INTRINSIC COMPOUND KERNEL ESTIMATES FOR THE TRANSITION PROBABILITY DENSITY OF LÉVY-TYPE PROCESSES AND THEIR APPLICATIONS*

BY

VICTORIA KNOPOVA (KIEV) AND ALEKSEI KULIK (KIEV)

Abstract. Starting with an integro-differential operator $(L, C^2_\infty(\mathbb{R}^n))$, we prove that its $C_\infty(\mathbb{R}^n)$-closure is the generator of a Feller process X, which admits a transition probability density. To construct this transition probability density, we develop a version of the parametrix method and a verification procedure, which proves that the constructed object is the claimed one. As a part of the construction, we prove the intrinsic upper and lower estimates on the density. As an application of the constructed estimates we state the necessary and (separately) sufficient conditions under which a given Borel measure belongs to the Kato and Dynkin classes with respect to the constructed transition probability density.

2010 AMS Mathematics Subject Classification: Primary: 60J35; Secondary: 60J75, 35S05, 35S10, 47G30.

Key words and phrases: Transition probability density, Lévy-type processes, pseudo-differential operator, generator, Levi's parametrix method.

1. INTRODUCTION

A Markov process $X = (X_t)_{t \geqslant 0}$ with values in $\mathbb{R}^n$ is called a *Lévy-type process* if its generator A is well defined on the space $C^2_\infty(\mathbb{R}^n)$ of twice continuously differentiable functions, vanishing at infinity together with their derivatives, and on this space A coincides with a *Lévy-type operator*

$$(1.1) \quad Lf(x) = a(x) \cdot \nabla f(x) + \sum_{j,k=1}^{n} Q_{jk}(x) \frac{\partial^2 f(x)}{\partial x_j \partial x_k}$$
$$+ \int_{\mathbb{R}^n \backslash \{0\}} \big(f(x+u) - f(x) - \nabla f(x) \cdot u \mathbb{1}_{\{\|u\| \leqslant 1\}}\big) \mu(x, du),$$

---

where $a(x) \in \mathbb{R}^n$, $Q(x) \equiv \left(Q_{jk}(x)\right)_{j,k=1}^n$ is a symmetric positive semidefinite matrix, and $\mu(x, \cdot)$ is a positive Borel measure, such that

$$\int\limits_{\mathbb{R}^n \setminus \{0\}} (1 \wedge \|u\|^2) \mu(x, du) < \infty$$

for any $x \in \mathbb{R}^n$. For an extensive survey on Lévy-type processes and Lévy-type operators we refer to [39]–[41] and [13]; here we briefly outline the items important for the exposition below.

In the "constant coefficients case", where a, Q, μ do not depend on x, (1.1) is just an expression for the generator of the semigroup of probability measures, which corresponds to a Lévy process. Hence, a Lévy-type process has a natural interpretation as a "process with locally independent increments" whose characteristic triplet depends on the spatial variable. This justifies the names "a Lévy-type process" and "a Lévy-type operator". On the other hand, let a Markov process X be *Feller*, that is, the respective semigroup $(S_t)_{t \geq 0}$,

$$(1.2) \qquad\qquad S_t f(x) := \mathbb{E}^x f(X_t),$$

maps the space $C_\infty(\mathbb{R}^n)$ of continuous functions vanishing at infinity into itself. The Courrège–Waldenfels theorem (cf. [39], Theorem 4.5.21, and [13], Theorem 2.21) states that if for a Feller process the generator of $(S_t)_{t \geq 0}$ is well defined on $C_c^\infty(\mathbb{R}^n)$ (the space of compactly supported infinitely differentiable functions), then on $C_c^\infty(\mathbb{R}^n)$ this generator admits representation (1.1). Heuristically, this means that a Lévy-type process is a generic form for a Feller process on $\mathbb{R}^n$, and a Feller process on $\mathbb{R}^n$ naturally gives rise to an integral-differential operator of the form (1.1).

The converse problem, i.e., how to show that an operator of the form (1.1) gives rise to a Feller process and, moreover, to investigate the distribution properties of this process, is highly non-trivial; this is the topic the current paper is focused on. Heuristically, the relation between Lévy-type processes and Lévy processes is similar to that between diffusions and the Brownian motion. Hence, the problem how to construct a Lévy-type process given a Lévy-type operator is similar (but technically is much more involved) to the problem of construction of a diffusion process with given coefficients.

There are several ways how to associate a Lévy-type operator (1.1) with a Markov process. One way is to solve the *martingale problem* for $\left(L, C_\infty^2(\mathbb{R}^n)\right)$, that is, to find a family of probability measures $\mathbb{P}^x$, $x \in \mathbb{R}^n$, such that $\mathbb{P}^x(X_0 = x) = 1$ and the process

$$f(X_t) - f(X_0) - \int\limits_0^t Lf(X_s) ds$$

is a $\mathbb{P}^x$-martingale for any $f \in C_\infty^2(\mathbb{R}^n)$, $x \in \mathbb{R}^n$. The difficult part in this problem is to show that the martingale problem is *well posed*, i.e., that the family $\mathbb{P}^x$, $x \in \mathbb{R}^n$, is unique. See [69], [70], [53], [4], [34], [35], as well as the survey paper

[5] and the monograph [41]. Note that although the martingale problem approach is an efficient tool for constructing the process, typically it does not give much information about its intrinsic distribution properties.

Another natural way is based on the general fact that at least in the simplest cases the transition probability density $p_t(x, y)$ of the process X is a *fundamental solution* to the Cauchy problem associated with the operator

$$(1.3) \qquad \qquad \partial_t - L,$$

see Section 2.1 below for the definition. In the parabolic case, i.e., when the integral part in (1.1) is absent, the classical *parametrix method* makes it possible both to construct the fundamental solution $p_t(x, y)$ to (1.3), and to give the explicit upper and lower bounds for $p_t(x, y)$; see the monograph by Friedman [29] for details. We also refer to the original paper by Levi [59] and to the paper by Feller [27], in which the parametrix construction of the transition probability density is given for continuous, purely discontinuous, and mixed processes. Since for L the *positive maximum principle* holds true, we can conclude that $p_t(x, y)$ is the transition probability density of a Feller process, and this process is the unique one associated with the generator $A = L$. We omit the details, since the same procedure will be discussed in detail below in a much more complicated setting.

The goal of our investigation is to extend the approach outlined above to the genuinely "Lévy type" case, where the diffusion part in (1.1) is absent, but instead the jump part is present: $Q \equiv 0$, $\mu(x, du)$ is non-trivial. In this case, the structural assumptions on the Lévy kernel $\mu(x, du)$ appear to be substantial. The case of $\mu(x, du)$ being comparable, in a sense, to the Lévy measure of an α-stable process $\mu(du) = c|u|^{-\alpha-n}du$ is well studied, see [22], [23], [51], [52], [19], and an extensive overview in the monograph [25]. Extending these results to more general classes of Lévy measures meets new serious difficulties, which we discuss in detail below.

Following the line of the classical "parabolic" parametrix method (cf. [59], [29]), in which the Gaussian kernel is taken as the zero order approximation of the solution to the respective Cauchy problem, the natural idea to develop a "Lévy type" parametrix method is to take as the zero order approximation $p_t^0(x, y)$ for the candidate for being the fundamental solution to (1.3) the transition probability density of some Lévy process. An important feature used in all the aforementioned papers is that the fundamental solution $g_t(x, y)$ to the respective constant coefficient Cauchy problem satisfies the upper estimate

$$(1.4) \qquad \qquad g_t(x, y) \leqslant C\rho_t^n f\big(\rho_t(y - x)\big),$$

where $\rho : (0, 1] \to (0, \infty)$ and $f \in L_1(\mathbb{R}^n)$ has the meaning of a "scaling function" and a "shape function", respectively (namely, one has $\rho_t = t^{-1/2}$, $f(x) = \exp\big(-\|x\|^2/(2c)\big)$ in the diffusive case, and $\rho_t = t^{-1/\alpha}$, $f(x) = 1 \wedge \|x\|^{-\alpha-n}$ in the symmetric α-stable case).

However, for general Lévy processes the estimate (1.4) does not hold, see [47] for a counterexample. In [47] and [46] it is shown that a natural analogue of (1.4) is the so-called *compound kernel* upper bound, see Proposition 3.4 below. An important ingredient in the approach which will be developed in the current paper is that in the parametrix construction of the kernel $p_t(x, y)$, the *single kernel-type* upper bound (1.4) can be successfully replaced by a *compound kernel* upper bound, valid under more mild structural assumptions on the model.

Our construction consists of the following three principal steps. First, we construct the kernel $p_t(x, y)$, which is a *candidate* for being the fundamental solution to the Cauchy problem associated with the operator (1.3) represented in the form of convergent series (2.4)–(2.7); see the detailed discussion in Section 3. We emphasize that in our case the verification of the fact that the constructed kernel $p_t(x, y)$ is indeed the fundamental solution hardly could be performed in the classical way described in [29]. The reason for this is that the space derivatives of the zero order approximation $p_t^0(x, y)$ have stronger singularity at $t = 0$, and one cannot prove directly that $p_t(\cdot, y)$ belongs to $C_\infty^2(\mathbb{R}^n)$, which is the domain of L.

Our second step is to prove that the constructed kernel $p_t(x, y)$ is the transition probability density of some Markov process, and the restriction of the generator of this process to $C_\infty^2(\mathbb{R}^n)$ equals L. The method we use to do this is described in [49] (see also [56]), and is based on the auxiliary notion of the *approximate fundamental solution*; see Section 4.

The final step is to identify uniquely the Markov process X obtained before in terms of the initial operator L. While spatial derivatives of $p_t(x, y)$ are hardly controllable, the time derivative is more manageable, which makes it possible to prove that the generator $\big(A, D(A)\big)$ of the $C_\infty(\mathbb{R}^n)$-semigroup of X is the closure of $\big(L, C_\infty^2(\mathbb{R}^n)\big)$. This yields that the martingale problem for $\big(L, C_\infty^2(\mathbb{R}^n)\big)$ is well posed, and shows the uniqueness of the Markov process X constructed in the first two steps. Also, we are able to show that $p_t(\cdot, y)$ belongs to the domain of the generator A ($\Leftrightarrow$ the closure of L), and that $p_t(x, y)$ is indeed the fundamental solution to the Cauchy problem for

$$(1.5) \qquad\qquad\qquad \partial_t - A,$$

which justifies our parametrix construction. Thus, starting with the Cauchy problem for (1.3), we construct the fundamental solution for (1.5), where A is the closure of L.

As an application of the estimates constructed for the kernel, we give the necessary and (separately) sufficient conditions for a finite Borel measure to belong to the Kato and Dynkin classes with respect to $p_t(x, y)$.

Let us give a brief overview of other existing results.

In the Lévy case the transition probability density is just the inverse Fourier transform of the characteristic function. This allows a lot of possibilities to estimate this transition probability density, see, for example, [64], [32], [33], [72] for the asymptotic behavior of an α-stable transition probability density, and [54], [42]–

[44], [61], [6], [47], [46], [65]–[68], [20] for the Lévy case. Of course, this list of publications is far from being complete.

In [7] the case of the fractional Laplacian perturbed by a gradient is treated, see also [8], [10] for the kernel estimates. The verification procedure presented in [7] shows that the integro-differential operator is the weak generator of the respective semigroup. In [45] the case of a singular perturbation of the fractional Laplacian is considered, and a different approach is used for the verification; see also [18] for an approach which relies on [7] and the martingale problem, as well as [19] and the references therein. In [63] and [62] the authors constructed the transition probability density of the process which is the weak solution to the SDE driven by a symmetric α-stable process with a drift. We refer to [28], [21], and [49], in which the gradient perturbations of an α-stable like operator with $0 < \alpha < 1$ are investigated.

Another approach to study the fundamental solution to the respective Cauchy problem relies on a different version of the parametrix method and the symbolic calculus technique, which allows us to prove the existence of the fundamental solution, and to construct it in the form of converging in a certain sense series. This approach uses the Hilbert space methods, and is developed in [71], [38], [57], [36], [37], [40], [11], and [12].

There is a large group of results devoted to the estimation of the transition probability density of a Markov process, associated with a Dirichlet form of a certain type. Under the assumption that the jump intensity measure of a Markov process is absolutely continuous and has certain regularity properties, estimates on the transition probability density are obtained in [14]–[17], [2], [3], [60]; of course, this list is far from being complete. The approach used in the above papers relies on the Dirichlet form technique and the Harnack principle. Note that in these papers the initially given object is a regular Dirichlet form, which already assumes the *existence* of the related Markov process.

The paper is organized as follows. In Section 2 we set the notation, outline the method, and formulate the results. The construction of the parametrix series is performed in Sections 3.1–3.4. Proofs of the continuity and smoothness properties are given in Section 3.5. Section 4 is devoted to the verification procedure. The uniqueness is studied in Section 6. Diagonal and lower bounds for the constructed fundamental solution are given in Section 7. Finally, Section 8 is devoted to the proof of the application result, that is, using the structure of the upper and lower bounds on $p_t(x, y)$, we provide the necessary and (separately) sufficient conditions for a measure to be in the Kato and Dynkin classes.

2. THE MAIN RESULTS: OUTLINE AND FORMULATION

Notation. For functions f, g we mean by $f \asymp g$ that there exist some constants $c_1, c_2 > 0$ such that $c_1 f(x) \leqslant g(x) \leqslant c_2 f(x)$ for all $x \in \mathbb{R}^n$. By $x \cdot y$ and $\|x\|$ we denote, respectively, the scalar product and the norm in $\mathbb{R}^n$; $\mathbb{S}^n$ denotes a unit

sphere in $\mathbb{R}^n$. We denote by c_i, c, C, etc., arbitrary positive constants. Denote by $(f * g)(t, x, y)$, $(f \circledast g)(t, x, y)$, $(F * G)_t(du)$, $(F \circledast G)_t(du)$ the respective convolutions of functions $f(t, x, y)$, $g(t, x, y)$, and of kernels $F_t(du)$ and $G_t(dv)$:

$$(f * g)(t, x, y) := \int_{\mathbb{R}^n} f(t, x, z) g(t, z, y) dz,$$

$$(f \circledast g)(t, x, y) := \int_0^t \int_{\mathbb{R}^n} f(t - s, x, z) g(s, z, y) dz ds,$$

$$(F * G)_t(du) = \int_{\mathbb{R}^n} F_t(du - z) G_t(dz),$$

$$(F \circledast G)_t(du) = \int_0^t \int_{\mathbb{R}^n} F_{t-s}(du - z) G_s(dz) ds.$$

By $B_b(\mathbb{R}^n)$ and $C_\infty^k(\mathbb{R}^n)$ we denote, respectively, the set of bounded Borel functions and the set of k-times differentiable functions, vanishing at infinity together with their derivatives. By $\| \cdot \|_\infty$ we define the sup-norm in C_∞.

2.1. The model and the outline of the method. In this section we describe in detail three steps of our approach, indicated in the Introduction.

Let L be an operator of the form (1.1) but with $Q \equiv 0$. Below we specify the assumptions on the drift and the kernel. In the first part we construct a candidate for the fundamental solution to the Cauchy problem

$$(2.1) \qquad \partial_t - L,$$

i.e. a function $p_t(x, y)$ such that

$$(2.2) \qquad p_t(x, \cdot) \to \delta_x \quad \text{as } t \to 0+, \, x \in \mathbb{R}^n,$$

and

$$(2.3) \qquad (\partial_t - L)p_t(x, y) = 0, \quad t > 0, \, x, y \in \mathbb{R}^n.$$

In order to simplify the further exposition, let us briefly outline the parametrix construction, see [29], pp. 310–311, or [40], pp. 144–145, for more information.

Consider *some* approximation $p_t^0(x, y)$ of $p_t(x, y)$, and denote by $r_t(x, y)$ the residue with respect to this approximation, that is,

$$(2.4) \qquad p_t(x, y) = p_t^0(x, y) + r_t(x, y).$$

Define

$$(2.5) \qquad \Phi_t(x, y) := (L - \partial_t)p_t^0(x, y), \quad t > 0, \, x, y \in \mathbb{R}^n.$$

Observe that since $p_t(x, y)$ is aimed to be the fundamental solution to the Cauchy problem for the operator (2.1), we should have

$$(\partial_t - L)r_t(x, y) = \Phi_t(x, y).$$

Therefore,

$$r_t(x, y) = (p \circledast \Phi)_t(x, y),$$

which by (2.4) allows us to write the equation for $r_t(x, y)$:

$$r_t(x, y) = (p^0 \circledast \Phi)_t(x, y) + (r \circledast \Phi)_t(x, y).$$

The formal solution to this equation is given by the convolution

(2.6)
$$r = p^0 \circledast \Psi,$$

where

(2.7)
$$\Psi_t(x, y) = \sum_{k=1}^{\infty} \Phi_t^{\circledast k}(x, y).$$

We can choose the zero order approximation $p_t^0(x, y)$ in the following way.

Consider the operator

(2.8)
$$L^z f(x) := a(z) \cdot \nabla f(x) + \int_{\mathbb{R}^n} \big(f(x + u) - f(x) - u \cdot \nabla f(x) \mathbb{1}_{\{\|u\| \leqslant 1\}}\big)\mu(z, du),$$

where $f \in C_\infty^2(\mathbb{R}^n)$. It is known that $\big(L^z, C_\infty^2(\mathbb{R}^n)\big)$ extends to the generator of a semigroup corresponding to a Lévy process, which under the condition **A1** (see below) has the transition probability density $\mathfrak{p}_t^z(x)$. Note that $\mathfrak{p}_t^z(y - x)$ is the fundamental solution to a Cauchy problem for the operator

$$\partial_t - L^z,$$

see [40], Example 2.7.14. Put

(2.9)
$$p_t^0(x, y) := \mathfrak{p}_t^z(y - x)\big|_{z=y}.$$

We prove that under such a choice of the zero order approximation $p_t^0(x, y)$, the series in (2.7) indeed converges, and that the expression (2.4) is well defined.

In the second step we associate with the constructed kernel $p_t(x, y)$ a Markov process. The keystone in this step is the usage of an auxiliary object, called the *approximate fundamental solution*, which is a certain approximation $p_{t,\epsilon}(x, y)$ of the constructed kernel $p_t(x, y)$. Using the expression for $p_t(x, y)$ and the estimates on $\Phi_t(x, y)$ and $\Psi_t(x, y)$, obtained in the first step, we show that for the operators $S_{t,\varepsilon}$ with kernel $p_{t,\varepsilon}(x, y)$ the following statements (a)–(d) hold true:

(a) For $f \in C_\infty(\mathbb{R}^n)$,

$$\lim_{\varepsilon \to 0} \|S_{t,\varepsilon} f - S_t f\|_\infty = 0,$$

uniformly on compact subsets of $(0, \infty)$.

(b) For $S_{t,\varepsilon} f$, $f \in C_\infty(\mathbb{R}^n)$, $\varepsilon > 0$, the identity (2.3) turns into the approximate identity $(\partial_t - L) S_{t,\varepsilon} f$, and

$$\lim_{t,\varepsilon \to 0} \|S_{t,\varepsilon} f - S_t f\|_\infty = 0;$$

here

(2.10) $S_t f(x) := \int\limits_{\mathbb{R}^n} f(y) p_t(x,y) dy, \quad t > 0, \ x \in \mathbb{R}^n.$

These properties of $S_{t,\varepsilon}$ allow us to develop a version of the positive maximum principle (see [26], p. 165, or [39], Corollary 4.5.14, for the classical positive maximum principle). This in turn enables us to show that $p_t(x,y)$ is the transition probability density of a Markov process, which is a solution to the martingale problem for $\left(L, C_\infty^2(\mathbb{R}^n)\right)$. In particular, the family of operators $(S_t)_{t \geqslant 0}$ forms a semigroup related to X by

$$S_t f(x) = \mathbb{E}^x f(X_t), \quad f \in C_\infty^2(\mathbb{R}^n).$$

The third step is devoted to the uniqueness problem for the constructed process. We show that the generator A of the semigroup $(S_t)_{t \geqslant 0}$ coincides on $C_\infty^2(\mathbb{R}^n)$ with L. Further, we employ the properties of the derivative $\partial_t S_{t,\varepsilon} f$:

(c) For any $f \in C_\infty(\mathbb{R}^n)$,

$$\lim_{\varepsilon \to 0} \|\partial_t S_{t,\epsilon} f - \partial_t S_t f\|_\infty = 0,$$

uniformly on compact subsets of $(0, \infty)$.

This property together with (b) allows us to control L on $S_{t,\varepsilon} f$, $f \in D(A)$, and show that A is the closure of L in C_∞. Consequently, the process X constructed in the previous step is the unique solution to the above martingale problem.

Finally, the property similar to (c) holds also for the kernels $p_{t,\varepsilon}(x,y)$ and $p_t(x,y)$:

(d) $\partial_t p_{t,\varepsilon}(x,y)$ approximates $\partial_t p_t(x,y)$ as $\varepsilon \to 0$, uniformly on compact subsets of $(0, \infty) \times \mathbb{R}^n \times \mathbb{R}^n$.

Therefore, in a similar way we show that $p_t(\cdot, y) \in D(A)$ for any fixed y, and that $p_t(x,y)$ is the fundamental solution to the Cauchy problem for $\partial_t - A$.

2.2. Main results. Consider

$$(2.11) \qquad q(\xi) := \int_{\mathbb{R}^n} \big(1 - \cos(\xi \cdot u)\big)\mu(du),$$

where $\mu(du)$ is a Lévy measure, i.e., a Borel measure satisfying

$$\int_{\mathbb{R}^n} (\|u\|^2 \wedge 1)\mu(du) < \infty,$$

and define

$$(2.12) \quad q^U(\xi) := \int_{\mathbb{R}^n} [(\xi \cdot u)^2 \wedge 1]\mu(du), \quad q^L(\xi) := \int_{|u \cdot \xi| \leqslant 1} (\xi \cdot u)^2 \mu(du).$$

The function $q(\xi)$ has the Lévy–Khinchin representation, and thus is the characteristic exponent of a Lévy process. It can be shown (cf. [47]) that the functions $q^L(\xi)$ and $q^U(\xi)$ satisfy

$$(2.13) \qquad (1 - \cos 1)q^L(\xi) \leqslant q(\xi) \leqslant 2q^U(\xi).$$

Note that in (2.11) and (2.12) we do not assume μ to be symmetric. Suppose that the measure μ satisfies the regularity assumption given below.

A1. There exists $\beta > 1$ such that

$$\sup_{l \in \mathbb{S}^n} q^U(rl) \leqslant \beta \inf_{l \in \mathbb{S}^n} q^L(rl) \quad \text{for all } r > 0 \text{ large enough.}$$

In what follows, we put

$$(2.14) \qquad\qquad\qquad \alpha := 2/\beta.$$

This notation is motivated by the particularly important example of a symmetric α-stable Lévy measure $\mu(du) := c(\alpha)\|u\|^{-n-\alpha}du$, $\alpha \in (0, 2)$: direct calculations show that in this case **A1** holds true with $\beta = 2/\alpha$. Note also that for any Lévy measure μ satisfying **A1** the respective Lévy exponent q admits a polynomial lower bound (see (2.19) below) which for the symmetric α-stable Lévy measure becomes an identity.

Throughout the paper we assume that the kernel $\mu(x, du)$ is of the form

$$(2.15) \qquad\qquad \mu(x, du) = m(x, u)\mu(du),$$

where $m(x, u)$ is some positive measurable function. We assume that the function $m(x, u)$ and the drift coefficient $a(x)$ satisfy the assumptions below.

A2. The functions $m(x, u)$ and $a(x)$ are measurable and satisfy with some constants b_1, b_2, $b_3 > 0$ the inequalities

$$b_1 \leqslant m(x, u) \leqslant b_2, \quad |a(x)| \leqslant b_3, \quad x, u \in \mathbb{R}^n.$$

A3. There exist constants $\gamma \in (0, 1]$ and $b_4 > 0$ such that

(2.16)
$$|m(x, u) - m(y, u)| + \|a(x) - a(y)\| \leqslant b_4(\|x - y\|^\gamma \wedge 1), \quad u, x, y \in \mathbb{R}^n.$$

A4. In the case $\alpha \in (0, 1]$ we assume that $a(x) = 0$ and the kernel $\mu(x, du)$ is symmetric with respect to u for all $x \in \mathbb{R}^n$.

Below we state the first main result of our paper.

THEOREM 2.1. *Suppose that assumptions* **A1–A4** *are satisfied, and the function* $p_t^0(x, y)$ *is given by* (2.9). *Then*:
(a) *The function* $p_t(x, y)$ *introduced in* (2.4)–(2.7) *is well defined in the sense that the series* (2.7) *and* (2.6) *converge absolutely for any* $t > 0$, $x, y \in \mathbb{R}^n$, *uniformly on compact subsets of* $(0, \infty) \times \mathbb{R}^n \times \mathbb{R}^n$.
(b) *The function* $p_t(x, y)$ *is continuous on* $(0, \infty) \times \mathbb{R}^n \times \mathbb{R}^n$.

Next we associate the constructed function $p_t(x, y)$ with the initial operator L. To make the structure the most transparent, we do this in two steps: we prove that $p_t(x, y)$ is a transition probability density of *some* Markov process, and then show that the $C_\infty(\mathbb{R}^n)$-generator of this process is an extension of $(L, C_\infty^2(\mathbb{R}^n))$. The second statement means that the semigroup (2.10) with $p_t(x, y)$ defined by (2.4) is in fact a *unique* Feller semigroup associated with the operator L.

THEOREM 2.2. *The family of operators* (2.10) *forms a strongly continuous conservative semigroup of nonnegative operators on* $C_\infty(\mathbb{R}^n)$, *which in turn defines a (strong) Feller Markov process* X. *Further, the set* $C_\infty^2(\mathbb{R}^n)$ *belongs to the domain* $D(A)$ *of the generator* A *of this semigroup, and*

$$Af(x) = Lf(x) \quad \text{for } f \in C_\infty^2(\mathbb{R}^n),$$

that is, $(A, D(A))$ *is an extension of* $(L, C_\infty^2(\mathbb{R}^n))$.

THEOREM 2.3. (a) *The generator* $(A, D(A))$ *is the closure of* $(L, C_\infty^2(\mathbb{R}^n))$.
(b) *The function* $p_t(\cdot, y)$ *belongs to the domain* $D(A)$ *of* A, *and is the fundamental solution to the Cauchy problem for the operator* $\partial_t - A$.

The first statement of Theorem 2.3 allows us to show the uniqueness of the solution to the martingale problem for $(L, C_\infty^2(\mathbb{R}^n))$.

THEOREM 2.4. *The Markov process* X *constructed in Theorem* 2.3 *is the unique solution to the martingale problem for* $(L, C_\infty^2(\mathbb{R}^n))$.

Finally, we give the upper and lower estimates for the constructed function $p_t(x, y)$ and its time derivative.
Let

(2.17)
$$q^*(r) := \sup_{l \in \mathbb{S}^n} q^U(rl), \quad r > 0,$$

and define

$$(2.18) \qquad \rho_t := \inf\{r : q^*(r) = 1/t\}, \quad t \in (0, T].$$

Since the function $q^*(r)$ is continuous and $\lim_{r \to \infty} q^*(r) = \infty$, the function ρ_t, $t \in (0, T]$, is well defined for any $T > 0$.

In [46], see also [47], we show that condition **A1** implies for r large enough the lower estimate

$$(2.19) \qquad q^*(r) \geqslant c_0 r^\alpha,$$

which in turn implies for any $T > 0$ the upper bound

$$(2.20) \qquad \rho_t \leqslant c_1 t^{-1/\alpha}, \quad t \in (0, T].$$

Note that for any $c > 1$ we have $q^U(c\xi) \leqslant (c^2 \wedge 1)q^U(\xi)$, implying $q^*(r) \leqslant c_2 r^2$, $r \geqslant 1$; then $\rho_t \geqslant c_3 t^{-1/2}$, $t \in (0, T]$. Denote by $\sigma \in [\alpha, 2]$ the minimal value for which there exists $c_4 > 0$ such that

$$(2.21) \qquad \rho_t \geqslant c_4 t^{-1/\sigma}, \quad t \in (0, T].$$

Denote by f_{up} and f_{low} the functions of the form

$$(2.22) \quad f_{up}(x) := d_1 e^{-d_2 \|x\|}, \quad f_{low}(x) := d_3 (1 - d_4 \|x\|)_+, \quad x \in \mathbb{R}^n,$$

where $d_i > 0$, $1 \leqslant i \leqslant 4$, are some constants which are yet to be chosen.

THEOREM 2.5. *For any $T > 0$ there exist constants $d_i > 0$, $1 \leqslant i \leqslant 4$, and a family of subprobability measures $\{Q_t, t \geqslant 0\}$ such that $p_t(x, y)$ satisfies the upper and lower estimates*

$$(2.23)$$
$$\rho_t^n f_{low}\big(\rho_t(y - x)\big) \leqslant p_t(x, y) \leqslant \rho_t^n \big(f_{up}(\rho_t \cdot) * Q_t\big)(y - x), \ t \in (0, T], \ x, y \in \mathbb{R}^n,$$

where f_{low} and f_{up} are of the form (2.22) with constants d_i, $1 \leqslant i \leqslant 4$.

THEOREM 2.6. (1) *There exists $\partial_t p_t(x, y)$ which is continuous in $(t, x, y) \in (0, \infty) \times \mathbb{R}^n \times \mathbb{R}^n$.*

(2) *For any $T > 0$ there exist constants $\tilde{d}_1, \tilde{d}_2 > 0$ and a family of subprobability measures $\{\tilde{Q}_t, t \geqslant 0\}$ such that*

$$|\partial_t p_t(x, y)| \leqslant t^{-1} \rho_t^n \big(f_{up}(\rho_t \cdot) * \tilde{Q}_t\big)(y - x), \quad t \in (0, T], \ x, y \in \mathbb{R}^n,$$

where f_{up} is of the form (2.22) with constants $\tilde{d}_1, \tilde{d}_2$.

To demonstrate an application of the above results, we need a bit more preparations.

Recall that a functional φ_t of a strong Markov process X is called a *W-functional* if it is additive, positive, continuous, almost surely homogeneous, and

$$v_t(x) := \mathbb{E}^x \varphi_t < \infty;$$

in this case the function $v_t(x)$ is called the *characteristic* of φ_t, see [24], §6.11. By [24], Theorem 6.3, the characteristic determines the W-functional uniquely up to equivalence. On the other hand, Theorem 6.3 of [24] gives a way how to check that a given function is a characteristic of some W-functional.

Recall (cf. [58], [1]) that a Borel measure ϖ is said to belong to

(i) the *Kato class* S_K with respect to $p_t(x, y)$ if

$$(2.24) \qquad \lim_{t \to 0} \sup_{x \in \mathbb{R}^n} \int_0^t \int_{\mathbb{R}^n} p_s(x, y) \varpi(dy) ds = 0;$$

(ii) the *Dynkin class* S_D with respect to $p_t(x, y)$ if there exists $t > 0$ such that

$$(2.25) \qquad \sup_{x \in \mathbb{R}^n} \int_0^t \int_{\mathbb{R}^n} p_s(x, y) \varpi(dy) ds < \infty.$$

Clearly, $S_K \subset S_D$. By [24], Theorem 6.6, the condition $\varpi \in S_K$ implies that the function

$$(2.26) \qquad \chi_t(x) := \int_0^t \int_{\mathbb{R}^n} p_s(x, y) \varpi(dy) ds$$

is the characteristic of some W-functional φ_t provided that the mapping $x \mapsto \chi_t(x)$ is measurable for each $t \geqslant 0$. Thus, to prove that $\chi_t(x)$ is the characteristic of some W-functional of X, we need to check whether the measure ϖ from (2.26) belongs to the Kato class with respect to $p_t(x, y)$. As an accompanying result, we get the condition under which ϖ belongs to the respective Dynkin class.

REMARK 2.1. *Up to our knowledge there are not many results on the necessary and sufficient conditions when a measure is in the Kato class. In the case of a symmetric α-stable process, $\alpha \in (0, 2)$, and a relativistic $1/2$-stable process, these conditions are stated in [73] (see also [30]). In the case of n-dimensional Brownian motion there is a one-to-one correspondence between the class of W-functions and so-called W-measures (see [24], Theorem 8.4); in our notation this theorem means that every measure from the Dynkin class is in one-to-one correspondence with a W-functional. An example of a measure which for a Brownian motion belongs to the class S_D but not to the S_K can be found, e.g., in [55].*

In the theorem below we present the necessary and (separately) sufficient conditions when a measure belongs to the Kato and Dynkin classes with respect to $p_t(x, y)$.

THEOREM 2.7. *Let ϖ be a finite Borel measure on $\mathbb{R}^n$.*

(a) *For $\varpi \in S_D$ with respect to $p_t(x, y)$ it is sufficient that*

$$(2.27) \qquad \int_0^\delta \frac{\sup_{x \in \mathbb{R}^n} \varpi\{y : \|x - y\| \leqslant s\}}{s^{n+1} q^*(1/s)} ds < \infty \quad \text{for some } \delta > 0,$$

and necessary that

$$(2.28) \qquad \sup_{x \in \mathbb{R}^n} \int_0^\delta \frac{\varpi\{y : \|x - y\| \leqslant s\}}{s^{n+1} q^*(1/s)} ds < \infty \quad \text{for some } \delta > 0.$$

(b) *For $\varpi \in S_K$ with respect to $p_t(x, y)$ it is sufficient that* (2.27) *holds true, and necessary that*

$$(2.29) \qquad \lim_{\delta \to 0} \sup_{x \in \mathbb{R}^n} \int_0^\delta \frac{\varpi\{y : \|x - y\| \leqslant s\}}{s^{n+1} q^*(1/s)} ds = 0.$$

3. CONSTRUCTION OF THE PARAMETRIX SERIES. PROOF OF THEOREM 2.1

3.1. Well-definiteness of $p_t^0(x, y)$. It is known that for any fixed $z \in \mathbb{R}^n$ the operator $\left(L^z, C_\infty^2(\mathbb{R}^n)\right)$ (see (2.8)) extends to the C_∞- generator of a Feller semigroup which corresponds to the Lévy process X_t^z with characteristic function

$$\mathbb{E} e^{i\xi \cdot X_t^z} = e^{-t q(z, \xi)},$$

where

$$(3.1) \qquad q(z, \xi) := -i a(z) \cdot \xi + \int_{\mathbb{R}^n} (1 - e^{i\xi \cdot u} + i\xi \cdot u \mathbb{1}_{\{\|u\| \leqslant 1\}}) \mu(z, du).$$

Note that due to the condition **A2** the kernels $\{\mu(z, du), \ z \in \mathbb{R}^n\}$ are comparable in the sense that for any $z, y \in \mathbb{R}^n$ and any Borel subset $A \subset \mathbb{R}^n \backslash \{0\}$ we have $\mu(z, A) \asymp \mu(y, A)$, implying that

$$(3.2) \qquad \operatorname{Re} q(z, \xi) \asymp \operatorname{Re} q(y, \xi) \quad \text{for all } z, y, \xi \in \mathbb{R}^n.$$

Condition **A3** implies that $\operatorname{Re} q(z, \xi)$ is continuous in z. Condition **A1** together with **A2** implies (cf. [47], [46]) that

$$(3.3) \qquad \min_z \operatorname{Re} q(z, \xi) \geqslant c \|\xi\|^\alpha \quad \text{for large } \|\xi\|,$$

where $\alpha = 2/\beta$. Thus, for any fixed z the process X_t^z admits a transition probability density, which will be denoted by $\mathfrak{p}_t^z(x)$. Note that by (3.3) we can write $\mathfrak{p}_t^z(x)$ as

$$(3.4) \qquad \mathfrak{p}_t^z(x) = (2\pi)^{-n} \int\limits_{\mathbb{R}^n} e^{-ix\cdot\xi - tq(z,\xi)} d\xi.$$

Thus, the function $p_t^0(x,y)$ given by (2.9) is well defined, and $p_t^0(\cdot, y) \in C_b^\infty(\mathbb{R}^n)$.

3.2. Estimate for $\Phi_t(x,y)$. In this subsection we derive the upper bound for $\Phi_t(x,y)$, see Lemma 3.1. In order to do this, we introduce some notation and state the auxiliary propositions, the proofs of which are deferred to Appendix A.

Let

$$(3.5) \qquad \Lambda_t(du) := t\mu(du)\mathbb{1}_{\{\rho_t\|u\| > 1\}}.$$

PROPOSITION 3.1. *For any $T > 0$ we have $\Lambda_t(\mathbb{R}^n) \leqslant n^2$, $t \in [0, T]$.*

PROPOSITION 3.2. *For any $\lambda \in [0, \alpha)$, $T > 0$, we have*

$$(3.6) \qquad \rho_t^\lambda \int\limits_{\mathbb{R}^n} (\|u\|^\lambda \wedge 1)\Lambda_t(du) \leqslant C, \quad t \in [0, T].$$

Define the probability measure

$$(3.7) \qquad P_t(du) := e^{-\Lambda_t(\mathbb{R}^n)} \sum_{k=0}^\infty \frac{1}{k!} \Lambda_t^{*k}(du).$$

Let α be the parameter defined in (2.14), and let $\gamma \in (0, 1]$ be the parameter of Hölder continuity from **A3**. Fix some $\epsilon \in (0, \alpha)$, and put

$$(3.8) \qquad \kappa := \begin{cases} \gamma & \text{if } \gamma \in (0, \alpha), \\ \alpha - \epsilon & \text{if } \gamma \geqslant \alpha. \end{cases}$$

Note that, by definition, $\kappa > 0$. Put

$$(3.9) \qquad P_{t,\kappa}(du) := \left(1 + \rho_t^\kappa(\|u\|^\kappa \wedge 1)\right) P_t(du).$$

PROPOSITION 3.3. *For any $T > 0$ there exists a constant $C \in (0, \infty)$ such that $P_{t,\kappa}(\mathbb{R}^n) \leqslant C$, $t \in [0, T]$.*

Finally, define

$$(3.10) \qquad G_t(du) := c_0\left(P_{t,\kappa}(du) + P_t * P_{t,\kappa}(du)\right),$$

where the constant $c_0 > 0$ is chosen in such a way that $G_t(\mathbb{R}^n) \leqslant 1$ for all $t \in [0, T]$. Such a choice of c_0 is possible due to Proposition 3.3.

Now we are ready to state the main result of this section. Recall that the parameter σ was defined in (2.21).

LEMMA 3.1. *For any $T > 0$ there exist constants $C, b > 0$ such that*

$$(3.11) \quad |\Phi_t(x, y)| \leqslant Ct^{-1+\eta}(g_t * G_t)(y - x), \quad t \in (0, T], \ x, y \in \mathbb{R}^n,$$

where $\eta = \frac{\kappa}{\sigma} \wedge \left(1 + \frac{\kappa - 1}{\alpha}\right)$,

$$(3.12) \qquad\qquad g_t(x) := \rho_t^n e^{-b\rho_t \|x\|}$$

with some $b > 0$, and $\{G_t(\cdot), \ t > 0\}$ is the family of subprobability measures given by (3.10).

In the proof of this lemma we use several auxiliary statements which will be formulated below.

Take f_{up} of the form (2.22), and put

$$(3.13) \qquad\qquad f_t(x) := \rho_t^n \big(f_{up}(\rho_t \cdot) * P_t\big)(x).$$

PROPOSITION 3.4. *Suppose that conditions **A1**, **A2** and **A4** hold true.*
Then for any $k = k_1 + \ldots + k_n \geqslant 0$, $T > 0$, there exist constants $A_k, a_k > 0$, such that

$$(3.14) \qquad \left| \frac{\partial^k}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}} \mathfrak{p}_t^z(x) \right| \leqslant \rho_t^k f_t(x), \quad t \in (0, T], \ x, z \in \mathbb{R}^n,$$

where $f_t(x)$ is the function of the form (3.13), *and the function f_{up} in the definition of f_t is of the form* (2.22), *with constants A_k, a_k in place of d_1, d_2, respectively.*
In particular,

$$(3.15) \quad \left| \frac{\partial^k}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}} p_t^0(x, y) \right| \leqslant \rho_t^k f_t(y - x), \quad t \in (0, T], \ x, y \in \mathbb{R}^n.$$

PROPOSITION 3.5. *Suppose that conditions **A1**, **A2** and **A4** hold true. For $T > 0$ there exist $d_3, d_4 > 0$ such that*

$$\mathfrak{p}_t^z(x) \geqslant \rho_t^n f_{low}(\|x\| \rho_t), \quad x, z \in \mathbb{R}^n, \ t \in (0, T],$$

where f_{low} is of the form (2.22) *with these constants d_3 and d_4. In particular,*

$$p_t^0(x, y) \geqslant \rho_t^n f_{low}(\|y - x\| \rho_t), \quad x \in \mathbb{R}^n, \ t \in (0, T].$$

REMARK 3.1. *Proceeding as in* [46] *and* [47] *one can show that it follows that $0 < a_k \leqslant a_{k-1}, k \geqslant 1$.*

The proof of this proposition repeats line by line the proof of a similar statement in [46], see also [47]. The only difference is that we need to check, using conditions **A1**, **A2** and **A4**, that the required estimates obtained in [46] hold true uniformly in z. We omit the details.

PROPOSITION 3.6. *For any $\theta \in (0,1)$, $T > 0$, we have*

$$(3.16) \quad (\|x\|^{\kappa} \wedge 1)f_t(x) \leqslant C\rho_t^{-\kappa}(g_{t,\theta} * P_{t,\kappa})(x), \quad t \in (0,T], \; x \in \mathbb{R}^n,$$

where κ is defined in (3.8), $P_{t,\kappa}(dw)$ is defined in (3.9), f_t is of the form (3.13) with some f_{up}, and

$$(3.17) \qquad\qquad g_{t,\theta}(x) = \rho_t^n f_{up}(\theta \rho_t x).$$

P r o o f o f L e m m a 3.1. By the definition of $p_t^0(x,y)$, for any $y \in \mathbb{R}^n$ we have

$$[\partial_t - L_x^y(D)]p_t^0(x,y) = 0.$$

Then

$$(3.18) \quad \Phi_t(x,y) = [L(x,D) - L_x^y(D)]p_t^0(x,y)$$

$$= \big(a(x) - a(y)\big) \cdot \nabla p_t^0(x,y)$$

$$+ \int_{\mathbb{R}^n} [p_t^0(x+u,y) - p_t^0(x,y) - u \cdot \nabla p_t^0(x,y)\mathbb{1}_{\{\|u\| \leqslant 1\}}]$$

$$\times [m(x,u) - m(y,u)]\mu(du)$$

$$= \big(a(x) - a(y)\big) \cdot \nabla p_t^0(x,y)$$

$$+ \Big[\int_{\rho_t\|u\| \leqslant 1} + \int_{\rho_t\|u\| > 1}\Big][p_t^0(x+u,y) - p_t^0(x,y) - u \cdot \nabla p_t^0(x,y)\mathbb{1}_{\{\|u\| \leqslant 1\}}]$$

$$\times [m(x,u) - m(y,u)]\mu(du) =: J_1 + J_2 + J_3.$$

We estimate the terms J_i, $i = 1,2,3$, separately. In what follows, f_t and $g_{t,\theta}$ are the functions appearing in Propositions 3.4 and 3.6.

Note that by **A4** we have $J_1 = 0$ for $\alpha \in (0,1]$. For $\alpha \in (1,2)$, by (3.15) and **A3** we have the estimates

$$(3.19) \quad |J_1| \leqslant \sqrt{n}\|a(x) - a(y)\|\rho_t f_t(y-x) \leqslant c_1(\|y-x\|^{\gamma} \wedge 1)\rho_t f_t(y-x).$$

Using Proposition 3.6, we obtain

$$(\|y-x\|^{\gamma} \wedge 1)f_t(y-x) \leqslant (\|y-x\|^{\kappa} \wedge 1)f_t(y-x) \leqslant c_2\rho_t^{-\kappa}(g_{t,\theta} * P_{t,\kappa})(y-x),$$

where κ is defined in (3.8), the semigroup $P_{t,\kappa}(dw)$ is defined in (3.9), and $\theta \in (0,1)$ is some constant. Note that since $\alpha \in (1,2)$, we have $\kappa = \gamma \leqslant 1$. Using (2.20), we get

$$\rho_t^{1-\kappa} \leqslant ct^{-(1-\kappa)/\alpha} = ct^{-1+\delta_1},$$

where

$$(3.20) \qquad\qquad \delta_1 := 1 + \frac{\kappa - 1}{\alpha}.$$

Note that since $\alpha \in (1,2)$, we have $\alpha + \kappa - 1 > 0$, which implies $\delta_1 > 0$. Thus,

$$(3.21) \qquad |J_1| \leqslant c_3 t^{-1+\delta_1}(g_{t,\theta} * P_{t,\kappa})(y - x).$$

To estimate J_2 recall that by the Taylor expansion we have

$$(3.22) \quad p_t^0(x + u, y) - p_t^0(x, y) - u \cdot \nabla p_t^0(x, y)$$
$$= \sum_{1 \leqslant i,j \leqslant n} u_i u_j \int_0^1 (1 - \vartheta) \frac{\partial^2}{\partial x_i \partial x_j} p_t^0(x + \vartheta u, y) d\vartheta.$$

Using (3.15), (3.22), and the definition of f_t, we obtain the estimates

$$(3.23) \quad |p_t^0(x + u, y) - p_t^0(x, y) - u \cdot \nabla p_t^0(x, y)|$$
$$\leqslant c_1 \Big| \sum_{1 \leqslant i,j \leqslant n} u_i u_j \Big| \rho_t^2 \int_0^1 (1 - \vartheta) f_t(y - x - \vartheta u) d\vartheta$$
$$\leqslant n^2 c_1 \|u\|^2 \rho_t^2 \int_0^1 f_t(y - x - \vartheta u) d\vartheta$$
$$\leqslant n^2 c_1 \|u\|^2 \rho_t^2 \Big[\int_0^1 e^{c_2 \vartheta \rho_t \|u\|} d\vartheta \Big] f_t(y - x)$$
$$\leqslant c_3 \|u\|^2 \rho_t^2 f_t(y - x),$$

where to get the last line we used the fact that in J_2 we have $\rho_t \|u\| \leqslant 1$. Observe that for any $r > 0$

$$\int_{r\|u\| \leqslant 1} (r\|u\|)^2 \mu(du) = \int_0^1 \mu\{u : v \leqslant (r\|u\|)^2 \leqslant 1\} dv$$
$$\leqslant \sum_{i=1}^n \int_0^1 \mu\{u : v/n \leqslant |r u_i|^2 \leqslant 1\} dv$$
$$\leqslant n^2 \max_{1 \leqslant i \leqslant n} \int_0^{1/n} \mu\{u : z \leqslant |r u_i|^2 \leqslant 1\} dz$$
$$\leqslant n^2 \max_{1 \leqslant i \leqslant n} \int_0^1 \mu\{u : z \leqslant |r u_i|^2 \leqslant 1\} dz$$
$$= n^2 \max_{1 \leqslant i \leqslant n} q^L(r\ell_i) \leqslant n^2 \max_{1 \leqslant i \leqslant n} q^U(r\ell_i)$$
$$\leqslant n^2 q^*(r),$$

where $\ell_i := (0, \ldots, \underset{i}{1}, \ldots, 0) \in \mathbb{S}^n$.

Thus, using **A3**, the above calculation and the equality $q^*(\rho_t) = 1/t$, we can estimate J_2 as follows:

$$|J_2| \leqslant c_1(\|y-x\|^\gamma \wedge 1)f_t(y-x) \int\limits_{\rho_t\|u\|\leqslant 1} (\|u\|\rho_t)^2\mu(du)$$

$$\leqslant c_2(\|y-x\|^\gamma \wedge 1)f_t(y-x)q^*(\rho_t) \leqslant c_2 t^{-1}(\|y-x\|^\kappa \wedge 1)f_t(y-x).$$

Moreover, by Proposition 3.6 and (2.21), we get

$$(3.24) \qquad |J_2| \leqslant c_3 t^{-1+\delta_2}(g_{t,\theta} * P_{t,\kappa})(y-x), \quad \text{where } \delta_2 := \kappa/\sigma.$$

Let us estimate J_3. We have

$$(3.25) \qquad |J_3| \leqslant \int\limits_{\rho_t\|u\|>1} \big(p_t^0(x+u,y) + p_t^0(x,y)\big)|m(y,u)-m(x,u)|\mu(du)$$

$$+ \Big| \int\limits_{1/\rho_t<\|u\|<1} u\cdot\nabla p_t^0(x,y)\,[m(y,u)-m(x,u)]\mu(du)\Big|$$

$$=: J_{31} + J_{32}.$$

For J_{31}, by **A3**, (3.15) and Proposition 3.6 we get the estimates

$$J_{31} \leqslant b_4 t^{-1}(\|y-x\|^\gamma \wedge 1)\Big\{ \int\limits_{\mathbb{R}^n} p_t^0(x+u,y)\Lambda_t(du) + p_t^0(x,y)\Lambda_t(\mathbb{R}^n)\Big\}$$

$$\leqslant c_1 t^{-1}(\|y-x\|^\gamma \wedge 1)\big((f_t * \Lambda_t)(y-x) + n^2 f_t(y-x)\big)$$

$$\leqslant c_2 t^{-1}(\|y-x\|^\kappa \wedge 1)[(f_t * \Lambda_t)(y-x) + f_t(y-x)]$$

$$\leqslant c_2 t^{-1}\Big\{ \int\limits_{\mathbb{R}^n} (\|y-x-u\|^\kappa \wedge 1)f_t(y-x-u)\Lambda_t(du)$$

$$+ \int\limits_{\mathbb{R}^n} f_t(y-x-u)(\|u\|^\kappa \wedge 1)\Lambda_t(du) + (\|x-y\|^\kappa \wedge 1)f_t(y-x)\Big\}$$

$$\leqslant c_3 t^{-1+\delta_2}\{(g_{t,\theta} * P_{t,\kappa} * \Lambda_t)(y-x) + (f_t * P_{t,\kappa})(y-x) + (g_{t,\theta} * P_{t,\kappa})(y-x)\}$$

$$\leqslant c_4 t^{-1+\delta_2}\big(g_{t,\theta} * (P_t * P_{t,\kappa} + P_{t,\kappa})\big)(y-x),$$

where in the last line we used the relation $f_t(x) \leqslant c(g_{t,\theta} * P_t)(x)$, and the fact that $\Lambda_t(du)$ is dominated by $P_t(du)$.

Finally, we estimate J_{32}. By **A4**, $J_{32} = 0$ in the case $\alpha \in (0,1]$. Assume that $\alpha \in (1,2)$. In this case, by **A3**, (3.15), Proposition 3.2 with $\lambda = 1$, and Proposition 3.6, we get

$$|J_{32}| \leqslant c_1 t^{-1}(\|y-x\|^\gamma \wedge 1)\rho_t f_t(y-x)\Big(\int\limits_{\mathbb{R}^n} (\|u\| \wedge 1)\Lambda_t(du)\Big)$$

$$\leqslant c_2 t^{-1}(\|y-x\|^\kappa \wedge 1)f_t(y-x) \leqslant c_3 t^{-1}\rho_t^{-\kappa}(g_{t,\theta} * P_{t,\kappa})(y-x)$$

$$\leqslant c_3 t^{-1+\delta_2}(g_{t,\theta} * P_{t,\kappa})(y-x),$$

where in the last line we used (2.21).

Thus, we arrive at

$$(3.26) \qquad J_3 \leqslant c_4 t^{-1+\delta_2}(g_{t,\theta} * G_t)(y - x).$$

Put

$$\eta := \delta_1 \wedge \delta_2 = \frac{\kappa}{\sigma} \wedge \left(1 + \frac{\kappa - 1}{\alpha}\right).$$

Thus, combining the estimates for J_1, J_2 and J_3, we obtain (3.11) with some constant $C > 0$, η, and $b = \theta a_2$, where $\theta \in (0, 1)$ is arbitrary, and a_2 is the constant from Proposition 3.4 (cf. Remark 3.1). ∎

3.3. Generic calculation. Let us rewrite the statement of Lemma 3.1 a bit differently. Although it might be seen as just some technical modification, it will become clear later that this new form allows us to write the estimate in a rather transparent way.

Put

$$(3.27) \qquad\qquad \delta := \eta/2,$$

and

$$(3.28) \qquad\qquad \tilde{g}_t(x) := t^\delta g_t(x).$$

Then the estimate (3.11) can be written as

$$(3.29) \quad |\Phi_t(x, y)| \leqslant C t^{-1+\delta}(\tilde{g}_t * G_t)(y - x), \quad t \in (0, T], \; x, y \in \mathbb{R}^n.$$

The next important step is to estimate iteratively the convolution powers $\Phi^{\circledast k}$, and this is the place where we encounter essential new difficulties. Below we explain this problem in detail, and give the generic calculation, which allows us to overcome these difficulties.

Let us define

$$(3.30) \qquad\qquad H_t(x, y) := (\tilde{g}_t * G_t)(y - x).$$

Observe that if this kernel would satisfy the following *subconvolution property*

$$(3.31) \qquad (H_{t-s} * H_s)(x, y) \leqslant c H_t(x, y), \quad 0 < s < t, \; x, y \in \mathbb{R}^n,$$

then the iterative estimation of the convolution powers $\Phi^{\circledast k}$ would be simple. For example, this is true for a perturbed α-stable noise: in this case we have

$$|\Phi_t(x, y)| \leqslant C t^{-1+\delta} H_t(x, y), \quad t \in (0, T], \; x, y \in \mathbb{R}^n,$$

with

$$H_t(x, y) = t^{-n/\alpha}(1 + \|y - x\|/t^{1/\alpha})^{-n-\alpha},$$

see [51], [52] and [7]; see also [49] for more involved kernels which appear for the gradient perturbations of an α-stable noise with $\alpha < 1$.

In our situation the kernel $H_t(x, y)$ has the more complicated structure: it is formed by the convolution of the function g_t and some measure G_t, which seems to be inevitable because of the "compound kernel" structure of the first approximation $p_t^0(x, y)$ given by (2.9). Moreover, in this case we cannot in general expect (3.31) to hold true.

To show what is going on, we give a calculation of the upper bound for the convolution of two "compound kernels".

LEMMA 3.2. *Suppose that the functions* Φ^i, $i = 1, 2$, *satisfy for any* $T > 0$ *the inequalities*

$$(3.32) \quad |\Phi_t^i(x, y)| \leqslant C_i t^{-1+\delta_i}(h_t^i * G_t^i)(y - x), \quad t \in (0, T], \ x, y \in \mathbb{R}^n,$$

with constants $C_i > 0$, $\delta_i > 0$, *some nonnegative and integrable functions* h^i, *and some subprobability measures* G_t^i, *respectively. Then* $\mathfrak{F} := \Phi^1 \circledast \Phi^2$ *satisfies*

$$(3.33) \quad |\mathfrak{F}_t(x, y)| \leqslant C t^{-1+\delta}(\mathfrak{h}_t * \mathfrak{G}_t)(y - x), \quad t \in (0, T], \ x, y \in \mathbb{R}^n,$$

with

$$\mathfrak{h}_t(x) = \sup_{s < t}(h_{t-s}^1 * h_s^2)(x),$$

$$\delta = \delta_1 + \delta_2, \quad C = C_1 C_2 B(\delta_1, \delta_2),$$

where $B(\cdot, \cdot)$ *is the beta function, and*

$$\mathfrak{G}_t(dw) := \frac{1}{B(\delta_1, \delta_2)} \int_0^1 \int_{\mathbb{R}^n} (1 - r)^{-1+\delta_1} r^{-1+\delta_2} G_{t(1-r)}^1(dw - u) G_{tr}^2(du) dr.$$

Moreover, $\mathfrak{G}_t(dw)$ *is the subprobability measure, i.e.,* $\mathfrak{G}_t(\mathbb{R}^n) \leqslant 1, \ t \in [0, T].$

P r o o f. Making the change of variables, we have

$$|\mathfrak{F}_t(x, y)| \leqslant C_1 C_2 \int_0^t \int_{\mathbb{R}^{3n}} h_{t-s}^1(z - w_1 - x) h_s^2(y - w_2 - z)$$

$$\times \frac{G_{t-s}^1(dw_1) G_s^2(dw_2)}{(t - s)^{1-\delta_1} s^{1-\delta_2}} dz ds$$

$$\leqslant C_1 C_2 B(\delta_1, \delta_2) t^{-1+\delta_1+\delta_2}$$

$$\times \int_{\mathbb{R}^n} \mathfrak{h}_t(y - x - w) \left[\int_0^1 \int_{\mathbb{R}^n} \frac{G_{t(1-r)}^1(dw - u) G_{tr}^2(du)}{B(\delta_1, \delta_2)(1 - r)^{1-\delta_1} r^{1-\delta_2}} dr \right],$$

which gives (3.33). Further, since $G_t^i(\mathbb{R}^n) \leqslant 1$, $i = 1, 2$, we get $\mathfrak{G}_t(\mathbb{R}^n) \leqslant 1$. ∎

By Lemma 3.2, we have the following estimates for the convolution powers of Φ:

$$(3.34) \quad |\Phi_t^{\circledast k}(x,y)| \leqslant \frac{C^k \Gamma^k(\delta)}{\Gamma(k\delta)} t^{-1+\delta k}(h_t^{(k)} * G_t^{(k)})(y-x), \quad t \in (0,T],$$

where $x, y \in \mathbb{R}^n$, $k \geqslant 1$, the constant $C > 0$ comes from (3.29), $h^{(1)} \equiv \tilde{g}$,

$$h_t^{(k+1)}(x) = \sup_{s<t}(h_{t-s}^{(k)} * h_s^{(1)})(x), \quad k \geqslant 1,$$

and

$$(3.35) \quad G_t^{(k)}(du)$$

$$:= \begin{cases} G_t(du), & k = 1, \\ \dfrac{1}{B\big(\delta, (k-1)\delta\big)} \int\limits_0^1 \int\limits_{\mathbb{R}^n} (1-r)^{-1+(k-1)\delta} r^{-1+\delta} G_{t(1-r)}^{(k-1)}(dw-u) G_{tr}(du) dr, \\ & k \geqslant 2, \end{cases}$$

Hence, to guarantee the convergence of the series of convolution powers (2.7), it is enough to derive a proper upper bound on the sequence of functions $h_t^{(k)}$. For this reason, we give the following lemma.

LEMMA 3.3. *Let $g_t(x)$ be defined in (3.12). Then for any $\theta \in (0,1)$, $T > 0$, we have*

$$(3.36) \quad (g_{t-s} * g_s)(x) \leqslant C_0(\theta) g_t(\theta x), \quad 0 < s < t, \ x \in \mathbb{R}^n, \ t \in (0,T],$$

where $C_0(\theta) = c(1-\theta)^{-n}$, and $c > 0$ is some constant.

P r o o f. Consider the integral

$$(3.37) \qquad\qquad I(t,x) := \int\limits_{\mathbb{R}^n} g_{t-s}(x-z) g_s(z) dz.$$

Suppose that $0 < s \leqslant t/2$. Note that for $s < t/2$ we infer by the monotonicity of ρ_t that $\rho_{t-s} \leqslant \rho_{t/2}$. Further, for $c_1 \geqslant 1$ we have $q^*(r) \leqslant q^*(c_1 r) \leqslant c_1^2 q^*(r)$ for all $r \geqslant 1$, which implies $\rho_t \asymp \rho_{c_1 t}$ for all $t \in (0,T]$. Therefore,

$$\rho_{t/2} \leqslant c_2 \rho_t, \quad t \in (0,T].$$

Since ρ_t is decreasing, the triangle inequality

$$\|x-z\|\rho_{t-s} + \|z\|\rho_s \geqslant \|x\|\rho_t$$

gives

$$(3.38) \qquad I(t,x) \leqslant e^{-b\theta\|x\|\rho_t} \int_{\mathbb{R}^n} \rho_{t-s}^n \rho_s^n e^{-b(1-\theta)[\rho_{t-s}\|x-z\|+\rho_s\|z\|]} dz$$

$$\leqslant \rho_{t/2}^n e^{-b\theta\rho_t\|x\|} \int_{\mathbb{R}^n} \rho_s^n e^{-b(1-\theta)\rho_s\|z\|} dz$$

$$= C(\theta)g_t(\theta x),$$

where $C(\theta) := c_2^n c_0 [b(1-\theta)]^{-n}$, $c_0 := \int_{\mathbb{R}^n} e^{-\|z\|} dz$. ∎

LEMMA 3.4. *For any $T > 0$ and any sequence $(\theta_k)_{k\geqslant 1}$ such that*

$$\theta_1 = 1, \quad and \quad \theta_{k+1} < \theta_k, \quad \theta_k > 0, \quad k \geqslant 1,$$

we have for $k \geqslant 2$ the estimate

$$(3.39) \quad |\Phi_t^{\circledast k}(x,y)| \leqslant C_k t^{-1+\delta k} (g_t^{(k)} * G_t^{(k)})(y-x), \quad t \in (0,T], \ x,y \in \mathbb{R}^n,$$

where the subprobability measures $(G^{(k)})_{k\geqslant 1}$ are defined in (3.35),

$$(3.40) \qquad\qquad g_t^{(k)}(x) := t^{\delta k} g_t(\theta_k x), \quad k \geqslant 1,$$

and

$$C_k := \frac{C^k c^{k-1} \Gamma^k(\delta)}{\Gamma(k\delta)} \prod_{j=2}^{k} \left(\frac{1}{\theta_{j-1} - \theta_j} \right)^n, \quad k \geqslant 2,$$

with the positive constants C and c coming from (3.29) and Lemma 3.3, respectively.

P r o o f. By the monotonicity of $g_t(x)$ in x we have $g_t(x) \leqslant g_t(\theta_{k-1}x)$. Therefore, using Lemma 3.3 with $\theta = \theta_k/\theta_{k-1}$, we get

$$(g_{t-s}^{(k-1)} * g_s^{(1)})(x) \leqslant t^{\delta k} \int_{\mathbb{R}^n} g_{t-s}(\theta_{k-1}x - \theta_{k-1}y) g_s(\theta_{k-1}y) \, dy$$

$$= t^{\delta k} \theta_{k-1}^{-n} \int_{\mathbb{R}^n} g_{t-s}(\theta_{k-1}x - y') g_s(y') \, dy'$$

$$\leqslant D_k t^{\delta k} g_t(\theta_k x)$$

$$= D_k g_t^{(k)}(x),$$

where

$$D_k = (\theta_{k-1})^{-n} C_0(\theta_k/\theta_{k-1}) = \frac{c}{\left(\theta_{k-1}(1 - \theta_k/\theta_{k-1})\right)^n} = \frac{c}{\left(\theta_{k-1} - \theta_k\right)^n}.$$

Then (3.39) follows from (3.34). ∎

The estimate (3.39) is still hardly applicable to verifying the convergence of the series of convolution powers (2.7): to keep this sequence of estimates consistent, one should choose the sequence $\{\theta_k\}$ such that $\inf_k \theta_k > 0$, and then it is difficult to bound properly the values of the constants C_k. In order to illustrate this, take, e.g., $\theta_k := 1/2 + 1/(2k)$. Then

$$\prod_{j=2}^{k} \left(\frac{1}{\theta_{j-1} - \theta_j} \right)^n = \prod_{j=2}^{k} \left(2j(j-1) \right)^n = \left(2^k k! (k-1)! \right)^n,$$

which increases faster than $\Gamma(k\delta)$ in the denominator in the definition of C_k.

In order to overcome this problem, we change after a finite number of steps the sequence $g_t^{(k)}(x)$. This change finally allows us to prove the convergence of the series (2.7).

Let

(3.41) $$k_0 = \left[\frac{n}{\alpha\delta} \right] + 1.$$

Note that for such k_0 we have $t^{\delta k_0} \rho_t^n \leqslant c(k_0)$ for $t \in (0, T]$. Then

$$(g_{t-s}^{(k_0)} * g_s^{(1)})(x) \leqslant c(k_0) \int_{\mathbb{R}^n} s^\delta \rho_s^n e^{-b\theta_{k_0}\rho_{t-s}\|x-z\| - b\rho_s\|z\|} dz$$

$$\leqslant c(k_0) M e^{-b\zeta\rho_t\|x\|} = c(k_0) M \rho_t^{-n} g_{t,\zeta}(x), \quad 0 < s < t,$$

where $\zeta = \theta_{k_0}$, $g_{t,\zeta}(x)$ is of the form (3.17), and

(3.42) $$M := T^\delta \int_{\mathbb{R}^n} e^{-b(1-\zeta)\|z\|} dz.$$

By induction, we get

(3.43)
$$\bar{g}_t^{(k_0+\ell+1)}(x) := \sup_{0<s<t} (\bar{g}_{t-s}^{(k_0+\ell)} * g_s^{(1)})(x) \leqslant c(k_0) M^{\ell+1} \rho_t^{-n} g_{t,\zeta}(x), \quad \ell \geqslant 0.$$

LEMMA 3.5. *For any $T > 0$ we have for $\ell \geqslant 1$, $t \in (0, T]$, $x, y \in \mathbb{R}^n$,*

(3.44) $$|\Phi_t^{\circledast(k_0+\ell)}(x,y)| \leqslant D_\ell t^{-1+\delta(k_0+\ell)} \rho_t^{-n} (g_{t,\zeta} * G_t^{(k_0+\ell)})(y-x),$$

where k_0 is given by (3.41), the family of subprobability measures $\{G_t^{(k)}, t > 0, k \geqslant 1\}$ is defined in (3.35),

(3.45) $$D_\ell := \frac{C(k_0)(CM)^\ell \Gamma^{k_0+\ell}(\delta)}{\Gamma\big((k_0+\ell)\delta\big)},$$

$C(k_0) > 0$ being some constant, and $C, M > 0$ coming from (3.29) and (3.42), respectively.

The proof follows by induction; we omit the details.

3.4. Proof of Theorem 2.1. (a) Using the relations (3.39), (3.44) and the inequality $g_t^{(k)}(x) \leqslant T^{k\delta} g_{t,\zeta}(x)$, $1 \leqslant k \leqslant k_0$, we get for $t \in (0, T]$

$$(3.46) \quad |\Psi_t(x, y)|$$

$$\leqslant \sum_{k=1}^{\infty} |\Phi_t^{\circledast k}(x, y)| \leqslant \sum_{k=1}^{k_0} C_k t^{-1+k\delta} (g_t^{(k)} * G_t^{(k)})(y - x)$$

$$+ \sum_{\ell=1}^{\infty} D_\ell t^{-1+\delta(k_0+\ell)} \rho_t^{-n} (g_{t,\zeta} * G_t^{(k_0+\ell)})(y - x)$$

$$\leqslant t^{-1+\delta} \Big(g_{t,\zeta} * \big(\sum_{k=1}^{k_0} T^{\delta(k-1)} C_k G_t^{(k)} + \sum_{\ell=1}^{\infty} T^{\delta(k_0+\ell-1)} D_\ell G_t^{(k_0+\ell)} \big) \Big)(y - x)$$

$$\leqslant \big(\sum_{k=1}^{k_0} T^{\delta(k-1)} C_k + \sum_{\ell=1}^{\infty} T^{\delta(k_0+\ell-1)} D_\ell \big) t^{-1+\delta} (g_{t,\zeta} * \Pi_t)(y - x),$$

$$(3.47)$$

$$\Pi_t(du) := \frac{\sum_{k=1}^{k_0} T^{\delta(k-1)} C_k G_t^{(k)}(du) + \sum_{\ell=1}^{\infty} T^{\delta(k_0+\ell-1)} D_\ell G_t^{(k_0+\ell)}(du)}{\sum_{k=1}^{k_0} T^{\delta(k-1)} C_k + \sum_{\ell=1}^{\infty} T^{\delta(k_0+\ell-1)} D_\ell}, \quad t \in (0, T].$$

Since $G_t^{(k)}$, $k \geqslant 1$, are the subprobability measures, we have

$$(3.48) \qquad\qquad\qquad \Pi_t(\mathbb{R}^n) \leqslant 1, \quad t \in (0, T].$$

Thus, we proved that the series $\Psi_t(x, y) = \sum_{k=1}^{\infty} \Phi_t^{\circledast k}(x, y)$ converges for any $t \in (0, T]$, $x, y \in \mathbb{R}^n$, uniformly on compact subsets of $(0, \infty) \times \mathbb{R}^n \times \mathbb{R}^n$.

Finally, let us show that $(p^0 \circledast \Psi)_t(x, y)$ is well defined. Using the upper bound for p^0 (cf. (3.15) with $k = 0$) and the estimate for Ψ (cf. (3.46)), by Lemmas 3.2 and 3.3 we get

$$(3.49) \quad |(p^0 \circledast \Psi)_t(x, y)| \leqslant t^\delta (g_{t,\chi} * \tilde{\Pi}_t)(y - x), \quad t \in (0, T], \, x, y \in \mathbb{R}^n,$$

where $\chi \in (0, \zeta)$, ζ comes from (3.46), $g_{t,\kappa}(x)$ is of the form (3.17), and

$$(3.50) \qquad \tilde{\Pi}_t(dw) := \delta \int_0^1 \int_{\mathbb{R}^n} (1 - r)^{-1+\delta} \Pi_{t(1-r)}(dw - u) P_{tr}(du) dr.$$

By the definition of Π_t and P_t, $\tilde{\Pi}_t(\mathbb{R}^n) \leqslant 1$. Thus, the expression (2.4) is well defined, and the series involved in this expression converges absolutely, uniformly on compact sets of $(0, \infty) \times \mathbb{R}^n \times \mathbb{R}^n$.

(b) By assumption **A2** we have

$$\|\xi\|^k |e^{i\xi(x-y) - tq(y,\xi)}| \leqslant e^{-ct_0 q(\xi)}, \quad x, y \in \mathbb{R}^n, \, t \in [t_0, \infty),$$

for any $t_0 > 0$ and $k \geqslant 0$. Therefore, since the function $|e^{i\xi(x-y)-tq(y,\xi)}|$ is continuous, has continuous derivatives, and $q(\xi) \geqslant c\|\xi\|^\alpha$ for $\|\xi\| \geqslant 1$ (see (2.13) and (2.19)), it follows by the dominated convergence theorem that the function

$$p_t^0(x,y) = (2\pi)^{-n} \int\limits_{\mathbb{R}^n} e^{i\xi(x-y)-tq(y,\xi)} d\xi$$

has continuous derivatives in $(t,x) \in (0,\infty) \times \mathbb{R}^n$.

Next we prove that the functions $\Phi_t^{\circledast k}(x,y)$, $k \geqslant 0$, are continuous in (t,x,y) on $(0,\infty) \times \mathbb{R}^n \times \mathbb{R}^n$. Let us show that $\Phi_t(x,y)$ is continuous. As we have just shown, for any $t_0 > 0$

$$|\partial_{x_i x_j}^2 p_t^0(x,y)| \leqslant c, \quad t \geqslant t_0,\ x,y \in \mathbb{R}^n,\ 1 \leqslant i,j \leqslant n,$$

we get

$$|p_t^0(x+u,y) - p_t^0(x,y) - \nabla_x p_t^0(x,y) \cdot u \mathbb{1}_{\{\|u\|\leqslant 1\}}| \leqslant C(\|u\|^2 \wedge 1),$$

where $t > t_0$, $x,y \in \mathbb{R}^n$. Therefore, from the dominated convergence theorem we derive that $L_x p_t^0(x,y)$ is continuous in (t,x,y) on $[t_0,\infty) \times \mathbb{R}^n \times \mathbb{R}^n$. Therefore, since $\partial_t p_t^0(x,y)$ is continuous in (t,x,y) on $[t_0,\infty) \times \mathbb{R}^n \times \mathbb{R}^n$ and $t_0 > 0$ is arbitrary, we obtain the desired continuity of $\Phi_t(x,y)$.

To show that the convolutions $\Phi_t^{\circledast k}(x,y)$ are continuous, we use induction.

Suppose that $\Phi_t^{\circledast(k-1)}(x,y)$ is continuous. Let $t_0 > \varepsilon > 0$, and suppose that $t \in [t_0,\infty)$. Write

$$(3.51) \qquad \Phi_t^{\circledast k}(x,y) = \int\limits_0^{t-\varepsilon} \Big[\int\limits_{\mathbb{R}^n} \Phi_{t-s}^{\circledast(k-1)}(x,z)\Phi_s(z,y)dz \Big] ds$$

$$+ \int\limits_{t-\varepsilon}^{t} \Big[\int\limits_{\mathbb{R}^n} \Phi_{t-s}^{\circledast(k-1)}(x,z)\Phi_s(z,y)dz \Big] ds$$

$$= \int\limits_0^{t-\varepsilon} \Big[\int\limits_{\mathbb{R}^n} \Phi_{t-s}^{\circledast(k-1)}(x,y-z)\Phi_s(y-z,y)dz \Big] ds$$

$$+ \int\limits_0^{\varepsilon} \Big[\int\limits_{\mathbb{R}^n} \Phi_{s}^{\circledast(k-1)}(x,x-z)\Phi_{t-s}(x-z,y)dz \Big] ds$$

$$=: I_1(t,x,y) + I_2(t,x,y).$$

We prove the continuity of $I_1(t,x,y)$; the continuity of $I_2(t,x,y)$ follows by the same argument.

By the induction assumption, the function $\Phi_{t-s}^{\circledast(k-1)}(x,z)\Phi_s(z,y)$ is continuous in $s \in (0,t-\varepsilon]$, $t \in [t_0,\infty)$, $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n$. Moreover, by (3.39) and (3.11) we obtain

$$|\Phi_{t-s}^{\circledast(k-1)}(x,y-w)\Phi_s(y-w,y)| \leqslant C(t-s)^{-1+2\delta(k-1)}s^{-1+\eta}\rho_{t-s}^n(g_s * G_s)(w)$$

$$\leqslant C(\varepsilon)s^{-1+\eta}(g_s * G_s)(w).$$

Since the right-hand side of the above inequality is integrable on $[0, t - \varepsilon] \times \mathbb{R}^n$, we infer by the dominated convergence theorem that $I_1(t, x, y)$ is continuous in $(t, x, y) \in [t_0, \infty) \times \mathbb{R}^n \times \mathbb{R}^n$. Finally, since t_0 and ε are arbitrary, we get the continuity in (t, x, y) on $(0, \infty) \times \mathbb{R}^n \times \mathbb{R}^n$.

Since the series $\sum_{k=1}^{\infty} \Phi_t^{\circledast k}(x, y)$ converges uniformly on compact subsets of $(0, \infty) \times \mathbb{R}^n \times \mathbb{R}^n$, the function $\Psi_t(x, y)$ is continuous in $(t, x, y) \subset (0, \infty) \times \mathbb{R}^n \times \mathbb{R}^n$.

The proof of the continuity of $p^0 \circledast \Psi$ follows by the same argument as the proof of continuity of $\Phi_t^{\circledast k}$, we only need to use estimates on p^0 and Ψ, see (3.15) and (3.46). ∎

3.5. Continuity properties of the operator S_t. Note that by construction we have for any subprobability measure $M_t(\cdot)$ on $\mathbb{R}^n$

$$(3.52) \qquad \int_{\mathbb{R}^n} (g_{\theta, t} * M_t)(y) dy \leqslant C, \quad t \in (0, T], \ x \in \mathbb{R}^n.$$

Therefore, by (3.15) and (3.49) we have

$$\int_{\mathbb{R}^n} p_t(x, y) dy \leqslant C, \quad t \in (0, T], \ x \in \mathbb{R}^n.$$

Then the operator $S_t f$, $t > 0$ (cf. (2.10)), is well defined for any bounded measurable function f.

LEMMA 3.6. (1) *For any $t > 0$ the operator S_t maps $C_\infty(\mathbb{R}^n)$ into $C_\infty(\mathbb{R}^n)$.* (2) *For every $f \in C_\infty(\mathbb{R}^n)$ we have $\lim_{t \to 0+} \|S_t f - f\|_\infty = 0$.*

The proof relies on the proposition below.

PROPOSITION 3.7. *For every $f \in C_\infty(\mathbb{R}^d)$*

$$(3.53) \qquad \lim_{|x| \to \infty} \int_{\mathbb{R}^d} p_t^0(x, y) f(y) dy = 0 \quad \textit{for any } t > 0,$$

$$(3.54) \qquad \sup_{x \in \mathbb{R}^d} \Big| \int_{\mathbb{R}^d} p_t^0(x, y) f(y) dy - f(x) \Big| \to 0, \quad t \to 0.$$

In order to keep the presentation as clear as possible, we defer the proof of this proposition to Appendix B.

P r o o f o f L e m m a 3.6. (1) The continuity of $S_t f$ follows from the continuity of $p_t(x, y)$. To prove that $S_t f(x)$ vanishes as $\|x\| \to \infty$, we use the representation for $p_t(x, y)$ (cf. (2.4) and (2.6)):

$$\Big| \int_{\mathbb{R}^n} p_t(x, y) f(y) dy \Big| \leqslant \int_{\mathbb{R}^n} p_t^0(x, y) |f(y)| dy + \int_{\mathbb{R}^n} |(p^0 \circledast \Psi)_t(x, y) f(y)| dy$$
$$=: I_1(t, x) + I_2(t, x).$$

By statement (3.53) of Proposition 3.7, the first term on the right-hand side tends to zero as $\|x\| \to \infty$. Using the upper estimate on $p^0 \circledast \Psi$ (cf. (3.49)), we get

$$I_2(t,x) \leqslant Ct^\delta \int\limits_{\mathbb{R}^n} \Big[\int\limits_{\mathbb{R}^n} g_{t,\chi}(y - x - w)|f(y)|dy \Big] \tilde{\Pi}_t(dw)$$
$$= Ct^\delta \int\limits_{\mathbb{R}^n} \Big[\int\limits_{\mathbb{R}^n} g_{t,\chi}(z - w)|f(z + x)|dy \Big] \tilde{\Pi}_t(dw),$$

and the right-hand side tends to zero as $|x| \to \infty$ by the dominated convergence theorem (recall that the parameter χ comes from (3.49)).

(2) By (3.54) it is enough to show that

(3.55) $$\sup_x \Big| \int\limits_0^t \int\limits_{\mathbb{R}^d} \int\limits_{\mathbb{R}^d} \Psi_{t-s}(x,z)p_s^0(z,y)f(y)dydzds \Big| \to 0, \quad t \to 0.$$

By (3.49) we have

$$\sup_x \Big| \int\limits_0^t \int\limits_{\mathbb{R}^d} \int\limits_{\mathbb{R}^d} \Psi_{t-s}(x,z)p_s^0(z,y)f(y)dydzds \Big| \leqslant c_1 t^\delta \int\limits_{\mathbb{R}^n} (g_{t,\chi} * \tilde{\Pi}_t)(y)\, dy \leqslant c_2 t^\delta,$$

which completes the proof. ■

4. APPROXIMATE POSITIVE MAXIMUM PRINCIPLE. PROOF OF THEOREM 2.2

We follow, with the necessary changes, the approach described in [49].

In Section 3 we constructed the function $p_t(x,y)$ under the assumption that $p_t(x,y)$ is a fundamental solution to the Cauchy problem for $\partial_t - L$. As in [49], the straightforward way to check (2.2) and (2.3) meets difficulties. To explain them, let us look at the behavior of the derivatives of $p_t^0(x,y)$ near the origin.

Let

(4.1) $$\mathcal{P}_t(du) := c\big(P_t(du) + (P_t * \Lambda_t)(du)\big),$$

where for any $T > 0$ the constant $c > 0$ is chosen such that $\mathcal{P}_t(\mathbb{R}^n) \leqslant 1$ for all $t \in (0, T]$, which is possible since $\Lambda_t(\mathbb{R}^n) \leqslant n^2$ (cf. Proposition 3.1).

PROPOSITION 4.1. *The function $p_t^0(x,y)$ is differentiable with respect to t, the derivative $\partial_t p_t^0(x,y)$ is continuous in $(t,x,y) \in (0,\infty) \times \mathbb{R}^n \times \mathbb{R}^n$, and for $k = k_1 + \ldots + k_n \geqslant 0$, $T > 0$, there exist constants $\tilde{A}_k, \tilde{a}_k > 0$ such that*

(4.2)
$$\left| \frac{\partial}{\partial t} \frac{\partial^k}{\partial x_1^{k_1} \ldots \partial x_n^{k_n}} p_t^0(x,y) \right| \leqslant t^{-1} \rho_t^k (f_{up} * \mathcal{P}_t)(y - x), \quad t \in (0,T], \ x, y \in \mathbb{R}^n,$$

where $\mathcal{P}_t(du)$ is the subprobability measure defined in (4.1), *and f_{up} is of the form* (2.22) *with constants $\tilde{A}_k$ and $\tilde{a}_k$ in place of d_1 and d_2, respectively.*

The proof of this proposition can be obtained by a simple modification of the proof of the respective statement in [48], see also [46]. We omit the details.

From (4.2) and (3.46) it is unclear why the function $p_t(x, y)$ given by (2.4) is in the domain of the operator L, which was so far defined on $C^2_\infty(\mathbb{R}^n)$-functions. Consequently, we cannot check straightforwardly if (2.3) holds true, and to verify in this way the correctness of the procedure performed in Section 3.

To avoid this difficulty, we introduce for $\varepsilon > 0$ the auxiliary function

$$(4.3) \qquad p_{t,\epsilon}(x, y) := p^0_{t+\epsilon}(x, y) + \int_0^t \int_{\mathbb{R}^n} p^0_{t-s+\varepsilon}(x, z) \Psi_s(z, y) dz ds.$$

Since the additional time shift by positive ε removes the singularity at the point $s = t$, the function $p_{t,\varepsilon}(x, y)$ has the following properties:

(i) $p_{\cdot,\epsilon}(x, y) \in C^1\big((0, \infty)\big)$ for any fixed $\epsilon > 0$, $x, y \in \mathbb{R}^n$;

(ii) $p_{t,\epsilon}(\cdot, y) \in C^2_\infty(\mathbb{R}^n)$ for any fixed $\epsilon > 0$, $t > 0$, $y \in \mathbb{R}^n$;

(iii) for any $0 < \tau < T$ we have $p_{t,\epsilon}(x, y) \to p_t(x, y)$ as $\epsilon \to 0$, uniformly in $(t, x) \in [\tau, T] \times \mathbb{R}^n \times \mathbb{R}^n$;

(iv) for any $0 < \tau < T$ we have

$$q_{t,\epsilon}(x, y) := (\partial_t - L_x) p_{t,\epsilon}(x, y) \to 0, \quad \varepsilon \to 0,$$

uniformly in $(t, x, y) \in [\tau, T] \times \mathbb{R}^n \times \mathbb{R}^n$.

The proofs of the above properties are analogous to the proofs of the respective properties for

$$(4.4) \qquad S_{t,\varepsilon} f(x) := \int_{\mathbb{R}^n} p_{t,\epsilon}(x, y) f(y) dy, \quad t > 0, \ x \in \mathbb{R}^n,$$

given in Lemmas 4.1 and 4.2. Here we only mention that properties (iii) and (iv) motivate to name $p_{t,\varepsilon}(x, y)$ the *approximate fundamental solution*.

LEMMA 4.1. (1) *For any $f \in C_\infty(\mathbb{R}^n)$, $\varepsilon > 0$, the function $S_{t,\varepsilon} f(x)$ belongs to $C^1\big((0, \infty)\big)$ as a function of t, and to $C^2_\infty(\mathbb{R}^n)$ as a function of x.*

(2) *For every $f \in C_\infty(\mathbb{R}^n)$, $T > 0$,*

$$(4.5) \qquad \lim_{\varepsilon \to 0} \| S_{t,\varepsilon} f - S_t f \|_\infty = 0,$$

uniformly in $t \in [0, T]$; and for every $\varepsilon > 0$

$$(4.6) \qquad S_{t,\varepsilon} f(x) \to 0, \quad \|x\| \to \infty,$$

uniformly in $t \in [0, T]$.

(3) *For $f \in C_\infty(\mathbb{R}^n)$ we have*

$$\lim_{t,\varepsilon \to 0+} \| S_{t,\varepsilon} f - f \|_\infty = 0.$$

P r o o f. The proof of the first statement follows from the upper estimate (3.46) on $\Psi_t(x, y)$, Proposition 4.1, and the dominated convergence theorem.

Observe that the function

$$[0, T] \ni t \mapsto \int_{\mathbb{R}^n} p_t^0(\cdot, y) f(y) \, dy \in C_\infty(\mathbb{R}^n)$$

is continuous, since the function $p_t^0(x, y)$ is continuous in t for $t > 0$, and the continuity of the integral at $t = 0$ follows from Proposition 3.7. Then

$$\int_{\mathbb{R}^n} p_{t+\epsilon}^0(x, y) f(y) \, dy \to \int_{\mathbb{R}^n} p_t^0(x, y) f(y) \, dy, \quad \epsilon \to 0,$$

uniformly in $t \in [0, T], x \in \mathbb{R}^n$. This together with Proposition 4.1, the estimate (4.2) and the dominated convergence theorem implies statement (2).

The proof of statement (3) is a slight modification of the proof of statement (2) in Lemma 3.6: we just need to substitute $p_{t+\varepsilon}^0(x, y)$ and $\Psi_{t+\varepsilon-s}(x, z)$ for $p_t^0(x, y)$ and $\Psi_{t-s}(x, z)$ in (3.53) and (3.55), respectively. ∎

Let us put

$$(4.7) \qquad V_{t,\varepsilon} f(x) = (\partial_t - L_x) S_{t,\varepsilon} f(x), \quad f \in C_\infty(\mathbb{R}^n).$$

LEMMA 4.2. *For any $f \in C_\infty(\mathbb{R}^n)$ the following statements hold true*:
(1) *We have*

$$(4.8) \qquad V_{t,\varepsilon} f(x) \to 0, \quad \epsilon \to 0,$$

uniformly in $(t, x) \in [\tau, T] \times \mathbb{R}^n$ for any $\tau > 0, T > \tau$.
(2) *We have*

$$(4.9) \qquad \int_0^t V_{s,\varepsilon} f(x) ds \to 0, \quad \epsilon \to 0,$$

uniformly in $(t, x) \in [0, T] \times \mathbb{R}^n$ for any $T > 0$.

P r o o f. Note that $S_{t,\varepsilon} f \in C_\infty^2(\mathbb{R}^n)$, and thus the expression

$$(4.10) \quad LS_{t,\varepsilon} f(x)$$

$$= L_x \int_{\mathbb{R}^n} p_{t+\epsilon}^0(x, y) f(y) dy + L_x \int_0^t \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} p_{t-s+\epsilon}^0(x, z) \Psi_s(z, y) f(y) dy dz ds$$

is well defined. Let us show that we can interchange L_x with the integrals in (4.10), i.e., that

$$(4.11) \quad LS_{t,\varepsilon} f(x)$$

$$= \int_{\mathbb{R}^n} L_x p_{t+\epsilon}^0(x, y) f(y) dy + \int_0^t \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} L_x p_{t-s+\epsilon}^0(x, z) \Psi_s(z, y) f(y) dy dz ds.$$

Recall the representation of L, cf. (1.1). By Proposition 4.1, the gradient term in (1.1) can be interchanged with the integral by the dominated convergence theorem. To do the same with the "integral part" $\mathcal{L}$ of L, observe that

$$\mathcal{L}f(x) = \lim_{v \to 0+} \mathcal{L}^{(v)}f(x),$$

$$\mathcal{L}^{(v)}f(x) := \int_{\|u\|>v} \big(f(x+u) - f(x) - \nabla f(x) \cdot u \mathbb{1}_{\{\|u\|\leqslant 1\}}\big)\mu(x,du).$$

Clearly, the operator $L^{(v)}$ can be interchanged with the integrals by the Fubini theorem. On the other hand,

$$|\mathcal{L}f(x) - \mathcal{L}^{(v)}f(x)| = \Big|\int_{\|u\|\leqslant v} \big(f(x+u) - f(x) - \nabla f(x) \cdot u \mathbb{1}_{\{\|u\|\leqslant 1\}}\big)\mu(x,du)\Big|$$

$$\leqslant C\|\nabla^2 f\|_\infty \sup_{x\in\mathbb{R}^n} \int_{\|u\|\leqslant v} \|u\|^2 \mu(x,du).$$

Using again Proposition 4.1, (3.46), and the dominated convergence theorem, we can pass to the limit in the expression

$$\int_0^t \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \mathcal{L}_x^{(v)} p_{t-s+\epsilon}^0(x,z)\Psi_s(z,y)f(y)dydzds$$

as $v \to 0$. Thus, (4.11) holds true.

Similarly, by Proposition 4.1 we get

(4.12)

$$\partial_t S_{t,\varepsilon}f(x) = \int_{\mathbb{R}^n} \partial_t p_{t+\epsilon}^0(x,y)f(y)dy + \int_0^t \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \partial_t p_{t-s+\epsilon}^0(x,z)\Psi_s(z,y)f(y)dydzds$$

$$+ \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} p_\epsilon^0(x,z)\Psi_t(z,y)f(y)dydz.$$

Since

$$(L_x - \partial_t)p_t^0(x,y) = \Phi_t(x,y),$$

combining (4.11) and (4.12), we obtain

(4.13) $$V_{t,\epsilon}f(x) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} p_\epsilon^0(x,z)\Psi_t(z,y)f(y)dydz - \int_{\mathbb{R}^n} \Phi_{t+\epsilon}(x,y)f(y)dy$$

$$- \int_0^t \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \Phi_{t-s+\epsilon}(x,z)\Psi_s(z,y)f(y)dydzds.$$

Since the function Ψ satisfies the equation

$$\Phi_t(x,y) = \Psi_t(x,y) - \int_0^t \int_{\mathbb{R}^n} \Phi_{t-s}(x,z)\Psi_s(z,y)\,dzds,$$

we can rewrite $V_{t,\epsilon}f(x)$ as follows:

$$V_{t,\epsilon}f(x) = \int\limits_{\mathbb{R}^n} \Big(\int\limits_{\mathbb{R}^n} p_\epsilon^0(x,z)\Psi_t(z,y)dz - \Psi_{t+\epsilon}(x,y)\Big)f(y)dy$$

$$+ \int\limits_{\mathbb{R}^n} \Big(\int\limits_t^{t+\epsilon} \int\limits_{\mathbb{R}^n} \Phi_{t-s+\epsilon}(x,z)\Psi_s(z,y)dzds\Big)f(y)dy$$

$$=: V_{t,\epsilon}^1 f(x) + V_{t,\epsilon}^2 f(x).$$

By the uniform continuity of Ψ on compact subsets of $(0,\infty) \times \mathbb{R}^n \times \mathbb{R}^n$ and the estimate (3.46), we have for $f \in C_\infty(\mathbb{R}^n)$

$$\sup_{t\in[\tau,T],x\in\mathbb{R}^n} \Big| \int\limits_{\mathbb{R}^n} \Psi_{t+\epsilon}(x,y)f(y)\,dy - \int\limits_{\mathbb{R}^n} \Psi_t(x,y)f(y)\,dy\Big| \to 0, \quad \varepsilon \to 0.$$

Similarly, (3.46), (3.54), and the uniform continuity of Ψ on compact subsets of $(0,\infty) \times \mathbb{R}^n \times \mathbb{R}^n$ give

$$\sup_{t\in[\tau,T],x,y\in\mathbb{R}^n\times\mathbb{R}^n} \Big| \int\limits_{\mathbb{R}^n} p_\varepsilon^0(x,z)\Psi_t(z,y)\,dz - \Psi_t(x,y)\Big| \to 0, \quad \varepsilon \to 0,$$

$$\sup_{t\in[\tau,T],x\in\mathbb{R}^n} \Big| \int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n} p_\varepsilon^0(x,z)\Psi_t(z,y)f(y)\,dzdy - \int\limits_{\mathbb{R}^n} \Psi_t(x,y)f(y)\,dy\Big| \to 0, \quad \varepsilon \to 0.$$

This proves (4.8) with $V_{t,\epsilon}^1 f(x)$ instead of $V_{t,\epsilon}f(x)$. Since by (3.49) we have

$$|V_{t,\epsilon}^1 f(x)| \leqslant Ct^{-1+\delta},$$

the convergence (4.9) for $V_{t,\varepsilon}^1 f(x)$ follows easily from (4.8).

Since $|f|$ is bounded, by (3.11) and (3.46) we obtain

$$(4.14) \quad \int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n}\int\limits_t^{t+\epsilon} |\Phi_{t-s+\epsilon}(x,z)\Psi_s(z,y)f(y)|dzdsdy$$

$$\leqslant c_1 \int\limits_{\mathbb{R}^n}\int\limits_{\mathbb{R}^n}\int\limits_t^{t+\varepsilon} (t-s+\varepsilon)^{-1+\eta}(g_{t-s+\varepsilon} * G_{t-s+\varepsilon})(z-x)$$

$$\times s^{-1+\delta}(g_{s,\zeta} * \Pi_s)(y-z)dsdzdy$$

$$\leqslant c_2 \int\limits_t^{t+\varepsilon}\int\limits_{\mathbb{R}^n} (t-s+\varepsilon)^{-1+\eta}s^{-1+\delta}g_{t-\varepsilon,\chi}(y-x-w)$$

$$\times \Big[\int\limits_{\mathbb{R}^n} G_{t-s+\varepsilon}(dw-u)\Pi_s(du)\Big]ds \leqslant c_3 \int\limits_t^{t+\varepsilon} (t-s+\varepsilon)^{-1+\eta}s^{-1+\delta}ds$$

$$\leqslant c_4 t^{-1+\delta}\varepsilon^\eta, \quad t \in [\tau,T],\ x,y \in \mathbb{R}^n.$$

This immediately gives (4.8) and (4.9) with $V_{t,\epsilon}^2 f(x)$ instead of $V_{t,\epsilon}f(x)$. $\blacksquare$

4.1. Positive maximum principle applied to the approximate fundamental solution. Proof of Theorem 2.2. The proof of Theorem 2.2 follows from Lemmas 4.3–4.5 given below. The arguments used in the proofs of these lemmas are literally the same as those used in [49]. In order to make our paper self-contained, we give the proof of Lemma 4.3, a hint of the proofs of Lemmas 4.4 and 4.5, and refer to [49] for the details.

LEMMA 4.3. *The operator S_t defined in (2.10) is positivity preserving, i.e., $S_t f \geqslant 0$ if $f \geqslant 0$.*

P r o o f. Take $f \in C_\infty(\mathbb{R}^n)$, $f \geqslant 0$, and suppose that

(4.15)
$$\inf_{t,x} S_t f(x) < 0.$$

Then there exists $T > 0$ such that

$$\inf_{t \leqslant T, x \in \mathbb{R}^n} S_t f(x) < 0.$$

Then by (4.5) there exist $\varsigma > 0, \theta > 0, \varepsilon_1 > 0$ such that

$$\inf_{t \leqslant T, x \in \mathbb{R}^n} \left(S_{t,\varepsilon} f(x) + \theta t \right) < -\varsigma, \quad \varepsilon < \varepsilon_1.$$

Let us put

$$u_\epsilon(t,x) = S_{t,\varepsilon} f(x) + \theta t,$$

and note that, by (4.6),

$$u_\epsilon(t,x) \to \theta t > 0, \quad \|x\| \to \infty,$$

uniformly in $t \in [0, T]$. Hence the above infimum is in fact attained at some point; in what follows we fix one such point for each ε, and denote it by $(t_\varepsilon, x_\varepsilon)$.

Since $f(x) \geqslant 0$, by statement (2) of Lemma 4.1 there exist $\varepsilon_0 > 0, \tau > 0$ such that

$$S_{t,\varepsilon} f(x) + \theta t \geqslant -\frac{\varsigma}{2}, \quad t \leqslant \tau, \ \varepsilon < \varepsilon_0, \ x \in \mathbb{R}^n.$$

Because

$$u_\epsilon(t_\varepsilon, x_\varepsilon) = \min_{t \in [0,T], x \in \mathbb{R}^n} u_\varepsilon(t,x) < -\varsigma < -\frac{\varsigma}{2},$$

we have $t_\varepsilon > \tau$ as soon as $\varepsilon < \varepsilon_0$.

The operator L satisfies the positive maximum principle; that is, if $f \in D(L)$, and $f(x_0) \geqslant 0$, where $x_0 = \arg\max f(x)$, then $Lf(x_0) \leqslant 0$, cf. [26], Chapter 4.2. Therefore,

$$L_x u_\varepsilon(t,x)\big|_{(t,x)=(t_\varepsilon,x_\varepsilon)} \geqslant 0.$$

In addition, for $\varepsilon < \varepsilon_0$

$$\partial_t u_\varepsilon(t, x)|_{(t,x)=(t_\varepsilon, x_\varepsilon)} \leqslant 0.$$

Note that the inequality sign may here appear if $t_\varepsilon = T$, and because we have excluded another "boundary case" $t_\varepsilon = \tau$, the inverse inequality is impossible.

Then

(4.16) $$(\partial_t - L_x)u_\varepsilon(t, x)|_{(t,x)=(t_\varepsilon, x_\varepsilon)} \leqslant 0.$$

On the other hand, because $t_\varepsilon \in [\tau, T]$, $\varepsilon < \varepsilon_0$, by the first statement of Lemma 4.2 we have

$$(\partial_t - L_x)u_\varepsilon(t, x)|_{(t,x)=(t_\varepsilon, x_\varepsilon)} = \theta + V_{t_\varepsilon, \varepsilon} f(x_\varepsilon) \to \theta > 0, \quad \varepsilon \to 0.$$

This gives a contradiction and shows that (4.15) fails. ∎

LEMMA 4.4. *The family of operators has the semigroup property*: $S_{t+s} = S_s S_t$.

P r o o f. Take $f \in C_\infty(\mathbb{R}^n)$. Applying the same argument as that used in the proof of Lemma 4.3 to the functions

$$u_\pm(t, x) = \pm S_{t+s}f(x) \mp S_t S_s f(x),$$

we have $u_\pm(t, x) \geqslant 0$, which implies that the identity $S_{t+s}f(x) - S_t S_s f(x) = 0$ holds true. ∎

LEMMA 4.5. *We have*

(4.17a) $$S_t f(x) - f(x) = \int_0^t S_s L f(x)\, ds, \quad f \in C_\infty^2(\mathbb{R}^n);$$

(4.17b) $$S_t 1 = 1.$$

P r o o f. Applying the same argument as that used in the proof of Lemma 4.3 to the functions

$$u_\pm(t, x) = \pm\big(S_t f(x) - f(x)\big) \mp \int_0^t S_s L f(x)\, ds, \quad f \in C_\infty^2(\mathbb{R}^n),$$

and using the statement (2) of Lemma 4.2, we get the identity (4.17a).

The identity (4.17b) follows from (4.17a) by taking $f_n \to 1$, $f_n \in C_\infty^2(\mathbb{R}^n)$, such that $L f_n(x) \to 0$. ∎

P r o o f   o f   T h e o r e m   2.2. By Lemmas 4.3–4.5, the family of operators $(S_t)_{t \geqslant 0}$ forms a strongly continuous contraction semigroup which is positivity preserving. Since the semigroup $(S_t)_{t \geqslant 0}$ has the continuous transition probability density $p_t(x, y)$, the respective Markov process X is the strong Feller process. Finally, the expression (4.17a) and the statement (2) of Lemma 3.6 imply that the restriction of the generator of $(S_t)_{t \geqslant 0}$ coincides with L on functions from $C_\infty^2(\mathbb{R}^n)$. ∎

5. TIME DERIVATIVES. PROOF OF THEOREM 2.6

Proposition 4.1 allows us to transfer the differentiability properties of $p_t^0(x, y)$ to $p_t(x, y)$. For this we need to establish the continuity and upper estimates on $\partial_t \Phi^{\circledast k}$ and $\partial_t \Psi^{\circledast k}$.

LEMMA 5.1. *The function* $\Psi_t(x, y)$ *is differentiable with respect to* t, *the derivative* $\partial_t \Psi_t(x, y)$ *is continuous in* $(t, x, y) \in (0, \infty) \times \mathbb{R}^n \times \mathbb{R}^n$, *and for any* $T > 0$ *there exists a family of subprobability measures* $\{\Theta_t, \ t \geqslant 0\}$ *such that*

$$(5.1) \quad |\partial_t \Psi_t(x, y)| \leqslant t^{-2+\delta}(g_{t,\zeta} * \Theta_t)(y - x), \quad t \in (0, T], \ x, y \in \mathbb{R}^n.$$

P r o o f. The proof follows the same strategy as that of Theorem 2.1. Using Proposition 4.1, we can obtain the estimate for $\partial_t \Phi_t(x, y)$ in the same way as it was done for $\Phi_t(x, y)$ in Lemma 3.1:

$$(5.2) \quad |\partial_t \Phi_t(x, y)| \leqslant C t^{-2+\delta}(\tilde{g}_t * \mathcal{G}_t)(y - x), \quad t \in (0, T], \ x, y \in \mathbb{R}^n,$$

where $C > 0$, $\delta \in (0, 1)$ are the same as in (3.29), $\tilde{g}_t$ is of the form (3.28), and the family of measures $\mathcal{G}_t(du)$ is given by

$$(5.3) \quad \mathcal{G}_t(du) := c\big(\mathcal{P}_{t,\kappa}(du) + (\Lambda_t * \mathcal{P}_{t,\kappa})(du)\big),$$

$$\mathcal{P}_{t,\kappa}(du) := \big(1 + \rho_t^\kappa(\|u\|^\kappa \wedge 1)\big)\mathcal{P}_t(du).$$

Here $c > 0$ is the normalizing constant such that $\mathcal{G}_t(\mathbb{R}^n) \leqslant 1$ for all $t \in (0, T]$.

Note that, by definition,

$$(5.4) \quad \mathcal{G}_t \geqslant G_t.$$

To show that $\partial_t \Phi_t(x, y)$ is continuous in (t, x, y), we follow line by line the proof of continuity of $\Phi_t(x, y)$ (cf. the proof of Theorem 2.1 (b)). Observe that the function $\partial_t p_t^0(x, y)$ is continuous in (t, x, y), and $\partial_t p_t^0(\cdot, y) \in C_\infty^2(\mathbb{R}^n)$. Then for any $t_0 > 0$

$$|\partial_t p_t^0(x + u, y) - \partial_t p_t^0(x, y) - \nabla_x \partial_t p_t^0(x, y) \cdot u \mathbb{1}_{\{\|u\| \leqslant 1\}}| \leqslant C(\|u\|^2 \wedge 1),$$

where $t > t_0$, $x, y \in \mathbb{R}^n$. Therefore, the function $\partial_t L_x p_t^0(x, y) = L_x \partial_t p_t^0(x, y)$ is continuous, which together with continuity of $\partial_t^2 p_t^0(x, y)$ implies the continuity of $\partial_t \Phi_t(x, y)$ in (t, x, y).

To show the continuity of $\partial_t \Phi_t^{\circledast k}(x, y)$ for $k \geqslant 2$ we use induction. Write

$$(5.5) \quad \Phi_t^{\circledast(k+1)}(x, y)$$
$$= \int\limits_0^{t/2} \int\limits_{\mathbb{R}^n} \Phi_{t-s}^{\circledast k}(x, z)\Phi_s(z, y)\, dz ds + \int\limits_0^{t/2} \int\limits_{\mathbb{R}^n} \Phi_s^{\circledast k}(x, z)\Phi_{t-s}(z, y)\, dz ds.$$

Observe that now the functions under the integrals do not have singularities in t. Differentiating the above expression in t, we get

$$(5.6) \quad \partial_t \Phi_t^{\circledast(k+1)}(x,y) = \int\limits_0^{t/2} \int\limits_{\mathbb{R}^n} (\partial_t \Phi^{\circledast k})_{t-s}(x,z) \Phi_s(z,y) \, dz ds$$

$$+ \int\limits_0^{t/2} \int\limits_{\mathbb{R}^n} \Phi_s^{\circledast k}(x,z)(\partial_t \Phi)_{t-s}(z,y) \, dz ds + \int\limits_{\mathbb{R}^n} \Phi_{t/2}^{\circledast k}(x,z) \Phi_{t/2}(z,y) \, dz.$$

Since by the induction assumption all functions under the integrals are continuous in (t,x,y), the above expression implies the continuity of $\partial_t \Phi_t^{(k+1)}(x,y)$ in $(t,x,y) \in (0,\infty) \times \mathbb{R}^n \times \mathbb{R}^n$.

Let us show by induction that

$$(5.7) \qquad |\partial_t \Phi_t^{\circledast k}(x,y)| \leqslant \tilde{C}_k t^{-2+k\delta} (g_t^{(k)} * \mathcal{G}_t^{(k)})(y-x), \quad k \geqslant 2,$$

where the sequence $g_t^{(k)}$ is given by (3.40), and

$$(5.8) \quad \mathcal{G}_t^{(k)}(dw) := \frac{1}{1 + B\big((k-1)\delta, \delta\big)} \Big(\int\limits_0^1 \int\limits_{\mathbb{R}^n} r^{-1+\delta}(1-r)^{-1+\delta(k-1)}$$

$$\times \, \mathcal{G}_{t(1-r)}^{(k-1)}(dw-u) \mathcal{G}_{tr}(du) dr + (\mathcal{G}_{t/2}^{(k-1)} * \mathcal{G}_{t/2})(dw) \Big), \quad k \geqslant 2.$$

Suppose that (5.7) holds true for some $k \geqslant 2$. Using (5.2), (5.6), (3.39) and Lemmas 3.2 and 3.3, we obtain

$$|\partial_t \Phi_t^{\circledast(k+1)}(x,y)| \leqslant \int\limits_0^{t/2} \int\limits_{\mathbb{R}^n} |(\partial_t \Phi^{\circledast k})_{t-s}(x,z) \Phi_s(z,y)| \, dz ds$$

$$+ \int\limits_0^{t/2} \int\limits_{\mathbb{R}^n} |\Phi_s^{\circledast k}(x,z)(\partial_t \Phi)_{t-s}(z,y)| \, dz ds + \int\limits_{\mathbb{R}^n} |\Phi_{t/2}^{\circledast k}(x,z) \Phi_{t/2}(z,y)| \, dz$$

$$\leqslant c_1(k) \int\limits_{\mathbb{R}^n} g_t^{(k+1)}(y-x-w) \Big[\int\limits_0^{t/2} \int\limits_{\mathbb{R}^n} (t-s)^{-2+k\delta} s^{-1+\delta} \mathcal{G}_{t-s}^{(k)}(dw-u) G_s(du) ds \Big]$$

$$+ c_2(k) \int\limits_{\mathbb{R}^n} g_t^{(k+1)}(y-x-w) \Big[\int\limits_0^{t/2} \int\limits_{\mathbb{R}^n} (t-s)^{-1+k\delta} s^{-2+\delta} G_{t-s}^{(k)}(dw-u) \mathcal{G}_s(du) ds \Big]$$

$$+ c_3(k) t^{-2+(k+1)\delta} \int\limits_{\mathbb{R}^n} g_t^{(k+1)}(y-x-w) \Big[\int\limits_{\mathbb{R}^n} G_{t/2}^{(k)}(dw-u) G_{t/2}(du) \Big]$$

$$\leqslant c_4(k) \int\limits_{\mathbb{R}^n} g_t^{(k+1)}(y-x-w) \Big[\int\limits_0^t \int\limits_{\mathbb{R}^n} (t-s)^{-1+k\delta} s^{-1+\delta} \mathcal{G}_{t-s}^{(k)}(dw-u) \mathcal{G}_s(du) ds$$

$$+ t^{-2+(k+1)\delta} (\mathcal{G}_{t/2}^{(k)} * \mathcal{G}_{t/2})(dw) \Big]$$

$$\leqslant \tilde{C}_k t^{-2+(k+1)\delta} (g_t^{(k+1)} * \mathcal{G}_t^{(k+1)})(y-x),$$

where we used (5.4). This proves (5.7).

Take as before $k_0 := [n/(\alpha\delta)] + 1$. Then, applying induction and (3.43) (cf. Lemma 3.5), we get

(5.9) $|\partial_t \Phi_t^{\circledast(k_0+\ell)}(x,y)| \leqslant \tilde{D}_\ell t^{-2+\delta(k_0+\ell)}(g_{t,\zeta} * \mathcal{G}_t^{(k_0+\ell)})(y-x), \quad \ell \geqslant 1,$

where

$$\tilde{D}_\ell := \frac{C(k_0)K^\ell \Gamma^{k_0+\ell}(\delta)}{\Gamma\big((k_0+\ell)\delta\big)}, \quad \ell \geqslant 1,$$

and $C(k_0), K > 0$ are some constants.

Finally, define

(5.10)

$$\Theta_t(du) := \frac{\sum\limits_{k=1}^{k_0} \tilde{C}_k T^{\delta(k-1)} \mathcal{G}_t^{(k)}(du) + \sum\limits_{\ell=1}^{\infty} \tilde{D}_\ell T^{\ell(k_0+\ell-1)} \mathcal{G}_t^{(k_0+\ell)}(du)}{\sum\limits_{k=1}^{k_0} \tilde{C}_k T^{\delta(k-1)} + \sum\limits_{\ell=1}^{\infty} \tilde{D}_\ell T^{\ell(k_0+\ell-1)}}, \quad t \in (0, T].$$

Then $\Theta_t(\mathbb{R}^n) \leqslant 1$, $t \in (0, T]$, and thus (5.1) follows from (5.7) and (5.9). ∎

P r o o f   o f   T h e o r e m   2.6. The proof of differentiability of $p_t(x, y)$ essentially follows from Proposition 4.1 and Lemma 5.1. Indeed, writing $p_t(x, y)$ in the form

(5.11) $p_t(x, y)$

$$= p_t^0(x, y) + \int\limits_0^{t/2} \int\limits_{\mathbb{R}^n} p_{t-s}^0(x, z)\Psi_s(z, y) \, dz ds + \int\limits_0^{t/2} \int\limits_{\mathbb{R}^n} p_s^0(x, z)\Psi_{t-s}(z, y) \, dz \, ds,$$

and applying the above lemmas, we get

$$|\partial_t p_t(x, y)| \leqslant Ct^{-1}(g_{t,\chi} * \tilde{Q}_t)(y - x), \quad t \in (0, T], \ x, y \in \mathbb{R}^n,$$

where $\chi \in (0, \zeta)$, ζ is coming from (5.9),

$$\tilde{Q}_t(du) := c\big(\mathcal{P}_t(du) + t^\delta \tilde{\mathcal{P}}_t(du)\big)$$

is a subprobability measure (here $c = c(T) > 0$ is the normalizing constant),

$$\tilde{\mathcal{P}}_t(du) = (\mathcal{P}_{t/2} * \Pi_{t/2})(du) + \int\limits_0^{1/2} \int\limits_{\mathbb{R}^n} r^{-1+\delta} \mathcal{P}_{t(1-r)}(dw - u)\Pi_{tr}(du)dr$$

$$+ \int\limits_0^{1/2} \int\limits_{\mathbb{R}^n} \Theta_{t(1-r)}(dw - u)P_{tr}(du)dr,$$

and the measures $\Pi_t(du)$ and $\Theta_t(du)$ are given in (3.47) and (5.10), respectively. ∎

We finish this section with a lemma, which plays an important role in the proof of Theorem 2.3.

LEMMA 5.2. (1) *For any $f \in C_\infty(\mathbb{R}^n)$,*

$$\|\partial_t S_{t,\varepsilon} f - \partial_t S_t f\|_\infty \to 0 \quad \text{as } \epsilon \to 0,$$

uniformly on compact subsets of $(0, \infty)$*. Also,* $\partial_t S_t f(x) = \int_{\mathbb{R}^n} \partial_t p_t(x, y) f(y) dy$*.*
  (2) *We have*

$$\partial_t p_{t,\epsilon}(x, y) \to \partial_t p_t(x, y) \quad \text{as } \epsilon \to 0,$$

uniformly on compact subsets of $(0, \infty) \times \mathbb{R}^n \times \mathbb{R}^n$*.*

The proof relies on the decomposition (5.11), and the estimates on $p^0, \Psi, \partial_t p^0$ and $\partial_t \Psi$ obtained above; see the proof of Lemma 6.4 in [49] for details.

6. PROOFS OF THEOREMS 2.3 AND 2.4

The proofs repeat literally the proofs of the respective statements in [49]. In order to make this paper self-contained, we sketch these proofs below.

P r o o f o f T h e o r e m 2.3. By Theorem 2.2 we know that $(L, C^2_\infty(\mathbb{R}^n))$ is the restriction of $(A, D(A))$. Since A is closed, this implies that $(L, C^2_\infty(\mathbb{R}^n))$ is closable. Let us show that its closure coincides with $(A, D(A))$.

Take $f \in C_\infty(\mathbb{R}^n) \cap D(A)$. Fix $t > 0$, and consider the functions $S_t f$ and $S_{t,\varepsilon} f$. Since $f \in D(A)$, we have $S_t f \in D(A)$, and

(6.1) $$AS_t f = \partial_t S_t f.$$

Recall that $S_{t,\varepsilon} f(\cdot) \in C^2_\infty(\mathbb{R}^n) \subset D(A)$, which implies

$$AS_{t,\varepsilon} f = LS_{t,\varepsilon} f = \partial_t S_{t,\varepsilon} f.$$

Further, statement (2) in Lemma 4.1 together with statements (1) of Lemmas 4.2 and 5.2 implies

$$LS_{t,\varepsilon} f \to AS_t f \quad \text{in } C_\infty(\mathbb{R}^n) \text{ as } \varepsilon \to 0,$$

and thus $S_t f$ belongs to the domain of the $C_\infty(\mathbb{R}^n)$-closure of $(L, C^2_\infty(\mathbb{R}^n))$. Consequently, this closure coincides with $(A, D(A))$.

In addition, applying the same argument to the function $p_{t,\varepsilon}(x, y)$ instead of $S_{t,\varepsilon} f(x)$ and using the second statement of Lemma 5.2, we infer that $p_t(x, y)$ belongs to $D(A)$, and is the fundamental solution to the Cauchy problem for $\partial_t - A$. ∎

P r o o f o f T h e o r e m 2.4. Using the Markov property of X, we deduce from (4.17a) and the semigroup property for $p_t(x, y)$ the following: For given $f \in C_\infty^2(\mathbb{R}^n)$, $t_2 > t_1$, and $x \in \mathbb{R}^n$, for any $m \geqslant 1$, $r_1, \ldots, r_m \in [0, t_1]$, and bounded measurable $G : (\mathbb{R}^n)^m \to \mathbb{R}^n$ the identity

$$\mathbb{E}_x \big[f(X_{t_2}) - f(X_{t_2}) - \int_{t_1}^{t_2} h_f(X_s)\, ds \big] G(X_{r_1}, \ldots, X_{r_m}) = 0$$

holds true. Thus, for every $f \in C_\infty^2(\mathbb{R}^n)$ the process

$$M_t^f = f(X_t) - \int_0^t h_f(X_s)\, ds, \quad t \geqslant 0,$$

is a $\mathbb{P}_x$-martingale for every $x \in \mathbb{R}^n$; that is, X is a solution to the martingale problem for $\big(L, C_\infty^2(\mathbb{R}^n)\big)$.

Note that the operator $\big(L, C_\infty^2(\mathbb{R}^n)\big)$ is dissipative, which follows from the positive maximum principle, see [26], Lemma 4.2.1, or [39], Lemma 4.5.2. Since its closure equals the generator A of the $C_\infty(\mathbb{R}^n)$-semigroup $\{S_t, t \geqslant 0\}$, for every $\lambda > 0$ the range of the resolvent $(\lambda - L)^{-1}$ in $C_\infty(\mathbb{R}^n)$ is dense. Hence the required uniqueness of the solution to the martingale problem $\big(L, C_\infty^2(\mathbb{R}^n)\big)$ follows by Theorem 4.4.1 in [26]. ∎

7. UPPER AND LOWER BOUNDS: PROOF OF THEOREM 2.5

P r o o f. U p p e r b o u n d. The upper bound is essentially contained in the proof of Theorem 2.1. Namely, we already obtained the upper estimate on $p^0 \circledast \Psi$, see (3.49). Combining this estimate with the estimate (3.15) (for $k = 0$) for p^0, we derive the upper bound in (2.23) with

$$(7.1) \qquad\qquad Q_t(du) := (1 + \delta)^{-1} \big(P_t(du) + t^\delta \tilde{\Pi}_t(du) \big).$$

Since P_t is the probability measure and $\tilde{\Pi}_t$ is the subprobability measure for $t \in [0, T]$, $Q_t(\mathbb{R}^n) \leqslant 1$ for all $t \in [0, T]$.

L o w e r b o u n d. By (3.49) and the fact that $\tilde{\Pi}$ is the subprobability measure we get

$$(7.2) \qquad\qquad |(p^0 \circledast \Psi)_t(x, y)| \leqslant c_1 \rho_t^n t^\delta, \quad x, y \in \mathbb{R}^n,\, t \in (0, T],$$

which implies the upper bound $p_t(x, x) \leqslant c_2 \rho_t^n$ for all $x \in \mathbb{R}^n$ and $t \in (0, T]$. Finally, using Proposition 3.5 and (7.2), we obtain for t small enough

$$p_t(x, y) \geqslant p_t^0(x, y) - |(p^0 \circledast \Psi)_t(x, y)| \geqslant \rho_t^n f_{low}(\|y - x\|\rho_t) - c_1 \rho_t^n t^\delta$$
$$\geqslant c_2 \rho_t^n f_{low}(\|y - x\|\rho_t). \quad \blacksquare$$

8. PROOF OF THEOREM 2.7

P r o o f   o f   T h e o r e m   2.7.   S u f f i c i e n c y. We use the upper bound constructed in Theorem 2.5.

Fix $\ell \in \mathbb{S}^n$, and define $\theta_t := \inf\{r : q^U(r\ell) \geqslant 1/t\}$. Note that by **A1** we have $\theta_t \asymp \rho_t$ for all $t \in (0, 1]$. For any $T \in (0, 1]$, making the change of variables $s = \theta_t$ and using (9.5) (see Appendix A) in the integration by parts, we get

$$(8.1) \quad \int\limits_0^T \int\limits_{\mathbb{R}^n} p_t(x, y)\varpi(dy)dt \leqslant c_0 \int\limits_0^T \int\limits_{\mathbb{R}^n} \int\limits_{\mathbb{R}^n} \theta_t^n e^{-c\|x-y-w\|\theta_t} \varpi(dy)Q_t(dw)dt$$

$$\leqslant c_1 \int\limits_{\theta_T}^\infty \int\limits_{\mathbb{R}^n} \int\limits_{\mathbb{R}^n} \frac{s^{n-1}q^L(\ell s)}{\left(q^U(s)\right)^2} e^{-c\|x-y-w\|s} \varpi(dy)\mathcal{Q}_s(dw)ds$$

$$\leqslant c_2 \int\limits_{\theta_T}^\infty \int\limits_{\mathbb{R}^n} \int\limits_{\mathbb{R}^n} \frac{s^{n-1}}{q^*(s)} e^{-c\|x-y-w\|s} \varpi(dy)\mathcal{Q}_s(dw)ds$$

$$= c_2 \int\limits_{\theta_T}^\infty \frac{s^{n-1}}{q^*(s)} \int\limits_{\mathbb{R}^n} \int\limits_0^\infty \varpi\{y : e^{-c\|x-y-w\|s} > r\}dr\mathcal{Q}_s(dw)ds$$

$$= c_3 \int\limits_{\theta_T}^\infty \frac{s^{n-1}}{q^*(s)} \int\limits_{\mathbb{R}^n} \int\limits_0^\infty \varpi\{y : \|x-y-w\| \leqslant v/s\}e^{-cv}dv\mathcal{Q}_s(dw)ds$$

$$\leqslant c_4 \int\limits_{\theta_T}^\infty \frac{s^{n-1}}{q^*(s)} \int\limits_0^\infty h(v/s)e^{-cv}dvds$$

$$= c_4 \int\limits_0^\infty \left[\int\limits_0^{1/\theta_T} \frac{h(sv)}{s^{n+1}q^*(1/s)}ds \right] e^{-cv}dv,$$

where $\mathcal{Q}_s(dw)$ is the image measure of $Q_t(dw)$ under the transformation $s = \theta_t$,

$$h(r) := \sup_{x \in \mathbb{R}^n} \varpi\{B(x, r)\},$$

and in the second line from below we made use of the relation $\mathcal{Q}_s(\mathbb{R}^n) \leqslant 1$ for all $s \in (\theta_T, \infty]$. Without loss of generality assume that $c = 1$. Split

$$I(T) := \int\limits_0^\infty \left[\int\limits_0^{1/\theta_T} \frac{h(sv)}{s^{n+1}q^*(1/s)}ds \right] e^{-v}dv$$

$$= \int\limits_0^1 \left[\int\limits_0^{1/\theta_T} \frac{h(sv)}{s^{n+1}q^*(1/s)}ds \right] e^{-v}dv + \int\limits_1^\infty \left[\int\limits_0^{1/\theta_T} \frac{h(sv)}{s^{n+1}q^*(1/s)}ds \right] e^{-v}dv$$

$$=: I_1(T) + I_2(T).$$

We show that under (2.27) (respectively, (2.28)) one has $I(T) \to 0$ as $T \to 0$ (respectively, $I(T) < \infty$).

By the monotonicity of $h(r)$ and (2.27) we have

$$I_1(T) \leqslant \int\limits_0^{1/\theta_T} \frac{h(s)}{s^{n+1}q^*(1/s)}ds \cdot \int\limits_0^1 e^{-v}dv \to 0 \quad \text{as } T \to 0.$$

Further, using the monotonicity of q^*, we get

$$I_2(T) \leqslant \int\limits_1^\infty \left[\int\limits_0^{v/\theta_T} \frac{h(u)}{u^{n+1}q^*(1/u)}du \right] v^n e^{-v}dv$$

$$= \left[\int\limits_1^\infty \int\limits_0^{1/\theta_T} + \int\limits_1^\infty \int\limits_{1/\theta_T}^{v/\theta_T} \right] \left[\frac{h(u)}{u^{n+1}q^*(1/u)}du \right] v^n e^{-v}dv =: I_{21}(T) + I_{22}(T).$$

For $I_{21}(T)$ we have

$$I_{21}(T) = \int\limits_1^\infty v^n e^{-v}ds \cdot \int\limits_0^{1/\theta_T} \frac{h(u)}{u^{n+1}q^*(1/u)}du \to 0, \quad T \to 0.$$

Further,

$$I_{22}(T) = \int\limits_{1/\theta_T}^\infty \left[\int\limits_{u\theta_T}^\infty v^n e^{-v}dv \right] \frac{h(u)}{u^{n+1}q^*(1/u)}du$$

$$\leqslant \int\limits_0^\infty e^{-\epsilon u\theta_T} \left[\int\limits_{u\theta_T}^\infty v^n e^{-(1-\epsilon)v}dv \right] \frac{h(u)}{u^{n+1}q^*(1/u)}du$$

$$\leqslant \int\limits_0^\infty e^{-\epsilon u} \left[\int\limits_{u\theta_T}^\infty v^n e^{-(1-\epsilon)v}dv \right] \frac{h(u)}{u^{n+1}q^*(1/u)}du.$$

Since by (2.27) the function

$$\phi(u) := \frac{e^{-\epsilon u}h(u)}{u^{n+1}q^*(1/u)}$$

is integrable on $(0, \infty)$, we infer, by the theorem on continuity with respect to a parameter, that $I_{22}(T) \to 0$ as $T \to 0$. Therefore, under (2.27) (respectively, (2.28)) we have $I(T) \to 0$ as $T \to 0$ (respectively, $I(T) < \infty$), and thus $\varpi \in S_K$ (respectively, $\varpi \in S_D$).

N e c e s s i t y. Using the lower bound for $p_t(x, y)$ and the inequality

$$(1 - \|x\|s)_+ \geqslant 2^{-1}\mathbb{1}_{\{2\|x\|s\leqslant 1\}},$$

we obtain

$$(8.2) \quad \int\limits_0^T \int\limits_{\mathbb{R}^n} p_t(x, y)\,\varpi(dy)dt \geqslant d_1 \int\limits_0^T \int\limits_{\mathbb{R}^n} \rho_t^n (1 - d_2\|x - y\|\rho_t)_+\,\varpi(dy)dt$$

$$\geqslant 2^{-1}d_1 \int\limits_0^T \int\limits_{\mathbb{R}^n} \rho_t^n \mathbb{1}_{\{2d_2\|x-y\|\rho_t\leqslant 1\}}\varpi(dy)dt.$$

Without loss of generality we assume that $\delta = \delta(T) := 1/\theta_T \in (0,1)$, and that $2d_2 = 1$. Therefore, using (9.5) and **A1**, we get

$$
(8.3) \qquad \int_0^T \rho_t^n \mathbb{1}_{\|x\|\rho_t \leqslant 1} ds \geqslant c_1 \int_{1/\delta}^{1/\|x\|} s^{n-1} \frac{q^L(\ell s)}{\left(q^U(\ell s)\right)^2} ds
$$

$$
\geqslant \beta^{-1} c_1 \int_{1/\delta}^{1/\|x\|} \frac{s^{n-1}}{q^*(s)} ds
$$

$$
= \beta^{-1} c_1 \big(U(\|x\|) - U(\delta)\big) \mathbb{1}_{\{\|x\| \leqslant \delta\}},
$$

where

$$
U(r) := \int_1^{1/r} \frac{s^{n-1}}{q^*(s)} ds, \quad r \in (0,1).
$$

Integrating by parts, we have

$$
\int_{\|x-y\|\leqslant \delta} \big(U(\|x-y\|) - U(\delta)\big) \varpi(dy)
$$

$$
= \int_0^{U(0)-U(\delta)} \varpi\{y : U(\|x-y\|) \geqslant r + U(\delta)\} dr
$$

$$
= \int_{U(\delta)}^{U(0)} \varpi\{y : U(\|x-y\|) \geqslant r\} dr = \int_0^\delta \frac{\varpi\{y : \|x-y\| \leqslant s\}}{s^{n+1} q^*(1/s)} ds.
$$

Note that $\delta_T \to 0$ if and only if $T \to 0$. Thus, if $\varpi \in S_K$ (respectively, $\varpi \in S_D$), then (2.29) (respectively, (2.28)) holds true. ∎

9. APPENDIX A

P r o o f   o f   P r o p o s i t i o n   3.1. Clearly, for $n=1$ the statement holds true. For $n \geqslant 2$ we have

$$
(9.1) \qquad \mu\{u : \|u\| \geqslant r\} \leqslant \sum_{i=1}^n \mu\{u : |u_i| \geqslant rn^{-1/2}\}
$$

$$
\leqslant n \max_{1\leqslant i\leqslant n} \mu\{u : |u_i| \geqslant rn^{-1/2}\}
$$

$$
\leqslant n \max_{1\leqslant i\leqslant n} q^U(\sqrt{n}r^{-1}\ell_i)
$$

$$
\leqslant n^2 \max_{1\leqslant i\leqslant n} q^U(r^{-1}\ell_i)
$$

$$
\leqslant n^2 q^*(1/r),
$$

where $\ell_i := (0, \ldots, \underset{i}{1}, \ldots, 0) \in \mathbb{S}^n$, and in the third line we used the fact that the inequality $q^U(\xi c) \leqslant (c^2 \wedge 1) q^U(\xi)$ holds true for any $c > 0$. Thus, for $n \geqslant 2$ we

have

$$(9.2) \qquad \Lambda_t(\mathbb{R}^n) = t\mu\{u : \|u\| \geqslant 1/\rho_t\} \leqslant n^2 tq^*(\rho_t) = n^2,$$

which completes the proof. ∎

Proof of Proposition 3.2. Using (2.20) and (9.1), we obtain

$$(9.3) \quad \int\limits_{\rho_t\|u\|\geqslant 1} (\|u\|^\lambda \wedge 1)\mu(du) = \int\limits_{1/\rho_t\leqslant\|u\|\leqslant 1} \|u\|^\lambda\mu(du) + \int\limits_{\|u\|\geqslant 1} \mu(du)$$

$$\leqslant \int\limits_{1/\rho_t\leqslant\|u\|\leqslant 1} \int\limits_0^{\|u\|^\lambda} dr\mu(du) + c_1$$

$$\leqslant \int\!\!\int \mathbb{1}_{1/\rho_t\leqslant\|u\|\leqslant 1} \mathbb{1}_{0<r<\|u\|^\lambda} dr\mu(du) + c_1$$

$$\leqslant \int\limits_0^{1/\rho_t^\lambda} \mu\{u : \|u\| \geqslant 1/\rho_t\}dr + \int\limits_0^1 \mu\{u : \|u\| \geqslant r^{1/\lambda}\}dr + c_1$$

$$\leqslant n^2\rho_t^{-\lambda}q^*(\rho_t) + \lambda\int\limits_1^{\rho_t} \frac{\mu\{u : \|u\| \geqslant 1/r\}}{r^{1+\lambda}}dr + c_1$$

$$\leqslant n^2 t^{-1}\rho_t^{-\lambda} + \lambda n^2\int\limits_1^{\rho_t} \frac{q^*(r)}{r^{1+\lambda}}dr + c_1, \quad t \in (0, T].$$

Note that condition **A1** implies for any $\ell \in \mathbb{S}^n$ the inequalities

$$(9.4) \qquad q^U(r\ell) \leqslant q^*(r) \leqslant \beta q^U(r\ell).$$

Let us estimate

$$I_\ell(r) := \int\limits_1^r \frac{q^U(v\ell)}{v^{1+\lambda}}dv,$$

where the vector $\ell \in \mathbb{S}^n$ is fixed, and $r > 1$. Note that for any $\ell \in \mathbb{S}^n$ the mapping $r \mapsto q^U(r\ell)$ is absolutely continuous, and for any $0 < r_1 < r_2$, $\ell \in \mathbb{S}^n$, we have

$$(9.5) \qquad q^U(r_2\ell) - q^U(r_1\ell) = \int\limits_{r_1}^{r_2} \frac{2q^L(v\ell)}{v}dv.$$

Recall that $\lambda \in [0, \alpha)$. Therefore, applying (9.5), we obtain

$$I_\ell(r) \leqslant \beta\int\limits_1^r \frac{q^L(v\ell)}{v^{1+\lambda}}dv = \frac{\beta}{2}\int\limits_1^r \frac{1}{v^\lambda}dq^U(v\ell) \leqslant \frac{\beta}{2}\left(\frac{q^U(r\ell)}{r^\lambda} + \lambda I_\ell(r)\right),$$

which gives $I_\ell(r) \leqslant (\alpha - \lambda)^{-1}q^U(r\ell)/r^\lambda$. Applying again (9.4), we get for any $\ell \in \mathbb{S}^n$

$$\int\limits_1^r \frac{q^*(v)}{v^{1+\lambda}}dv \leqslant \beta I_\ell(r) \leqslant \frac{2}{\alpha(\alpha - \lambda)} \frac{q^*(r)}{r^\lambda},$$

which together with the last line in (9.3) finally gives

$$(9.6) \qquad \int_{\mathbb{R}^n} (\|u\|^\lambda \wedge 1)\Lambda_t(du) \leqslant n^2 \rho_t^{-\lambda} + c_2 t \rho_t^{-\lambda} q^*(\rho_t) + c_1 t \leqslant c_3 \rho_t^{-\lambda} + c_1 t$$

$$\leqslant c_4 \rho_t^{-\lambda}, \quad t \in (0, T],$$

where for the last inequality we again used the fact that $\lambda < \alpha$, and hence $t\rho_t^\lambda \leqslant c$, $t \in [0, T]$. This proves the statement of Proposition 3.2. ∎

Proof of Proposition 3.3. By Proposition 3.2 we have for any $T > 0$

$$(9.7) \qquad \rho_t^\kappa \int_{\mathbb{R}^n} (\|u\|^\kappa \wedge 1)\Lambda_t(du) \leqslant C, \quad t \in [0, T].$$

Then by (9.7) and Proposition 3.1 we have

$$\rho_t^\kappa \int_{\mathbb{R}^n} (\|u\|^\kappa \wedge 1)\Lambda_t^{*2}(du) \leqslant \rho_t^\kappa \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} (\|u\|^\kappa \wedge 1)\Lambda_t(du - w)\Lambda_t(dw)$$

$$\leqslant 2^\kappa \Big[\rho_t^\kappa \int_{\mathbb{R}^n} (\|u - w\|^\kappa \wedge 1)\Lambda_t(du - w)\Lambda_t(dw)$$

$$+ \rho_t^\kappa \int_{\mathbb{R}^n} (\|w\|^\kappa \wedge 1)\Lambda_t(du - w)\Lambda_t(dw) \Big]$$

$$\leqslant 2^\kappa \Big[\rho_t^\kappa \int_{\mathbb{R}^n} (\|v\|^\kappa \wedge 1)\Lambda_t(dv)\Lambda_t(\mathbb{R}^n) + \rho_t^\kappa \int_{\mathbb{R}^n} (\|w\|^\kappa \wedge 1)\Lambda_t(dw)\Lambda_t(\mathbb{R}^n) \Big]$$

$$\leqslant 2^{\kappa+1} n^2 C, \quad t \in [0, T],$$

where in the second line we applied the inequality

$$(9.8) \qquad (a + b)^\kappa \leqslant 2^\kappa (a^\kappa + b^\kappa), \quad a, b \geqslant 0.$$

Let us check that

$$(9.9) \quad \rho_t^\kappa \int_{\mathbb{R}^n} (\|u\|^\kappa \wedge 1)\Lambda_t^{*m}(du) \leqslant C(2^{\kappa+1} n^2)^{m-1}, \quad m \geqslant 2, \, t \in [0, T].$$

Indeed, by induction we have

$$\rho_t^\kappa \int_{\mathbb{R}^n} (\|u\|^\kappa \wedge 1)\Lambda_t^{*m}(du) = \rho_t^\kappa \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} (\|u\|^\kappa \wedge 1)\Lambda_t^{*(m-1)}(du - w)\Lambda_t(dw)$$

$$\leqslant 2^\kappa \Big[\rho_t^\kappa \int_{\mathbb{R}^n} (\|u - w\|^\kappa \wedge 1)\Lambda_t^{*(m-1)}(du - w)\Lambda_t(dw)$$

$$+ \rho_t^\kappa \int_{\mathbb{R}^n} (\|w\|^\kappa \wedge 1)\Lambda_t^{*(m-1)}(du - w)\Lambda_t(dw) \Big]$$

$$\leqslant 2^\kappa [C \cdot 2^{(m-2)(\kappa+1)} n^{2(m-1)} + Cn^{2(m-1)}]$$

$$\leqslant C(2^{\kappa+1} n^2)^{m-1}.$$

Finally, by (9.9) we have for all $t \in [0, T]$

(9.10)
$$\rho_t^{\kappa} \int\limits_{\mathbb{R}^n} (\|u\|^{\kappa} \wedge 1) P_t(du) = e^{-\Lambda_t(\mathbb{R}^n)} \sum_{m=1}^{\infty} \frac{\rho_t^{\kappa}}{m!} \int\limits_{\mathbb{R}^n} (\|u\|^{\kappa} \wedge 1) \Lambda_t^{*m}(du) \leqslant \frac{Ce^{2^{\kappa+1}n^2}}{2^{\kappa+1}n^2},$$

which completes the proof. ∎

Proof of Proposition 3.6. Take an arbitrary $\theta \in (0, 1)$. Using (9.8) and the inequality $z^{\kappa}e^{-z} \leqslant c_1 e^{-\theta z}$, $z \geqslant 0$, where $c_1 > 0$ is some constant, we obtain, for $t \in (0, T]$,

$$(\|x\|^{\kappa} \wedge 1) f_t(x) \leqslant c_2 2^{\kappa} \rho_t^{-\kappa} \Big[\int\limits_{\mathbb{R}^n} \big(\|\rho_t(x-w)\| \wedge \rho_t\big)^{\kappa} g_t(x-w) P_t(dw)$$
$$+ \int\limits_{\mathbb{R}^n} g_t(x-w)(\|\rho_t w\| \wedge \rho_t)^{\kappa} P_t(dw) \Big]$$
$$\leqslant c_2 \cdot 2^{\kappa} \rho_t^{-\kappa} \Big[c_3 \int\limits_{\mathbb{R}^n} g_{t,\theta}(x-w) P_t(dw) + \int\limits_{\mathbb{R}^n} g_t(x-w)(\|\rho_t w\| \wedge \rho_t)^{\kappa} P_t(dw) \Big]$$
$$\leqslant c_4 \rho_t^{-\kappa} \int\limits_{\mathbb{R}^n} g_{t,\theta}(x-w)\big(1+(\|\rho_t w\| \wedge \rho_t)^{\kappa}\big) P_t(dw) = c_4 \rho_t^{-\kappa}(g_{t,\theta} * P_{t,\kappa})(x). \quad \blacksquare$$

10. APPENDIX B

Proof of Proposition 3.7. 1. Using (3.15), we get

$$\Big| \int\limits_{\mathbb{R}^n} p_t^0(x,y) f(y) dy \Big| \leqslant C \int\limits_{\mathbb{R}^n} \int\limits_{\mathbb{R}^n} g_t(y-x-w)|f(y)| P_t(dw) dz$$
$$= C \int\limits_{\mathbb{R}^n} \int\limits_{\mathbb{R}^n} g_t(z-w)|f(x+z)| P_t(dw) dz.$$

Then the right-hand side follows by the inequality (3.52) and the dominated convergence theorem.

2. By the definition of $p_t^0(x,y)$ (cf. (2.9)) we have

$$\Big| \int\limits_{\mathbb{R}^n} p_t^0(x,y) f(y) dy - f(x) \Big| \leqslant \Big| \int\limits_{\mathbb{R}^n} \mathfrak{p}_t^x(y-x)\big(f(y) - f(x)\big) dy \Big|$$
$$+ \Big| \int\limits_{\mathbb{R}^n} \big(\mathfrak{p}_t^y(y-x) - \mathfrak{p}_t^x(y-x)\big) f(y) dy \Big|$$
$$=: I_1(t,x) + I_2(t,x).$$

Fix $\varepsilon > 0$. Then, since f is continuous, we have $|f(x) - f(y)| < \varepsilon$ as soon as $\|x - y\| \leqslant \delta$ for some $\delta = \delta(\varepsilon, x)$. Then

$$I_1(t,x) \leqslant \Big(\int\limits_{\|x-y\| \leqslant \delta} + \int\limits_{\|x-y\| > \delta} \Big) \mathfrak{p}_t^x(y-x)|f(y) - f(x)| dy$$
$$\leqslant C_1 \Big(\varepsilon + \int\limits_{\|z\| > \delta} (g_t * P_t)(z) dz \Big),$$

where we used Proposition 3.4 (see also (3.12) and (3.13)) and the fact that $f(x)$ is bounded. Note that

$$
\begin{aligned}
I_{11}(t,x) := \int\limits_{\|z\|>\delta} (g_t * P_t)(z)dz &\leqslant C_2 \int\limits_{\|u\|\geqslant\delta\rho_t} \int\limits_{\mathbb{R}^n} e^{-c\|u-\rho_t w\|} P_t(dw)du \\
&= C_2 \int\limits_{\|u\|\geqslant\delta\rho_t} \int\limits_{\mathbb{R}^n} e^{-c\|u-v\|} P_t^\sharp(dv)du,
\end{aligned}
$$

where $P_t^\sharp(dv)$ is the measure obtained from $P_t(dw)$ by the change of variables $\rho_t w = v$. Observe that $P_t^\sharp(dv)$ is a subprobability measure. Therefore, we have $\sup_x I_{11}(t,x) \to 0$ as $t \to 0$, which in turn implies that $\lim_{t\to 0} \sup_x I_1(t,x) \leqslant \varepsilon$.

Let us estimate $I_2(t,x)$. Since, for $a,b > 0$, $|e^{-a} - e^{-b}| \leqslant |a-b|e^{-(a\wedge b)}$, we get, by the Hölder continuity of $m(x,u)$ (cf. the representation of $q(x,\xi)$),

$$
\begin{aligned}
\left|\mathfrak{p}_t^x(y-x) - \mathfrak{p}_t^y(y-x)\right| &= (2\pi)^{-n}\Big| \int\limits_{\mathbb{R}^n} e^{-i\xi(y-x)}(e^{-tq(x,\xi)} - e^{-tq(y,\xi)|})d\xi\Big| \\
&\leqslant c_1(|y-x|^\gamma \wedge 1)\Big| \int\limits_{\mathbb{R}^n} tq^U(\xi)e^{-ctq^U(\xi)}d\xi\Big| \\
&\leqslant c_2|y-x|^\gamma \rho_t^n, \quad t \in (0,1], \; x,y \in \mathbb{R}^n.
\end{aligned}
$$

Take now $\varsigma > n/(n+\gamma)$. Then

$$
\begin{aligned}
I_2(t,x) &\leqslant \Big(\int\limits_{\|y-x\|\leqslant\rho_t^{-\varsigma}} + \int\limits_{\|y-x\|>\rho_t^{-\varsigma}} \Big)\left|\mathfrak{p}_t^x(y-x) - \mathfrak{p}_t^y(y-x)\right||f(y)|dy \\
&\leqslant C\Big(\rho_t^{n-(n+\gamma)\varsigma} + \int\limits_{\|u\|\geqslant\delta\rho_t^{1-\varsigma}} \int\limits_{\mathbb{R}^n} e^{-c\|u-w\|} P_t^\sharp(dw)du\Big).
\end{aligned}
$$

By our choice of ς, both terms tend to zero as $t \to 0$, uniformly in x. Thus,

$$
\lim_{t\to 0} \sup_x \Big| \int\limits_{\mathbb{R}^n} p_t^0(x,y)f(y)dy - f(x)\Big| < \varepsilon.
$$

Since $\varepsilon > 0$ is arbitrary, this implies the convergence (3.54). ∎

REFERENCES

[1] S. Albeverio and Z. Ma, *Additive functionals, nowhere Radon and Kato class smooth measures associated with Dirichlet forms*, Osaka J. Math. 29 (1992), pp. 247–265.

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

[3] M. T. Barlow, A. Grigor'yan, and T. Kumagai, *Heat kernel upper bounds for jump processes and the first exit time*, J. Reine Angew. Math. 626 (2009), pp. 135–157.

[4] R. F. Bass, *Uniqueness in law for pure jump Markov processes*, Probab. Theory Related Fields 79 (2) (1988), pp. 271–287.

[5] R. F. Bass, *Stochastic differential equations with jumps*, Probab. Surv. 1 (2004), pp. 1–19.

[6] K. Bogdan, T. Grzywny, and M. Ryznar, *Density and tails of unimodal convolution semigroups*, J. Funct. Anal. 266 (6) (2014), pp. 3543–3571.

[7] K. Bogdan and T. Jakubowski, *Estimates of heat kernel of fractional Laplacian perturbed by gradient operators*, Comm. Math. Phys. 271 (1) (2007), pp. 179–198.

[8] K. Bogdan, T. Jakubowski, and S. Sydor, *Estimates of perturbation series for kernels*, J. Evol. Equ. 12 (2012), pp. 973–984.

[9] K. Bogdan, V. Knopova, and P. Sztonyk, *Construction of transition densities for Markov processes with local scaling*, in preparation.

[10] K. Bogdan and S. Sydor, *On nonlocal perturbations of integral kernels*, in: *Semigroups of Operators – Theory and Applications*, Springer Proc. Math. Stat. 113 (2015), pp. 7–42.

[11] B. Böttcher, *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), pp. 1235–1241.

[12] B. Böttcher, *Construction of time-inhomogeneous Markov processes via evolution equations using pseudo-differential operators*, J. London Math. Soc. 78 (2008), pp. 605–621.

[13] B. Böttcher, R. L. Schilling, and J. Wang, *Lévy-Type Processes: Construction, Approximation and Sample Path Properties*, Lecture Notes in Math., Vol. 2099: Lévy Matters III, Springer, Berlin 2014.

[14] E. A. Carlen, S. Kusuoka, and D. W. Stroock, *Upper bounds for symmetric Markov transition functions*, Ann. Inst. Henri Poincaré Probab. Statist. 23 (2) (1987), pp. 245–287.

[15] Z.-Q. Chen, P. Kim, and T. Kumagai, *Weighted Poincaré inequality and heat kernel estimates for finite range jump processes*, Math. Ann. 342 (4) (2008), pp. 833–883.

[16] Z.-Q. Chen, P. Kim, and T. Kumagai, *Global heat kernel estimates for symmetric jump processes*, Trans. Amer. Math. Soc. 363 (2011), pp. 5021–5055.

[17] Z. Q. Chen and T. Kumagai, *Heat kernel estimates for jump processes of mixed types on metric measure spaces*, Probab. Theory Related Fields 140 (1–2) (2008), pp. 277–317.

[18] Z.-Q. Chen and L. Wang, *Uniqueness of stable processes with drift*, Proc. Amer. Math. Soc. 144 (2016), pp. 2661–2675.

[19] Z.-Q. Chen and X. Zhang, *Heat kernels and analyticity of non-symmetric jump diffusion semigroups*, Probab. Theory Related Fields 165 (2016), pp. 267–312.

[20] W. Cygan, T. Grzywny, and B. Troyan, *Asymptotic behaviour of densities of unimodal convolution semigroups*, available at http://arxiv.org/pdf/1504.08358v2.pdf; Trans. Amer. Math. Soc. (to appear).

[21] A. Debussche and N. Fournier, *Existence of densities for stable-like driven SDE's with Hölder continuous coefficients*, J. Funct. Anal. 264 (8) (2013), pp. 1757–1778.

[22] Ja. M. Drin', *Fundamental solution of the Cauchy problem for a class of parabolic pseudo-differential equations* (in Ukrainian), Dopov. Akad. Nauk Ukraïn. RSR Ser. A (1977), pp. 198–203.

[23] Ja. M. Drin' and S. D. Eidelman, *Construction and investigation of classical fundamental solution of the Cauchy problem for uniformly parabolic pseudo-differential equations* (in Russian), Mat. Issled. 63 (1981), pp. 18–33.

[24] E. B. Dynkin, *Markov Processes*, vols. 1–2, Springer, Berlin 1965.

[25] S. D. Eidelman, S. D. Ivasyshen, and A. N. Kochubei, *Analytic Methods in the Theory of Differential and Pseudo-Differential Equations of Parabolic Type*, Birkhäuser, Basel 2004.

[26] S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, Wiley, New York 1986.

[27] W. Feller, *Zur Theorie der stochastischen Prozesse. Existenz- und Eindeutigkeitssätze*, Math. Ann. 113 (1936), pp. 113–160.

[28] N. Fournier and J. Printems, *Absolute continuity of some one-dimensional processes*, Bernoulli 16 (2) (2010), pp. 343–360.

[29] A. Friedman, *Partial Differential Equations of Parabolic Type*, Prentice-Hall, New York 1964.

[30] M. Fukushima, Y. Oshima, and M. Takeda, *Dirichlet Forms and Symmetric Markov Processes*, Walter de Gruyter, Berlin 1994.

[31] I. I. Gikhman and A. V. Skorokhod, *The Theory of Stochastic Processes*, vol. II, Springer, Berlin 1974.

[32] S. Hiraba, *Asymptotic behaviour of densities of multi-dimensional stable distributions*, Tsukuba J. Math. 18 (1) (1994), pp. 223–246.

[33] S. Hiraba, *Asymptotic estimates for densities of multi-dimensional stable distributions*, Tsukuba J. Math. 27 (2) (2003), pp. 261–287.

[34] W. Hoh, *The martingale problem for a class of pseudo-differential operators*, Math. Ann. 300 (1) (1994), pp. 121–147.

[35] W. Hoh, *Pseudodifferential operators with negative definite symbols and the martingale problem*, Stoch. Stoch. Rep. 55 (3–4) (1995), pp. 225–252.

[36] W. Hoh, *Pseudo differential operators generating Markov processes*, Habilitationsschrift, Bielefeld 1998.

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

[38] Ch. Iwasaki (Tsutsumi), *The fundamental solution for pseudo-differential operators of parabolic type*, Osaka Math. J. 14 (3) (1977), pp. 569–592.

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

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

[41] N. Jacob, *Pseudo-Differential Operators and Markov Processes*, vol. III: *Markov Processes and Applications*, Imperial College Press, London 2005.

[42] K. Kaleta and P. Sztonyk, *Upper estimates of transition densities for stable-dominated semigroups*, J. Evol. Equ. 13 (3) (2013), pp. 633–650.

[43] K. Kaleta and P. Sztonyk, *Estimates of transition densities and their derivatives for jump Lévy processes*, J. Math. Anal. Appl. 431 (1) (2015), pp. 260–282.

[44] K. Kaleta and P. Sztonyk, *Small time sharp bounds for kernels of convolution semigroups*, J. Anal. Math. (to appear).

[45] P. Kim and R. Song, *Stable process with singular drift*, Stochastic Process. Appl. 124 (2014), pp. 2479–2516.

[46] V. Knopova, *Compound kernel estimates for the transition probability density of a Lévy process in $\mathbb{R}^n$*, Theory Probab. Math. Statist. 89 (2014), pp. 57–70.

[47] V. Knopova and A. Kulik, *Intrinsic small time estimates for distribution densities of Lévy processes*, Random Oper. Stoch. Equ. 21 (4) (2013), pp. 321–344.

[48] V. Knopova and A. Kulik, *Parametrix construction for certain Lévy-type processes and applications*, Random Oper. Stoch. Equ. 23 (2) (2015), pp. 111–136.

[49] V. Knopova and A. Kulik, *The parametrix method and the weak solution to an SDE driven by an α-stable noise*, http://arxiv.org/abs/1412.8732; Ann. Inst. Henri Poincaré (to appear).

[50] V. Knopova and R. L. Schilling, *Transition density estimates for a class of Lévy and Lévy-type processes*, J. Theoret. Probab. 24 (1) (2012), pp. 144–170.

[51] A. N. Kochubei, *Parabolic pseudodifferential equations, hypersingular integrals, and Markov processes*, Math. USSR-Izv. 33 (1989), pp. 233–259.

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

[53] T. Komatsu, *On the martingale problem for generators of stable processes with perturbations*, Osaka J. Math. 21 (1) (1984), pp. 113–132.

[54] T. Kulczycki and M. Ryznar, *Gradient estimates of harmonic functions and transition densities for Lévy processes*, Trans. Amer. Math. Soc. 368 (2016), pp. 281–318.

[55] A. Kulik, *Difference approximation of the local times of multidimensional diffusions*, Theory Probab. Math. Statist. 78 (2009), pp. 97–114.

[56] A. Kulik, *On weak uniqueness and distributional properties of a solution to an SDE with α-stable noise*, available at http://arxiv.org/abs/1511.00106

[57] H. Kumano-go, *Pseudo-Differential Operators*, MIT Press, Cambridge, Mass., 1981.

[58] K. Kuwae and M. Takahashi, *Kato class measures of symmetric Markov processes under heat kernel estimates*, J. Funct. Anal. 250 (2007), pp. 86–113.

[59] E. E. Levi, *Sulle equazioni lineari totalmente ellittiche alle derivate parziali*, Rend. Circ. Mat. Palermo 24 (1907), pp. 275–317.

[60] A. Mimica, *Heat kernel upper estimates for symmetric jump processes with small jumps of high intensity*, Potential Anal. 36 (2) (2012), pp. 203–222.

[61] A. Mimica, *Heat kernel estimates for subordinate Brownian motions*, available at http://arxiv.org/abs/1503.01600

[62] S. I. Podolynny and N. I. Portenko, *On multidimensional stable processes with locally unbounded drift*, Random Oper. Stoch. Equ. 3 (2) (1995), pp. 113–124.

[63] N. I. Portenko, *Some perturbations of drift-type for symmetric stable processes*, Random Oper. Stoch. Equ. 2 (3) (1994), pp. 211–224.

[64] W. E. Pruitt and S. J. Taylor, *The potential kernel and hitting probabilities for the general stable process in $\mathbb{R}^n$*, Trans. Amer. Math. Soc. 146 (1969), pp. 299–321.

[65] J. Rosinski and J. L. Sinclair, *Generalized tempered stable processes*, in: *Stability in Probability*, J. K. Misiewicz (Ed.), Banach Center Publ. 90 (2010), pp. 153–170.

[66] P. Sztonyk, *Estimates of tempered stable densities*, J. Theoret. Probab. 23 (1) (2010), pp. 127–147.

[67] P. Sztonyk, *Transition density estimates for jump Lévy processes*, Stochastic Proc. Appl. 121 (2011), pp. 1245–1265.

[68] P. Sztonyk, *Estimates of densities for Lévy processes with lower intensity of large jumps*, available at http://prac.im.pwr.edu.pl/ sztonyk/

[69] M. Tsuchiya, *On a small drift of Cauchy process*, J. Math. Kyoto Univ. 10 (1970), pp. 475–492.

[70] M. Tsuchiya, *On some perturbations of stable processes*, in: *Proceedings of the Second Japan-USSR Symposium on Probability Theory (Kyoto, August 2–9, 1972)*, Springer, 1973, pp. 490–497.

[71] Ch. Tsutsumi, *The fundamental solution for a degenerate parabolic pseudo-differential operator*, Proc. Japan Acad. 50 (1974), pp. 11–15.

[72] T. Watanabe, *Asymptotic estimates of multi-dimensional stable densities and their applications*, Trans. Amer. Math. Soc. 359 (6) (2007), pp. 2851–2879.

[73] Z. Zhao, *A probabilistic principle and generalized Schrödinger perturbation*, J. Funct. Anal. 101 (1991), pp. 162–176.

Victoria Knopova   Aleksei Kulik
V. M. Glushkov Institute of Cybernetics   Institute of Mathematics NAS of Ukraine
NAS of Ukraine   3, Tereshchenkivska str.
40, Acad. Glushkov Ave.   01601 Kiev, Ukraine
03187 Kiev, Ukraine   *E-mail*: kulik@imath.kiev.ua
E-mail: vicknopova@googlemail.com