

Coupling property and gradient estimates of Lévy processes via the symbol

RENÉ L. SCHILLING¹, PAWEŁ SZTONYK² and JIAN WANG^{1,3}

¹*TU Dresden, Institut für Mathematische Stochastik, 01062 Dresden, Germany.*

E-mail: rene.schilling@tu-dresden.de

²*Institute of Mathematics and Computer Science, Wrocław University of Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland. E-mail: sztonyk@pwr.wroc.pl*

³*School of Mathematics and Computer Science, Fujian Normal University, 350007 Fuzhou, P.R. China. E-mail: jianwang@fjnu.edu.cn*

We derive explicitly the coupling property for the transition semigroup of a Lévy process and gradient estimates for the associated semigroup of transition operators. This is based on the asymptotic behaviour of the symbol or the characteristic exponent near zero and infinity, respectively. Our results can be applied to a large class of Lévy processes, including stable Lévy processes, layered stable processes, tempered stable processes and relativistic stable processes.

Keywords: coupling; gradient estimates; Lévy process; symbol

1. Introduction and main results

Let X_t be a pure jump Lévy process on \mathbb{R}^d with the symbol (or characteristic exponent)

$$\Phi(\xi) = \int_{z \neq 0} (1 - e^{i\xi \cdot z} + i\xi \cdot z \mathbb{1}_{B(0,1)}(z)) \nu(dz),$$

where ν is the Lévy measure, that is, a σ -finite measure on $\mathbb{R}^d \setminus \{0\}$ such that the integral $\int_{z \neq 0} (1 \wedge |z|^2) \nu(dz) < \infty$. There are many papers studying regularity properties of Lévy processes in terms of the symbol Φ . For example, recently [15], Theorem 1, points out the relations between the classic Hartman–Wintner condition (see [9] or (1.1) below) and some smoothness properties of the transition density for Lévy processes. In particular, the condition that the symbol $\Phi(\xi)$ of the Lévy process X_t satisfies

$$\liminf_{|\xi| \rightarrow \infty} \frac{\operatorname{Re} \Phi(\xi)}{\log(1 + |\xi|)} = \infty \tag{1.1}$$

is *equivalent* to the statement that for all $t > 0$ the random variables X_t have a transition density $p_t(y)$ such that $\nabla p_t \in L_1(\mathbb{R}^d) \cap C_\infty(\mathbb{R}^d)$, where $C_\infty(\mathbb{R}^d)$ denotes the set of all continuous functions which vanish at infinity. The main purpose of this paper is to derive an explicit coupling property and gradient estimates of Lévy processes directly from the corresponding symbol Φ .

Let $(X_t)_{t \geq 0}$ be a Markov process on \mathbb{R}^d with transition probability function $\{P_t(x, \cdot)\}_{t \geq 0, x \in \mathbb{R}^d}$. An \mathbb{R}^{2d} -valued process $(X'_t, X''_t)_{t \geq 0}$ is called a *coupling of the Markov process* $(X_t)_{t \geq 0}$, if both

$(X'_t)_{t \geq 0}$ and $(X''_t)_{t \geq 0}$ are Markov processes which have the same transition functions $P_t(x, \cdot)$ but possibly different initial distributions. In this case, $(X'_t)_{t \geq 0}$ and $(X''_t)_{t \geq 0}$ are called the *marginal processes* of the coupling process; the *coupling time* is defined by $T := \inf\{t \geq 0: X'_t = X''_t\}$. The coupling $(X'_t, X''_t)_{t \geq 0}$ is said to be *successful* if T is a.s. finite. If for any two initial distributions μ_1 and μ_2 , there exists a successful coupling with marginal processes starting from μ_1 and μ_2 , respectively, we say that X_t has the *coupling property* (or admits *successful couplings*). According to [19] and the proof of [27], Theorem 4.1, the coupling property is equivalent to the statement that

$$\lim_{t \rightarrow \infty} \|P_t(x, \cdot) - P_t(y, \cdot)\|_{\text{Var}} = 0 \quad \text{for any } x, y \in \mathbb{R}^d,$$

where $P_t(x, \cdot)$ is the transition function of the Markov process $(X_t)_{t \geq 0}$. By $\|\mu\|_{\text{Var}}$ we denote the total variation norm of the signed measure μ . We know from [27], Theorem 4.1, that every Lévy process has the coupling property if the transition functions have densities for all sufficiently large $t > 0$. In this case, the transition probability function satisfies

$$\|P_t(x, \cdot) - P_t(y, \cdot)\|_{\text{Var}} \leq \frac{C(1 + |x - y|)}{\sqrt{t}} \wedge 2 \quad \text{for } t > 0 \text{ and } x, y \in \mathbb{R}^d. \tag{1.2}$$

It is clear that for any $x, y \in \mathbb{R}^d$ and $t \geq 0$, $\|P_t(x, \cdot) - P_t(y, \cdot)\|_{\text{Var}} \leq 2$, and that the norm $\|P_t(x, \cdot) - P_t(y, \cdot)\|_{\text{Var}}$ is decreasing with respect to t . This shows that it is enough to estimate $\|P_t(x, \cdot) - P_t(y, \cdot)\|_{\text{Var}}$ for large values of t . We will call any estimate for $\|P_t(x, \cdot) - P_t(y, \cdot)\|_{\text{Var}}$ an *estimate of the coupling time*. The rate $1/\sqrt{t}$ in (1.2) is not optimal for general Lévy processes which admit successful couplings. For example, for rotationally invariant α -stable Lévy processes we can prove, see [2], Example 2.3, that

$$\|P_t(x, \cdot) - P_t(y, \cdot)\|_{\text{Var}} \asymp \frac{1}{t^{1/\alpha}} \quad \text{as } t \rightarrow \infty,$$

where for any two non-negative functions g and h , the notation $g \asymp h$ means that there are two positive constants c_1 and c_2 such that $c_1 g \leq h \leq c_2 g$.

Let $P_t(x, \cdot)$ and P_t be the transition function and the semigroup of the Lévy process X_t , respectively. We begin with coupling time estimates of Lévy processes which satisfy the following Hartman–Wintner condition for some $t_0 > 0$:

$$\liminf_{|\xi| \rightarrow \infty} \frac{\text{Re } \Phi(\xi)}{\log(1 + |\xi|)} > \frac{d}{t_0}; \tag{1.3}$$

this condition actually ensures that the transition function of the Lévy process X_t is, for all $t > t_0$, absolutely continuous, see, for example, [9] or [15]. Note that (1.3) becomes (1.1) if $t_0 \rightarrow 0$.

Theorem 1.1. *Suppose that (1.3) holds and*

$$\text{Re } \Phi(\xi) \asymp f(|\xi|) \quad \text{as } |\xi| \rightarrow 0,$$

where $f : [0, \infty) \rightarrow \mathbb{R}$ is a strictly increasing function which is differentiable near zero and which satisfies

$$\liminf_{r \rightarrow 0} f(r) |\log r| < \infty$$

and

$$\limsup_{s \rightarrow 0} f^{-1}(2s)/f^{-1}(s) < \infty.$$

Then the corresponding Lévy process X_t has the coupling property, and there exist two constants $c, t_1 > 0$ such that for any $x, y \in \mathbb{R}^d$ and $t \geq t_1$,

$$\|P_t(x, \cdot) - P_t(y, \cdot)\|_{\text{var}} \leq cf^{-1}(1/t).$$

It can be seen from the above remark on rotationally invariant α -stable Lévy processes that the estimate in Theorem 1.1 is sharp.

Remark 1.2. (1) In our earlier paper [2], in particular [2], Theorem 1.1 and (1.3), we showed that the following condition on the Lévy measure ν ensures that a (pure jump) Lévy process admits a successful coupling:

$$\nu(dz) \geq |z|^{-d} g(|z|^{-2}) dz \tag{1.4}$$

for some Bernstein function g . In the present paper, we use a different condition in terms of the characteristic exponent $\Phi(\xi)$. Let us briefly compare [2], Theorem 1.1, and Theorem 1.1. If (1.4) holds, then we know that

$$\Phi(\xi) = \Phi_\rho(\xi) + \Phi_\mu(\xi),$$

where Φ_ρ and Φ_μ denote the (pure-jump) characteristic exponents with Lévy measures $\rho(dz) = |z|^{-d} g(|z|^{-2}) dz$ and $\mu = \nu - \rho$, respectively. Note that (1.4) guarantees that μ is a nonnegative measure. By [13], Lemma 2.1, and some tedious, but otherwise routine, calculations one can see that $\Phi_\rho(\xi) \asymp g(|\xi|^2)$ as $|\xi| \rightarrow 0$.

If g satisfies [2], (2.10) and (2.11), – these conditions coincide with the asymptotic properties required of f in Theorem 1.1 –, we can apply Theorem 1.1 to the symbol $\Phi_\rho(\xi)$ with $f(s) = g(s^2)$, and follow the argument of [2], Proposition 2.9 and Remark 2.10, to get a new proof of [2], Theorem 1.1. Note that this argument uses the fact that, we can (in law) decompose the Lévy process with exponent $\Phi(\xi)$ into two independent Lévy processes with characteristic exponents $\Phi_\rho(\xi)$ and $\Phi_\mu(\xi)$, respectively.

(2) The considerations from (1) can be adapted to show that we may replace the two-sided estimate $\text{Re } \Phi(\xi) \asymp f(|\xi|)$ in Theorem 1.1 by $\text{Re } \Phi(\xi) \geq cf(|\xi|)$; this, however, requires that we know in advance that $\Phi(\xi) - cf(|\xi|)$ is a characteristic exponent of some Lévy process. While this was obvious under (1.4) and for the difference of two Lévy measures being again a nonnegative measure, there are no good conditions in general when the difference of two characteristic exponents is again an characteristic exponent of some Lévy process.

(3) The present result, Theorem 1.1, trivially applies to most subordinate stable Lévy processes: here the characteristic exponent is of the form $f(|\xi|^\alpha)$, $0 < \alpha \leq 2$, but the corresponding

Lévy measures cannot be given in closed form. In Example 1.5 below, we have a situation where the Lévy measure is known. Nevertheless, the methods of [2] are only applicable in *one* particular case, while Theorem 1.1 applies to *all* non-degenerate settings.

Now we turn to explicit gradient estimates for the semigroup of a Lévy process. For a function $u \in B_b(\mathbb{R}^d)$ we define

$$|\nabla u(x)| := \limsup_{y \rightarrow x} \frac{|u(y) - u(x)|}{|y - x|}, \quad x \in \mathbb{R}^d.$$

If u is differentiable at x , then $|\nabla u(x)|$ is just the norm of the gradient of u at x . We are interested in sub-Markov semigroups P_t on $B_b(\mathbb{R}^d)$ which satisfy that for some positive function ϕ on $(0, \infty)$

$$\|\nabla P_t u\|_\infty \leq \|u\|_\infty \phi(t), \quad t > 0, u \in B_b(\mathbb{R}^d).$$

Similar uniform gradient estimates for Markov semigroups have attracted a lot of attention in analysis and probability, for example, see [21] and references therein. Because of the Markov property of the semigroup P_t , $\phi(t)$ is decreasing with respect to t . Thus, it is enough to obtain sharp estimates for $\phi(t)$ both as $t \rightarrow 0$ and $t \rightarrow \infty$. For Lévy processes, we have the following theorem.

Theorem 1.3. *Assume that (1.1) holds. If there is a strictly increasing function f which is differentiable near infinity and which satisfies*

$$\limsup_{s \rightarrow \infty} f^{-1}(2s)/f^{-1}(s) < \infty,$$

and

$$\operatorname{Re} \Phi(\xi) \asymp f(|\xi|) \quad \text{as } |\xi| \rightarrow \infty,$$

then there exists a constant $c > 0$ such that for $t > 0$ small enough,

$$\|\nabla P_t u\|_\infty \leq c \|u\|_\infty f^{-1}(1/t), \quad u \in B_b(\mathbb{R}^d). \tag{1.5}$$

Similarly, let f be a strictly increasing function which is differentiable near zero and which satisfies

$$\liminf_{r \rightarrow 0} f(r)|\log r| < \infty, \quad \limsup_{s \rightarrow 0} f^{-1}(2s)/f^{-1}(s) < \infty$$

and

$$\operatorname{Re} \Phi(\xi) \asymp f(|\xi|) \quad \text{as } |\xi| \rightarrow 0.$$

Then there exists a constant $c > 0$ such that (1.5) holds for $t > 0$ large enough.

We will see in Remark 3.3 below that Theorem 1.3 is also sharp for rotationally invariant α -stable Lévy processes. Roughly speaking, Theorems 1.1 and 1.3 show that the gradient estimate (1.5) for a Lévy process for small $t \ll 1$ depends on the asymptotic behaviour of the symbol

Φ near infinity, while (1.5) for large $t \gg 1$ relies on the asymptotic behaviour of the symbol Φ near zero. This situation is familiar from estimates of the coupling time of Lévy processes. More details can be found in the examples given below.

In order to illustrate the power of Theorems 1.1 and 1.3, we present two examples.

Example 1.4. Let X_t be a subordinate Brownian motion with symbol $f(|\xi|^2)$, where $f(\lambda) = \lambda^{\alpha/2}(\log(1 + \lambda))^{\beta/2}$, $\alpha \in (0, 2)$ and $\beta \in (-\alpha, 2 - \alpha)$. To see that f is indeed a Bernstein function we observe that λ , $\log(1 + \lambda)$ and $\lambda/\log(1 + \lambda)$ are complete Bernstein functions, and that for $\alpha, \beta \geq 0$

$$\lambda^{\alpha/2} \cdot (\log(1 + \lambda))^{\beta/2} \quad \text{is a complete Bernstein function if } \frac{\alpha}{2} + \frac{\beta}{2} \leq 1,$$

while for $-\alpha \leq \beta \leq 0 \leq \alpha$

$$\lambda^{(\alpha-\beta)/2} \cdot \left(\frac{\lambda}{\log(1 + \lambda)}\right)^{\beta/2} \quad \text{is a complete Bernstein function if } \frac{\alpha}{2} + \frac{\beta}{2} \leq 1.$$

This follows easily from [26], (Proof of) Proposition 7.10, see also [28], Examples 5.15, 5.16.

There are two constants $c_1, t_0 > 0$ such that for all $x, y \in \mathbb{R}^d$ and $t \geq t_0$,

$$\|P_t(x, \cdot) - P_t(y, \cdot)\|_{\text{Var}} \leq c_1 t^{-1/(\alpha+\beta)},$$

and there exists a constant $c_2 > 0$ such that for all $u \in B_b(\mathbb{R}^d)$,

$$\|\nabla P_t u\|_{\infty} \leq \begin{cases} c_2 [t^{-1}(\log(1 + t^{-1}))^{-\beta/2}]^{1/\alpha} \|u\|_{\infty} & \text{for small } t \ll 1; \\ c_2 t^{-1/(\alpha+\beta)} \|u\|_{\infty} & \text{for large } t \gg 1. \end{cases}$$

Example 1.5. Let μ be a finite nonnegative measure on the unit sphere \mathbb{S} and assume that μ is nondegenerate in the sense that its support is not contained in any proper linear subspace of \mathbb{R}^d . Let $\alpha \in (0, 2)$, $\beta \in (0, \infty]$ and assume that the Lévy measure ν satisfies that for some constant $r_0 > 0$ and any $A \in \mathcal{B}(\mathbb{R}^d)$,

$$\nu(A) \geq \int_0^{r_0} \int_{\mathbb{S}} \mathbb{1}_A(s\theta) s^{-1-\alpha} \, ds \mu(d\theta) + \int_{r_0}^{\infty} \int_{\mathbb{S}} \mathbb{1}_A(s\theta) s^{-1-\beta} \, ds \mu(d\theta).$$

Then, by Theorem 1.1, there are two constants $c_1, t_0 > 0$ such that for all $x, y \in \mathbb{R}^d$ and $t \geq t_0$,

$$\|P_t(x, \cdot) - P_t(y, \cdot)\|_{\text{Var}} \leq c_1 t^{-1/(\beta \wedge 2)}.$$

In the present situation, the methods of [2] only apply if μ is (essentially) the uniform measure on \mathbb{S} ; this is not the case for the condition of Theorem 1.1.

Moreover, Theorem 1.3 shows that there exists a constant $c_2 > 0$ such that for all $u \in B_b(\mathbb{R}^d)$,

$$\|\nabla P_t u\|_{\infty} \leq \begin{cases} c_2 \|u\|_{\infty} t^{-1/\alpha} & \text{for small } t \ll 1; \\ c_2 \|u\|_{\infty} t^{-1/(\beta \wedge 2)} & \text{for large } t \gg 1. \end{cases}$$

Coupling techniques for Lévy-driven SDEs and Lévy-type processes have been considered in the literature before, see, for example, [16–18,32]. As far as we know, however, only the papers [30,31] by F.-Y. Wang deal with couplings of Lévy-driven SDEs. Wang shows the existence of successful couplings and gradient estimates for Ornstein–Uhlenbeck processes driven by Lévy processes.

The remaining part of this paper is organized as follows. In Section 2, we first present estimates for the derivatives of the density for infinitely divisible distributions in terms of the corresponding Lévy measure; this part is of some interest on its own. Then we use these estimates to investigate derivatives of the density for Lévy processes, whose Lévy measures have (modified) bounded support. In Section 3, we give the proofs of all the theorems and examples stated in Section 1, by using the results of Section 2. Some remarks and examples are also included here to illustrate the optimality and the efficiency of Theorems 1.1 and 1.3.

2. Derivatives of densities for infinitely divisible distributions

Let π be an infinitely divisible distribution. It is well known that its characteristic function $\widehat{\pi}(\xi) := \int_{\mathbb{R}^d} e^{i\xi \cdot y} \pi(dy)$ is of the form $\exp(-\Phi(\xi))$, where

$$\Phi(\xi) = \int_{y \neq 0} (1 - e^{i\xi \cdot y} + i\xi \cdot y \mathbb{1}_{B(0,1)}(y)) \nu(dy),$$

and ν is a Lévy measure on $\mathbb{R}^d \setminus \{0\}$ such that $\int_{y \neq 0} (1 \wedge |y|^2) \nu(dy) < \infty$. In this section, we first aim to study estimates for derivatives of the density of π . As usual, we denote for every $n \in \mathbb{N}_0$ by $C_b^n(\mathbb{R}^d)$ the set of all n -times continuously differentiable functions on \mathbb{R}^d which are, together with all their derivatives, bounded; for $n = 0$ we use the convention that $C_b^0(\mathbb{R}^d) = C_b(\mathbb{R}^d)$ denotes the set of continuous and bounded functions on \mathbb{R}^d .

Proposition 2.1. *If for some $n, m \in \mathbb{N}_0$,*

$$\int e^{-\operatorname{Re} \Phi(\xi)} (1 + |\xi|)^{n+m} d\xi < \infty, \quad (2.6)$$

and

$$\int_{|y|>1} |y|^{2\nu n} \nu(dy) < \infty, \quad (2.7)$$

then π has a density $p \in C_b^{m+n}(\mathbb{R}^d)$ such that for every $\beta \in \mathbb{N}_0^d$ with $|\beta| \leq m$,

$$|\partial^\beta p(y)| \leq \psi(n, m, \nu) (1 + |y|)^{-n}, \quad y \in \mathbb{R}^d,$$

where for $n \geq 0$

$$\psi(n, m, \nu) = C(n, d) \left(1 + \int (|y|^2 + |y|^{2\nu n}) \nu(dy) \right)^n \int e^{-\operatorname{Re} \Phi(\xi)} (1 + |\xi|)^{n+m} d\xi.$$

Proof. The existence of the density $p \in C_b^{m+n}(\mathbb{R}^d)$ is a consequence of (2.6) and [23], Proposition 28.1, or [20], Proposition 0.2.

To prove the second assertion, we recall some necessary facts and notations. Given a function $f \in L^1(\mathbb{R}^d)$, its Fourier transform is given by

$$\widehat{f}(\xi) = \int_{\mathbb{R}^d} f(y)e^{i\xi \cdot y} dy.$$

For $\xi \in \mathbb{R}^d$ and a multiindex $\beta = (\beta_1, \beta_2, \dots, \beta_d) \in \mathbb{N}_0^d$, we set $M_\beta(\xi) := \xi^\beta = \xi_1^{\beta_1} \xi_2^{\beta_2} \dots \xi_d^{\beta_d}$. If $\widehat{f} \in C^N(\mathbb{R}^d)$ and $\partial^\gamma(M_\beta \widehat{f}) \in L^1(\mathbb{R}^d)$ for $N \in \mathbb{N}_0$ and every $\gamma \in \mathbb{N}_0^d$ such that $|\gamma| \leq N$, then, using the inverse Fourier transform and the integration by parts formula, we obtain that for every $\delta \in \mathbb{N}_0^d$ with $|\delta| \leq N$

$$y^\delta \partial^\beta f(y) = (2\pi)^{-d} (-1)^{|\beta|} (i)^{|\beta| - |\delta|} \int \partial^\delta [M_\beta \widehat{f}](\xi) e^{-iy \cdot \xi} d\xi.$$

This yields

$$|y^\delta \partial^\beta f(y)| \leq (2\pi)^{-d} \int |\partial^\delta [M_\beta \widehat{f}](\xi)| d\xi. \tag{2.8}$$

In particular, for every $n \in \mathbb{N}_0$,

$$|y_k|^n |\partial^\beta f(y)| \leq (2\pi)^{-d} \int \left| \frac{\partial^n}{\partial \xi_k^n} [M_\beta \widehat{f}](\xi) \right| d\xi. \tag{2.9}$$

For $n = 0$, the required assertion immediately follows from (2.8) if we use $f = p$ and $\delta = 0$. If $n > 0$, then for every $\beta \in \mathbb{N}_0^d$ such that $|\beta| = 1$ we have

$$\partial^\beta \Phi(\xi) = -i \int y^\beta (e^{i\xi \cdot y} - \mathbb{1}_{B(0,1)}(y)) \nu(dy).$$

By the Hölder inequality,

$$\begin{aligned} & |\partial^\beta \Phi(\xi)| \\ & \leq \left[\int |y|^2 \nu(dy) \right]^{1/2} \left[2 \int_{B(0,1)} (1 - \cos(\xi \cdot y)) \nu(dy) + \nu(B(0,1)^c) \right]^{1/2} \\ & \leq \left[\int |y|^2 \nu(dy) \right]^{1/2} \left[|\xi|^2 \int_{B(0,1)} |y|^2 \nu(dy) + \nu(B(0,1)^c) \right]^{1/2} \\ & \leq \int |y|^2 \nu(dy) \cdot (|\xi|^2 + 1)^{1/2} \\ & \leq (1 + |\xi|) \int |y|^2 \nu(dy). \end{aligned} \tag{2.10}$$

On the other hand, for $1 < |\beta| \leq n$, we have

$$\partial^\beta \Phi(\xi) = -(\text{i})^{|\beta|} \int y^\beta e^{i\xi \cdot y} \nu(\text{d}y),$$

and so

$$|\partial^\beta \Phi(\xi)| \leq \int |y|^{|\beta|} \nu(\text{d}y). \tag{2.11}$$

For symmetric Lévy measures ν similar estimates are due to Hoh [10], see also [12], Theorem 3.7.13.

Let $k \in \{1, \dots, d\}$ and $M \in \mathbb{N}$ with $M \leq n$. We use Faa di Bruno’s formula, see [6], to obtain

$$\frac{\partial^M}{\partial \xi_k^M} \widehat{p}(\xi) = M! \exp(-\Phi(\xi)) \sum_{j=1}^M \sum_{u(M,j)} \prod_{l=1}^M \left(\frac{\partial^l (-\Phi)}{\partial \xi_k^l}(\xi) \right)^{\lambda_l} / ((\lambda_l!)(l!)^{\lambda_l}),$$

where

$$u(M, j) = \left\{ (\lambda_1, \dots, \lambda_M) : \lambda_l \in \mathbb{N}_0, \sum_{l=1}^M \lambda_l = j, \sum_{l=1}^M l\lambda_l = M \right\}.$$

This, (2.10) and (2.11) yield

$$\begin{aligned} \left| \frac{\partial^M}{\partial \xi_k^M} \widehat{p}(\xi) \right| &\leq |\exp(-\Phi(\xi))| \sum_{j=1}^M \sum_{u(M,j)} \prod_{l=1}^M \frac{M!}{(\lambda_l!)(l!)^{\lambda_l}} \left[(1 + |\xi|) \int_{\mathbb{R}^d} |y|^{l\vee 2} \nu(\text{d}y) \right]^{\lambda_l} \\ &\leq e^{-\text{Re} \Phi(\xi)} \sum_{j=1}^M \left[(1 + |\xi|) \int_{\mathbb{R}^d} (|y|^2 + |y|^{2\nu n}) \nu(\text{d}y) \right]^j \sum_{u(M,j)} \prod_{l=1}^M \frac{M!}{(\lambda_l!)(l!)^{\lambda_l}} \\ &\leq e^{-\text{Re} \Phi(\xi)} (1 + |\xi|)^M \sum_{j=1}^M \left[\int_{\mathbb{R}^d} (|y|^2 + |y|^{2\nu n}) \nu(\text{d}y) \right]^j \sum_{u(M,j)} \prod_{l=1}^M \frac{M!}{(\lambda_l!)(l!)^{\lambda_l}} \\ &\leq c_1(n) e^{-\text{Re} \Phi(\xi)} (1 + |\xi|)^n \left[1 + \int_{\mathbb{R}^d} (|y|^2 + |y|^{2\nu n}) \nu(\text{d}y) \right]^n. \end{aligned}$$

We note that this inequality remains valid for $M = 0$.

For $\beta \in \mathbb{N}_0^d$ with $|\beta| \leq m$, we can use the Leibniz rule to get

$$\begin{aligned} \left| \frac{\partial^n}{\partial \xi_k^n} (M_\beta \widehat{p})(\xi) \right| &\leq \sum_{j=0}^n \binom{n}{j} \left| \frac{\partial^j}{\partial \xi_k^j} M_\beta(\xi) \frac{\partial^{n-j}}{\partial \xi_k^{n-j}} \widehat{p}(\xi) \right| \\ &\leq (1 + |\xi|)^{|\beta|} \sum_{j=0}^n \binom{n}{j} \left| \frac{\partial^{n-j}}{\partial \xi_k^{n-j}} \widehat{p}(\xi) \right| \end{aligned}$$

$$\leq c_2(n)e^{-\operatorname{Re} \Phi(\xi)}(1 + |\xi|)^{n+m} \left[1 + \int_{\mathbb{R}^d} (|y|^2 + |y|^{2\nu_n})\nu(dy) \right]^n.$$

By (2.9), we see

$$|y_k|^n |\partial^\beta p(y)| \leq c_2(n) \left[1 + \int_{\mathbb{R}^d} (|y|^2 + |y|^{2\nu_n})\nu(dy) \right]^n \int e^{-\operatorname{Re} \Phi(\xi)} (1 + |\xi|)^{n+m} d\xi.$$

Finally,

$$(1 + |y|)^n \leq 2^{n-1} (1 + |y|^n) \leq 2^{n-1} d^{n/2} \left(1 + \sum_{k=1}^d |y_k|^n \right),$$

and the required assertion follows with $C(n, d) = 2^{n-1} d^{n/2} (d + 1) c_2(n)$. □

We will now study the derivatives of transition densities for Lévy processes with (modified) bounded support. For this, we need Proposition 2.1. Let Φ be the symbol (i.e., the characteristic exponent) of a Lévy process and consider for every $r > 0$ the semigroup of infinitely divisible measures $\{P_t^r, t \geq 0\}$ whose Fourier transform is of the form $\widehat{P}_t^r(\xi) = \exp(-t\Phi_r(\xi))$, where

$$\Phi_r(\xi) = \int_{|y| \leq r} (1 - e^{i\xi \cdot y} + i\xi \cdot y)\nu(dy)$$

(ν is the Lévy measure of the symbol Φ). For $\rho > 0$ and $t > 0$, we define

$$\varphi(\rho) = \sup_{|\eta| \leq \rho} \operatorname{Re} \Phi(\eta) \quad \text{and} \quad h(t) := \frac{1}{\varphi^{-1}(1/t)}. \tag{2.12}$$

Proposition 2.2. *Assume that (1.3) holds, and there exist $m \in \mathbb{N}_0$ and $c, t_1 > 0$ such that for all $t \geq t_1$,*

$$\int \exp(-t \operatorname{Re} \Phi(\xi)) |\xi|^m d\xi \leq c \left(\varphi^{-1} \left(\frac{1}{t} \right) \right)^{m+d}. \tag{2.13}$$

Then there is a constant $t_2 = t_2(m, d) > 0$ such that for any $t \geq t_2$, there exists a density $p_t^{h(t)} \in C_b^m(\mathbb{R}^d)$ of $P_t^{h(t)}$, and for every $n \in \mathbb{N}_0$ and $\beta \in \mathbb{N}_0^d$ with $|\beta| \leq m - n$,

$$|\partial_y^\beta p_t^{h(t)}(y)| \leq C(m, n, |\beta|, \Phi) (\varphi^{-1}(t^{-1}))^{d+|\beta|} (1 + \varphi^{-1}(t^{-1})|y|)^{-n}, \quad y \in \mathbb{R}^d.$$

Proof. *Step 1.* For $\xi \in \mathbb{R}^d$,

$$\begin{aligned} |\widehat{P}_t^r(\xi)| &= \exp\left(-t \int_{|y| < r} (1 - \cos(\xi \cdot y))\nu(dy)\right) \\ &= \exp\left(-t \left(\operatorname{Re} \Phi(\xi) - \left(\int_{|y| \geq r} (1 - \cos(\xi \cdot y))\nu(dy) \right) \right)\right) \\ &\leq \exp(-t(\operatorname{Re} \Phi(\xi))) \exp(2t\nu(B(0, r)^c)). \end{aligned} \tag{2.14}$$

By (1.3) and [23], Proposition 28.1, it follows that there exists $t_3 := t_3(d) > 0$ such that for all $r > 0$ and for any $t \geq t_3$, the measure P_t^r has a density $p_t^r \in C_b(\mathbb{R}^d)$.

Step 2. For $t \geq t_3$, we define $g_t(y) = h(t)^d p_t^{h(t)}(h(t)y)$. We consider the infinitely divisible distribution $\pi_t(dy) = g_t(y) dy$. Its Fourier transform is given by

$$\begin{aligned} \widehat{\pi}_t(\xi) &= (h(t))^d \int e^{i\xi \cdot y} p_t^{h(t)}(h(t)y) dy \\ &= \int e^{i\xi \cdot y/h(t)} p_t^{h(t)}(y) dy \\ &= \exp\left(-t \int_{|y| \leq h(t)} \left(1 - e^{i\xi \cdot y/h(t)} + \frac{i\xi \cdot y}{h(t)}\right) \nu(dy)\right) \\ &= \exp\left(-\int_{|y| \leq 1} (1 - e^{i\xi \cdot y} + i\xi \cdot y) \lambda_t(dy)\right) \\ &= \exp(-G_t(\xi)), \end{aligned} \tag{2.15}$$

where λ_t is the Lévy measure of π_t , that is, for any Borel set $B \subset \mathbb{R}^d \setminus \{0\}$,

$$\lambda_t(B) = t \int_{|y| \leq h(t)} \mathbb{1}_B(y/h(t)) \nu(dy).$$

For $n \geq 2$, we have

$$\begin{aligned} \int |y|^n \lambda_t(dy) &= t \int_{|y| \leq h(t)} \left(\frac{|y|}{h(t)}\right)^n \nu(dy) \\ &\leq t \int_{|y| \leq h(t)} \left(\frac{|y|}{h(t)}\right)^2 \nu(dy) \\ &\leq 2t \int_{|y| \leq h(t)} \frac{(|y|/h(t))^2}{1 + (|y|/h(t))^2} \nu(dy) \\ &\leq 2t \int \frac{(|y|/h(t))^2}{1 + (|y|/h(t))^2} \nu(dy) \\ &= 2t \iint (1 - \cos(y/h(t) \cdot \xi)) f_d(\xi) d\xi \nu(dy) \\ &= 2t \iint (1 - \cos(y \cdot \xi/h(t))) \nu(dy) f_d(\xi) d\xi \\ &= 2t \int \operatorname{Re} \Phi\left(\frac{\xi}{h(t)}\right) f_d(\xi) d\xi, \end{aligned}$$

where

$$f_d(\xi) = \frac{1}{2} \int_0^\infty (2\pi\rho)^{-d/2} e^{-|\xi|^2/(2\rho)} e^{-\rho/2} d\rho.$$

Obviously, $f_d(\xi)$ possesses all moments, see, for example, [25], (2.5) and (2.6). By using several times the subadditivity of $\eta \mapsto \sqrt{\operatorname{Re} \Phi(\eta)}$, we can easily find, see, for example, the proof of [24], Lemma 2.3,

$$\operatorname{Re} \Phi\left(\frac{\xi}{h(t)}\right) \leq 2(1 + |\xi|^2) \sup_{|\eta| \leq 1/h(t)} \operatorname{Re} \Phi(\eta) = 2(1 + |\xi|^2) \frac{1}{t}.$$

So,

$$\int \operatorname{Re} \Phi\left(\frac{\xi}{h(t)}\right) f_d(\xi) \, d\xi \leq 2 \sup_{|\eta| \leq 1/h(t)} \operatorname{Re} \Phi(\eta) \int (1 + |\xi|^2) f_d(\xi) \, d\xi =: \frac{c_0}{t}.$$

According to the definition of $h(t)$, we get that for any $t > 0$,

$$\int |y|^n \lambda_t(dy) \leq 2c_0. \tag{2.16}$$

Step 3. It is easily seen from (2.15) that the characteristic exponent of π_t is $G_t(\xi)$, and

$$\operatorname{Re} G_t(\xi) = t \operatorname{Re}(\Phi_{h(t)}(h(t)^{-1}\xi)).$$

Thus,

$$\begin{aligned} \operatorname{Re} G_t(\xi) &= t \left[\operatorname{Re}(\Phi(h(t)^{-1}\xi)) - \int_{|y|>h(t)} (1 - \cos(h(t)^{-1}\xi \cdot y)) \nu(dy) \right] \\ &\geq t \operatorname{Re}(\Phi(h(t)^{-1}\xi)) - 2t\nu(B(0, h(t))^c). \end{aligned}$$

For any $t > 0$,

$$\begin{aligned} \nu(B(0, h(t))^c) &\leq 2 \int_{|y|>h(t)} \frac{(|y|/h(t))^2}{1 + |y|^2/h(t)^2} \nu(dy) \\ &\leq 2 \int \frac{(|y|/h(t))^2}{1 + (|y|/h(t))^2} \nu(dy) \\ &= 2 \iint (1 - \cos(h(t)^{-1}\xi \cdot y)) \nu(dy) f_d(\xi) \, d\xi \\ &= 2 \int \operatorname{Re} \Phi\left(\frac{\xi}{h(t)}\right) f_d(\xi) \, d\xi \\ &\leq 2c_0 \sup_{|\eta| \leq 1/h(t)} \operatorname{Re} \Phi(\eta), \end{aligned}$$

where the last two lines follow from the same arguments as those leading to (2.16). Hence, for any $t > 0$, we have

$$t\nu(B(0, h(t))^c) \leq 2c_0t \sup_{|\eta| \leq 1/h(t)} \operatorname{Re} \Phi(\eta) = 2c_0.$$

By (2.13), for $m \in \mathbb{N}_0$ and $c_1 > 0$, there exists $t_4 := t_4(m, \Phi, c_1) \geq t_3$ such that for any $t \geq t_4$,

$$\int \exp(-(t \operatorname{Re} \Phi(\xi))) |\xi|^m d\xi \leq c_1 h(t)^{-(m+d)}.$$

Therefore, we obtain

$$\begin{aligned} & \int \exp[-\operatorname{Re}(G_t(\xi))] |\xi|^m d\xi \\ & \leq e^{4c_0} \int \exp[-t \operatorname{Re}(\Phi(\xi/h(t)))] |\xi|^m d\xi \\ & = e^{4c_0} h(t)^{m+d} \int \exp[-(t \operatorname{Re} \Phi(\xi))] |\xi|^m d\xi \\ & = c_1 e^{4c_0} < \infty. \end{aligned} \tag{2.17}$$

Step 4. According to (2.16), (2.17) and Proposition 2.1, $g_t \in C_b^m(\mathbb{R}^d)$ for any $t \geq t_4$, and for every $n \in \mathbb{N}_0$ and $\beta \in \mathbb{N}_0^d$ with $|\beta| \leq m - n$ we get

$$|\partial_y^\beta g_t(y)| \leq C(m, n, |\beta|, \Phi)(1 + |y|)^{-n}, \quad y \in \mathbb{R}^d.$$

This finishes the proof since $\partial_y^\beta g_t(y) = h(t)^{d+|\beta|} \partial_y^\beta p_t^{h(t)}(h(t)y)$. \square

The following result is the counterpart of Proposition 2.2, which presents estimates for the derivatives of the densities $p_t^{h(t)}$ for small time. Recall the definitions of φ and h from (2.12).

Proposition 2.3. *Assume that (1.1) is satisfied, and there exist constants $m \in \mathbb{N}_0$ and $c, t_0 > 0$ such that for every $0 < t \leq t_0$,*

$$\int \exp[-(t \operatorname{Re} \Phi(\xi))] |\xi|^m d\xi \leq c(\varphi^{-1}(1/t))^{m+d}. \tag{2.18}$$

Then there is a constant $t_1 > 0$ such that for all $0 < t \leq t_1$, there exists a density $p_t^{h(t)} \in C_b^m(\mathbb{R}^d)$ of $P_t^{h(t)}$. Moreover, for every $n \in \mathbb{N}_0$ and $\beta \in \mathbb{N}_0^d$ with $|\beta| \leq m - n$,

$$|\partial_y^\beta p_t^{h(t)}(y)| \leq C(m, n, |\beta|, \Phi, t_0)(\varphi^{-1}(t^{-1}))^{d+|\beta|} (1 + \varphi^{-1}(t^{-1})|y|)^{-n}, \quad y \in \mathbb{R}^d.$$

Proof. The proof is similar to that of Proposition 2.2, and we only sketch some key differences. We continue to use the notations of the proof of Proposition 2.2. According to (2.14) and (1.1), for all $r > 0$ and $t > 0$, the measure P_t^r is absolutely continuous with respect to Lebesgue measure. Since $t > 0$ may be arbitrarily small, we need (1.1) rather than (1.3). Denote by p_t^r its density. Following the argument of Proposition 2.2, we find that (2.16) is still valid, and according to

(2.18), there exists some $t_2 > 0$ such that (2.17) holds for all $0 < t \leq t_2$. The required assertion follows now from Proposition 2.1. \square

3. Proofs of the main theorems and further examples

We will now give the proofs for Theorem 1.1 and 1.3. For this, we need to estimate the coupling time of a general Lévy process. We will use the functions φ and h defined in (2.12).

Theorem 3.1. *Assume that (1.3) holds, and there are $c, t_1 > 0$ such that for any $t \geq t_1$,*

$$\int \exp[-(t \operatorname{Re} \Phi(\xi))] |\xi|^{d+2} d\xi \leq c(\varphi^{-1}(1/t))^{2d+2}. \tag{3.19}$$

Then, the Lévy process X_t has the coupling property, and there exist $t_2, C > 0$ such that for any $x, y \in \mathbb{R}^d$ and $t \geq t_2$,

$$\|P_t(x, \cdot) - P_t(y, \cdot)\|_{\text{var}} \leq C|x - y|\varphi^{-1}(1/t). \tag{3.20}$$

Proof. Set

$$\Phi_r(\xi) = \int_{|y| \leq r} (1 - e^{i\xi \cdot y} + i\xi \cdot y) \nu(dy)$$

and

$$\Psi_r(\xi) := \Phi(\xi) - \Phi_r(\xi) = \int_{|y| > r} (1 - e^{i\xi \cdot y}) \nu(dy) - i\xi \cdot \int_{1 < |y| \leq r} y \nu(dy).$$

Let Y_t and Z_t be two independent Lévy processes whose symbols are $\Phi_r(\xi)$ and $\Psi_r(\xi)$, respectively. Denote by Q_t and $Q_t(x, \cdot)$ the semigroup and the transition function of Y_t . Similarly, R_t and $R_t(x, \cdot)$ stand for the semigroup and the transition function of Z_t . Then,

$$\begin{aligned} \|P_t(x, \cdot) - P_t(y, \cdot)\|_{\text{var}} &= \sup_{\|f\|_{\infty} \leq 1} |P_t f(x) - P_t f(y)| \\ &= \sup_{\|f\|_{\infty} \leq 1} |Q_t R_t f(x) - Q_t R_t f(y)| \\ &\leq \sup_{\|g\|_{\infty} \leq 1} |Q_t g(x) - Q_t g(y)| \\ &= \|Q_t(x, \cdot) - Q_t(y, \cdot)\|_{\text{var}}. \end{aligned} \tag{3.21}$$

Now, we take $r = h(t)$. Then, according to (3.19) and Proposition 2.2, there exists $t_3 > 0$ such that for any $t \geq t_3$, the kernel Q_t has a density $q_t \in C_b^{d+2}(\mathbb{R}^d)$, and for all $y \in \mathbb{R}^d$,

$$|\nabla q_t(y)| \leq c(d, \Phi) h(t)^{-(d+1)} (1 + h(t)^{-1} |y|)^{-(d+1)}. \tag{3.22}$$

Thus, for any $t \geq t_3$,

$$\begin{aligned} \|Q_t(x, \cdot) - Q_t(y, \cdot)\|_{\text{Var}} &= \sup_{\|f\|_{\infty} \leq 1} \left| \int f(z) Q_t(x, dz) - \int f(z) Q_t(y, dz) \right| \\ &= \sup_{\|f\|_{\infty} \leq 1} \left| \int f(z) q_t(z-x) dz - \int f(z) q_t(z-y) dz \right| \quad (3.23) \\ &= \int |q_t(z-x) - q_t(z-y)| dz. \end{aligned}$$

Let $t \geq t_3$. Assume that $|x-y| > h(t)$. Then,

$$\int |q_t(z-x) - q_t(z-y)| dz \leq 2 \leq \frac{2|x-y|}{h(t)}.$$

If $|x-y| \leq h(t)$, then, by (3.22),

$$\begin{aligned} &\int |q_t(z-x) - q_t(z-y)| dz \\ &= \int_{|z-x| > 2h(t)} |q_t(z-x) - q_t(z-y)| dz + \int_{|z-x| \leq 2h(t)} |q_t(z-x) - q_t(z-y)| dz \\ &\leq c(d, \Phi) \frac{|x-y|}{h(t)^{d+1}} \left[\int_{|z-x| > 2h(t)} \left[h(t)^{d+1} \sup_{w \in B(z-x, |y-x|)} |\nabla q_t(w)| \right] dz + \int_{|z-x| \leq 2h(t)} dz \right] \\ &\leq c(d, \Phi) \frac{|x-y|}{h(t)^{d+1}} \left[\int_{|z-x| > 2h(t)} \left[1 + \frac{|z-x|}{2h(t)} \right]^{-d-1} dz + c_d (2h(t))^d \right] \\ &= c(d, \Phi) \frac{|x-y|}{h(t)^{d+1}} \int \left[1 + \frac{|z-x|}{2h(t)} \right]^{-d-1} dz + 2^d c_d c(d, \Phi) \frac{|x-y|}{h(t)} \\ &\leq \frac{C|x-y|}{h(t)}. \end{aligned}$$

Therefore, there exists $C > 0$ such that for all $x, y \in \mathbb{R}^d$ and $t \geq t_3$,

$$\int |q_t(z-x) - q_t(z-y)| dz \leq \frac{C|x-y|}{h(t)}. \quad (3.24)$$

The assertion follows now from (3.21), (3.23) and (3.24). \square

Next, we turn to the proof of Theorem 1.1.

Proof of Theorem 1.1. Under the conditions of Theorem 1.1, we can suppose that for any $\xi \in \mathbb{R}^d$,

$$\operatorname{Re} \Phi(\xi) \geq F(|\xi|),$$

where $F(r)$ is a strictly increasing and differentiable function on $(0, \infty)$ such that

$$F(r) = \begin{cases} c_1 f(r) & \text{if } r \in (0, c_2); \\ c_3 \log(c_4 + c_5 r) & \text{if } r \in [c_2, \infty) \end{cases}$$

for some constants $c_i > 0, i = 1, 2, 3, 4, 5$. Thus,

$$\begin{aligned} \int \exp[-t \operatorname{Re} \Phi(\xi)] |\xi|^{d+2} d\xi &\leq \int \exp[-t F(|\xi|)] |\xi|^{d+2} d\xi \\ &= c_d \int_0^\infty e^{-tr} [F^{-1}(r)]^{2d+1} dF^{-1}(r) \\ &= \frac{c_d}{2(d+1)} \int_0^\infty e^{-tr} d[F^{-1}(r)]^{2(d+1)}. \end{aligned}$$

Since $\liminf_{r \rightarrow 0} f(r) |\log r| < \infty$ and $\limsup_{s \rightarrow 0} f^{-1}(2s)/f^{-1}(s) < \infty$, we have

$$\liminf_{r \rightarrow 0} F(r) |\log r| < \infty$$

and

$$\limsup_{s \rightarrow 0} F^{-1}(2s)/F^{-1}(s) < \infty.$$

Note that the function F also satisfies

$$\lim_{s \rightarrow \infty} F(s) / \log s = c_3.$$

Then, following the proof of [2], Theorem 2.1, we obtain that for $t \rightarrow \infty$,

$$\begin{aligned} \int_0^\infty e^{-tr} d[F^{-1}(r)]^{2(d+1)} &\asymp [F^{-1}(1/t)]^{2(d+1)} \\ &= [f^{-1}(1/t)]^{2(d+1)} \\ &\asymp [\varphi^{-1}(1/t)]^{2(d+1)}. \end{aligned}$$

In the last step we used, in particular, the upper bound of the two-sided comparison $\operatorname{Re} \Phi(\xi) \asymp f(|\xi|)$ as $|\xi| \rightarrow 0$. The desired assertion follows from Theorem 3.1. \square

The following result is the short-time analogue of Theorem 3.1 which gives, additionally, gradient estimates for general Lévy processes.

Theorem 3.2. Assume that (1.1) holds and let ϕ and h be as in (2.12). If there is a constant $c_0 > 0$ such that

$$\int \exp(-t \operatorname{Re} \Phi(\xi)) |\xi|^{d+2} d\xi \leq c_0 (\varphi^{-1}(1/t))^{2d+2} \quad \text{for all } t \ll 1 [t \gg 1], \quad (3.25)$$

then there exists a constant $c > 0$ such that for all $u \in B_b(\mathbb{R}^d)$

$$\|\nabla P_t u\|_\infty \leq c \|u\|_\infty \varphi^{-1}(1/t) \quad \text{for all } t \ll 1 [t \gg 1]. \quad (3.26)$$

Proof. We will treat the short- and large-time cases separately.

Recall the notations used in the proof of Theorem 3.1: Q_t and R_t are the semigroups corresponding to $\Phi_r(\xi)$ and $\Psi_r(\xi)$, respectively. According to (3.25) and Proposition 2.3, for small enough $t \ll 1$, and $r = h(t)$, the measure Q_t has a density $q_t \in C_b^{d+2}(\mathbb{R}^d)$ such that for any $y \in \mathbb{R}^d$,

$$|\nabla q_t(y)| \leq c(d, \Phi) (\varphi^{-1}(t^{-1}))^{d+1} (1 + \varphi^{-1}(t^{-1})|y|)^{-(d+1)}. \quad (3.27)$$

Then, for all $u \in B_b(\mathbb{R}^d)$,

$$\begin{aligned} \sup_{\|u\|_\infty \leq 1} \|\nabla Q_t u\|_\infty &= \sup_{\|u\|_\infty \leq 1} \sup_{x \in \mathbb{R}^d} |\nabla Q_t u(x)| \\ &= \sup_{\|u\|_\infty \leq 1} \sup_{x \in \mathbb{R}^d} \left| \nabla \int q_t(z-x) \cdot u(z) dz \right| \\ &= \sup_{x \in \mathbb{R}^d} \sup_{\|u\|_\infty \leq 1} \left| \int \nabla q_t(z-x) \cdot u(z) dz \right| \\ &= \sup_{x \in \mathbb{R}^d} \int |\nabla q_t(z-x)| dz \\ &= \int |\nabla q_t(z)| dz \\ &\leq c \varphi^{-1}(t^{-1}), \end{aligned} \quad (3.28)$$

where we used (3.27) and dominated convergence. This calculation shows

$$\|\nabla Q_t u\|_\infty \leq c \varphi^{-1}(t^{-1}) \|u\|_\infty.$$

Therefore,

$$\|\nabla P_t u\|_\infty = \|\nabla Q_t(R_t u)\|_\infty \leq c \varphi^{-1}(t^{-1}) \|R_t u\|_\infty \leq c \varphi^{-1}(t^{-1}) \|u\|_\infty, \quad (3.29)$$

which finishes the proof for small $t \ll 1$.

If $t \gg 1$ is sufficiently large, we can apply (3.20), to find for any $u \in B_b(\mathbb{R}^d)$ with $\|u\|_\infty = 1$,

$$\begin{aligned} |\nabla P_t u(x)| &\leq \limsup_{y \rightarrow x} \frac{|P_t u(x) - P_t u(y)|}{|y - x|} \\ &\leq \limsup_{y \rightarrow x} \frac{\sup_{\|w\|_\infty \leq 1} |P_t w(x) - P_t w(y)|}{|y - x|} \\ &\leq \limsup_{y \rightarrow x} \frac{\|P_t(x, \cdot) - P_t(y, \cdot)\|_{\text{var}}}{|y - x|} \\ &\leq C\varphi^{-1}(t^{-1}). \end{aligned}$$

This finishes the proof for large $t \gg 1$. □

Remark 3.3. Let X_t be a rotationally invariant α -stable Lévy process on \mathbb{R}^d , and p_t be its density function. By the scaling property, for any $t > 0$ and $x \in \mathbb{R}^d$, $p_t(x) = t^{-d/\alpha} p_1(t^{-1/\alpha}x)$. On the other hand, it is well known that, see, for example, [8],

$$|\nabla p_1(x)| \leq \frac{c}{1 + |x|^{d+\alpha}}.$$

Denote by P_t the semigroup of X_t . Then, according to the proof of (3.28), we have

$$\sup_{\|u\|_\infty \leq 1} \|\nabla P_t u\|_\infty = t^{-1/\alpha} \int |\nabla p_1(z)| \, dz.$$

This implies that Theorem 1.3 is optimal.

We can now use Theorem 3.2 to prove Theorem 1.3.

Proof of Theorem 1.3. The second assertion easily follows from Theorem 1.1 and the proof of Theorem 3.2. It is therefore enough to consider the first conclusion. Under the conditions assumed in Theorem 1.3, we know that for any $\xi \in \mathbb{R}^d$,

$$\text{Re } \Phi(\xi) \geq F(|\xi|),$$

where $F(r)$ is an increasing function on $(0, \infty)$ such that

$$F(r) = \begin{cases} 0 & \text{if } r \in (0, c_1); \\ c_2 f(r) & \text{if } r \in (c_1, \infty) \end{cases}$$

for some constants $c_i > 0$, $i = 1, 2$, and f is strictly increasing and differentiable on (c_1, ∞) . Therefore,

$$\int \exp[-t \text{Re } \Phi(\xi)] |\xi|^{d+2} \, d\xi \leq \int \exp[-t F(|\xi|)] |\xi|^{d+2} \, d\xi.$$

Since $\limsup_{s \rightarrow \infty} f^{-1}(2s)/f^{-1}(s) < \infty$, we can choose $c > 2$ and $s_0 > 0$ such that we have $f^{-1}(2s) \leq cf^{-1}(s)$ for all $s \geq s_0$. For any $k \geq 1$ the monotonicity of f^{-1} gives

$$f^{-1}(2^k s) \leq c^k f^{-1}(s) = 2^{k\alpha} f^{-1}(s),$$

where we use $\alpha = \log_2 c$. Then, for sufficiently small $t \ll 1$,

$$\begin{aligned} & \int \exp[-tF(|\xi|)]|\xi|^{d+2} d\xi \\ &= \int_{|\xi| < c_1} |\xi|^{d+2} d\xi + c_d \int_{c_1}^{\infty} e^{-c_2 t f(r)} r^{2d+1} dr \\ &\leq C_1 + c_d \int_0^{\infty} e^{-c_2 s} d_s [f^{-1}(s/t)]^{2d+2} \\ &\leq C_1 + c_d \left\{ \int_0^1 + \sum_{n=1}^{\infty} \int_{2^{n-1}}^{2^n} \right\} e^{-c_2 s} d_s [f^{-1}(s/t)]^{2d+2} \\ &\leq C_1 + c_d [f^{-1}(1/t)]^{2d+2} + c_d \sum_{n=1}^{\infty} \exp[-c_2 2^{n-1}] [f^{-1}(2^n/t)]^{2d+2} \\ &\leq C_1 + c_d \left(1 + \sum_{n=1}^{\infty} \exp[-c_2 2^{n-1}] 2^{n\alpha(2d+2)} \right) [f^{-1}(1/t)]^{2d+2} \\ &\leq C_1 + C_2 [f^{-1}(1/t)]^{2d+2}. \end{aligned}$$

Because of (1.1) and $\text{Re } \Phi(\xi) \asymp f(|\xi|)$ as $|\xi| \rightarrow \infty$, we find $f^{-1}(1/t) \rightarrow \infty$ as $t \rightarrow 0$. Thus,

$$\begin{aligned} \int \exp[-t \text{Re } \Phi(\xi)]|\xi|^{d+2} d\xi &\leq \int \exp[-tF(|\xi|)]|\xi|^{d+2} d\xi \\ &\leq C_3 [f^{-1}(1/t)]^{2d+2}. \end{aligned}$$

In the last step we used, in particular, the upper bound of the two-sided comparison $\text{Re } \Phi(\xi) \asymp f(|\xi|)$ as $|\xi| \rightarrow \infty$. Now the assertion follows from Theorem 3.2. \square

Remark 3.4. If we assume in the statement of Theorem 1.3 that $f^{-1}(s) = s^\alpha \ell(s)$ for some $\alpha > 0$ and some positive function ℓ which is slowly varying at ∞ – that is, $\lim_{s \rightarrow \infty} \ell(\lambda s)/\ell(s) = 1$ for every $\lambda > 0$, then standard Abelian and Tauberian arguments (see, e.g., [1], Theorems 1.7.1 and 1.7.1', or [7], Chapter XIII.5, Theorems 1 and 3), we can obtain that

$$\int \exp[-tF(|\xi|)]|\xi|^{d+2} d\xi \asymp [f^{-1}(1/t)]^{2d+2} \quad \text{as } t \rightarrow 0.$$

Let us finally turn to the examples from Section 1.

Proof of Example 1.4. The symbol of the subordinate Brownian motion here satisfies

$$\operatorname{Re} \Phi(\xi) \asymp |\xi|^{\alpha+\beta} \quad \text{as } |\xi| \rightarrow 0,$$

and

$$\operatorname{Re} \Phi(\xi) \asymp |\xi|^\alpha (\log(1 + |\xi|))^{\beta/2} \quad \text{as } |\xi| \rightarrow \infty.$$

For $r > 0$, set $f(r) = r^\alpha (\log(1 + r))^{\beta/2}$ and $g(r) = (r(\log(1 + r))^{-\beta/2})^{1/\alpha}$. Then, for $r \rightarrow \infty$, we have

$$\begin{aligned} f(g(r)) &= r(\log(1 + r))^{-\beta/2} [\log(1 + (r(\log(1 + r))^{-\beta/2})^{1/\alpha})]^{\beta/2} \\ &\asymp r(\log r)^{-\beta/2} \left[\frac{2 \log r - \beta \log \log r}{\alpha} \right]^{\beta/2} \\ &= r \left[\frac{2 \log r - \beta \log \log r}{\alpha \log r} \right]^{\beta/2} \\ &\asymp r. \end{aligned}$$

This shows that $f^{-1}(r) \asymp g(r)$ for $r \rightarrow \infty$, and now Theorems 1.1 and 1.3 apply. □

Proof of Example 1.5. Let Y_t and Z_t be Lévy processes whose Lévy measures are given by

$$\nu^Y(A) := \int_0^{r_0} \int_{\mathbb{S}} \mathbb{1}_A(s\theta) s^{-1-\alpha} \, ds \mu(d\theta) + \int_{r_0}^\infty \int_{\mathbb{S}} \mathbb{1}_A(s\theta) s^{-1-\beta} \, ds \mu(d\theta)$$

and

$$\nu^Z(dz) := \nu(dz) - \nu^Y(dz) \geq 0,$$

respectively. After some elementary calculations, we see that the symbol Φ^Y of Y_t satisfies $\operatorname{Re} \Phi^Y(\xi) \asymp |\xi|^\alpha$ as $|\xi| \rightarrow \infty$ and $\operatorname{Re} \Phi^Y(\xi) \asymp |\xi|^{\beta \wedge 2}$ as $|\xi| \rightarrow 0$. Let $P_t^Y(x, \cdot)$ and P_t^Z denote the transition function and the semigroup of Y_t . According to Theorems 1.1 and 1.3, we can prove the claim first for Y (if we replace in these Theorems Φ , P_t and $P_t(x, \cdot)$ by the corresponding objects Φ^Y , P_t^Y and $P_t^Y(x, \cdot)$). To come back to the original process resp. semigroup, we can now use (3.21) and (3.29). □

Example 1.5 applies to a large number of interesting and important Lévy processes, whose Lévy measures are of the following polar coordinates form:

$$\nu(A) = \int_0^\infty \int_{\mathbb{S}} \mathbb{1}_A(s\theta) Q(\theta, s) \, ds \mu(d\theta);$$

$Q(\theta, s)$ is a nonnegative function on $\mathbb{S} \times (0, \infty)$. For instance, Example 1.5 is applicable for the following processes:

1. Stable Lévy processes [29]:

$$Q(\theta, s) \asymp s^{-\alpha-1},$$

where $\alpha \in (0, 2)$.

2. Layered stable processes [11]:

$$Q(\theta, s) \asymp s^{-\alpha-1} \mathbb{1}_{(0,1]}(s) + s^{-\beta-1} \mathbb{1}_{[1,\infty)}(s),$$

where $\alpha \in (0, 2)$ and $\beta \in (0, \infty)$.

3. Tempered stable processes [22]:

$$Q(\theta, s) \asymp s^{-\alpha-1} e^{-cs},$$

where $\alpha \in (0, 2)$ and $c > 0$.

4. Relativistic stable processes [3,5]:

$$Q(\theta, s) \asymp s^{-\alpha-1} (1+s)^{(d+\alpha-1)/2} e^{-s},$$

where $\alpha \in (0, 2)$.

5. Lamperti stable processes [4]:

$$Q(\theta, s) = s^{-\alpha-1} \frac{\exp(sf(\theta))s^{1+\alpha}}{(e^s - 1)^{1+\alpha}},$$

where $\alpha \in (0, 2)$ and $f: \mathbb{S} \rightarrow \mathbb{R}$ such that $\sup_{\theta \in \mathbb{S}} f(\theta) < 1 + \alpha$.

6. Truncated stable processes [14]:

$$Q(\theta, s) \asymp s^{-\alpha-1} \mathbb{1}_{(0,1]}(s),$$

where $\alpha \in (0, 2)$.

Motivated by Example 1.5, we can present a short proof of (one part of) F.-Y. Wang's result on explicit gradient estimates for the semigroup of a general Lévy processes.

Theorem 3.5 (F.-Y. Wang [31], Theorem 1.1). *Let X_t be a Lévy process on \mathbb{R}^d with Lévy measure ν . Assume that there exists some $r \in (0, \infty]$ such that*

$$\nu(dz) \geq |z|^{-d} f(|z|^{-2}) \mathbb{1}_{\{|z| \leq r\}} dz,$$

where f is Bernstein function such that

$$\liminf_{r \rightarrow \infty} \frac{f(r)}{\log r} = \infty \quad \text{and} \quad \limsup_{s \rightarrow \infty} \frac{f^{-1}(2s)}{f^{-1}(s)} \in (0, \infty).$$

Then there exists a constant $c > 0$ such that for any $t > 0$,

$$\|\nabla P_t u\|_\infty \leq c \|u\|_\infty f^{-1}\left(\frac{1}{t \wedge 1}\right), \quad u \in B_b(\mathbb{R}^d).$$

Proof. According to the proofs of Example 1.5 and [31], Theorem 1.1, we see that X_t can be decomposed into two independent Lévy processes Y_t and Z_t , such that the symbol $\Phi^Y(\xi)$ of Y_t satisfies $\Phi^Y(\xi) \asymp f(|\xi|^2)$ as $|\xi| \rightarrow \infty$. Now we can apply Theorem 1.3 and the claim follows. \square

Acknowledgements

The authors would like to thank the referee and the Associate Editor for their helpful comments on earlier versions of the paper. Financial support through DFG (Grant Schi 419/5-1) and DAAD (PPP Kroatien) (for René L. Schilling), MNiSW, Poland (Grant N N201 397137) (for Paweł Sztonyk) and the Alexander-von-Humboldt Foundation and the Natural Science Foundation of Fujian (No. 2010J05002) (for Jian Wang, the corresponding author) is gratefully acknowledged.

References

- [1] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation. Encyclopedia of Mathematics and Its Applications* **27**. Cambridge: Cambridge Univ. Press. [MR0898871](#)
- [2] Böttcher, B., Schilling, R.L. and Wang, J. (2011). Constructions of coupling processes for Lévy processes. *Stochastic Process Appl.* **121** 1201–1216.
- [3] Byczkowski, T., Małecki, J. and Ryznar, M. (2009). Bessel potentials, hitting distributions and Green functions. *Trans. Amer. Math. Soc.* **361** 4871–4900. [MR2506430](#)
- [4] Caballero, M.E., Pardo, J.C. and Pérez, J.L. (2010). On the Lamperti stable processes. *Probab. Math. Stat.* **30** 1–28.
- [5] Carmona, R., Masters, W.C. and Simon, B. (1990). Relativistic Schrödinger operators: Asymptotic behavior of the eigenfunctions. *J. Funct. Anal.* **91** 117–142. [MR1054115](#)
- [6] Constantine, G.M. and Savits, T.H. (1996). A multivariate Faà di Bruno formula with applications. *Trans. Amer. Math. Soc.* **348** 503–520. [MR1325915](#)
- [7] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications, Vol. 2*, 2nd ed. New York: Wiley.
- [8] Głowacki, P. (1986). Stable semigroups of measures as commutative approximate identities on non-graded homogeneous groups. *Invent. Math.* **83** 557–582. [MR0827368](#)
- [9] Hartman, P. and Wintner, A. (1942). On the infinitesimal generators of integral convolutions. *Amer. J. Math.* **64** 273–298. [MR0006635](#)
- [10] Hoh, W. (1998). A symbolic calculus for pseudo-differential operators generating Feller semigroups. *Osaka J. Math.* **35** 789–820. [MR1659620](#)
- [11] Houdré, C. and Kawai, R. (2007). On layered stable processes. *Bernoulli* **13** 252–278. [MR2307406](#)
- [12] Jacob, N. (2001). *Pseudo Differential Operators and Markov Processes, Vol. I: Fourier Analysis and Semigroups*. London: Imperial College Press. [MR1873235](#)
- [13] Jacob, N. and Schilling, R.L. (2005). Function spaces as Dirichlet spaces (about a paper by W. Maz'ya and J. Nagel). *Z. Anal. Anwendungen* **24** 3–28. [MR2146549](#)
- [14] Kim, P. and Song, R. (2008). Boundary behavior of harmonic functions for truncated stable processes. *J. Theoret. Probab.* **21** 287–321. [MR2391246](#)
- [15] Knopova, V. and Schilling, R.L. (2011). A note on the existence of transition probability densities for Lévy processes. *Forum Math.* To appear. Available at [arXiv:1003.1419v2](#).
- [16] Kolokoltsov, V.N. (2011). The Lévy–Khintchine type operators with variable Lipschitz continuous coefficients generate linear or nonlinear Markov processes and semigroups. *Probab. Theory Related Fields* **151** 95–123.

- [17] Kolokoltsov, V.N. (2011). Stochastic integrals and SDE driven by nonlinear Lévy noise. In *Stochastic Analysis in 2010* (D. Crisan, ed.) 227–242. London: Springer.
- [18] Kulik, A.M. (2009). Exponential ergodicity of the solutions to SDE's with a jump noise. *Stochastic Process. Appl.* **119** 602–632. [MR2494006](#)
- [19] Lindvall, T. (1992). *Lectures on the Coupling Method. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. New York: Wiley. [MR1180522](#)
- [20] Picard, J. (1996). On the existence of smooth densities for jump processes. *Probab. Theory Related Fields* **105** 481–511. [MR1402654](#)
- [21] Priola, E. and Wang, F.Y. (2006). Gradient estimates for diffusion semigroups with singular coefficients. *J. Funct. Anal.* **236** 244–264. [MR2227134](#)
- [22] Rosiński, J. (2007). Tempering stable processes. *Stochastic Process. Appl.* **117** 677–707. [MR2327834](#)
- [23] Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions. Cambridge Studies in Advanced Mathematics* **68**. Cambridge: Cambridge Univ. Press. [MR1739520](#)
- [24] Schilling, R.L. (1998). Conservativeness and extensions of Feller semigroups. *Positivity* **2** 239–256. [MR1653474](#)
- [25] Schilling, R.L. (1998). Growth and Hölder conditions for the sample paths of Feller processes. *Probab. Theory Related Fields* **112** 565–611. [MR1664705](#)
- [26] Schilling, R.L., Song, R. and Vondraček, Z. (2010). *Bernstein Functions: Theory and Applications. de Gruyter Studies in Mathematics* **37**. Berlin: de Gruyter. [MR2598208](#)
- [27] Schilling, R.L. and Wang, J. (2011). On the coupling property of Lévy processes. *Ann. Inst. Henri Poincaré. Probab. Stat.* **47** 1147–1159.
- [28] Song, R. and Vondraček, Z. (2009). Potential theory of subordinate Brownian motion. In *Potential Analysis of Stable Processes and Its Extensions. Lect. Notes Math.* **1980** 87–176. Berlin: Springer.
- [29] Sztonyk, P. (2010). Regularity of harmonic functions for anisotropic fractional Laplacians. *Math. Nachr.* **283** 289–311. [MR2604123](#)
- [30] Wang, F.Y. (2011). Coupling for Ornstein–Uhlenbeck jump processes. *Bernoulli* **17** 1136–1158.
- [31] Wang, F.Y. (2011). Gradient estimate for Ornstein–Uhlenbeck jump processes. *Stochastic Process. Appl.* **121** 466–478. [MR2763092](#)
- [32] Wang, J. (2010). Regularity of semigroups generated by Lévy type operators via coupling. *Stochastic Process. Appl.* **120** 1680–1700. [MR2673970](#)

Received March 2011 and revised April 2011