Transition density estimates for a class of Lévy and Lévy-type processes

Viktorya Knopova* and René L. Schilling**

Abstract

We show on-and off-diagonal upper estimates for the transition densities of symmetric Lévy and Lévy-type processes. To get the on-diagonal estimates we prove a Nash type inequality for the related Dirichlet form. For the off-diagonal estimates we assume that the characteristic function of a Lévy (type) process is analytic, which allows to apply the complex analysis technique.

Keywords: Bernstein function; carré du champ operator; Dirichlet form; Feller process; Lévy process; large deviations.

MSC 2010: Primary: 60J35. Secondary: 31C25; 32A10; 47D07; 60G51; 60J75.

1 Introduction

Transition density estimates for jump processes received much attention during the recent years. Two-sided heat kernel estimates for a class of stable-like processes in $\mathbb{R}^n$ were obtained by Kolokoltsov [29]; Bass and Levin [4] used a completely different approach to get transition density estimates for discrete time Markov chains in $\mathbb{Z}^d$ with certain conductance. For the transition density estimates for a tempered stable process see [34] and the references therein. Chen and Kumagai [12] obtained two-sided heat kernel estimates and a parabolic Harnack inequality on d-sets, and further extended these results to symmetric jump-type processes on metric measure spaces [13]. Further results on two-sided heat kernel estimates and a version of the parabolic Harnack inequality for symmetric jump processes are contained in [2] and [3]. Pure jump processes whose jump kernel is comparable to the one of truncated stable-like process are considered by Chen, Kim and Kumagai [10]. See [11] for the heat kernel estimates for a class of symmetric jump processes in $\mathbb{R}^n$ with exponentially decaying jump kernels.

*V.M. Glushkov Institute of Cybernetics, NAS of Ukraine, 40, Acad. Glushkov Ave., 03187, Kiev, Ukraine, vic_knopova@gmx.de

**Institut für Mathematische Stochastik, Technische Universität Dresden, D-01062 Dresden, Germany, rene.schilling@tu-dresden.de

The techniques of getting upper bounds are often based on the paper [9] by Carlen, Kusuoka and Stroock where on- and off-diagonal upper bounds for a class of symmetric Markov processes are obtained. The off-diagonal upper bound is obtained in [9] in terms of the so-called carré du champ operator which is uniquely determined by the Dirichlet form related to the Markov process. For Lévy processes the expression of the carré du champ operator is hidden in the representation of the transition density provided the latter exists. Such observations suggest that in the case of Lévy processes the off-diagonal upper bound can be obtained in a similar form as it was done in [9].

Our goal is to get the upper on- and off-diagonal transition density estimates for certain classes of Lévy and Lévy-type processes. In Section 3 we prove the equivalence of the Nash type inequality to the existence of the transition density of the related semigroup of probability measures, and find an on-diagonal upper bound for the density. Then we apply these results to study the estimates for the transition density of Lévy processes whose characteristic exponent is of the form $f(|\xi|^2)$, where f is a (complete) Bernstein function. The main result, Theorem 6, is contained in Section 4: if the Lévy process has no diffusion part and if it has exponential moments, we can use the complex analysis technique to get the off-diagonal upper estimates for the transition density provided it exists.

As a by-product we show that the function which controls the off-diagonal behaviour of the transition density is exactly the rate function of the Lévy process, coming from the theory of large deviations. As an application in Section 5 we show that under the conditions of Theorem 6 the transition density satisfies the large deviation principle.

In Section 6 we show how the results obtained in the previous sections can be generalized to a large class of Lévy-type processes.

Notation. $\mathbb{N}$, $\mathbb{R}$, $\mathbb{C}$ are the sets of positive integers, real numbers and complex numbers, respectively; $\mathbb{R}^n$ is the n-dimensional Euclidean space; $C_0^\infty(\mathbb{R}^n)$ are the test functions, $S(\mathbb{R}^n)$ is the Schwartz space of rapidly decreasing functions on $\mathbb{R}^n$. By $f \asymp g$ we indicate that there exist some positive constants c_1, c_2 such that $c_1 f \leqslant g \leqslant c_2 f$. We write $B(x, r)$ for the open ball with centre x and radius r.

2 Preliminary results

In this section we will briefly recall some basic facts on continuous negative definite functions and Dirichlet forms which we will need later on.

An n-dimensional Lévy process $X = (X_t)_{t\geqslant 0}$ is a stochastic process with values in $\mathbb{R}^n$ with stationary and independent increments and stochastically continuous sample paths. It is well known that such processes are—up to indistinguishability—characterized by the characteristic function

$$\mathbb{E}e^{i\xi\cdot X_t} = e^{-t\psi(\xi)}, \quad \xi \in \mathbb{R}^n,\ t \geqslant 0,$$

and the characteristic exponent $\psi : \mathbb{R}^n \to \mathbb{C}$. The exponent $\psi : \mathbb{R}^n \to \mathbb{C}$ is a *continuous negative definite function* (in the sense of Schoenberg) which is equivalent to saying that ψ

enjoys a *Lévy-Khintchine representation*,

$$\psi(\xi) = i\ell \cdot \xi + \frac{1}{2}\,\xi \cdot Q\xi + \int_{\mathbb{R}^n\setminus\{0\}} \left(1 - e^{i\xi\cdot y} + \frac{i\xi \cdot y}{1+|y|^2}\right) \nu(dy). \tag{1}$$

Here $\ell \in \mathbb{R}^n$, $Q \in \mathbb{R}^{n\times n}$ is a positive semi-definite matrix, and $\nu(dy)$ is the Lévy measure, i.e. a measure on $\mathbb{R}^n \setminus \{0\}$ such that $\int_{\mathbb{R}^n\setminus\{0\}} (1 \wedge |y|^2)\,\nu(dy) < \infty$. The triplet (ℓ, q, ν) uniquely characterizes ψ, hence X.

Note that negative definiteness of ψ does not mean that $-\psi$ is positive definite. Note also, that continuous negative definite functions are always polynomially bounded:

$$|\psi(\xi)| \leqslant c_\psi\,(1 + |\xi|^2), \qquad c_\psi = \sup_{|\eta|\leqslant 1} |\psi(\eta)|.$$

A standard reference for these and further properties is the monograph [22].

In this paper we will only consider real-valued ψ with $Q \equiv 0$. Thus, (1) becomes

$$\psi(\xi) = \int_{\mathbb{R}^n\setminus\{0\}} \big(1 - \cos(\xi \cdot y)\big)\,\nu(dy) \tag{2}$$

and the associated Lévy process X is symmetric in the sense that $\operatorname{law}(X_t) = \operatorname{law}(-X_t)$.

Let $\psi : \mathbb{R}^n \to \mathbb{R}$ and write $\widehat{u}(\xi) = (2\pi)^{-n} \int u(x)\, e^{-ix\cdot\xi}\,dx$ for the Fourier transform. Then

$$-\,\psi(D)u(x) = -\int_{\mathbb{R}^n} e^{ix\cdot\xi}\,\psi(\xi)\,\widehat{u}(\xi)\,d\xi, \quad u \in C_0^\infty(\mathbb{R}^n), \tag{3}$$

is a pseudo-differential operator with symbol ψ. It is not hard to see that $-\psi(D)$ coincides on the test functions $C_0^\infty(\mathbb{R}^n)$ with the infinitesimal generator A of the Lévy process X. More precisely, if we denote by $P_t u(x) := \mathbb{E}u(X_t + x)$ the convolution semigroup associated with X, $P_t|_{C_0^\infty(\mathbb{R}^n)}$ can be extended to a strongly continuous, symmetric sub-Markovian semigroup on each $L_p(\mathbb{R}^n, dx)$, $p \in [1, \infty)$, as well as on the space of continuous functions vanishing at infinity: $C_\infty(\mathbb{R}^n) := \overline{C_0^\infty(\mathbb{R}^n)}^{\|\cdot\|_\infty}$. In any case, $C_0^\infty(\mathbb{R}^n)$ is an operator core, and the closure of $-\psi(D)|_{C_0^\infty(\mathbb{R}^n)}$ is the L_p-generator. On $C_0^\infty(\mathbb{R}^n)$ it is also possible to express the semigroup $(P_t)_{t\geqslant 0}$ as pseudo-differential operators,

$$P_t u(x) = \int_{\mathbb{R}^n} e^{ix\cdot\xi}\,e^{-t\psi(\xi)}\,\widehat{u}(\xi)\,d\xi, \quad u \in C_0^\infty(\mathbb{R}^n), \tag{4}$$

see e.g. [22, Chapter 4].

We are mostly interested in the L_2 setting. The domain of the L_2 generator is a concrete function space

$$H^{\psi,2}(\mathbb{R}^n) := \overline{C_0^\infty(\mathbb{R}^n)}^{\|\cdot\|_{\psi,2}} \quad \text{where} \quad \|u\|_{\psi,2}^2 = \|\psi(D)u\|_{L_2}^2 + \|u\|_{L_2}^2, \tag{5}$$

see [22, Chapter 4.1]. Denote by $\mathcal{E}^\psi(u,v)$ the Dirichlet form associated with $-\psi(D)$, see [22, Example 4.7.28]. It is given by

$$\begin{aligned}\mathcal{E}^\psi(u,v) &= \int_{\mathbb{R}^n} \psi(\xi)\widehat{u}(\xi)\overline{\widehat{v}}(\xi)\,d\xi \\ &= \frac{1}{2}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} (u(x+y)-u(x))(v(x+y)-v(x))\,\nu(dy)\,dx.\end{aligned} \tag{6}$$

The domain $D(\mathcal{E}^\psi) = H^{\psi,1}(\mathbb{R}^n)$ is the closure of $C_0^\infty(\mathbb{R}^n)$ with respect to the norm induced by the scalar product

$$\mathcal{E}_1^\psi(u,v) := \mathcal{E}^\psi(u,v) + \langle u,v\rangle_{L_2}.$$

In Section 4 we will need the notion of a *carré du champ operator* to show the off-diagonal upper bound for the transition density. Below we give the definition of the carré du champ operator in the sense of Bouleau and Hirsch for a general Dirichlet form $(\mathcal{E}, D(\mathcal{E}))$ on $L_2(\Omega, m)$, see [8, Definition 4.1.2]. As usual, m is a positive σ-finite Radon measure on a locally compact measurable space $(\Omega, \mathcal{F})$.

Definition 1. We say that a Dirichlet form $(\mathcal{E}, D(\mathcal{E}))$ admits a *carré du champ* if there exists a subspace $\mathcal{C}$ of $D(\mathcal{E}) \cap L_\infty(\Omega, m)$ which is dense in $D(\mathcal{E})$ such that for all $f \in \mathcal{C}$ there exists a function $\tilde{f}$ such that

$$2\mathcal{E}(fh, f) - \mathcal{E}(h, f^2) = \int h\tilde{f}\,dm \quad \text{for all } h \in D(\mathcal{E}) \cap L_\infty(\Omega, m).$$

If the assumptions of Definition 1 are satisfied, then there exists a unique positive symmetric and continuous bilinear operator $\Gamma : D(\mathcal{E}) \times D(\mathcal{E}) \to L_1(\Omega, m)$, called the *carré du champ operator*, such that for all $f, g, h \in D(\mathcal{E}) \cap L_\infty(\Omega, m)$

$$\mathcal{E}(fh, g) + \mathcal{E}(gh, f) - \mathcal{E}(h, fg) = \int h\,\Gamma(f,g)\,dm \tag{7}$$

cf. [8, Proposition 4.1.3]. For a Lévy process we have an explicit representation of the carré du champ operator

$$\Gamma(u,u) = \Gamma^\psi(u,u) = \frac{1}{2}\int_{\mathbb{R}^n} (u(x+y)-u(x))^2\,\nu(dy) \tag{8}$$

where ν is the Lévy measure and ψ is given by (2).

We are particularly interested in radially symmetric negative definite functions $\psi : \mathbb{R}^n \to \mathbb{R}$. In this case $\psi(\xi) = f(|\xi|^2)$ for some function $f : [0,\infty) \to [0,\infty)$. In order that $\xi \mapsto f(|\xi|^2)$, $\xi \in \mathbb{R}^n$, is negative definite for *every* dimension $n \geqslant 1$ it is necessary and sufficient that f is a Bernstein function. This is a classic result by Schönberg and Bochner, see [33] or [25].

Definition 2. a) A real-valued function on $[0,\infty)$ is called a *Bernstein function*, if

$$f(t) = a + bt + \int_{(0,\infty)} (1 - e^{-st})\,\mu(dt)$$

with $a, b \geqslant 0$ and a measure μ on $(0,\infty)$ satisfying $\int_{(0,\infty)} (1 \wedge t)\,\mu(dt) < \infty$.

b) A Bernstein function is said to be a *complete Bernstein function* if the representing measure $\mu(dt)$ is of the form $m(t)\,dt$ with a completely monotone density $m(t)$, i.e. where $m(t)$ is the Laplace transform of a positive measure on $[0,\infty)$.

Bernstein functions are equivalently characterized by the requirement that

$$f \in C^\infty(0,\infty), \quad f \geqslant 0, \quad \text{and} \quad (-1)^k f^{(k)}(x) \leqslant 0 \quad \text{for all} \quad k \geqslant 1.$$

For the properties of Bernstein and complete Bernstein functions we refer to [22, Chapter 3.9] and [33]. For our purposes we will only mention the following estimate

$$\left| f^{(k)}(x) \right| \leqslant \frac{k!}{x^k} f(x), \quad k \geqslant 0,\ x > 0. \tag{9}$$

3 An on-diagonal estimate for the transition density

In this section we will prove the equivalence of the Nash type inequality for general Dirichlet forms and the absolute continuity of the related probability measure with respect to Lebesgue measure on $\mathbb{R}^n$; then we will establish an upper bound for this transition density. For the special case where $f(x) = x^\alpha$, $0 < \alpha < 1$, we refer to [9] and to the discussion in [5]. The importance of Nash and Sobolev type inequalities for (symmetric) Markovian semigroups has been pointed out in the monograph [35].

Let f be a Bernstein function, $(\mathcal{E}, D(\mathcal{E}))$ be a symmetric regular Dirichlet form with generator $(A, D(A))$ and associated $L_2(\mathbb{R}^n)$– sub-Markovian semigroup $(P_t)_{t\geqslant 0}$. Since the operators P_t are symmetric, we may (and do) extend P_t to the whole scale L_p, see e.g. [15]. By $\|P_t\|_{L_p \to L_q}$ we indicate the corresponding L_p–L_q operator norm.

Proposition 3. *Let $(P_t)_{t\geqslant 0}$ be the semigroup associated with the regular symmetric Dirichlet form $(\mathcal{E}, D(\mathcal{E}))$ and let $f : (0,\infty) \to [0,\infty)$ be some differentiable function such that f is subadditive,*

$$f'(x) > 0 \quad \text{and} \quad f'(x) \leqslant \frac{1}{x} f(x) \quad \text{for all} \quad x > 0.$$

Then the following Nash inequality

$$\|u\|_{L_2}^2\, f\left(\left[\frac{\|u\|_{L_2}}{\|u\|_{L_1}} \right]^{4/n} \right) \leqslant C_0 \left[\mathcal{E}(u,u) + \delta \|u\|_{L_2}^2 \right], \tag{10}$$

($C_0 > 0$ and $\delta \geqslant 0$ are some constants) holds if, and only if,

$$\|P_t\|_{L_1 \to L_\infty} \leqslant \left[f^{-1}\left(\frac{1}{\gamma t} \right) \right]^{n/2} e^{2\delta t}. \tag{11}$$

where the constant $\gamma > 0$ depends only on C_0, and δ is as in (10).

Proof. Assume that (10) holds. Fix some $u \in D(\mathcal{E}) \cap L_\infty(\mathbb{R}^n) \cap L_1^+(\mathbb{R}^n)$ with $\|u\|_{L_1} = 1$ and set $u_t(x) := P_t u(x)$ and $\phi(t) := e^{-2\delta t}\|u_t\|_{L_2}^2$. Since $\mathcal{E}(u,u) = \|(-A)^{1/2}u\|_{L_2}^2$ and $\frac{d}{dt}P_t = AP_t$ we find

$$\begin{aligned} -\frac{d\phi(t)}{dt} = 2e^{-2\delta t}\left[\mathcal{E}(u_t,u_t) + \delta\|u_t\|_{L_2}^2\right] &\geqslant e^{-2\delta t}\frac{2}{C_0}\|u_t\|_{L_2}^2\, f\left(\|u_t\|_{L_2}^{4/n}\right) \\ &\geqslant \frac{2}{C_0}\,\phi(t)\, f\left(\phi(t)^{2/n}\right). \end{aligned}$$

For the last estimate we used the fact that $f(e^{4\delta t/n}\phi^2(t)) \geqslant f(\phi^2(t))$. It is not hard to see that for each $n \in \mathbb{N}$ the function $f_1(x) := f(x^{1/n})$ enjoys the same properties as f: it is strictly increasing and

$$f_1'(x) = \frac{1}{n}\, f'(x^{1/n})\, x^{1/n-1} \leqslant \frac{1}{n}\, f(x^{1/n})\, x^{-1/n}\, x^{1/n-1} \leqslant \frac{1}{x}\, f_1(x).$$

Combining this with the above estimate for ϕ' we get

$$\left(\frac{1}{f_1(\phi^2)} \right)' = -\frac{2\, f_1'(\phi^2)\,\phi\,\phi'}{f_1^2(\phi^2)} \geqslant -\frac{2\,\phi'}{\phi\, f_1(\phi^2)} \geqslant \frac{4\,\phi\, f_1(\phi^2)}{C_0\,\phi\, f_1(\phi^2)} = \frac{4}{C_0}. \tag{12}$$

Now we integrate (12) from 0 to t and conclude

$$\frac{1}{f_1(\phi^2(t))} \geqslant \frac{1}{f_1(\phi^2(t))} - \frac{1}{f_1(\|u\|_{L_2}^{4/n})} \geqslant \frac{4}{C_0}\, t$$

or

$$\phi(t) \leqslant \left[f^{-1}\left(\frac{1}{\gamma t} \right) \right]^{n/2}, \quad \gamma = \frac{4}{C_0}$$

This implies that

$$\|P_t\|_{L_1 \to L_2} \leqslant \left[f^{-1}\left(\frac{1}{\gamma t} \right) \right]^{n/4} e^{\delta t},$$

and, by the usual semigroup and symmetry arguments, see e.g. [35], we finally get

$$\|P_t\|_{L_1 \to L_\infty} \leqslant \|P_{t/2}\|_{L_1 \to L_2} \cdot \|P_{t/2}\|_{L_2 \to L_\infty} = \|P_{t/2}\|_{L_1 \to L_2}^2 \leqslant \left[f^{-1}\left(\frac{1}{\gamma t} \right) \right]^{n/2} e^{2\delta t}$$

which is (11) with $\gamma = 4/C_0$.

Conversely, assume that (11) holds. Let $v \in D(A) \cap L_\infty(\mathbb{R}^n) \cap L_1^+(\mathbb{R}^n)$ and set $v_t(x) := e^{-\delta t} P_t v(x)$. Then $\|v_t\|_{L_\infty} \leqslant \|v\|_{L_1}[f^{-1}(1/\gamma t)]^{n/2}$, and

$$v_t = v - \int_0^t (\delta \operatorname{id} + A) v_s \, ds.$$

Since P_t and A are self-adjoint operators,

$$\langle v, P_t v \rangle_{L_2} \leqslant \|v\|_{L_2}^2 \quad \text{and} \quad \langle v, A v_t \rangle_{L_2} = e^{-\delta t} \langle v, A P_t v \rangle_{L_2} = e^{-\delta t} \|(A P_t)^{1/2} v\|_{L_2}^2 \leqslant \mathcal{E}(v, v).$$

Therefore

$$\begin{aligned} \|v\|_{L_1}^2 \left[f^{-1}\left(\frac{1}{\gamma t}\right)\right]^{n/2} \geqslant \|v\|_{L_1} \|v_t\|_{L_\infty} \geqslant \langle v, v_t \rangle_{L_2} &= \|v\|_{L_2}^2 - \int_0^t \langle v, (\delta \operatorname{id} + A) v_s \rangle_{L_2} \, ds \\ &\geqslant \|v\|_{L_2}^2 - t \left[\mathcal{E}(v, v) + \delta \|v\|_{L_2}^2 \right], \end{aligned}$$

and we arrive at

$$\delta \|v\|_{L_2}^2 + \mathcal{E}(v, v) \geqslant \frac{\|v\|_{L_2}^2}{t} \left[1 - \frac{\|v\|_{L_1}^2}{\|v\|_{L_2}^2} \left[f^{-1}\left(\frac{1}{\gamma t}\right)\right]^{n/2} \right].$$

Choosing $t = t_0 = 1 \Big/ \gamma f\Big(\big(\frac{\|v\|_{L_2}}{\sqrt{2}\, \|v\|_{L_1}} \big)^{4/n} \Big)$, we get

$$\|u\|_{L_2}^2 \, f \left(\left[\frac{\|v\|_{L_2}}{\sqrt{2}\, \|v\|_{L_1}} \right]^{4/n} \right) \leqslant \frac{2}{\gamma} \left[\mathcal{E}(u, u) + \delta \|u\|_{L_2}^2 \right].$$

Since f is subadditive, we see that $\frac{1}{4} f(y) \leqslant f(y/4) \leqslant f(y/4^{1/n})$ and this yields (10) with $C_0 = 8/\gamma$. □

We can apply Proposition 3 to get on-diagonal estimates for a wide class of Lévy processes which are subordinate to Brownian motion. It is well-known that the symbols of such Lévy processes are of the form $\xi \mapsto f(|\xi|^2)$ where $f(x)$ is a Bernstein function. The corresponding Dirichlet form can be expressed as

$$\mathcal{E}^{f(|\cdot|^2)}(u, u) = \int_{\mathbb{R}^n} f(|\xi|^2) \, |\widehat{u}(\xi)|^2 \, d\xi, \quad u \in C_0^\infty(\mathbb{R}^n).$$

From a function-space point of view this representation is rather complicated and it is desirable to get an equivalent representation involving differences of the function u, cf. [25]. Let us, for simplicity, assume that $f(x)$ has no linear term. Under the additional condition that there exists some $0 < \kappa < 1$ such that

$$t \mapsto f(t)\, t^{-\kappa} \quad \text{is increasing as} \quad t \to \infty$$

it was shown in [28] and [30] that

$$\mathcal{E}^{\psi_1}(u,u) = \frac{1}{2}\int_{\mathbb{R}^n}\int_{B(0,1)} |u(x)-u(y)|^2 f\left(\frac{1}{|x-y|^2}\right)\frac{dy\,dx}{|x-y|^n},$$

where

$$\psi_1(\xi) := \int_{B(0,1)} (1-\cos(\xi\cdot y)) f\left(\frac{1}{|y|^2}\right)\frac{dy}{|y|^n}, \tag{13}$$

is a negative definite function, defines an equivalent Dirichlet form. This is to say that $D(\mathcal{E}^{f(|\cdot|^2)}) = D(\mathcal{E}^{\psi_1})$ and that

$$\mathcal{E}^{f(|\cdot|^2)}(u,u) + \langle u,u\rangle_{L_2} \asymp \mathcal{E}^{\psi_1}(u,u) + \langle u,u\rangle_{L_2}$$

for all $u \in D(\mathcal{E}^{f(|\cdot|^2)}) = D(\mathcal{E}^{\psi_1})$.

Obviously, if the Nash inequality (10) holds for $\mathcal{E}^{\psi_1}$, it also holds for $\mathcal{E}^{f(|\cdot|^2)}$ and vice versa. Note that the Bernstein function f satisfies the assumptions of Proposition 3.

Lemma 4. *Let f be a Bernstein function without linear term such that for some $\kappa \in (0,1)$ the function $t \mapsto f(t)\,t^{-\kappa}$ increases as $t \to \infty$. Then the transition density of the Lévy process with Lévy exponent $f(|\xi|^2)$ (or $\psi_1(\xi)$) exists and there exist suitable constants $c > 0, \gamma \in (0,1]$ such that*

$$p_t(x) \leqslant c\left[f^{-1}\left(\frac{1}{\gamma t}\right)\right]^{n/2}, \quad \textit{for all } \ 0 < t \leqslant 1 \textit{ and } x \in \mathbb{R}^n. \tag{14}$$

Proof. We check (10) for the Dirichlet form $\mathcal{E}^{\psi_1}$. Let $u \in L_\infty(\mathbb{R}^n) \cap L_1(\mathbb{R}^n)$ and fix some $0 < r < 1$. Then

$$\begin{aligned}
\mathcal{E}^{\psi_1}(u,u) &\geqslant \int_{\mathbb{R}^n}\int_{B(x,r)} \frac{|u(x)-u(y)|^2}{|x-y|^n} f\left(\frac{1}{|x-y|^2}\right) dy\,dx \\
&\geqslant f\left(\frac{1}{r^2}\right)\frac{1}{r^n}\int_{\mathbb{R}^n}\int_{B(x,r)} |u(x)-u(y)|^2\,dy\,dx \\
&\geqslant c\,f\left(\frac{1}{r^2}\right)\frac{1}{r^{2n}}\int_{\mathbb{R}^n}\left|\int_{B(x,r)} (u(x)-u(y))\,dy\right|^2 dx \\
&\geqslant c'\,f\left(\frac{1}{r^2}\right)\int_{\mathbb{R}^n} |u(x)-u_r|^2\,dx \\
&= c' f\left(\frac{1}{r^2}\right)\|u(\cdot)-u_r\|_{L_2}^2,
\end{aligned}$$

where $u_r := \tau_n^{-1} r^{-n}\int_{B(x,r)} u(y)\,dy$ and $\tau_n = \pi^{\frac{n}{2}}/\Gamma(1+\frac{n}{2})$ is the volume of the unit ball $B(0,1)$ in $\mathbb{R}^n$. Now observe that

$$\|u_r\|_{L_1} \leqslant \int_{\mathbb{R}^n}\frac{1}{\tau_n\,r^n}\int_{B(x,r)} |u(y)|\,dy\,dx = \frac{1}{\tau_n\,r^n}\int_{B(0,r)}\int_{\mathbb{R}^n} |u(x+y)|\,dx\,dy = \|u\|_{L_1}$$

and that $\|u_r\|_{L_\infty} \leqslant \tau_n^{-1} r^{-n} \|u\|_{L_1}$. By Hölder's inequality we get

$$\|u_r\|_{L_2}^2 \leqslant \|u_r\|_{L_\infty} \cdot \|u_r\|_{L_1} \leqslant \tau_n^{-1} \frac{\|u\|_{L_1}^2}{r^n},$$

and so

$$\|u\|_{L_2}^2 = \|u - u_r + u_r\|_{L_2}^2 \leqslant 2\left(\|u - u_r\|_{L_2}^2 + \|u_r\|_{L_2}^2\right) \leqslant 2\,\|u - u_r\|_{L_2}^2 + 2\tau_n^{-1} \frac{\|u\|_{L_1}^2}{r^n}.$$

Thus, we arrive at

$$\|u\|_{L_2}^2 \leqslant \frac{c}{f(1/r^2)}\, \mathcal{E}^{\psi_1}(u,u) + c\, r^{-n}\, \|u\|_{L_1}^2. \tag{15}$$

We will now distinguish between two cases. Let $h(r) := r^n / f(r^{-2})$.

Case 1. Suppose that $\|u\|_{L_1}^2 < h\big(1\big)\, \mathcal{E}^{\psi_1}(u,u)$. Since

$$\lim_{r\to 0} h(r) = \lim_{x\to\infty} (x^n\, f(x^2))^{-1} = 0,$$

the equation $h(r) = \|u\|_{L_1}^2 / \mathcal{E}^{\psi_1}(u,u)$ has a solution $r_0 \in (0,1)$. Substituting $\mathcal{E}^{\psi_1}(u,u) = \|u\|_{L_1}^2 / h(r_0)$ into (15) with $r \equiv r_0$ yields

$$\|u\|_{L_2}^2 \leqslant 2c\, r_0^{-n}\, \|u\|_{L_1}^2 = \frac{\|u\|_{L_1}^2}{\left[h^{-1}\left(\frac{\|u\|_{L_1}^2}{\mathcal{E}^{\psi_1}(u,u)}\right)\right]^n}.$$

A few elementary rearrangements give

$$\mathcal{E}^{\psi_1}(u,u) \geqslant \frac{\|u\|_{L_1}^2}{h\left(\left(\frac{\|u\|_{L_1}}{\|u\|_{L_2}}\right)^{2/n}\right)}$$

which becomes (10) if we express h in terms of f.

Case 2. $\|u\|_{L_1}^2 \geqslant h(1)\, \mathcal{E}^{\psi_1}(u,u)$ where c is the constant appearing in (15). Take $r = 1$ in (15). Then

$$\|u\|_{L_2}^2 \leqslant 2c\, \|u\|_{L_1}^2,$$

and by monotonicity $f\Big(\Big(\frac{\|u\|_{L_2}^2}{\|u\|_{L_1}^2}\Big)^{4/n}\Big) \leqslant C$. This implies (10), and by Proposition 3 we get (14). □

There is a related situation where we can apply Proposition 3. Consider a Lévy process with the symbol

$$\psi_\infty(\xi) := \int_{\mathbb{R}^n} (1 - \cos(\xi \cdot y))\, g\left(\frac{1}{|y|^2}\right) \frac{dy}{|y|^n}$$

where g is a complete Bernstein function. By [25, Theorem 3.5], the corresponding Dirichlet form $\mathcal{E}^{\psi_\infty}(u,u)$ is equivalent to the Dirichlet form

$$\mathcal{E}^{\mu(|\cdot|^2)}(u,u) := \int_{\mathbb{R}^n} \mu(|\xi|^2)|\widehat{u}(\xi)|^2\,d\xi,$$

where

$$\mu(t) := \int_0^t \int_r^\infty \frac{f(s)}{s^2}\,ds\,dr$$

is also a complete Bernstein function. It is easy to see that the Dirichlet form $\mathcal{E}^{\psi_\infty}(u,u)$ satisfies (10) with $\delta = 0$. Hence we get

Lemma 5. *Let g and μ be as above and let X be the Lévy process with symbol $\psi_\infty(\xi)$ (or $\mu(|\xi|^2)$). Then X has a transition density $p_t(x)$ and there exist suitable constants $c, \gamma > 0$ such that*

$$p_t(x) \leqslant c\left[g^{-1}\left(\frac{1}{\gamma t}\right)\right]^{n/2}, \quad \textit{for all } 0 < t < \infty,\ x \in \mathbb{R}^n. \tag{16}$$

4 An off-diagonal estimate for the transition density of Lévy processes

In this section we prove an off-diagonal upper estimate for the transition density of a class of Lévy processes. Let X_t be a Lévy process with symbol $\psi(\xi)$ and convolution semigroup $\mu_t(dy)$, $t \geqslant 0$. It is known that $\mu_t(dy)$ has a density $p_t(dy)$ with respect to Lebesgue measure if, and only if, $T_t f := f \star \mu_t$ is continuous for all Borel measurable functions f, cf. [22, Lemmas 4.8.19, 4.8.20]. Apart from this criterion there are no good necessary and sufficient criteria for the existence of a transition density $p_t(x)$. Assume that $p_t(x)$ exists; since $\widehat{p_t}(\xi) = e^{-t\psi(\xi)}$, we have necessarily

$$p_t(x) = \int_{\mathbb{R}^n} e^{i\xi\cdot x - t\psi(\xi)}\,d\xi, \quad x \in \mathbb{R}^n,\ t > 0. \tag{17}$$

This shows that the growth of ψ as $|\xi| \to \infty$, e.g. $\psi(\xi) \geqslant |\xi|^\kappa$ for some $0 < \kappa < 2$, guarantees that (17) converges absolutely. More generally, one has sufficient conditions due to Hartman-Wintner [18] and Kallenberg [26], see Bodnarchuk and Kulik [6] for the n-dimensional situation.

Throughout we will assume that the Lévy process X_t has a transition density $p_t(x)$ for all $t > 0$ and that the Lévy measure $\nu(dy)$ satisfies

$$\int_{|y|\geqslant 1} e^{\alpha\cdot y}\,\nu(dy) < \infty \quad \text{for all } \alpha \in \mathbb{R}^n. \tag{\textbf{A1}}$$

Note that (**A1**) is equivalent to saying that the Lévy process has exponential moments, cf. Sato [31, Theorem 25.3]. Under (**A1**) the function

$$w(z) := \frac{1}{2}\int_{\mathbb{R}^n} \big(e^{z\cdot y} - 1\big)\big(e^{-z\cdot y} - 1\big)\,\nu(dy) = \int_{\mathbb{R}^n} \big(1 - \cosh(z\cdot y)\big)\,\nu(dy),$$

exists for any $z \in \mathbb{C}^n$, $z_j = \xi_j + i\eta_j$, $j = 1, \dots, n$. In what follows, we write $\xi = (\xi_1, \dots, \xi_n)$, $\eta = (\eta_1, \dots, \eta_n)$. Observe, that for $\xi = 0$, i.e. $z = i\eta$, we have

$$\begin{aligned} w(i\eta) &= \frac{1}{2} \int_{\mathbb{R}^n} \left(e^{i\eta\cdot y} - 1\right)\left(e^{-i\eta\cdot y} - 1\right) \nu(dy) \\ &= \int_{\mathbb{R}^n} (1 - \cos(y \cdot \eta))\, \nu(dy) = \psi(\eta). \end{aligned} \tag{18}$$

This means that we have $w(\xi) = \Gamma(e^{\xi \cdot}, e^{-\xi \cdot})$ for $\xi \in \mathbb{R}^n$; here $\Gamma(u, u)$ is the carré du champ operator associated with $\psi(\xi)$ by (8).

Let us briefly recall Carlen, Kusuoka and Stroock's upper bound for the transition density $p_t(x, y)$ of a general symmetric Markov process; for details we refer to [9, Theorem 3.25].

Let $(X_t)_{t\geqslant 0}$ be a symmetric Markov process given by a regular Dirichlet form $(\mathcal{E}, D(\mathcal{E}))$ which admits a carré du champ operator $\Gamma(\cdot, \cdot)$. Assume that there exists some $\phi \in D(\mathcal{E}) \cap L_\infty(\mathbb{R}^n) \cap C_b(\mathbb{R}^n)$ such that

$$\gamma(\phi) := \sqrt{\|e^{-2\phi}\, \Gamma(e^{\phi}, e^{\phi})\|_{L_\infty}} \vee \sqrt{\|e^{2\phi}\, \Gamma(e^{-\phi}, e^{-\phi})\|_{L_\infty}} < \infty.$$

If the Dirichlet form $(\mathcal{E}, D(\mathcal{E}))$ satisfies a Nash inequality with $f(x) = x^\alpha$, i.e. if

$$\|u\|_{L_2}^{2\alpha/n} \leqslant A\left(\mathcal{E}(u, u) + \delta\|u\|_{L_2}^2\right) \|u\|_{L_1}^{2\alpha/n},$$

then the $(X_t)_{t\geqslant 0}$ has a transition density $p_t(x, y)$ and

$$p_t(x, y) \leqslant c_1\, t^{-n/\alpha}\, e^{c_2 t \gamma^2(\phi) - |\phi(x) - \phi(y)| + c_3 t} \tag{19}$$

for all $t > 0$ and almost all $x, y \in \mathbb{R}^n$.

Even in simple situations it is a non-trivial task to find concrete functions ϕ which also lead to reasonable estimates in (19). Therefore we aim for a different way to derive a concrete off-diagonal upper bound for the transition density $p_t(x)$ of a Lévy process. Let

$$Q_t(\xi, x) := i\xi \cdot x - tw(i\xi) = i\xi \cdot x - t\psi(\xi).$$

Observe that $Q_t(i\xi, x)$ resembles the exponent of the upper bound (19). It is therefore a natural question whether it is possible to apply complex analysis techniques to get off-diagonal upper bounds similar to those in (19).

Let us look closely at the properties of w. For $\xi \in \mathbb{R}^n$ we have:

i) $w|_{\mathbb{R}^n}$ is even and $\nabla w(\xi)|_{\xi=0} = 0$;

ii) $w(z)$ is an analytic function in each variable $z_j \in \mathbb{C}$, $j = 1, \dots, n$; by Hartogs' theorem it is analytic in $\mathbb{C}^n$.

iii) $w|_{\mathbb{R}^n}$ is concave, i.e. for all $\xi, \xi' \in \mathbb{R}^n$ and $t \in (0,1)$ one has $w(t\xi + (1-t)\xi') \geqslant t\,w(\xi) + (1-t)w(\xi')$. This follows directly from Hölder's inequality.

iv) $w|_{\mathbb{R}^n}$ attains its maximum at $\xi = 0$. This follows immediately from i) and iii).

Consider the function

$$v_t(\xi, x) := -\xi \cdot x - tw(\xi), \quad \xi \in \mathbb{R}^n. \tag{20}$$

Due to iii) it attains its minimum in ξ at some point $\xi_0 = \arg\min_\xi v_t(\xi, x)$. Since $v_t(0, x) = 0$, we have $v_t(\xi_0, x) \leqslant 0$ for all $x \in \mathbb{R}^n$, and $v_t(\xi_0, 0) = 0$. The following function plays a key role in the off-diagonal upper estimate. Write

$$D_t^2(x) := -v_t(\xi_0, x) \quad \text{where} \quad \xi_0 = \xi_0(t, x) = \arg\min_\xi v_t(\xi, x). \tag{21}$$

Theorem 6. *Let $(X_t)_{t\geqslant 0}$ be a Lévy process with symbol $\psi(\xi)$. Assume that X_t has for all $t > 0$ a transition density and assume that ψ satisfies the Assumption* (**A1**). *Then*

$$p_t(x) \leqslant e^{-D_t^2(x)} p_t(0) \quad \textit{for all} \quad x \in \mathbb{R}^n,\ t > 0, \tag{22}$$

where $D_t^2(x)$ is given by (21).

For the proof of Theorem 6 we need a few properties of the function D_t^2.

Lemma 7. i) *For all $t > 0$, the function $x \mapsto D_t^2(x), x \in \mathbb{R}^n$ is even and increasing as $|x| \to \infty$;*

ii) *$D_t(x) = 0$ if and only if $x = 0$.*

iii) *$D_t^2(x) \leqslant \dfrac{|x|^2}{4ct}$ for all $x \in \mathbb{R}^n$ and $t > 0$.*

Proof. i) Since $x \mapsto \Gamma(e^{\xi\cdot x}, e^{-\xi\cdot x})$ is even,

$$\min_\xi v_t(\xi, x) = \min_\xi v_t(-\xi, -x) = \min_{-\xi} v_t(\xi, -x),$$

and therefore $\xi_0(t, x) = -\xi_0(t, -x)$. This can be used to show

$$\begin{aligned} v_t(\xi_0(t,x), x) &= -\xi_0(t,x)\cdot x - tw(\xi_0(t,x)) \\ &= \xi_0(t,-x)\cdot x - tw(-\xi_0(t,x)) \\ &= -\xi_0(t,-x)\cdot(-x) - tw(\xi_0(t,-x)) \\ &= v_t(\xi_0(t,-x), -x), \end{aligned}$$

which implies that $x \mapsto D_t^2(x)$ is even. That $D_t^2(x)$ increases as $|x| \to \infty$ follows directly from the definition.

ii) Since $v_t(\xi_0, x)$ is non-positive and since $v_t(\xi_0, 0) = 0$ it is obvious that $D_t(x) = 0$ if, and only if, $x = 0$.

iii) By Taylor's theorem there exists a constant $c > 0$ such that

$$\int_{\mathbb{R}^n} \big(\cosh(\xi \cdot y) - 1\big)\nu(dy) \geqslant c|\xi|^2 \quad \text{for all } \xi \in \mathbb{R}^n.$$

Then for all $x \in \mathbb{R}^n$

$$v_t(\xi, x) = -\xi \cdot x - tw(\xi) \geqslant -\xi \cdot x + ct|\xi|^2;$$

if we minimize this expression, we get $-D_t^2(x) = v_t(\xi_0, x) \geqslant -|x|^2/(4ct)$. □

We are now ready for the

Proof of Theorem 6. In order to estimate $\int_{\mathbb{R}^n} e^{Q_t(\xi,x)}\, d\xi$ we apply the Cauchy-Poincaré theorem. Since w can be extended analytically to $\mathbb{C}^n$, the function $Q_t(z, x)$ is analytic in $z \in \mathbb{C}^n$. Without loss of generality we may assume that $\xi_0 > 0$, in the case $\xi_0 < 0$ the arguments are similar. Consider the domain

$$G := \left\{ z \in \mathbb{C}^n \ : \ \operatorname{Im} z = t\xi_0,\ 0 \leqslant t \leqslant 1,\ \operatorname{Re} z \in \prod_{j=1}^n [-M_j, M_j],\ M_j > 0,\ 1 \leqslant j \leqslant n \right\}.$$

This is an $n+1$-dimensional cube with base $\left\{ z \in \mathbb{C}^n \ : \ \operatorname{Re} z \in \prod_{j=1}^n [-M_j, M_j],\ \operatorname{Im} z = 0 \right\}$ and lid $\left\{ z \in \mathbb{C}^n \ : \ \operatorname{Re} z \in \prod_{j=1}^n [-M_j, M_j],\ \operatorname{Im} z = \xi_0 \right\}$. Since the number of sides of G is even, we can fix some orientation on ∂G such that base and lid have opposite orientation. By the Cauchy-Poincaré theorem

$$\int_{\partial G} e^{Q_t(z,x)}\, dz_1 \wedge dz_2 \wedge \cdots \wedge dz_n = 0. \tag{23}$$

Consider the integrals over the sides (except the base and the lid)

$$\int_0^1 e^{Q_t(M \pm is\xi_0, x)} ds, \quad \text{where} \quad M = (\pm M_1, \ldots, \pm M_n) \tag{24}$$

and recall that $Q_t(\xi, x) = i\xi \cdot x - tw(i\xi)$. After some rearrangements we get

$$\begin{aligned}
&\operatorname{Re} Q_t(M + i\eta, x) \\
&= \operatorname{Re}\left[i(M + i\eta) \cdot x - \frac{t}{2} \int_{\mathbb{R}^n} \big(e^{i(M+i\eta)\cdot y} - 1\big)\big(e^{-i(M+i\eta)\cdot y} - 1\big)\, \nu(dy) \right] \\
&= -\eta \cdot x - t \int_{\mathbb{R}^n} \big(1 - \cosh(\eta \cdot y)\big)\, \nu(dy) - t \int_{\mathbb{R}^n} \cosh(\eta \cdot y)(1 - \cos(M \cdot y))\, \nu(dy) \\
&\leqslant -t\psi(M) - \eta \cdot x - t\psi(i\eta).
\end{aligned}$$

Therefore,

$$\left| e^{Q_t(M \pm is\xi_0, x)} \right| \leqslant e^{-t\psi(M) - \eta \cdot x - t\psi(is\xi_0)},$$

which means that the integrands in (24) tend, uniformly in s, to 0 as $|M| \to \infty$. Therefore, (23) becomes, as $|M| \to \infty$,

$$\int_{\mathbb{R}^n} e^{Q_t(\xi,x)}\,d\xi = \int_{\mathbb{R}^n} e^{Q_t(\xi+i\xi_0,x)}\,d\xi.$$

Since

$$\begin{aligned}\operatorname{Re} Q_t(\xi+i\xi_0,x) &= v_t(\xi_0,x) - t\int_{\mathbb{R}^n} \cosh(\xi_0\cdot y)(1-\cos(\xi\cdot y))\,\nu(dy)\\ &\leqslant v_t(\xi_0,x) - t\psi(\xi),\end{aligned}$$

we finally get

$$\int_{\mathbb{R}^n} e^{Q_t(\xi,x)}\,d\xi \leqslant e^{v_t(\xi_0,x)}\int_{\mathbb{R}^n} e^{-t\psi(\xi)}\,d\xi = e^{-D_t^2(x)}p_t(0). \qquad \square$$

We can combine Theorem 6 with the on-diagonal estimates from Section 3, e.g. with Lemma 4. Note that all processes satisfying the assumptions of Lemma 4 automatically have transition densities.

Corollary 8. *Let $(X_t)_{t\geqslant 0}$ be a Lévy process with symbol ψ_1 as in (13) where f is a Bernstein function without linear term and such that $f(0)=0$. Then there exist suitable constants $c,\gamma>0$ such that*

$$p_t(x) \leqslant c e^{-D_t^2(x)} \left[f^{-1}\left(\frac{1}{\gamma t}\right)\right]^{n/2}, \quad \textit{for all}\ \ 0<t\leqslant 1,\ x\in\mathbb{R}^n. \tag{25}$$

Since the representing measure $\nu(dy)$ has compact support, (**A1**) is satisfied and the proof follows by Lemma 4 and Theorem 6. Note that we do not require the growth condition imposed on f in Lemma 4, since this was only used to show the equivalence of the Dirichlet forms $\mathcal{E}^{\psi_1}$ and $\mathcal{E}^{f(|\cdot|^2)}$.

Example 9. Let us indicate a few generic examples for our estimates. Since we require exponential moments, our Lévy measures ν must satisfy $\int_{|y|\geqslant 1} e^{\xi\cdot y}\,\nu(dy) < \infty$ for all $\xi\in\mathbb{R}^n$. Typically, this can be achieved if $\operatorname{supp}\nu$ is bounded, see Examples i) and ii) below, or if ν has an exponentially fast decaying density, as in Example iii). For simplicity we restrict ourselves to the one-dimensional setting $n=1$ and estimates for $p_t(x)$ with $x/t\to\infty$, e.g. where $x \gg 1$ and $t>0$ is fixed.

i) Assume that $\operatorname{supp}\nu \subset [-1,1]$ and that $\nu(dy) = g(y)\,dy$ for some density function $g(y)$. This covers the situation of (13) and Corollary 8.

   For our heat kernel estimate we need to control the behaviour of the function $v_t(\xi_0,x)$ from (20) at the point $\xi_0 = \arg\min_\xi v_t(\xi,x)$. For symmetry reasons we only have to consider the case where $\xi>0$ and $x>0$. Clearly,

$$v_t(\xi,x) \leqslant -\xi x + c_0 t\xi^2 e^{\xi}\int_{-1}^{1} y^2\,g(y)\,\nu(dy) \leqslant -\xi x + c_1 t e^{\xi(1+\varepsilon)}$$

for some for some $\varepsilon > 0$ and suitable constants c_0 and c_1. We will now minimize the expression on the right-hand side. From

$$\frac{\partial}{\partial \xi}\left(-\xi x + c_1 t e^{\xi(1+\varepsilon)}\right) = -x + tc_1(1+\varepsilon)\, e^{\xi(1+\varepsilon)} \stackrel{!}{=} 0$$

we find the critical point for the minimum of the right-hand side

$$\xi = \frac{1}{1+\varepsilon} \ln\left(\frac{x}{tc_1(1+\varepsilon)}\right),$$

and so

$$v_t(\xi_0, x) \leqslant \min_{\xi}\Big(-\xi x + c_1 t e^{\xi(1+\varepsilon)}\Big) = -\frac{x}{1+\varepsilon}\ln\left(\frac{x}{tc_1(1+\varepsilon)}\right) + \frac{x}{1+\varepsilon}.$$

If, as in Corollary 8, $g(y) = f\big(|y|^{-2}\big)/|y|^n$ the upper bound becomes for $x \neq 0$

$$\begin{aligned} p_t(x) &\leqslant c\, e^{-D_t^2(x)} \left[f^{-1}\left(\frac{1}{\gamma t}\right)\right]^{1/2} \\ &\leqslant e^{-\frac{|x|}{1+\varepsilon}\ln\left(\frac{|x|}{tc_1(1+\varepsilon)}\right) + \frac{|x|}{1+\varepsilon}} \left[f^{-1}\left(\frac{1}{\gamma t}\right)\right]^{1/2} \end{aligned}$$

for suitable constants $\gamma, \varepsilon > 0$.

ii) Assume that $\operatorname{supp} \nu \subset [-1, 1]$ and that ν is discrete. Let us consider the case where $\nu(dy) = \sum_{n=0}^{\infty} 2^{\alpha n}\,(\delta_{2^{-n}} + \delta_{-2^{-n}})$ and $0 < \alpha < 2$. Since $\nu_0 = \sum_{n=-\infty}^{\infty} 2^{\alpha n}(\delta_{2^{-n}} + \delta_{-2^{-n}})$ corresponds to the so-called *α-semi-stable process*, see [31, Example 13.3], we can use [31, Proposition 24.20] to get $\psi_0(\xi) \geqslant c\,|\xi|^{\alpha}$ for some $c > 0$ and all $|\xi| \geqslant 1$. Since the characteristic exponents ψ and ψ_0 corresponding to ν and ν_0 satisfy $\psi(\xi) \asymp \psi_0(\xi)$ for large values of $|\xi|$, we also have $\psi(\xi) \geqslant c'\,|\xi|^{\alpha}$, $|\xi| \geqslant 1$. This means, in particular, that a transition density $p_t(x)$ exists.

Note that all our calculations for the upper estimate for $v_t(\xi, x)$ from the first example remain valid and, what is more, up to the constant $c_1 = c_1(\nu)$ they depend only on the (size of the) support of the Lévy measure, but not on the particular form of ν. This means that we can also in this case estimate the transition density by

$$p_t(x) \leqslant c_2\, t^{-\frac{1}{\alpha}}\, e^{-\frac{|x|}{1+\varepsilon}\ln\left(\frac{|x|}{tc_1(1+\varepsilon)}\right) + \frac{|x|}{1+\varepsilon}}.$$

iii) Assume that $\nu(dy) = \nu_0(dy) + \mathbb{1}_{|y| \geqslant 1}\, e^{-|y|^{\beta}}\, dy$ where $\beta > 1$ and where ν_0 is as in Example i) or ii).

In this case the behaviour of $v_t(\xi, x)$ is determined by the tail of the measure. As before, we first estimate $v_t(\xi, x)$ from above for $x > 0$ and $\xi > 0$:

$$v_t(\xi, x) \leqslant -\xi x + c_1 t\, e^{\xi(1+\varepsilon)} + c_2 t \int_1^{\infty} e^{\xi y - |y|^{\beta}}\, dy.$$

In order to find the asymptotics of $I_1(\xi) := \int_1^\infty e^{\xi y - |y|^\beta}\,dy$ we use the Laplace method, see [14, §18, p. 58]. It is known that for sufficiently smooth functions h the integral $I_h(\xi) := \int_a^b e^{h(\xi,y)}\,dy$ where $a, b \in [-\infty, +\infty]$ satisfies

$$I_h(\xi) \sim \sqrt{\frac{\pi}{2|h''(\xi,y_0)|}}\, e^{h(\xi,y_0)}, \quad \text{as} \quad \xi \to \infty.$$

In the expression above y_0 is the (unique) point where h reaches its maximum. If we use $h(\xi,y) = \xi y - y^\beta$ and $y_0 = \left(\frac{\xi}{\beta}\right)^{\frac{1}{\beta-1}}$, we get

$$\begin{aligned} h(\xi,y_0) &= -c_{\beta,1}\,\xi^{\frac{\beta}{\beta-1}} \quad \text{where} \quad c_{\beta,1} = (\beta-1)\beta^{\frac{\beta}{1-\beta}} \\ |h''(\xi,y_0)| &= \beta(\beta-1)\beta^{\frac{1}{1-\beta}}\,\xi^{\frac{\beta-2}{\beta-1}}, \\ I_1(\xi) &\asymp c_{\beta,2}\,\xi^{\frac{2-\beta}{2(\beta-1)}}\,e^{c_{\beta,1}\xi^{\frac{\beta}{\beta-1}}}, \quad \xi \to \infty. \end{aligned}$$

Observe that for $\xi \to \infty$

$$v_t(\xi,x) \leqslant f(\xi,x,t) = -\xi x + t\,c_{\beta,2}\,\xi^{\frac{2-\beta}{2(\beta-1)}}\,e^{c_{\beta,1}\xi^{\frac{\beta}{\beta-1}}}.$$

The point where $f(\cdot,x,t)$ becomes extremal satisfies the equation

$$\begin{aligned} &0 \overset{!}{=} f'_\xi(\xi,x,t) \\ &= -x + t\,c_{\beta,2}\,\frac{2-\beta}{2(\beta-1)}\,\xi^{\frac{2-\beta}{2(\beta-1)}-1}e^{c_{\beta,1}\,\xi^{\frac{\beta}{\beta-1}}} + t\,c_{\beta,2}c_{\beta,1}\,\frac{\beta}{\beta-1}\,\xi^{\frac{\beta}{\beta-1}+\frac{2-\beta}{2(\beta-1)}-1}\,e^{c_{\beta,1}\,\xi^{\frac{\beta}{\beta-1}}}. \end{aligned}$$

Rather than solving this equation for ξ explicitly, we determine the asymptotic behaviour of the solution as $x/t \to \infty$. Note that $f'_\xi(\xi,x,t) = 0$ if, and only if,

$$c_{\beta,2}\,\frac{2-\beta}{2(\beta-1)}\,\xi^{\frac{2-\beta}{2(\beta-1)}-1}\,e^{c_{\beta,1}\xi^{\frac{\beta}{\beta-1}}} + c_{\beta,2}c_{\beta,1}\,\frac{\beta}{\beta-1}\,\xi^{\frac{\beta}{\beta-1}+\frac{2-\beta}{2(\beta-1)}-1}\,e^{c_{\beta,1}\xi^{\frac{\beta}{\beta-1}}} = \frac{x}{t}.$$

Taking logarithms on both sides we arrive at

$$\xi = \left(\frac{1}{c_{\beta,1}}\log\left(\frac{x}{t}\right)\right)^{\frac{\beta-1}{\beta}} + o\left(\log\frac{x}{t}\right) \quad \text{as} \quad \frac{x}{t} \to \infty,$$

hence,

$$v_t(\xi_0,x) \leqslant -(1-\varepsilon)x\,\left(\frac{1}{c_{\beta,1}}\,\log\left(\frac{x}{t}\right)\right)^{\frac{\beta-1}{\beta}}.$$

as well as

$$p_t(x) \leqslant p_t(0)e^{-(1-\varepsilon)x\left(\frac{1}{c_{\beta,1}}\,\log\left(\frac{x}{t}\right)\right)^{\frac{\beta-1}{\beta}}}.$$

With considerably more effort it is possible to obtain the exact asymptotics of $v_t(\xi_0,x)$ and $p_t(x)$, see [27, Proposition 6.1] by Kulik and one of the present authors.

5 An application to large deviations

In this section we will show an application of Theorem 6 to the theory of large deviations. We show that the transition density (17) satisfies the *large deviation principle* (LDP) with the rate function $D_t^2(x)$.

Let us briefly recall the LDP. Let $(X_t)_{t\geqslant 0}$ be a Lévy process associated with transition function $\mu_t(dx)$. Moreover, we assume that X_t has exponential moments, i.e.

$$\int_{\mathbb{R}^n} e^{y\cdot\lambda}\mu_t(dy) < \infty \quad \text{for all } \lambda \in \mathbb{R}^n \text{ and } t > 0. \tag{\textbf{A2}}$$

By [31, Theorem 23.5], this is equivalent to our assumption (**A1**). Therefore, we can extend ψ analytically from $\mathbb{R}^n$ to $\mathbb{C}^n$.

Let

$$\Lambda_\mu^*(x,t) := \sup_\xi \{\xi \cdot x - \Lambda_\mu(\xi,t)\},$$

where

$$\Lambda_\mu(\xi,t) := \log \int_{\mathbb{R}^n} e^{\xi\cdot y}\mu_t(dy) = t\psi(i\xi).$$

By $\mu_t^{(\ell)}(dx)$ we denote the probability measure related to $Y_t^{(\ell)} := \frac{1}{\ell}\sum_{j=1}^{\ell} X_t^j$, where X_t^j are independent copies of X_t. It is known, see e.g. [17, Chapter 3], that under (**A1**) the sequence of measures $(\mu_t^{(\ell)}(dx))_{\ell\geqslant 1}$ is exponentially tight and, by Cramer's theorem, it satisfies the LDP with *good rate function* $\Lambda_\mu^*(x,t)$, i.e. for all measurable subsets $B \subset \mathbb{R}^n$ the inequalities

$$-\inf_{x\in B^\circ} \Lambda_\mu^*(x,t) \leqslant \varliminf_{\ell\to\infty} \frac{1}{\ell} \log \mu_t^{(\ell)}(B) \leqslant \varlimsup_{\ell\to\infty} \frac{1}{\ell} \log \mu_t^{(\ell)}(B) \leqslant -\inf_{x\in\overline{B}} \Lambda_\mu^*(x,t).$$

hold; B° and $\overline{B}$ are the open interior and the closure of B, respectively. Clearly, $\Lambda_\mu^*(x,t) = -v_t(\xi_0,x) = D_t^2(x)$. By Theorem 6 we have the analogue of the LDP for the transition density $p_t(x)$ from Theorem 6.

Proposition 10. *Let $(X_t)_{t\geqslant 0}$ be a Lévy process with symbol $\psi(\xi)$ satisfying*

$$\lim_{|\xi|\to\infty} \frac{\psi(\xi)}{\log(1+|\xi|)} > C. \tag{26}$$

*Assume that (**A1**), or equivalently (**A2**), holds. Then X_t has for all $t > t_0 := n/C$ a transition density $p_t(x)$ and for all $t > t_0$*

$$\lim_{\ell\to\infty} \frac{\log p_{\ell t}(\ell x)}{\ell} = -D_t^2(x). \tag{27}$$

Proof. Condition (26) is the Hartman-Wintner condition which ensures that X_t has a (continuous) transition density for all $t > t_0 = n/C$, see [18]. Clearly, we may follow the arguments of the proof of Theorem 6 whenever $p_t(x)$ exists, i.e. for $t > t_0$.

As in the proof of Theorem 6, we can write the transition density $p_t(x)$ as

$$p_t(x) = e^{v_t(x,\xi_0)} \int_{\mathbb{R}^n} e^{A_{t,x,\xi_0}(\xi)}\, d\xi,$$

where ξ_0, $v_t(x,\xi_0)$ are as before, and $A_{t,x,\xi_0} := Q_t(\xi + i\xi_0, x) - v_t(\xi_0, x)$. Note that

$$\left| \int_{\mathbb{R}^n} e^{A_{t,x,\xi_0}(\xi)}\, d\xi \right| \leqslant \left| \int_{\mathbb{R}^n} e^{A_{t,0,\xi_0}(\xi)}\, d\xi \right| \leqslant p_t(0). \tag{28}$$

By (26) we obtain for $k \geqslant t_0$

$$\begin{aligned} \frac{\log p_k(0)}{k} \leqslant \log \left(\int_{\mathbb{R}^n} e^{-kC' \log(1+|\xi|)} d\xi \right)^{1/k} &= \log \left(\int_{\mathbb{R}^n} \frac{d\xi}{(1+|\xi|)^{kC'}} \right)^{1/k} \\ &\asymp \frac{1}{k} \log \frac{1}{kC'} \xrightarrow{k\to\infty} 0. \end{aligned}$$

Because of (28)

$$\lim_{\ell\to\infty} \frac{\log \int_{\mathbb{R}^n} e^{A_{\ell t,x,\xi_0}(\xi)}\, d\xi}{\ell} = 0,$$

and since $v_{\ell t}(\xi, \ell x) = -\ell\xi \cdot x - \ell t w(\xi)$, it is clear that $\xi_0 = \arg\min_\xi v_t(\xi, x)$ does not depend on ℓ. Hence,

$$-D^2_{\ell t}(x) = v_{\ell t}(\xi_0, \ell x) = \ell v_t(\xi_0, x) = -\ell D^2_t(x).$$

Combining the last two formulae we get

$$\lim_{\ell\to\infty} \frac{\log p_{\ell t}(\ell x)}{\ell} = \lim_{\ell\to\infty} \frac{\log \int_{\mathbb{R}^n} e^{A_{\ell t,x,\xi_0}(\xi)}\, d\xi}{\ell} + \lim_{\ell\to\infty} \frac{\log e^{-\ell D^2_t(x)}}{\ell} = -D^2_t(x). \qquad \square$$

6 Estimates for Lévy-type processes

In this section we generalize the results obtained in Section 4 to the case of pseudo-differential operators with continuous negative definite symbol. We show that under some conditions one can construct an upper bound in the form similar to (22) for the transition density of a Markov process related to a pseudo-differential operator. Typical examples are Feller processes such that the test functions C_0^∞ are in the domain $D(A)$ of their generator, cf. [22, 24].

Let $u \in C_0^\infty(\mathbb{R}^n)$. Consider the operator

$$q(x, D)u(x) = \int_{\mathbb{R}^n} e^{i\xi\cdot x} q(x,\xi)\, \widehat{u}(\xi)\, d\xi, \tag{29}$$

where $q : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is locally bounded and for each $x \in \mathbb{R}^n$ the function $q(x,\cdot)$ is continuous negative definite. This means that $q(x,\xi)$ admits a Lévy-Khintchine representation:

$$q(x,\xi) = \frac{1}{2}\, \xi \cdot Q(x)\xi + \int_{\mathbb{R}^n\setminus\{0\}} (1 - \cos(y\cdot\xi))\; N(x, dy), \quad x \in \mathbb{R}^n, \tag{30}$$

where $Q(x) = (Q_{jk}(x)) \in \mathbb{R}^{n\times n}$ is a positive semi-definite matrix, and $N(x, dy)$ is a Lévy kernel, i.e. for fixed $x \in \mathbb{R}^n$ it is a Borel measure on $\mathbb{R}^n \setminus \{0\}$, such that

$$\int_{\mathbb{R}^n\setminus\{0\}} (|y|^2 \wedge 1)\, N(x, dy) < \infty.$$

Such an operator $q(x, D)$ is called a pseudo-differential operator with real-valued continuous negative definite symbol. Using Fourier inversion the following integro-differential representation for $u \in C_0^\infty(\mathbb{R}^n)$ is easily derived:

$$-q(x, D)u(x) = \frac{1}{2} \sum_{j,k=1}^{n} Q_{jk}(x)\partial_j\partial_k u(x) + \int_{\mathbb{R}^n\setminus\{0\}} (u(x + y) - u(x))\, N(x, dy). \tag{31}$$

In the sequel we will need a few further assumptions.

$$(-q(x, D), C_0^\infty(\mathbb{R}^n)) \text{ extends to the generator } (A, D(A)) \text{ of a Feller semigroup.} \tag{B1}$$

Remark 11. For sufficient conditions when (**B1**) is satisfied we refer to [21, Theorem 5.2], [20, Theorem 4.14], [23, Section 2.6] or the survey paper [24] and the references given there.

Denote by $(T_t)_{t\geqslant 0}$ the Feller semigroup and by $(X_t)_{t\geqslant 0}$ the Feller process generated by (the extension of) $-q(x, D)$. We set

$$\lambda_t(x, \xi) := e^{-i\xi\cdot x}T_t(e^{i\xi\bullet})(x). \tag{32}$$

Since $q(x, \xi)$ is real-valued, $\lambda_t(x, \xi) = \lambda_t(x, -\xi)$ for all $x, \xi \in \mathbb{R}^n$ and $t > 0$. Writing $p_t(x, dy)$ for the transition function of the process, it is easy to see that each T_t is a pseudo-differential operator with symbol $\lambda_t(x, \xi)$:

$$T_t u(x) = \int_{\mathbb{R}^n} u(y)p_t(x, dy) = \int_{\mathbb{R}^n} e^{i\xi\cdot x}\lambda_t(x, \xi)\, \widehat{u}(\xi)\, d\xi, \quad u \in C_0^\infty(\mathbb{R}^n). \tag{33}$$

From (33) we see that if $\lambda_t(x, \cdot) \in L_1(\mathbb{R}^n)$, the probability measures $p_t(x, dy)$ have densities $p_t(x, y)$ w.r.t. Lebesgue measure dy and we see, cf. [22, Theorem 3.2.1], that

$$\|T_t u\|_{L_\infty} \leqslant \|\lambda_t(x, \cdot)\|_{L_1} \cdot \|\widehat{u}\|_{L_\infty} \leqslant \|\lambda_t(x, \cdot)\|_{L_1} \cdot \|u\|_{L_1}.$$

Thus,

$$p_t(x, y) \leqslant \|\lambda_t(x, \cdot)\|_{L_1}, \quad \text{for all} \quad x, y \in \mathbb{R}^n,\ t > 0, \tag{34}$$

and we can express $p_t(x, y)$ in terms of the symbol $\lambda_t(x, \xi)$:

$$p_t(x, y) = \int_{\mathbb{R}^n} e^{i\xi\cdot(x-y)}\lambda_t(x, \xi)\, d\xi. \tag{35}$$

In addition to (**B1**) we will also need

$$\lambda_t(x,\cdot) \in L_1(\mathbb{R}^n) \quad \text{for all } x \in \mathbb{R}^n,\ t > 0. \tag{\textbf{B2}}$$

$$\sup_{z\in\mathbb{R}^n} \int_{0<|y|\leqslant 1} |y|^2 N(z,dy) + \sup_{z\in\mathbb{R}^n} \int_{|y|>1} e^{\zeta\cdot y} N(z,dy) < \infty \quad \text{for all } \zeta \in \mathbb{R}^n. \tag{\textbf{B3}}$$

$$\left| \frac{\lambda_t(x,\eta+i\xi)}{\lambda_t(x,i\xi)} \right| \leqslant |\lambda_t(x,\eta)| \quad \text{for all } \xi,\eta \in \mathbb{R}^n,\ t>0. \tag{\textbf{B4}}$$

Note that (**B3**) entails that the process $(X_t)_{t\geqslant 0}$ has exponential moments.

Lemma 12. *Let $-q(x,D)$ be a pseudo-differential operator satisfying* (**B1**) *and* (**B3**)*. Then the Feller process $(X_t)_{t\geqslant 0}$ generated by (the extension of) $-q(x,D)$ admits exponential moments:*

$$\mathbb{E}^x e^{\zeta\cdot X_t} = \int_{\mathbb{R}^n} e^{\zeta\cdot y}\, p_t(x,y)\,dy < \infty \tag{36}$$

for all $\zeta \in \mathbb{R}^n$ and $x \in \mathbb{R}^n$.

Proof. Write $(A, D(A))$ for the generator of the process and its semigroup $(T_t)_{t\geqslant 0}$. Since $(X_t)_{t\geqslant 0}$ is a Feller process we know that for all $u \in D(A)$ the process $M_t^u := u(X_t) - u(X_0) - \int_0^t Au(X_s)\,ds$ is a martingale. In particular,

$$T_t u(x) - u(x) = -\int_0^t T_s q(x,D)u(x)\,ds, \tag{37}$$

for all $u \in C_0^\infty(\mathbb{R}^n)$.

Let us show that (37) still holds for $u(x) := u_{z,\zeta}(x) := \cosh\big(\zeta\cdot(x-z)\big)$, $\zeta \in \mathbb{R}^n$. For this pick a sequence of test functions $\chi_\ell \in C_0^\infty(\mathbb{R}^n)$, $\ell \in \mathbb{N}$, with $\mathbb{1}_{B(0,\ell)} \leqslant \chi_\ell \leqslant \mathbb{1}_{B(0,2\ell)}$; observe that $x \mapsto \chi_\ell(x-z)\cosh\big(\zeta\cdot(x-z)\big)$ is in $D(A)$. Therefore we can rewrite (37) in the following way

$$\begin{aligned}
&\mathbb{E}^x\Big[\chi_\ell(X_t-z)\cosh\big(\zeta\cdot(X_t-z)\big)\Big]\Big|_{z=x} - 1\\
&= \mathbb{E}^x \int_0^t \int_{\mathbb{R}^n\setminus\{0\}} \Big[\chi_\ell(X_s-x+y)\cosh\big(\zeta\cdot(X_s-x+y)\big) - \chi_\ell(X_s-x)\cosh\big(\zeta\cdot(X_s-x)\big)\Big]\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \times N(X_s,dy)\,ds
\end{aligned}$$

Since we are integrating with respect to Lebesgue measure ds and since $s \mapsto X_s$ has almost surely at most countably many jumps, we may replace X_s in the above formula by its left limit X_{s-}. Moreover, set

$$\tau_k := \inf\{s \geqslant 0 \,:\, |X_s - x| \geqslant k\}.$$

Using optional stopping for the martingale M_t^u and writing $X_s^{\tau_k} := X_{\tau_k \wedge s}$ for the stopped process, we get from the identity above

$$
\begin{aligned}
&\mathbb{E}^x\Big[\chi_\ell(X_t^{\tau_k} - x)\cosh\big(\zeta\cdot(X_t^{\tau_k}-x)\big)\Big] - 1\\
&= \left|\mathbb{E}^x \int_0^{t\wedge\tau_k}\int_{\mathbb{R}^n\setminus\{0\}} (\cdots)\, N(X_{s-},dy)\,ds\right|\\
&\leqslant \mathbb{E}^x \int_0^{t\wedge\tau_k}\int_{\mathbb{R}^n\setminus\{0\}} |(\cdots)|\, N(X_{s-},dy)\,ds\\
&\leqslant \mathbb{E}^x \int_0^{t\wedge\tau_k}\int\limits_{0<|y|\leqslant 1} |(\cdots)|\, N(X_{s-},dy)\,ds + \mathbb{E}^x \int_0^{t\wedge\tau_k}\int\limits_{|y|>1} |(\cdots)|\, N(X_{s-},dy)\,ds.
\end{aligned} \tag{38}
$$

Let us estimate the last two integrals separately. For $\ell \geqslant k+1$, $|y|\leqslant 1$ and $s \leqslant \tau_k \wedge t$ we have $\chi_\ell(X_{s-}-x-y) = \chi_\ell(X_{s-}-x) = 1$. Thus,

$$
\begin{aligned}
&\mathbb{E}^x \int_0^{t\wedge\tau_k}\int_{0<|y|\leqslant 1} |(\cdots)|\, N(X_{s-},dy)\,ds\\
&= \mathbb{E}^x \int_0^{t\wedge\tau_k}\int_{0<|y|\leqslant 1} \big|\cosh\big(\zeta\cdot(X_{s-}-x+y)\big) - \cosh\big(\zeta\cdot(X_{s-}-x)\big)\big|\, N(X_{s-},dy)\,ds\\
&\leqslant \mathbb{E}^x \int_0^{t\wedge\tau_k}\int_{0<|y|\leqslant 1} \cosh\big(\zeta\cdot(X_{s-}-x)\big)\cdot\big(\cosh(\zeta\cdot y)-1\big)\, N(X_{s-},dy)\,ds\\
&\leqslant \frac{1}{2}|\,\zeta|^2 e^{|\zeta|}\, \mathbb{E}^x \int_0^{t\wedge\tau_k}\int_{0<|y|\leqslant 1} \cosh\big(\zeta\cdot(X_{s-}-x)\big)\, |y|^2\, N(X_{s-},dy)\,ds\\
&\leqslant \frac{1}{2}\,|\zeta|^2 e^{|\zeta|}\, t\, \sup_{s<t}\mathbb{E}^x\Big(\cosh\big(\zeta\cdot(X_{s-}^{\tau_k}-x)\big)\Big) \sup_{z\in\mathbb{R}^n}\int_{0<|y|\leqslant 1} |y|^2\, N(z,dy)\,ds
\end{aligned}
$$

where we used the elementary inequalities $|\cosh(a+b)-\cosh(a)| \leqslant \cosh(a)(\cosh(b)-1)$ and $\cosh(b)-1 \leqslant \frac{1}{2}\, b^2\, e^{|b|}$.

A similar calculation using the fact that $|\chi_\ell| \leqslant 1$ and $\cosh(a+b) \leqslant e^{|a|}\cosh(b)$ yields for the second integral term in (38)

$$
\begin{aligned}
&\mathbb{E}^x \int_0^{t\wedge\tau_k}\int_{|y|>1} |(\cdots)|\, N(X_{s-},dy)\,ds\\
&\leqslant \mathbb{E}^x \int_0^{t\wedge\tau_k}\int_{|y|>1} \Big(\big|\cosh\big(\zeta\cdot(X_{s-}-x+y)\big)\big| + \big|\cosh\big(\zeta\cdot(X_{s-}-x)\big)\big|\Big) N(X_{s-},dy)\,ds\\
&\leqslant t\, \sup_{s<t}\mathbb{E}^x\Big(\cosh\big(\zeta\cdot(X_{s-}^{\tau_k}-x)\big)\Big) \sup_{z\in\mathbb{R}^n}\int_{|y|>1} \big(e^{|\zeta\cdot y|}+1\big)\, N(z,dy).
\end{aligned}
$$

Because of (**B3**) we find a constant $C = C_\zeta$ not depending on k or x such that for all $t > 0$

$$\begin{aligned}\mathbb{E}^x\left[\cosh\left(\zeta\cdot(X_t^{\tau_k}-x)\right)\right] &= \sup_{\ell\in\mathbb{N}}\mathbb{E}^x\left[\chi_\ell(X_t^{\tau_k}-x)\cosh\left(\zeta\cdot(X_t^{\tau_k}-x)\right)\right]\\ &\leqslant 1+t\,C\sup_{s\leqslant t}\mathbb{E}^x\left[\cosh\left(\zeta\cdot(X_{\tau_k\wedge s-}-x)\right)\right]\\ &\leqslant 1+t\,C\sup_{s\leqslant t}\left(\gamma_k\wedge\mathbb{E}^x\left[\cosh\left(\zeta\cdot(X_s^{\tau_k}-x)\right)\right]\right).\end{aligned}$$

where $\gamma_k > \sup_{|y|\leqslant k}\cosh\left(\zeta\cdot(y-x)\right)$. Since $t > 0$ was arbitrary and since the right-hand side depends monotonically on t, the above calculation also gives

$$\sup_{s\leqslant t}\mathbb{E}^x\cosh\left(\zeta\cdot(X_s^{\tau_k}-x)\right)\leqslant 1+t\,C\sup_{s\leqslant t}\left(\gamma_k\wedge\mathbb{E}^x\cosh\left(\zeta\cdot(X_s^{\tau_k}-x)\right)\right).$$

Estimating the left-hand side trivially from below, and choosing $t < t_0 < 1/(2C)$, we find

$$\sup_{s\leqslant t}\Big(\gamma_k\wedge\mathbb{E}^x\cosh\left(\zeta\cdot(X_s^{\tau_k}-x)\right)\Big)-\frac{1}{2}\sup_{s\leqslant t}\Big(\gamma_k\wedge\mathbb{E}^x\cosh\left(\zeta\cdot(X_s^{\tau_k}-x)\right)\Big)\leqslant 1.$$

Note that (**B3**) entails that the generator $-q(x, D)$ has bounded 'coefficients', i.e. the life-time of X_t is a.s. infinite and $\lim_{k\to\infty}\tau_k = \infty$, see [32]. By Fatou's Lemma we get

$$\sup_{s\leqslant t}\mathbb{E}^x\cosh\left(\zeta\cdot(X_s-x)\right)\leqslant 2\quad\text{for all}\quad x\in\mathbb{R}^n,\ t\in(0,t_0].$$

Note that $t_0 = t_0(\zeta)$. Now it is a simple exercise using the Markov property to show that

$$\mathbb{E}^x\cosh\left(\zeta\cdot(X_t-x)\right)<\infty\quad\text{for all}\quad t>0. \qquad\square$$

Here are some examples of Markov processes for which the conditions (**B1**)–(**B4**) are satisfied.

Example 13. Let

$$\mathcal{L} := \sum_{j,k=1}^n\frac{\partial}{\partial x_j}\left(a_{jk}(x)\frac{\partial}{\partial x_k}\right),$$

where $(a_{jk}(x))_{j,k=1}^n$ is a symmetric positive definite matrix with bounded measurable coefficients. It is known that there exists the transition density $p_t^{(0)}(x, y)$ of the diffusion semigroup associated with $\mathcal{L}$, which satisfies Aronson's estimates, see [1] and [15], and which is jointly Hölder continuous in x and y. Let $\psi^{(0)}$ be a continuous negative definite function satisfying the conditions of Theorem 6. Then the operator $\mathcal{L} + \psi^{(0)}(D)$ satisfies (**B1**)–(**B4**). Indeed, in this case the symbol $\lambda_t(x, \xi)$ associated with $\mathcal{L} + \psi^{(0)}(D)$ is the product of two symbols, associated with $\mathcal{L}$ and with $\psi^{(0)}(D)$, both satisfying (**B1**)–(**B4**).

Example 14. Consider a Markov process in $\mathbb{R}_+^m \times \mathbb{R}^n$, such that for every $t > 0$ the characteristic function $\lambda_t(x,\xi)$ has exponential affine dependence on x. That is, for every $(t,\xi) \in \mathbb{R}_+ \times i\mathbb{R}^{n+m}$ there exist $\Phi(t,\xi) \in \mathbb{C}$, $\Psi(t,\xi) = (\Psi^Y(t,\xi), \Psi^Z(t,\xi)) \in \mathbb{C}^m \times \mathbb{C}^n$, such that for all $x \in \mathbb{R}_+^m \times \mathbb{R}^n$

$$\lambda_t(x,\xi) = e^{\Phi(t,\xi)+(\Psi(t,\xi),x)}. \tag{39}$$

A Markov process X_t with such a characteristic function is called an *affine process*. Such processes have been recently considered in mathematical finance, cf. [16].

If an affine process is regular—i.e. $q(x,\xi) := \partial_t \lambda_t(x,\xi)\big|_{t=0}$ exists for all x and all $\xi \in \{z = (z_1,\dots,z_m) \in \mathbb{C}^m \;:\; \operatorname{Im} z_j \geqslant 0, j = 1,\dots,m\} \times \mathbb{R}^n$ and is continuous at $\xi = 0$—, then X_t is a Feller process, see [16, §8], hence (**B1**) is satisfied. Since for affine processes we explicitly know the representation of the characteristic function, it is easy to find conditions in terms of Φ and Ψ such that (**B2**)–(**B4**) hold. For example, (**B2**)–(**B4**) are satisfied for $\Phi = \Psi^Y = 0$, $\Psi^Z(\xi) = (\psi_{t,1}(\xi),\dots,\psi_{t,n}(\xi))$ where $\psi_{t,j}(\xi)$, $j = 1\dots n$, are continuous negative definite functions satisfying the conditions of Theorem 6 for all values of the parameter t.

Note that (36) is equivalent to the finiteness of the following integral:

$$w_t(x,\zeta) := \ln\left[e^{x\cdot\zeta}\int_{\mathbb{R}^n} e^{-\zeta\cdot y} p_t(x,y)\,dy\right] = \ln\lambda_t(x,i\zeta) \quad \text{for all} \quad \zeta \in \mathbb{R}^n, \tag{40}$$

which is a convex function of ζ. Indeed, let $0 < \alpha < 1$, $\zeta,\xi \in \mathbb{R}^n$. Then

$$\begin{aligned}
&\alpha \ln \int_{\mathbb{R}^n} e^{\zeta\cdot y} p_t(x,y)\,dy + (1-\alpha)\ln\int_{\mathbb{R}^n} e^{\xi\cdot y}\, p_t(x,y)\,dy \\
&\quad= \ln\left[\left(\int_{\mathbb{R}^n} e^{\zeta\cdot y} p_t(x,y)\,dy\right)^{\alpha}\left(\int_{\mathbb{R}^n} e^{\xi\cdot y}\, p_t(x,y)\,dy\right)^{1-\alpha}\right] \\
&\quad\geqslant \ln\int_{\mathbb{R}^n} e^{\alpha\zeta\cdot y+(1-\alpha)\xi\cdot y}\, p_t(x,y)\,dy.
\end{aligned}$$

Since $\nabla_\zeta \ln\int_{\mathbb{R}^n} e^{\zeta\cdot y}\, p_t(x,y)\,dy\Big|_{\zeta=0} = 0$, the function

$$v_t(x-y,x,\zeta) := -\zeta\cdot(x-y) + \ln\lambda_t(x,i\zeta) = -\zeta\cdot y + \ln\int_{\mathbb{R}^n} e^{-\zeta\cdot h}\, p_t(x,h)\,dh \tag{41}$$

has a minimum, see [36, Proposition 47.12]. Define

$$\zeta_0 = \zeta_0(t,x,y) := \arg\min_{\zeta} v_t(x-y,x,\zeta). \tag{42}$$

By construction, there exists the analytic extension of $v_t(x-y,x,\cdot)$ to $\mathbb{C}^n$. This means that the arguments of Section 4 can be used to show the next theorem.

Theorem 15. *Let $q(x,D)$ be defined by (29), and suppose that (**B1**)–(**B4**) are satisfied. Then the transition density of the probability measure associated with $q(x,D)$ exists and satisfies*

$$p_t(x,y) \leqslant e^{v_t(x-y,x,\zeta_0)}\|\lambda_t(x,\cdot)\|_{L_1}, \quad \textit{for all} \quad x,y \in \mathbb{R}^n,\ t > 0, \tag{43}$$

where $v_t(x-y,x,\zeta)$ and ζ_0 are defined by (41) and (42), respectively.

Proof. Since

$$p_t(x,y) = \int_{\mathbb{R}^n} e^{i\xi\cdot(x-y)}\lambda_t(x,\xi)\,d\xi, \tag{44}$$

we can use, under our assumptions on $\lambda_t(x,\xi)$, the same approach as in Section 4. By Cauchy's theorem we get

$$p_t(x,y) = \int_{\mathbb{R}^n} e^{-\zeta_0\cdot(x-y)+i\eta\cdot(x-y)}\lambda_t(x,\eta+i\zeta_0)\,d\eta. \tag{45}$$

Hence, by condition (**B4**),

$$\begin{aligned} p_t(x,y) &\leqslant e^{v_t(x-y,x,\zeta_0)} \int_{\mathbb{R}^n} \left| \frac{\lambda_t(x,\eta+i\zeta_0)}{\lambda_t(x,i\zeta)} \right| d\eta \\ &\leqslant e^{v_t(x-y,x,\zeta_0)} \int_{\mathbb{R}^n} |\lambda_t(x,\eta)|\,d\eta, \end{aligned}$$

which proves (43). □

We can get an upper bound for the transition density $p_t(x,y)$ in terms of the symbol $q(x,\xi)$ and some remainder term. For this we need further assumptions, e.g. that the symbol of the generator belongs to a certain symbol class introduced in [19], see also [23, Definitions 2.4.3 and 2.4.4].

To state the corollary of Theorem 15 we need a few basic facts of symbol classes. Let $\rho(|\alpha|) := |\alpha| \wedge 2$.

Definition 16. i) A continuous negative definite function $\psi : \mathbb{R}^n \to \mathbb{R}$ belongs to the class Λ if for all $\alpha \in \mathbb{N}_0^n$ there exists a constant $c_{|\alpha|} \geqslant 0$ such that

$$\left|\partial_\xi^\alpha(1+\psi(\xi))\right| \leqslant c_{|\alpha|}(1+\psi(\xi))^{\frac{2-\rho(|\alpha|)}{2}}. \tag{46}$$

holds for all $\xi \in \mathbb{R}^n$

ii) Let $m \in \mathbb{R}$, $\psi \in \Lambda$. A C^∞-function $q : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}$ is a symbol in the class $S_\rho^{m,\psi}(\mathbb{R}^n)$ if for all $\alpha,\beta \in \mathbb{N}_0^n$ there are constants $c_{\alpha\beta} \geqslant 0$ such that

$$\left|\partial_\xi^\alpha \partial_x^\beta q(x,\xi)\right| \leqslant c_{\alpha\beta}(1+\psi(\xi))^{\frac{m-\rho(|\alpha|)}{2}} \tag{47}$$

holds for all $\xi, x \in \mathbb{R}^n$.

If $\rho \equiv 0$ we will simply write $S_0^{m,\psi}(\mathbb{R}^n)$. From now on we will also assume that the symbol $q \in S_\rho^{2,\psi}(\mathbb{R}^n)$ satisfies

$$q(x,\xi) \geqslant c_r\,(1+\psi(\xi)) \quad \text{for } x \in \mathbb{R}^n \text{ and sufficiently large} \quad |\xi| \geqslant r. \tag{\textbf{B5}}$$

where $\psi \in \Lambda$ and $\psi(\xi) \geqslant c_1|\xi|^\kappa$ for some $\kappa > 0$.

Note that (**B5**) implies (**B1**), i.e. $(q(x,D), C_0^\infty(\mathbb{R}^n))$ extends to the generator of a Feller semigroup, see [23, Theorem 2.6.9]. By Theorem 2.8 from [7] we can decompose $\lambda_t(x,\xi)$ in the following way.

$$\lambda_t(x,\xi) = e^{-tq(x,\xi)} + r(t,x,\xi), \tag{48}$$

where $r(t,x,\xi) \to 0$ weakly in $S_0^{-1,\psi}(\mathbb{R}^n)$ as $t \to 0$.

Assume that (**B2**), (**B3**) and (**B5**) hold. Then $r(t,x,\xi) \in L_1(\mathbb{R}^n)$, and by (35)

$$p_t(x,y) = \int_{\mathbb{R}^n} e^{i\xi\cdot(x-y)-tq(x,\xi)}\,d\xi + \int_{\mathbb{R}^n} e^{i\xi\cdot(x-y)} r(t,x,\xi)\,d\xi. \tag{49}$$

This allows us to formulate the following corollary of Theorem 15.

Corollary 17. *Let $q \in S_\rho^{2,\psi}(\mathbb{R}^n)$, and suppose that* (**B2**), (**B3**) *and* (**B5**) *are satisfied. Then the transition function of the Feller process generated by $-q(x,D)$ has a density $p_t(x,y)$, and for all $x, y \in \mathbb{R}^n$, $t > 0$,*

$$p_t(x,y) \leqslant e^{v_t(x-y,x,\zeta_0)} \int_{\mathbb{R}^n} e^{-tq(x,\xi)}\,d\xi + \int_{\mathbb{R}^n} e^{i\xi\cdot(x-y)}\, r(t,x,\xi)\,d\xi, \tag{50}$$

where $v_t(x-y,x,\zeta) := -\zeta\cdot(x-y) - tq(x,i\zeta)$, and $\zeta_0 := \arg\min_\zeta v_t(x-y,x,\zeta)$.

Acknowledgement. Financial support through INTAS (grant YSF 06-1000019-6024), DAAD (research grant June-August 2009), and the Ministry of Science of Ukraine (grant no. M/7-2008) (for V.K.) and DFG (grant Schi 419/5-1) (for R.L.S.) is gratefully acknowledged. We would like to thank two anonymous referees for their careful reading; their comments helped to improve the quality of this paper.

References

[1] Aronson, D.G.: Bounds for the fundamental solution of a parabolic equation. *Bull. Am. Math. Soc.* **73** (1967), 890–896.

[2] Barlow, M.T., Bass, R.B., Chen, Z.-Q. and Kassmann, M.: Non-local Dirichlet forms and symmetric jump processes. *Trans. Am. Math. Soc.* **361** (2009), 1963–1999.

[3] Barlow, M.T., Grigoryan, A. and Kumagai, T.: Heat kernel upper bounds for jump processes and the first exit time. *J. Reine Angew. Math.* **626** (2009), 135–157.

[4] Bass, R.F. and Levin, D.A.: Transition probabilities for symmetric jump processes. *Trans. Am. Math. Soc.* **354** (2002), 2933–2953.

[5] Bendikov, A. and Maheux, P.: Nash type inequalities for fractional powers of non-negative self-adjoint operators. *Trans. Am. Math. Soc.* **359** (2007), 3085–3097.

[6] Bodnarchuk, S. and Kulik, A.: Conditions for existence and smoothness of the distribution density for an Ornstein-Uhlenbeck process with Lévy noise. *Probab. Math. Stat.* **79** (2008), 20–33. (Ukrainian; English preprint is available at http://arXiv:0806.0442).

[7] Böttcher, B.: A parametrix construction for the fundamental solution of the evolution equation associated with a pseudo-differential operator generating a Markov process. *Math. Nachr.* **278** (2005), 1235–1241.

[8] Bouleau, N. and Hirsch, F.: *Dirichlet Forms and Analysis on Wiener Space*. De Gruyter, Studies in Math. vol. **14**, Berlin 1991.

[9] Carlen, E.A., Kusuoka, S. and Stroock, D.W.: Upper bounds for symmetric Markov transition functions. *Ann. Inst. Henri Poincaré, Probab. Stat.* suppl. t. **23** (1987), 245–287.

[10] Chen Z.-Q., Kim, P. and Kumagai, T.: Weighted Poincaré inequality and heat kernel estimates for finite range jump processes. *Math. Ann.* **342** (2008), 833–883.

[11] Chen Z.-Q., Kim, P. and Kumagai, T.: Global heat kernel estimates for symmetric jump processes. Preprint 2010.

[12] Chen, Z.-Q. and Kumagai, T.: Heat kernel estimates for stable-like processes on d-sets. *Stoch. Proc. Appl.* **108** (2003), 27–62.

[13] Chen, Z.-Q. and Kumagai, T.: Heat kernel estimates for jump processes of mixed types on metric measure spaces. *Probab. Theory Relat. Fields* **140** (2008), 277–317.

[14] Copson, E.T.: *Axamptotic Expansions*. Cambridge University Press, Tracts in Mathematics vol. **55**, Cambridge 2004.

[15] Davies, E.B.: *Heat Kernels and Spectral Theory*. Cambridge Univ. Press, Tracts in Math. vol. **92**, 1998.

[16] Duffie, D., Filipovic, D. and Schachermayer, W.: Affine processes and applications in finance. *Ann. Appl. Probab.* **13** (2003), 984–1053.

[17] Feng, J. and Kurtz, T.G.: *Large Deviations For Stochastic Processes*. American Mathematical Society, Math. Surv. Monogr. vol. **131**, Providence (RI) 2006.

[18] Hartman, P. and Wintner, A.: On the infinitesimal generators of integral convolutions. *Am. J. Math.* **64** (1942), 273–298.

[19] Hoh, W.: A symbolic calculus for pseudo-differential operators generating Feller semigroups. *Osaka J. Math.* **35** (1998), 789–820.

[20] Hoh, W.: *Pseudo-Differential Operators Generating Markov Processes*. Habilitationsschrift, Universität Bielefeld 1998.

[21] Jacob, N.: A class of Feller semigroups generated by pseudo-differential operators. *Math. Z.* **215** (1994), 151–166.

[22] Jacob, N.: *Pseudo Differential Operators and Markov Processes. Vol. I: Fourier Analysis and Semigroups*. Imperial College Press, London 2001.

[23] Jacob, N.: *Pseudo Differential Operators and Markov Processes. Vol. II: Generators and Their Potential Theory*. Imperial College Press, London, 2002.

[24] Jacob, N. and R. L. Schilling: Lévy-type processes and pseudo-differential operators. In: Barndorff-Nielsen, O. E. et al. (eds.): *Lévy processes: Theory and Applications*, Birkhäuser, Boston 2001, 139–168.

[25] Jacob, N. and Schilling, R.L.: Function spaces as Dirichlet spaces (about a paper by V. Maz'ya and J. Nagel). *Z. Anal. Anwendungen* **24** (2005), 3–28.

[26] Kallenberg, O.: Splitting at backward times in regenerative sets. *Ann. Probab.* **9** (1981), 781–799.

[27] Knopova, V. and Kulik, A.M.: Exact asymptotics for a distribution denisty of certain Lévy functionals. *Preprint*, Kiev 2009. (arXiv:0911.4683v1)

[28] Knopova, V. and Zähle, M.: Spaces of generalized smoothness on h-sets and related Dirichlet forms. *Studia Math.* **174** (2006), 277–308.

[29] Kolokoltsov, V.: Symmetric stable laws and stable-like jump-diffusions. *Proc. London Math. Soc.* **80** (2000), 725–768.

[30] Moura, S.: On some characterization of Besov spaces of generalized smoothness. *Preprint* **05-30**, Coimbra University 2005.

[31] Sato, K.: *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, Studies adv. Math. vol. **69**, Cambridge 1999.

[32] Schilling, R. L.: Conservativeness and Extensions of Feller Semigroups. *Positivity* **2** (1998), 239–256.

[33] Schilling, R.L., Song, R. and Vondraček, Z.: *Bernstein Functions–Theory and Applications.* De Gruyter, Studies in Math. vol. **37**, Berlin 2010.

[34] Sztonyk, P.: Estimates of tempered stable densities. *J. Theor. Probab* **23(1)** (2010), 127–147.

[35] Varopoulos, N.Th., Saloff-Coste, L. and Coulhon, T.: *Analysis and Geometry on Groups*, Cambridge University Press, Camb. Tracts Math. vol. **100**, Cambridge 1992.

[36] Zeidler, E.: *Nonlinear Functional Analysis and Its Applications. Vol. III: Variational Methods and Optimization.* Springer, New York 1985.