P_t^0 be generated by Δ on $\mathbb{R}^d$. We have

$$P_s^0 f(x) = \frac{1}{(4\pi s)^{d/2}} \int_{\mathbb{R}^d} e^{-|x-y|^2/(4s)} f(y) dy.$$

Take

$$f(x) = 1_{[0,\infty)}(x_1) - 1_{(-\infty,0)}(x_1).$$

We have $\|f\|_\infty = 1$ and

$$\begin{aligned} P_s^0 f(x) &= \frac{1}{2\sqrt{\pi s}}\left\{\int_0^\infty \mathrm{e}^{-(r-x_1)^2/(4s)}\mathrm{d}r - \int_{-\infty}^0 \mathrm{e}^{-(r-x_1)^2/(4s)}\mathrm{d}r\right\} \\ &= \frac{1}{2\sqrt{\pi s}}\left\{\int_{-x_1}^\infty \mathrm{e}^{-r^2/(4s)}\mathrm{d}r - \int_{-\infty}^{-x_1} \mathrm{e}^{-r^2/(4s)}\mathrm{d}r\right\}. \end{aligned}$$

So,

$$\frac{\mathrm{d}}{\mathrm{d}x_1} P_s^0 f(x) = \frac{1}{\sqrt{\pi s}}\mathrm{e}^{-x_1^2/(4s)} \le \frac{1}{\sqrt{\pi s}}, \quad s > 0,\ x \in \mathbb{R}^d.$$

Combining this with (2.3) and using the dominated convergence theorem, we arrive at

$$\frac{\mathrm{d}}{\mathrm{d}x_1} P_t^S f(x)\bigg|_{x=0} = \frac{1}{\sqrt{\pi}}\int_0^\infty \frac{1}{\sqrt{s}}\mu_t^S(\mathrm{d}s) = \frac{1}{\sqrt{2}\pi}\int_0^\infty \frac{1}{\sqrt{r}}\mathrm{e}^{-tS(r)}\mathrm{d}r. \quad \square$$

3. A derivative formula

Let $\nu(\mathrm{d}z) \ge \rho_0(z)\mathrm{d}z =: \nu_0(\mathrm{d}z)$ for some non-negative measurable function ρ_0 on $\mathbb{R}^d$ such that

$$\lambda_0 := \int_{\mathbb{R}^d} \rho_0(z)\mathrm{d}z \in (0, \infty). \tag{3.1}$$

Let $(L_t^0)_{t\ge0}$ be the compound Poisson process with Lévy measure ν_0. Then L_t^0 can be realized as

$$L_t^0 = \sum_{i=1}^{N_t} \xi_i, \quad t \ge 0, \tag{3.2}$$

where N_t is the Poisson process with rate λ_0 and $\{\xi_i\}$ are i.i.d. random variables on $\mathbb{R}^d$ which are independent of $(N_t)_{t\ge0}$ and have common distribution ν_0/λ_0. Here, we set $\sum_{i=1}^0 \xi_i = 0$ by convention. Let $(L_t^1)_{t\ge0}$ be the Lévy process which is independent of $(L_t^0)_{t\ge0}$ and has Lévy measure $\nu - \nu_0$, such that

$$L_t := L_t^1 + L_t^0, \quad t \ge 0 \tag{3.3}$$

is the Lévy process with symbol η. As we explained in the Introduction, to ensure the strong Feller property for a jump process, it is essential to restrict on the event that the process jumps before a fixed time. Thus, instead of P_t, it is natural for us to investigate the gradient estimate for P_t^1 defined by

$$P_t^1 f(x) = \mathbb{E}\Big\{f(X_t^x)1_{\{N_t\ge1\}}\Big\}, \quad f \in \mathscr{B}_b(\mathbb{R}^d), t > 0,$$

where X_t^x solves (1.2) with initial data x. The following result provides a derivative formula for this operator, which can be regarded as the jump counterpart of the Bismut–Elworthy–Li formula for diffusion processes [3,6].

Theorem 3.1. *Let ρ_0 be non-negative and differentiable such that $\nu(\mathrm{d}z) \geq \rho_0(z)\mathrm{d}z$, $\lambda_0 := \int_{\mathbb{R}^d} \rho_0(z)\mathrm{d}z \in (0, \infty)$, and*

$$\int_{\mathbb{R}^d} \Big\{ \sup_{x:|x-z|\leq\varepsilon} |\nabla\rho_0|(x) \Big\} \mathrm{d}z < \infty \tag{3.4}$$

holds for some $\varepsilon > 0$. Then for any $t > 0$ and $f \in \mathscr{B}_b(\mathbb{R}^d)$,

$$\nabla P_t^1 f(x) = -\mathbb{E}\Big\{ f(X_t^x) 1_{\{N_t \geq 1\}} \frac{1}{N_t} \sum_{i=1}^{N_t} \mathrm{e}^{A^*\tau_i} \nabla \log \rho_0(\xi_i) \Big\}, \tag{3.5}$$

where τ_i is the i-th jump time of $(N_t)_{t\geq 0}$ and A^* is the transposition of A. Consequently, if $A \leq -\theta I$ then

$$\|\nabla P_t^1 f\|_\infty \leq \|f\|_\infty \frac{\mathrm{e}^{\theta^- t}(1 - \mathrm{e}^{-\lambda_0 t})}{\lambda_0} \int_{\mathbb{R}^d} |\nabla\rho_0|(z)\mathrm{d}z, \quad t > 0,\, f \in \mathscr{B}_b(\mathbb{R}^d).$$

Proof. We shall make use of a formula for random shifts of the compound Poisson process derived in [15]. Let $\Lambda(\mathrm{d}w)$ be the distribution of $L^0 := (L_t^0)_{t\geq 0}$ which is a probability measure on the path space

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

equipped with the σ-algebra induced by $\{w \mapsto w_t : t \geq 0\}$.

Let (τ, ξ) be a $[0, t] \times \mathbb{R}^d$-valued random variable such that the joint distribution of (L^0, τ, ξ) is

$$g(w, s, z)\Lambda(\mathrm{d}w)\mathrm{d}s\,\nu_0(\mathrm{d}z).$$

Let $\Delta w_t = w_t - w_{t-}$ and

$$U(w) = \sum_{\Delta w_t \neq 0} g(w - \Delta w_t 1_{[t,\infty)}, t, \Delta w_t).$$

By [15, Corollary 2.3], for any bounded measurable function F on the path space of L^0, one has

$$\mathbb{E}\big(F 1_{\{U>0\}}\big)(L^0) = \mathbb{E}\Big\{ \frac{F 1_{\{U>0\}}}{U} \Big\}(L^0 + \xi 1_{[\tau,\infty)}). \tag{3.6}$$

Now, let (τ, ξ) be independent of $(L_t^1, L_t^0)_{t\geq 0}$ with distribution

$$\frac{1}{t\lambda_0} 1_{[0,t]}(s)\mathrm{d}s\,\nu_0(\mathrm{d}z).$$

We have $g(w, s, z) = \frac{1}{\lambda_0 t} 1_{[0,t]}(s)$. Since τ is independent of L^0 so that with probability one τ $(\leq t)$ is not a jump time of L^0, and since $\xi \neq 0$ a.s., we have

$$U(L^0 + \xi 1_{[\tau,\infty)}) = \frac{N_t + 1}{\lambda_0 t}.$$

Since $Y_t := \int_0^t \mathrm{e}^{(t-s)A} \mathrm{d}L_s^1$ is independent of

$$\mathrm{e}^{At} x + \int_0^t \mathrm{e}^{A(t-s)} \mathrm{d}L_s^0,$$

it follows from (3.6) that for any $z_0 \in \mathbb{R}^d$ and $\varepsilon \in (-1, 1)$,

$$\begin{aligned} P_t^1 f(x+\varepsilon z_0) &= \mathbb{E}\bigg\{ f\bigg(Y_t + \mathrm{e}^{At}(x+\varepsilon z_0) + \int_0^t \mathrm{e}^{A(t-s)} \mathrm{d}L_s^0 \bigg) 1_{\{N_t\geq 1\}} \bigg\} \\ &= \lambda_0 t \mathbb{E}\bigg\{ \frac{f\big(Y_t + \mathrm{e}^{At}(x+\varepsilon z_0) + \int_0^t \mathrm{e}^{A(t-s)} \mathrm{d}\{L^0 + \xi 1_{[\tau,\infty)}\}_s\big)}{N_t+1} \bigg\} \\ &= \lambda_0 t \mathbb{E}\bigg\{ \frac{f\big(Y_t + \mathrm{e}^{At}x + \int_0^t \mathrm{e}^{A(t-s)} \mathrm{d}\{L^0 + (\xi + \varepsilon \mathrm{e}^{A\tau} z_0) 1_{[\tau,\infty)}\}_s\big)}{N_t+1} \bigg\}. \end{aligned} \tag{3.7}$$

On the other hand, since the joint distribution of $(L^0, \tau, \xi + \varepsilon \mathrm{e}^{A\tau} z_0)$ is

$$\frac{1}{\lambda_0 t} 1_{[0,t]}(s) \frac{\rho_0(z - \varepsilon \mathrm{e}^{As} z_0)}{\rho_0(z)} \Lambda(\mathrm{d}w)\mathrm{d}s\, \nu_0(\mathrm{d}z),$$

(3.6) holds for $\xi' := \xi + \varepsilon \mathrm{e}^{\tau A} z_0$ in place of ξ with

$$U(L^0) = \frac{1}{\lambda_0 t} \sum_{i=1}^{N_t} \frac{\rho_0(\xi_i - \varepsilon \mathrm{e}^{\tau_i A} z_0)}{\rho_0(\xi_i)}.$$

Consequently, for any $F \geq 0$, using FU in place of F in (3.6) one obtains

$$\mathbb{E}\big\{F(L^0)U(L^0)1_{\{N_t\geq 1\}}\big\} = \mathbb{E}F(L^0 + \xi' 1_{[\tau,\infty)}).$$

Taking $n_t(w) = \sum_{s\leq t} 1_{\{\Delta w_s \neq 0\}}$ and

$$F(w) = \frac{f\big(z + \int_0^t \mathrm{e}^{(t-s)A} \mathrm{d}w_s\big)}{n_t(w)} 1_{\{n_t(w)\geq 1\}}, \quad w \in W$$

for $z \in \mathbb{R}^d$, we arrive at

$$\begin{aligned} &\frac{1}{\lambda_0 t} \mathbb{E}\left\{ f\bigg(z + \int_0^t \mathrm{e}^{(t-s)A} \mathrm{d}L_s^0 \bigg) \frac{1_{\{N_t\geq 1\}}}{N_t} \sum_{i=1}^{N_t} \frac{\rho_0(\xi_i - \varepsilon \mathrm{e}^{A\tau_i} z_0)}{\rho_0(\xi_i)} \right\} \\ &\quad = \mathbb{E}\left\{ \frac{f\big(z + \int_0^t \mathrm{e}^{A(t-s)} \mathrm{d}\{L^0 + (\xi + \varepsilon \mathrm{e}^{A\tau} z_0)1_{[\tau,\infty)}\}_s\big)}{N_t+1} \right\}, \quad z \in \mathbb{R}^d. \end{aligned}$$

Combining this with (3.7), we obtain

$$P_t^1 f(x+\varepsilon z_0) = \mathbb{E}\left\{ f(X_t^x) 1_{\{N_t\geq 1\}} \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{\rho_0(\xi_i - \varepsilon \mathrm{e}^{A\tau_i} z_0)}{\rho_0(\xi_i)} \right\}.$$

Therefore, for any $\varepsilon \neq 0$ we have

$$\frac{P_t^1 f(x+\varepsilon z_0) - P_t^1 f(x)}{\varepsilon} = \mathbb{E}\left\{ f(X_t^x) 1_{\{N_t\geq 1\}} \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{\rho_0(\xi_i - \varepsilon \mathrm{e}^{A\tau_i} z_0) - \rho_0(\xi_i)}{\varepsilon \rho_0(\xi_i)} \right\}. \tag{3.8}$$

Noting that for $i \leq N_t$ one has $\tau_i \leq t$ so that $\mathrm{e}^{A\tau_i} z_0$ is bounded, and noting that for each i one has

$$\lim_{\varepsilon\downarrow 0} \frac{\rho_0(\xi_i - \varepsilon \mathrm{e}^{A\tau_i} z_0) - \rho_0(\xi_i)}{\varepsilon \rho_0(\xi_i)} = -\langle \mathrm{e}^{A\tau_i} z_0, \nabla \log \rho_0(\xi_i)\rangle = -\langle z_0, \mathrm{e}^{A^*\tau_i} \nabla \log \rho_0(\xi_i)\rangle,$$

by (3.4) we are able to use the dominated convergence theorem to derive (3.5) by letting $\varepsilon \to 0$ in (3.8). □

4. Proof of Theorem 1.1

4.1. Proof of (1.7) for $A = 0$

We shall first consider the case where $r_0 = \infty$ then pass to finite r_0 by using Theorem 3.1.
(I) For $r_0 = \infty$, i.e.

$$\nu(\mathrm{d}z) \geq |z|^{-d} S(|z|^{-2})\mathrm{d}z. \tag{4.1}$$

Then

$$\begin{aligned}\eta_1(u) &:= \int_{\mathbb{R}^d} \big(\mathrm{e}^{\mathrm{i}\langle u,z\rangle} - 1 - \mathrm{i}\langle u, z\rangle 1_{\{|z|<1\}}\big)|z|^{-d} S(|z|^{-2})\mathrm{d}z \\ \eta_2(u) &:= \eta(u) - \eta_1(u) \\ &= \mathrm{i}\langle u, b\rangle - \langle Qu, u\rangle + \int_{\mathbb{R}^d} \big(\mathrm{e}^{\mathrm{i}\langle u,z\rangle} - 1 - \mathrm{i}\langle u, z\rangle 1_{\{|z|<1\}}\big)\big\{\nu(\mathrm{d}z) - |z|^{-d} S(|u|^2)\mathrm{d}z\big\}\end{aligned}$$

provide two Lévy symbols. Noting that $S(|z|^{-2}) \geq 1_{\{|z|\leq|u|^{-1}\}} S(|u|^2)$ and

$$\begin{aligned}&-\int_{\{|z|\leq|u|^{-1}\}} \big(\mathrm{e}^{\mathrm{i}\langle u,z\rangle} - 1 - \mathrm{i}\langle u, z\rangle 1_{\{|z|<1\}}\big)|z|^{-d}\mathrm{d}z = \int_{\{|z|\leq|u|^{-1}\}} (1 - \cos\langle u, z\rangle)|z|^{-d}\mathrm{d}z \\ &= \int_{\{|z|\leq 1\}} \Big(1 - \cos\Big\langle \frac{u}{|u|}, z\Big\rangle\Big)|z|^{-d}\mathrm{d}z \\ &= \int_{\{|z|\leq 1\}} (1 - \cos z_1)|z|^{-d}\mathrm{d}z = c_0 \in (0, \infty),\end{aligned}$$

we see that

$$\begin{aligned}&u \mapsto \eta(u) + c_0 S(|u|^2) \\ &= \eta_2(u) + \int_{\mathbb{R}^d} \big(\mathrm{e}^{\mathrm{i}\langle u,z\rangle} - 1 - \mathrm{i}\langle u, z\rangle 1_{\{|z|<1\}}\big)|z|^{-d}\big\{S(|z|^{-2}) - S(|u|^2)1_{\{|z|\leq|u|^{-1}\}}\big\}\mathrm{d}z\end{aligned}$$

is also a Lévy symbol. Let P_t^S be the semigroup of the Lévy process with Lévy symbol $-c_0 S(|\cdot|^2)$, and let $\tilde{P}_t^S$ be the one with Lévy symbol $\eta + c_0 S(|\cdot|^2)$. We have

$$P_t = P_t^S \tilde{P}_t^S. \tag{4.2}$$

Since P_t^S is the $c_0 S$-subordination of the semigroup generated by Δ on $\mathbb{R}^d$, according to Corollary 2.2 for $E = \mathbb{R}^d$ and $Z = 0$,

$$\|\nabla P_t^S f\|_\infty \leq \|f\|_\infty \int_0^\infty \frac{1}{\sqrt{2\pi r}} \mathrm{e}^{-c_0 t S(r)}\mathrm{d}r = \frac{1}{\sqrt{2\pi}}\alpha(c_0 t)\|f\|_\infty. \tag{4.3}$$

Combining this with (4.2) we derive

$$\|\nabla P_t f\|_\infty \leq \frac{1}{\sqrt{2\pi}}\alpha(c_0 t)\|f\|_\infty. \tag{4.4}$$

Thus, the desired assertion holds if $r_0 = \infty$.

(II) For $r_0 \in (0, \infty)$. Take

$$\rho_0(z) = (r_0 \vee |z|)^{-d} S((r_0 \vee |z|)^{-2}).$$

Then

$$\bar{\nu}(\mathrm{d}z) := \nu(\mathrm{d}z) + \rho_0(z)\mathrm{d}z \geq |z|^{-d} S(|z|^{-2})\mathrm{d}z. \tag{4.5}$$

Let $\bar{L}_t^0$ be the compound Poisson process with Lévy measure $\rho_0(z)\mathrm{d}z$, and let

$$\bar{P}_t^1 f(x) = \mathbb{E}\big\{1_{\{\bar{\tau}_1 \leq t\}} f(x + \bar{L}_t^0)\big\},$$

where $\bar{\tau}_1$ is the first jump time of $\bar{L}_t^0$. Let L_t be the Lévy process with Lévy symbol η which is independent of $\bar{L}_t^0$. Then $\bar{L}_t := L_t + \bar{L}_t^0$ is the Lévy process with Lévy symbol

$$u \mapsto \eta(u) + \int_{\mathbb{R}^d} (\cos\langle u, z\rangle - 1)\rho_0(z)\mathrm{d}z.$$

Therefore,

$$\begin{aligned} \bar{P}_t f(x) &:= \mathbb{E} f(x + \bar{L}_t) \\ &= \mathbb{E}\big\{f(x + L_t)1_{\{\bar{\tau}_1 > t\}}\big\} + \mathbb{E}\big\{f(x + L_t + \bar{L}_t^0)1_{\{\bar{\tau}_1 \leq t\}}\big\} \\ &= \mathrm{e}^{-\lambda_0 t} P_t f(x) + \bar{P}_t^1 P_t f(x). \end{aligned}$$

This implies that

$$P_t f(x) = \mathrm{e}^{\lambda_0 t}\big(\bar{P}_t f - \bar{P}_t^1 P_t f\big)(x). \tag{4.6}$$

According to (4.5) and (I), (4.4) holds for $\bar{P}_t$ in place of P_t, i.e.

$$\|\nabla \bar{P}_t f\|_\infty \leq \frac{1}{\sqrt{2\pi}} \alpha(c_0 t)\|f\|_\infty. \tag{4.7}$$

On the other hand, we have

$$|\nabla \rho_0(z)| \leq 1_{\{|z| \geq r_0\}}\big\{d|z|^{-d-1} S(r_0^{-2}) + 2|z|^{-d-3} S'(|z|^{-2})\big\}.$$

Since S' is decreasing, S is increasing and $S(0) = 0$, from this we may find a constant c depending only on d such that

$$\begin{aligned} &\int_{\mathbb{R}^d} \Big\{\sup_{x:|x-z|<r_0/2} |\nabla \rho_0(x)|\Big\}\mathrm{d}z \leq c\int_{r_0}^\infty r^{-2}\big\{S(r_0^{-2}) + r^{-2} S'(r^{-2}/4)\big\}\mathrm{d}r \\ &\quad = c\int_{r_0}^\infty \Big\{\frac{S(r_0^{-2})}{r^2} - \frac{2}{r}\frac{\mathrm{d}}{\mathrm{d}r} S(r^{-2}/4)\Big\}\mathrm{d}r \leq \frac{c}{r_0} S(r_0^{-2}) + \frac{2c}{r_0} S(r_0^{-2}/4) \leq \frac{3c}{r_0} S(r_0^{-2}). \end{aligned}$$

Therefore, it follows from Theorem 3.1 with $\theta = 0$ that

$$\begin{aligned} \|\nabla \bar{P}_t^1 f\|_\infty &\leq \frac{3cS(r_0^{-2})(1 - \mathrm{e}^{-\lambda_0 t})}{r_0 \lambda_0}\|f\|_\infty \\ &\leq \frac{3cS(r_0^{-2})t}{r_0}\|f\|_\infty, \quad t > 0. \end{aligned} \tag{4.8}$$

Combining this with (4.6) and (4.7) we obtain the desired gradient estimate (1.7).

4.2. Proof of (1.6) for $A \neq 0$

(III) We first observe that it suffices to prove (1.6) for $t \in (0, 1]$. Assume that (1.6) holds for $t \in (0, 1]$. By the semigroup property we have

$$|\nabla P_t f| \le |\nabla P_{t\wedge 1}(P_{(t-1)^+} f)| \le c_1 \alpha(c_0(t \wedge 1))\|f\|_\infty, \quad t > 0$$

for some constant $c_0, c_1 \in (0, \infty)$. So, the desired inequality (1.6) holds for $\theta \le 0$. Next, since $A \le -\theta I$ implies that $|X_t^x - X_t^y| \le \mathrm{e}^{-\theta t}|x - y|$, we have

$$\begin{aligned}\frac{|P_t f(x) - P_t f(y)|}{|x-y|} &\le \frac{|\mathbb{E}P_1 f(X_{t-1}^x) - \mathbb{E}P_1 f(X_{t-1}^y)|}{|x-y|}\\ &\le \mathrm{e}^{-\theta(t-1)}\mathbb{E}\left\{\frac{|P_1 f(X_{t-1}^x) - P_1 f(X_{t-1}^y)|}{|X_{t-1}^x - X_{t-1}^y|}\right\}.\end{aligned}$$

Letting $y \to x$ and using the assertion for $t = 1$ and the dominated convergence theorem, we arrive at

$$|\nabla P_t f(x)| \le \mathrm{e}^{-\theta(t-1)}|\nabla P_1 f(X_{t-1}^x)| \le c_1 \mathrm{e}^{-\theta(t-1)}\alpha(c_0(t\wedge 1))\|f\|_\infty, \quad t > 1.$$

That is, (1.6) holds also for $t > 1$ with a different constant c_1.

(IV) For $r_0 = \infty$ and $t \in (0, 1]$. Let

$$\begin{aligned}\eta_1(u) &= \int_{\mathbb{R}^d}\big(\mathrm{e}^{\mathrm{i}\langle u,z\rangle} - 1 - \mathrm{i}\langle u, z\rangle 1_{\{|z|<1\}}\big)|z|^{-d}S(|z|^{-2})\mathrm{d}z\\ &= \int_{\mathbb{R}^d}\big(\cos\langle u, z\rangle - 1\big)|z|^{-d}S(|z|^{-2})\mathrm{d}z,\end{aligned}$$

and $\eta_2 = \eta - \eta_1$. By (4.1), both η_1 and η_2 are Lévy symbols. We have

$$\begin{aligned}\eta_1(\mathrm{e}^{sA^*}u) + c_0 S(|u|^2) &= \int_{\mathbb{R}^d}\big(\cos\langle z, \mathrm{e}^{sA^*}u\rangle - 1\big)|z|^{-d}S(|z|^{-2})\mathrm{d}z + c_0 S(|u|^2)\\ &= \int_{\mathbb{R}^d}\left(\cos\left\langle z, \frac{\mathrm{e}^{sA^*}u}{|\mathrm{e}^{sA^*}u|}\right\rangle - 1\right)|z|^{-d}S(|z|^{-2}|\mathrm{e}^{sA^*}u|^2)\mathrm{d}z + c_0 S(|u|^2)\\ &= \int_{\mathbb{R}^d}\big(\cos z_1 - 1\big)|z|^{-d}\big\{S(|z|^{-2}|\mathrm{e}^{sA^*}u|^2) - S(|u|^2)1_{\{|z|\le \mathrm{e}^{-\|A\|}\}}\big\}\mathrm{d}z\\ &= \int_{\mathbb{R}^d}\big(\mathrm{e}^{\mathrm{i}\langle u,z\rangle} - 1 - \mathrm{i}\langle u,z\rangle 1_{\{|z|<1\}}\big)|z|^{-d}\big\{S(|z|^{-2}|\mathrm{e}^{sA^*}u|^2) - S(|u|^2)1_{\{|z|\le\mathrm{e}^{-\|A\|}\}}\big\}\mathrm{d}z.\end{aligned}$$

Since for $s \in [0, 1]$

$$S(|z|^{-2}|\mathrm{e}^{sA^*}u|^2) \ge S(|u|^2)1_{\{|u|\le \mathrm{e}^{-\|A\|}\}},$$

this implies that

$$u \mapsto \eta_1(\mathrm{e}^{sA^*}u) + c_0 S(|u|^2)$$

is a Lévy symbol. In particular, there exists a probability measure π_t on $\mathbb{R}^d$ with log-characteristic function

$$\log\hat{\pi}_t(u) = \int_0^t \eta(\mathrm{e}^{sA^*}u)\mathrm{d}s + tc_0 S(|u|^2)$$

$$= \int_0^t \eta_2(e^{sA^*}u)ds + \int_0^t \{\eta_1(e^{sA^*}u)ds + c_0 S(|u^2|)\}ds.$$

Now, letting P_t^S be the semigroup for the Lévy process with Lévy symbol $-c_0 S(|\cdot|^2)$, and letting

$$\tilde{P}_t f(x) = \int_{\mathbb{R}^d} f(x+z)\pi_t(dz),$$

we obtain from (1.3), (1.4) and the definition of π_t that

$$P_t f(x) = P_t^S \tilde{P}_t f(e^{tA}x).$$

Combining this with (4.3) we obtain

$$\|\nabla P_t f\|_\infty \le \|f\|_\infty \alpha(c_0 t).$$

(V) For $t \in (0,1]$ and $r_0 \in (0,\infty)$. Let ρ_0, $\bar{L}_t^0$ and $\bar{L}_t$ be in (II). Let

$$\bar{P}_t^1 f(x) = \mathbb{E}\left\{ f\left(e^{tA}x + \int_0^t e^{(t-s)A} d\bar{L}_s^0\right) 1_{\{\bar{\tau}_1 \le t\}}\right\},$$

$$\bar{P}_t f(x) = \mathbb{E} f\left(e^{tA}x + \int_0^t e^{(t-s)A} d\bar{L}_s\right).$$

Then (4.6) holds. Since (4.1) holds for $\bar{\nu}$ in place of ν, according to (IV) and the argument leading to (4.8) using Theorem 3.1, there exists a constant $c \in (0,\infty)$ depending only on d and θ such that

$$\|\nabla \bar{P}_t\|_\infty \le \|f\|_\infty \alpha(c_0 t), \qquad \|\nabla \bar{P}_t^1 f\|_\infty \le \frac{cS(r_0^{-2})t}{r_0}\|f\|_\infty.$$

Combining this with (4.6) we derive the desired gradient estimate (1.6).

5. Proofs of Corollary 1.2 and Example 1.3

Proof of Corollary 1.2. Since the gradient estimate $\|\nabla P_t^{+\sigma} f\|_\infty \le c(t)\|f\|_\infty$ is equivalent to

$$|P_t^{+\sigma} f(x) - P_t^{+\sigma} f(y)| \le c(t)\|f\|_\infty |x-y|, \quad x, y \in \mathbb{R}^d,$$

by the monotone class theorem it suffices to prove for $f \in C_b^2(\mathbb{R}^d)$. By (1.8), in this case we have

$$\frac{d}{ds} P_s P_{t-s}^{+\sigma} f = P_s(\mathscr{L} - \mathscr{L}^{+\sigma}) P_{t-s}^{+\sigma} f = -P_s(\sigma P_{t-s}^{+\sigma} f), \quad s \in [0,t].$$

Consequently,

$$P_t^{+\sigma} f = P_t f + \int_0^t P_s(\sigma P_{t-s}^{+\sigma} f)ds.$$

Combining this with Theorem 1.1, we finish the proof. □

Proof of Example 1.3. (1) follows immediately from Theorem 1.1 and Corollary 1.2 by taking $S(r) = cr^{\alpha/2}$. To prove (2), we take

$$S_\varepsilon(r) = \log^{1+\varepsilon}(1 + r^{1/(1+\varepsilon)}).$$

According to [13], for any Bernstein function S and any $\delta > 1$, $r \mapsto S^\delta(r^{1/\delta})$ is again a Bernstein function. In this case we have

$$\nu(\mathrm{d}z) \geq c1_{\{|z|\leq r_0\wedge 1\}}|z|^{-d}S_\varepsilon(|z|^{-2})\mathrm{d}z.$$

Then the desired gradient estimate follows immediately from Theorem 1.1. □

Acknowledgement

The author would like to thank Dr. Jian Wang for a number of corrections and comments on earlier versions of the paper, and to thank Professors M. Röckner, R. Schilling, R. Song, and L. Wu for useful conversations. He would also like to thank the referees for careful comments and corrections.

References

[1] D. Applebaum, Lévy Processes and Stochastic Calculus, Cambridge University Press, 2004.
[2] D. Bakry, On Sobolev and logarithmic Sobolev inequalities for Markov semigroups, in: K.D. Elworthy, S. Kusuoka, I. Shigekawa (Eds.), New Trends in Stochastic Analysis, World Scientific, Singapore, 1997.
[3] J.M. Bismut, Large Deviations and the Malliavin Calculus, Birkhäuser, MA, Boston, 1984.
[4] V.I. Bogachev, M. Röckner, B. Schmuland, Generalized Mehler semigroups and applications, Probab. Theory Related Fields 105 (1996) 193–225.
[5] K. Bogdan, T. Jakubowski, Estimate of heat kernel of fractional Laplacian perturbed by gradient operators, Comm. Math. Phys. 271 (2007) 179–198.
[6] K.D. Elworthy, X.-M. Li, Formulae for the derivatives of heat semigroups, J. Funct. Anal. 125 (1994) 252–286.
[7] M. Gordina, M. Röckner, F.-Y. Wang, Harnack inequalities for subordinate semigroups, Potential Analysis (in press). arXiv:1004.3016.
[8] N. Jacob, Pseudo Differential Operators and Markov Processes (Volume I), Imperial College Press, London, 2001.
[9] T. Jakubowski, K. Szczypkowski, Time-dependent gradient perturbations of fractional Laplacian, J. Evol. Equ. 10 (2010) 319–339.
[10] V. Knopova, R.L. Schilling, A note on the existence of transition probability densities for Lévy processes, arXiv:1003.1419.
[11] V. Knopova, R.L. Schilling, Transition density estimates for a class of Lévy and Lévy-type processes, arXiv:0912.1482v1.
[12] E. Priola, J. Zabczyk, Densities for Ornstein–Uhlenbeck processes with jumps, Bull. Lond. Math. Soc. 41 (2009) 41–50.
[13] L.R. Schilling, Subordination in the sense of Bochner and a related functional calculus, J. Austral. Math. Soc. Ser. A 64 (1998) 368–396.
[14] R. Song, Z. Vondraček, Potential theory of subordinate Brownian motion, in: Potential Analysis of Stable Processes and its Extensions, in: Lecture Notes in Math., vol. 1980, Springer, Berlin, 2009, pp. 87–176.
[15] F.-Y. Wang, Coupling for Ornstein–Uhlenbeck Processes with Jumps, Bernoulli (in press). arXiv:1002.2890.