SMOOTHNESS OF HARMONIC FUNCTIONS FOR PROCESSES WITH JUMPS

Jean Picard* and Catherine Savona

Laboratoire de Mathématiques Appliquées (CNRS – UMR 6620)

Université Blaise Pascal

63177 Aubière Cedex, France

Abstract. We consider a non local operator L associated to a Markov process with jumps, we stop this process when it quits a domain D, and we study the C^j smoothness on D of the functions which are harmonic for the stopped process. A previous work was devoted to the existence of a C^∞ transition density; here, the smoothness of harmonic functions is deduced by applying a duality method and by estimating the density in small time.

Keywords. Harmonic functions, Diffusions with jumps, Excessive measures, Malliavin calculus.

Mathematics Subject Classification (1991). 60J75 60H07

*E-mail: `picard@ucfma.univ-bpclermont.fr`

1 Introduction

Consider a partial differential equation (PDE) $Lh = 0$ on some open subset D of $\mathbb{R}^d$, where we suppose that L is a second order differential operator associated to a continuous diffusion X_t. The classical Hörmander theorem gives a condition under which L is hypoelliptic on D; this means that if h is a solution of the PDE in distribution sense, then h is a C^∞ function. In particular, we can consider bounded functions h which are solutions in the probabilistic sense as they were introduced in [19]; if τ is the first exit time of D for the diffusion X_t, we require the process $h(X_t)$ stopped at time τ to be a martingale. Such a probabilistic solution will be said to be harmonic on D, and it is C^∞ under Hörmander's condition. The aim of this work is to extend the study of harmonic functions to some non local operators L associated to Markov processes X_t with jumps.

Classical probabilistic proofs of Hörmander's theorem [12, 4] are based on the smoothness of the probability transition density $y \mapsto p(t, x, y)$ of X_t and on estimates in small time. More precisely, the smoothness of p with respect to y is obtained from the Malliavin calculus, and a duality method shows that p is also smooth with respect to x; in particular, functions h which are harmonic on the whole space $\mathbb{R}^d$ are smooth. Then estimates in small time enable to localize the problem and prove the smoothness of functions which are harmonic on D. If now L is a non local operator associated to a Markov process with jumps, methods have been worked out in [3, 13, 15, 2, 16] for proving the C^∞ smoothness of the probability transition density; more precisely we follow the framework of [16]. The general scheme for studying harmonic functions is then similar to the continuous case, but the localization is much more delicate; in order to obtain it, we will make more precise the duality method, and in particular use the relation between non negative excessive functions and excessive measures. With this method, it appears that the smoothness of harmonic functions is directly related (as in the local case) to estimates for the density in small time which themselves are related to the number of jumps needed to quit D (see [17, 18] for some estimates). In particular, the harmonic functions do not always inherit the C^∞ smoothness of the transition density; we obtain the C^j smoothness when the process needs a large enough number of jumps to quit D, and we obtain the C^∞ smoothness only when the process cannot quit D with jumps. However, we also check that this assumption on the number of jumps can be removed under additional smoothness assumptions on L.

We will consider more generally functions $h(t,x)$ which are solutions of the heat equation $\partial h/\partial t = Lh$ on $\mathbb{R}_+ \times D$; harmonic functions $h(x)$ are then a particular case.

In Section 2, we state the problem and the main result. We obtain some preliminary estimates in small time in Section 3, and complete the proof of the main result in Section 4. We derive some extensions in Section 5, and consider the case of smooth jumps in Section 6.

2 The main result

We first introduce the class of Markov processes with jumps for which one can apply the result of [16] for the existence of smooth densities and of [18] for their behaviour in small time; other Malliavin calculus techniques ([3, 13, 15, 2]) can probably be also applied by modifying the subsequent proofs. The advantage of the approach of [16] is that it can be applied to singular Lévy measures. Thus, following [16], we suppose that the operator L on $\mathbb{R}^d$ does not contain a second order part and that its Markov process can be interpreted as the solution of an equation driven by a Lévy process. We let

$$\begin{aligned} Lf(x) &= f'(x)b(x) \\ &\quad + \int \Big(f(x+\gamma(x,\lambda)) - f(x) - f'(x)\gamma_0(x)\lambda 1_{\{|\lambda|\le 1\}} \Big) \mu(d\lambda), \end{aligned}$$

with the following assumptions.

Assumptions on μ. We suppose that $\mu(d\lambda)$ is a measure on $\mathbb{R}^m \setminus \{0\}$ which integrates $|\lambda|^2 \wedge 1$, and that there exists an index $\beta \in (0,2)$ such that

$$c\rho^{2-\beta} I \le \int_{\{|\lambda|\le\rho\}} \lambda\lambda^\star \mu(d\lambda) \le C\rho^{2-\beta} I \tag{1}$$

as $\rho \to 0$ (this is an inequality between symmetric matrices). This condition can also be written as

$$c\rho^{2-\beta} \le \int_{\{|\lambda|\le\rho\}} (\lambda.u)^2 \mu(d\lambda) \le C\rho^{2-\beta}$$

for unit vectors u. If $\beta \in (0,1)$, we replace (1) by the stronger condition

$$c\rho^{2-\beta} \leq \int_{\{|\lambda|\leq\rho\}} (\lambda.u)^2 1_{\{\lambda.u>0\}}\mu(d\lambda) \leq C\rho^{2-\beta} \tag{2}$$

for unit vectors u; if $\beta = 1$, we suppose in addition to (1) that

$$\limsup_{\varepsilon\to 0}\Big|\int_{\{\varepsilon<|\lambda|\leq 1\}} \lambda\,\mu(d\lambda)\Big| < \infty. \tag{3}$$

We associate to μ a m-dimensional Lévy process Λ_t with characteristic function given by the Lévy-Khintchine formula

$$\mathbb{E}[e^{iw.\Lambda_1}] = \exp\int (e^{iw.\lambda} - 1 - iw.\lambda 1_{\{|\lambda|\leq 1\}})\mu(d\lambda) \tag{4}$$

(the measure μ is called the Lévy measure of Λ_t). We recall that

$$\Lambda_t - \sum_{s\leq t} 1_{\{|\Delta\Lambda_s|>1\}}\Delta\Lambda_s$$

is a martingale.

Remark. The assumption (1) implies that

$$\int_{\{|\lambda|\leq 1\}} |\lambda|^\alpha \mu(d\lambda) < \infty \tag{5}$$

for any $\alpha > \beta$. This property will be important subsequently.

Remark. If μ is symmetric, then conditions (2) and (1) are equivalent, and (3) is always satisfied.

Examples. Let μ be a measure satisfying the scaling property

$$\mu(rB) = r^{-\beta}\mu(B) \tag{6}$$

for $r > 0$; if $\mathcal{S}^{m-1}$ is the unit sphere of $\mathbb{R}^m$, the map $x \mapsto (|x|, x/|x|)$ enables to identify $\mathbb{R}^m \setminus \{0\}$ and $\mathbb{R}_+^\star \times \mathcal{S}^{m-1}$, and measures μ satisfying (6) can be written as

$$\mu(dr, dS) = r^{-1-\beta} dr\, \mu_0(dS)$$

for a measure μ_0 on $\mathcal{S}^{m-1}$. Then

$$\int_{\{|\lambda|\leq\rho\}} \lambda\lambda^\star \mu(d\lambda) = \rho^{2-\beta}\int_{\{|\lambda|\leq 1\}} \lambda\lambda^\star\mu(d\lambda),$$

so the assumption (1) is satisfied as soon as μ is not supported by an hyperplane; if $\beta \in (0,1)$, the condition (2) is satisfied if μ is not supported by a closed half space, and if $\beta = 1$, the condition (3) is satisfied if

$$\int_{\mathcal{S}^{m-1}} S\,\mu_0(dS) = 0.$$

The Lévy process Λ_t is then a β-stable process (plus possibly a drift). An example of such a measure is the measure with density

$$\frac{\mu(d\lambda)}{d\lambda} = |\lambda|^{-m-\beta} \tag{7}$$

so that μ_0 is a uniform measure on $\mathcal{S}^{m-1}$; the Lévy process Λ_t is a rotation-invariant β-stable process. Another example is to take for μ_0 a purely atomic measure

$$\mu_0 = \sum_{k=1}^{K} \alpha_k \delta_{S_k}, \qquad \alpha_k > 0, \qquad S_k \in \mathcal{S}^{m-1}, \tag{8}$$

satisfying the above assumptions. The Lévy process is then the sum of K independent β-stable processes X_t^k with values in the lines $\mathbb{R}S_k$. One can also consider the case where the scaling property (6) is not satisfied for all r, but only for some geometric sequence, for instance $r_i = 2^{-i}$; then (1) (or (2)) again holds as soon as μ is not supported by an hyperplane (or a closed half space); this enables the study of purely atomic measures such as

$$\mu = \sum_{k=1}^{K} \sum_{i \in \mathbb{Z}} 2^{i\beta} \delta_{2^{-i} S_k}. \tag{9}$$

More details about the Lévy process associated to this measure in the case $m = K = 1$ can be found in [17], where the behaviour in small time is studied.

In view of these examples, the condition (1) can be viewed as an approximate scaling and non degeneracy condition. The additional conditions (2) or (3) are required in [18] for the derivation of small time estimates; they imply that the influence of the drift is negligible in the small time behaviour of the semigroup associated to L.

Assumptions on the coefficients b, γ and γ_0. We suppose that

$$\gamma(x,\lambda) = \gamma_0(x)\lambda + O(|\lambda|^\alpha) \tag{10}$$

for some $\alpha > 1 \vee \beta$, as $\lambda \to 0$, uniformly in x, that the coefficients b and γ_0 are C_b^∞, that γ is C_b^∞ with respect to x uniformly in (x, λ), and that the relation (10) also holds for the derivatives with respect to x. We also suppose that the function $x \mapsto x + \gamma(x, \lambda)$ is invertible and that its inverse can be written as $x \mapsto x + \overline{\gamma}(x, \lambda)$ for a function $\overline{\gamma}$ which is C_b^∞ in x uniformly in (x, λ) (notice that $\overline{\gamma}$ has a decomposition of type (10) with $\overline{\gamma}_0 = -\gamma_0$). If $\beta \in (0, 1)$, then $|\lambda| \wedge 1$ is μ-integrable, and we suppose moreover that

$$b(x) = \gamma_0(x) \int \lambda 1_{\{|\lambda| \le 1\}} \mu(d\lambda) \tag{11}$$

so that

$$Lf(x) = \int \Big(f(x + \gamma(x, \lambda)) - f(x) \Big) \mu(d\lambda).$$

The assumption (10) and the property (5) imply that $Lf(x)$ is well defined for any C_b^2 function f.

Notice that no smoothness assumption is made with respect to λ, except as $\lambda \to 0$. In particular, if μ is a singular measure, it is not possible to use an integration by parts with respect to λ. However, in Section 6, we will see what can be said when μ and the coefficients are smooth with respect to λ.

When the measure $|\lambda|^2 \mu(d\lambda)$ converges to the Dirac mass at 0, then the operator L converges to a second order differential operator L_0 with diffusion coefficient $\gamma_0 \gamma_0^\star$; in particular, the Hörmander theorem gives a condition under which L_0 is hypoelliptic. However, in this work, in order to apply [16], we will not assume Hörmander's condition on γ_0, but a more restrictive condition, namely the non degeneracy of $\gamma_0 \gamma_0^\star$. We want to prove the "hypoellipticity" of L, and more generally of $L - \partial/\partial t$. However this will not be a genuine hypoellipticity since the operator is not local and we will not always obtain the C^∞ smoothness.

We now give the probabilistic interpretation of the operator L. By using the Lévy process Λ_t of (4), it is the generator of the process $X_t = X_t(x)$ solution of

$$X_t = x + \int_0^t b(X_s) ds + \int_0^t \gamma(X_{s-}, d\Lambda_s), \tag{12}$$

where the stochastic integral is defined by

$$\int_0^t \gamma(X_{s-}, d\Lambda_s) = \int_0^t \gamma_0(X_{s-}) d\Lambda_s + \sum_{s \le t} \Big(\gamma(X_{s-}, \Delta\Lambda_s) - \gamma_0(X_{s-}) \Delta\Lambda_s \Big)$$

and converges from (5) and (10). Our assumptions are sufficient to ensure the existence and the uniqueness of a solution to (12) (Theorem IV.9.1 of [10]); the smoothness of the coefficients and the invertibility of $x \mapsto x + \gamma(x, \lambda)$ imply that $x \mapsto X_t(x)$ has a modification consisting in a stochastic flow of diffeomorphisms ([8]). If $\beta \in (0, 1)$, then X_t has finite variation, and the additional condition (11) means that X_t is a pure jump process. Notice that γ is supposed to be bounded, so the jumps of X_t are bounded.

Definition 1 *Let D be an open subset of $\mathbb{R}^d$ and let*

$$\tau = \tau(x) = \inf\Big\{t \geq 0; X_t(x) \notin D\Big\}$$

be the first exit time of D.

1. *A locally bounded function $h(x)$ defined on $\mathbb{R}^d$ is said to be harmonic on D if the stopped process $h(X_{t\wedge\tau})$ is a local martingale for any initial condition $x \in D$.*

2. *A locally bounded function $h(t, x)$ defined on $\mathbb{R}_+ \times \mathbb{R}^d$ is said to be solution of the heat equation $\partial h/\partial t = Lh$ on $\mathbb{R}_+ \times D$ if the process $(h(r - t, X_t); 0 \leq t \leq r)$ stopped at time τ is a local martingale for any $r > 0$ and any initial condition $x \in D$.*

Notice that the local martingales involved in this definition are (up to a negligible event) right continuous. This is because the process X_t is a strong Markov process, see XVII.5 in [7]; the set of times where the martingale differs from its right continuous modification is optional, and at any optional time one can apply the strong Markov property.

Example. If one is given a bounded function ϕ on D^c, one can consider the Dirichlet problem $Lh = 0$ in D, and $h = \phi$ in D^c. A bounded solution can be constructed by

$$h(x) = \mathbb{E}[\phi(X_\tau)1_{\{\tau<\infty\}}]$$

for $X_t = X_t(x)$ and we obtain a bounded harmonic function. To see if the solution is unique, one has to see if τ is finite with probability 1. The process can quit D either with a jump, or continuously (X continuous at τ); in the symmetric case, it is shown in [5] that this question can be translated in terms of Dirichlet spaces. Notice that if it may quit with a jump, then the function ϕ should be defined on D^c and not only on ∂D as in the continuous case. If $D = \mathbb{R}^\star \times \mathbb{R}^{d-1}$ and X_t is a Lévy process, then τ is the hitting time of 0 by the real Lévy process X_t^1; one knows that this hitting time is finite with positive probability in some situations such as $\beta > 1$ (this question is related to the potential analysis of the process, see for instance [1]), and it is finite with probability 1 if moreover X_t^1 is recurrent. This situation can be extended to more general Markov processes X_t and more general sets D such that D^c is an hypersurface, see [18]. Notice that a more analytic study of the probabilistic Dirichlet problem can be found in [9].

Definition 2 *Let $\mathcal{A}_1$ be the set of points (x, y) of $\mathbb{R}^d \times \mathbb{R}^d$ such that*

$$\mu\Big\{\lambda; |y - x - \gamma(x, \lambda)| \leq \varepsilon\Big\} > 0$$

for any $\varepsilon > 0$. Let $\mathcal{A}_n$ be the set of points (x, y) for which there exists a chain $x = y_0, y_1, \ldots, y_n = y$ such that (y_j, y_{j+1}) is in $\mathcal{A}_1$. We also let $\mathcal{A}_n(x)$ be the set of y such that (x, y) is in $\mathcal{A}_n$. If y is in $\mathcal{A}_n(x)$, we say that y is accessible from x (by the process X_t) in n jumps; similarly, if A and B are subsets of $\mathbb{R}^d$, we say that B is accessible from A in n jumps if $A \times B$ intersects $\mathcal{A}_n$.

We deduce from (1) that 0 is an accumulation point of the support of μ, so $\mathcal{A}_n$ is included in $\mathcal{A}_{n+1}$. It follows from the smoothness of the coefficients that $\mathcal{A}_n(x)$ and $\mathcal{A}_n$ are closed. We can now state the main result of this work.

Theorem 1 *Assume that $\gamma_0\gamma_0^\star$ is elliptic at any point. Let h be a locally bounded function which is harmonic on an open set D, and let $x_0 \in D$. For any integer j, there exists an integer n depending only on j, the dimension d and the scaling index β such that if D^c is not accessible from x_0 in n jumps, then h is C^j on a neighbourhood of x_0.*

Example. If $D = (0, \infty) \times \mathbb{R}^{d-1}$ and if the jumps of the first component X_t^1 of X_t are positive, then D^c is not accessible with any number of jumps, so we can conclude that h is C^∞ in D. Other cases may be more complicated; for instance, if $D = \mathbb{R}^\star \times \mathbb{R}^{d-1}$ and if the jumps of X_t^1 take their values in the set of $\pm 2^i$, $i \in \mathbb{Z}$, each of them being possible with positive probability (see the example (9), take for S_k, $1 \leq k \leq d$, the canonical basis of $\mathbb{R}^d$, $S_{d+k} = -S_k$

and $X_t = \Lambda_t$), then the number of jumps which are needed to reach D^c is the number of 1's in the dyadic decomposition of the first component x_1 of x; one can say that h is smooth at points x such that x_1 is not a dyadic number.

By putting $h(t,x) = h(x)$, an harmonic function becomes a solution of the heat equation, so Theorem 1 becomes a particular case of the following result.

Theorem 2 *Assume that $\gamma_0\gamma_0^\star$ is elliptic at any point and let D be an open subset of $\mathbb{R}^d$. Let $h(t,x)$ be a locally bounded solution of the heat equation $\partial h/\partial t = Lh$ on $\mathbb{R}_+ \times D$, and let $x_0 \in D$. For any integer j, there exists an integer n depending only on j, d and β such that if D^c is not accessible from x_0 with n jumps, then there exists a neighbourhood B_0 of x_0 such that h is C^j on $(0,\infty) \times B_0$.*

Example. Let $X_t(x)$ be real-valued and $D = \mathbb{R}^\star$; then one has to consider the hitting time $\tau = \tau(x)$ of 0 by $X_t(x)$. Consider the function

$$h(t,x) = \mathbb{P}[\tau(x) < t] \tag{13}$$

defined on $\mathbb{R} \times \mathbb{R}$; in particular $h(t,x) = 0$ for $t \le 0$. If $\mathcal{F}_t$ is the filtration of Λ_t, one has

$$h(r-t, X_t(x)) = \mathbb{P}[\tau(x) < r \mid \mathcal{F}_t]$$

on $\{t \le \tau(x)\}$, so h is solution of the heat equation on $\mathbb{R} \times D$ (and not only on $\mathbb{R}_+ \times D$). Theorem 2 says that the smoothness of h is related to the number of jumps needed to reach 0 from x; in particular, the function $t \mapsto h(t,x)$ is C^j on $\mathbb{R}$ if this number is large enough. On the other hand, in the case $\beta > 1$, if 0 is accessible from x in a finite number of jumps and under some other assumptions (see [18]), the function $h(t,x)$ is of order t^η as $t \downarrow 0$ for some positive $\eta = \eta(x)$. This implies that the function $t \mapsto h(t,x)$ is not C^j for $j \ge \eta$. This shows the importance of the assumption of inaccessibility.

The next two sections are devoted to the proof of Theorem 2. We have assumed that the jumps are bounded (γ bounded), so the problem can be reduced to the case where D is bounded; to this end, one intersects D with a ball with large enough radius, so that the complement of this ball is not accessible from x_0 with n jumps (n is chosen from the proof of the theorem). We also suppose in the proof that $\gamma_0\gamma_0^\star$ is uniformly elliptic and h is bounded.

3 Estimates in small time

We first give a large deviation result for the law of X_t as $t \downarrow 0$; this result was proved in [17] in the case of real-valued Lévy processes, and we extend it to our class of Markov processes.

Lemma 1 *Fix $n \geq 0$ and $x_0 \in \mathbb{R}^d$. Let B_n be a neighbourhood of $\mathcal{A}_n(x_0)$. Then*

$$\mathbb{P}[X_t(x) \notin B_n] \leq C\, t^{n+1}$$

for x in a neighbourhood of x_0.

Proof. Let $B'_n \subset B_n$ be another neighbourhood of $\mathcal{A}_n(x_0)$, such that the distance between B'_n and $(B_n)^c$ is $\varepsilon > 0$. We deduce from the smoothness of the coefficients that B'_n contains $\mathcal{A}_n(x)$ for x in a neighbourhood of x_0, and we are going to prove the lemma for these x. Fix $\rho > 0$ (it will be chosen small enough later), and consider for each $t > 0$ the decomposition (depending on t)

$$\Lambda_s = \Lambda_s^1 + \Lambda_s^2 + \Lambda_s^3, \qquad s \leq t$$

into independent Lévy processes defined as follows; the process Λ_s^1 is the sum of jumps $\Delta\Lambda_u$, $u \leq s$, such that $|\Delta\Lambda_u| > \rho$, and the process Λ_s^2 is the sum of jumps such that $\rho t^{1/4} < |\Delta\Lambda_u| \leq \rho$. We denote by J_s^1 and J_s^2 the number of jumps of Λ^1 and Λ^2 up to time s. Let $X_s^1 = X_s^1(x)$, $s \leq t$ be the step process defined by

$$\Delta X_s^1 = \gamma(X_{s-}^1, \Delta\Lambda_s^1), \qquad X_0^1 = x.$$

The support of the law of $X_t^1(x)$ on $\{J_t^1 = n\}$ is $\mathcal{A}_n(x)$, so the support on $\{J_t^1 \leq n\}$ is also $\mathcal{A}_n(x)$, and it is therefore included in B'_n. Thus

$$\mathbb{P}[X_t(x) \notin B_n] \leq \mathbb{P}[|X_t - X_t^1| \geq \varepsilon] + \mathbb{P}[J_t^1 > n]. \tag{14}$$

On the other hand,

$$\begin{aligned} X_s - X_s^1 &= \int_0^s b(X_u)du + \int_0^s \gamma(X_{u-}, d\Lambda_u^3) + \sum_{u \leq s} \gamma(X_{u-}, \Delta\Lambda_u^2) \\ &\quad + \sum_{u \leq s} \Big(\gamma(X_{u-}, \Delta\Lambda_u^1) - \gamma(X_{u-}^1, \Delta\Lambda_u^1) \Big), \end{aligned}$$

so

$$
\begin{aligned}
|X_s - X^1_s| \;\; &\leq \;\; C\,s + \Big|\int_0^s \gamma(X_{u-}, d\Lambda^3_u)\Big| + C\sum_{u\leq s}|\Delta\Lambda^2_u| \\
&\quad + C\sum_{u\leq s}|X_{u-} - X^1_{u-}|\Delta J^1_u. \qquad (15)
\end{aligned}
$$

The integral with respect to Λ^3 is a semimartingale with bounded jumps, and we look for its decomposition into a local martingale and a predictable process with finite variation. We have

$$
\int_0^s \gamma(X_{u-}, d\Lambda^3_u) = \int_0^s \gamma_0(X_{u-})d\Lambda^3_u + \sum_{u\leq s}\Big(\gamma(X_{u-}, \Delta\Lambda^3_u) - \gamma_0(X_{u-})\Delta\Lambda^3_u\Big).
$$

The process

$$
\Lambda_s - \sum_{u\leq s} 1_{\{|\Delta\Lambda_u|>1\}}\,\Delta\Lambda_u = \Lambda^3_s + \sum_{u\leq s} 1_{\{\rho t^{1/4}<|\Delta\Lambda_u|\leq 1\}}\,\Delta\Lambda_u
$$

is a local martingale, so we can deduce that

$$
\begin{aligned}
M_s \;\; = \;\; & t^{-1/4}\Big(\int_0^s \gamma(X_{u-}, d\Lambda^3_u) - \int_0^s\int_{\{|\lambda|\leq\rho t^{1/4}\}}(\gamma(X_u,\lambda) - \gamma_0(X_u)\lambda)\mu(d\lambda)du \\
& + \int_0^s\int_{\{\rho t^{1/4}<|\lambda|\leq 1\}}\gamma_0(X_u)\lambda\,\mu(d\lambda)du\Big)
\end{aligned}
$$

is a local martingale. If we estimate the two last integrals, we obtain

$$
\begin{aligned}
& \Big|\int_0^s\int_{\{|\lambda|\leq\rho t^{1/4}\}}(\gamma(X_u,\lambda) - \gamma_0(X_u)\lambda)\mu(d\lambda)du\Big| \\
& \leq C\,s\int_{\{|\lambda|\leq\rho t^{1/4}\}}|\lambda|^\alpha\mu(d\lambda) \leq C'\,s \leq C'\,t,
\end{aligned}
$$

and

$$
\begin{aligned}
& \Big|\int_0^s\int_{\{\rho t^{1/4}<|\lambda|\leq 1\}}\gamma_0(X_u)\lambda\,\mu(d\lambda)du\Big| \\
& \leq C\,s\int_{\{\rho t^{1/4}<|\lambda|\leq 1\}}|\lambda|\mu(d\lambda) \leq C'\,s\,t^{-1/4} \leq C'\,t^{3/4},
\end{aligned}
$$

where the second inequality follows from the μ-integrability of $(|\lambda|^2\wedge 1)$. Thus

$$
\Big|\int_0^s \gamma(X_{u-}, d\Lambda^3_u)\Big| \leq t^{1/4}|M_s| + C\,t^{3/4}
$$

and (15) becomes

$$|X_s - X_s^1| \le C\, t^{3/4} + t^{1/4}|M_s| + C\rho J_s^2 + C\sum_{u\le s}|X_{u-} - X_{u-}^1|\Delta J_u^1.$$

If one defines

$$R_s = \sup_{u\le s}|X_u - X_u^1|$$

for $s \le t$, then

$$R_s \le C\Big(t^{3/4} + t^{1/4}\sup_{u\le t}|M_u| + \rho J_t^2\Big) + C\sum_{u\le s} R_{u-}\Delta J_u^1.$$

We deduce that

$$|X_t - X_t^1| \le R_t \le C\Big(t^{3/4} + t^{1/4}\sup_{s\le t}|M_s| + \rho J_t^2\Big)e^{C\,J_t^1}. \tag{16}$$

The constant C does not depend on ρ, and we now choose ρ small enough so that

$$C(2n+1)\rho\, e^{C\,n} \le \varepsilon/2, \tag{17}$$

where ε was introduced in the beginning of the proof. The variable J_t^1 is a Poisson variable with mean $t\,\mu(|\lambda| > \rho)$, so the probability that it is greater than n is $O(t^{n+1})$. Similarly, the Poisson variable J_t^2 has mean

$$t\,\mu(\rho t^{1/4} < |\lambda| \le \rho) = O(\sqrt{t})$$

because $|\lambda|^2 \wedge 1$ is μ integrable, so the probability that J_t^2 is greater than $2n+1$ is also $O(t^{n+1})$. Thus (16) and (17) imply that

$$\begin{aligned}
|X_t - X_t^1| &\le C\Big(t^{3/4} + t^{1/4}\sup_{s\le t}|M_s| + (2n+1)\rho\Big)e^{C\,n} \\
&\le C\Big(t^{3/4} + t^{1/4}\sup_{s\le t}|M_s|\Big)e^{C\,n} + \varepsilon/2 \qquad (18)
\end{aligned}$$

except on an event of probability $O(t^{n+1})$. The jumps of each component M_s^i of M_s are bounded in absolute value by $1/2$ if ρ is small enough, so the process

$$\mathcal{E}_s^i = e^{M_s^i}\prod_{u\le s}(1+\Delta M_u^i)e^{-\Delta M_u^i}$$

is a positive local martingale and its expectation is therefore at most 1. Thus

$$\begin{aligned}
\mathbb{E}[e^{M_t^i/2}] &= \mathbb{E}\Big[(\mathcal{E}_t^i)^{1/2}\prod_{u\le t}(1+\Delta M_u^i)^{-1/2}e^{\Delta M_u^i/2}\Big]\\
&\le \mathbb{E}\Big[\prod_{u\le t}(1+\Delta M_u^i)^{-1}e^{\Delta M_u^i}\Big]^{1/2}\\
&\le \mathbb{E}\Big[\exp(C\,[M^i,M^i]_t)\Big]^{1/2}\\
&\le \mathbb{E}\Big[\exp(C'\,t^{-1/2}[\Lambda^3,\Lambda^3]_t)\Big]^{1/2}.
\end{aligned}$$

The variable $t^{-1/2}[\Lambda^3,\Lambda^3]_t$ is an infinitely divisible variable; its Lévy measure has bounded support (uniformly as $t\to 0$), and its expectation and variance are bounded; one can deduce from the Lévy-Khintchine formula that it has bounded exponential moments. Thus $e^{M_t^i/2}$ has bounded expectation; in particular, M_t has bounded moments, and also $\sup_{s\le t}|M_s|$ by Doob's inequality; this variable is therefore less than $t^{-1/8}$ except on an event of probability $O(t^k)$ for any k. By using this estimate in (18), we obtain

$$|X_t - X_t^1| \le C(t^{3/4}+t^{1/8})e^{C\,n}+\varepsilon/2<\varepsilon$$

for t small enough, expect on an event of probability $O(t^{n+1})$. Thus, since $\mathbb{P}[J_t^1>n]$ is $O(t^{n+1})$, we can conclude from (14). $\square$

Lemma 2 *The solution $X_t(x)$ of (12) has a C^∞ density $y\mapsto p(t,x,y)$ for $t>0$. Let $p^{(0)}=p$ and for $k\ge 1$, denote by $p^{(k)}$ the vector consisting of all the derivatives of order k with respect to y. For any k, there exists an integer n satisfying the following property; if y_0 is not accessible from x_0 in n jumps, then $p^{(k)}(t,x,y)$ converges to 0 as $t\downarrow 0$ uniformly for (x,y) in a neighbourhood of (x_0,y_0).*

Remark. Estimates in small time for the density of X_t can be found in [14, 11, 17, 18], but here we also need estimates on the derivatives of the density.

Proof. We know from [16, 18] that X_t has a smooth density, and $p^{(k)}(t,x,y)$ is uniformly dominated by $t^{-(k+d)/\beta}$ (Theorem 1 of [18]). On the other hand, if ϕ is a smooth function with compact support in $\mathbb{R}^d$, its Fourier transform

$\widehat{\phi}$ can be estimated by means of the L^1 norm of ϕ or its derivatives; one has for any $j \geq 0$

$$|\widehat{\phi}(u)| \leq C\,|u|^{-j}\|\phi^{(j)}\|_1.$$

One deduces from the Fourier inversion formula that

$$\begin{aligned}
|\phi^{(k)}(y)| &\leq C\int |u|^k\,|\widehat{\phi}(u)|du \\
&\leq C'\|\phi\|_1\int_{\{|u|\leq M\}}|u|^k du + C'\|\phi^{(j)}\|_1\int_{\{|u|>M\}}|u|^{k-j}du \\
&\leq C''M^{k+d}\|\phi\|_1 + C''M^{-1}\|\phi^{(j)}\|_1
\end{aligned}$$

where we have taken $j = k + d + 1$ in the last inequality. We apply this relation by letting ϕ be the function $y \mapsto p(t,x,y)$ multiplied by a localization function (a smooth function which is 1 in a neighbourhood of y_0 and has a compact support which is disjoint from $\mathcal{A}_n(x_0)$). The L^1 norm of this function is $O(t^{n+1})$ from Lemma 1, the L^1 norm of $\phi^{(j)}$ is $O(t^{-(j+d)/\beta})$, so we obtain

$$|p^{(k)}(t,x,y)| \leq C\,M^{k+d}t^{n+1} + C\,M^{-1}t^{-(k+2d+1)/\beta}$$

for (x,y) in a neighbourhood of (x_0,y_0). We choose $M = t^{-i}$ for $i > (k+2d+1)/\beta$, and then $n > i(k+d) - 1$ so that the lemma holds. $\square$

We now consider the process $\overline{X}_t = \overline{X}_t(x)$ which is the process X_t killed when it quits D; this means that $\overline{X}_t = X_t$ on $\{t < \tau\}$, and $\overline{X}_t$ is a cemetery point ∂ on $\{t \geq \tau\}$. This process has a transition density $\overline{p}(t,x,y)$.

Lemma 3 *For any k, there exists an integer n satisfying the following property.*

1. *If y_0 is not accessible from $\{x_0\} \cup D^c$ in n jumps, then $y \mapsto \overline{p}(t,x,y)$ is C_b^k, uniformly for $t > 0$ and (x,y) in a neighbourhood of (x_0,y_0).*
2. *If y_0 is not accessible from D^c in n jumps, the same property holds for $t \geq t_0 > 0$, x in D and y in a neighbourhood of y_0.*

Proof. We write

$$\mathbb{E}[f(\overline{X}_t)] = \mathbb{E}[f(X_t)1_{\{t<\tau\}}] = \mathbb{E}[f(X_t)] - \mathbb{E}[\mathbb{E}[f(X_t)|\mathcal{F}_\tau]1_{\{\tau\leq t\}}].$$

If $\nu(x, ds, dz)$ is the law of (τ, X_τ) on $\{\tau < \infty\}$, we deduce from the strong Markov property that

$$\overline{p}(t, x, y) = p(t, x, y) - \int_0^t \int_{D^c} p(t - s, z, y)\nu(x, ds, dz).$$

Since the jumps are bounded, the integral with respect to z is actually on a compact subset of D^c. One can then apply Lemma 2 to estimate the derivatives with respect to y of this integral. The smoothness of $p(t, x, y)$ is obtained from Lemma 2 for the first statement of the lemma, and from [16] for the second statement. □

4 Proof of the main result

We consider a function $h(t, x)$ which is a solution of the heat equation and study its smoothness in order to prove Theorem 2. We first consider its smoothness with respect to x. To this end, we have to reduce the problem to a more tractable one.

Lemma 4 *The problem can be reduced to the case where the Lebesgue measure is almost surely invariant by the stochastic flow $x \mapsto X_t(x)$; in particular it is invariant by the semigroup of X_t.*

Proof. Let J_t be the Jacobian determinant of $x \mapsto X_t(x)$. Differentiation of (12) yields

$$dJ_t = B_0(X_t)J_t dt + J_{t-}B_1(X_{t-}, d\Lambda_t), \qquad J_0 = 1,$$

where B_0 is the divergence of b,

$$B_1(x, \lambda) = \det(I + \gamma'(x, \lambda)) - 1$$

and γ' is the derivative with respect to x. Let H_t be an independent real-valued Lévy process with Lévy measure

$$|x|^{-1-\beta} 1_{\{|x| \le 1\}}$$

(this is a truncated β-stable process), let $V_t = V_t(x, v)$ be the solution of

$$dV_t = -B_0(X_t)V_t dt - V_{t-}B_2(X_{t-}, d\Lambda_t) + dH_t, \qquad V_0 = v, \tag{19}$$

with $B_2 = B_1(1+B_1)^{-1}$, and consider the process $\widetilde{X}_t(x,v) = (X_t(x), V_t(x,v))$ which is the solution of an equation driven by the Lévy process (Λ_t, H_t); it satisfies the assumptions of Section 2 (the non degeneracy comes from the introduction of H_t). The Jacobian of the map $(x,v) \mapsto \widetilde{X}_t(x,v)$ is $J_t W_t$ with

$$W_t(x,v) = \partial V_t(x,v) \;/\; \partial v.$$

The differentiation of (19) shows that W_t is solution of

$$dW_t = -B_0(X_t)W_t dt - W_{t-}B_2(X_{t-}, d\Lambda_t), \qquad W_0 = 1,$$

so by writing the equation satisfied by $J_t W_t$, we obtain $J_t W_t = 1$. Thus the Lebesgue measure is invariant by the stochastic flow of $\widetilde{X}_t$. Moreover, the function $h(t,x,v) = h(t,x)$ is a solution of the heat equation for the process $\widetilde{X}_t$. $\square$

In Section 3, we have studied the smoothness with respect to y of the density $\overline{p}(t,x,y)$. We actually need the smoothness with respect to x, and as it has been said in the introduction, this will be made with a duality method that we now describe.

Let $X_t^\star = X_t^\star(x)$ be the solution of

$$dX_t^\star = -b(X_t^\star)dt + \overline{\gamma}(X_{t-}^\star, d\Lambda_t), \qquad X_0^\star = x,$$

where we recall that $x \mapsto x + \overline{\gamma}(x,\lambda)$ is the inverse of $x \mapsto x + \gamma(x,\lambda)$. Let $\mathcal{C} = C(\mathbb{R}^d, \mathbb{R}^d)$ be the space of continuous functions from $\mathbb{R}^d$ into itself, endowed with the topology of uniform convergence on compact subsets. Then X_t and $X_t^\star$ can be viewed as $\mathcal{C}$-valued variables, and if X_t^{-1} is the inverse of X_t, we have the following result.

Lemma 5 *The variables $X_t^\star$ and X_t^{-1} have the same law.*

Proof. This result will be proved by approximating X_t by processes with finitely many jumps, so let us first suppose that the Lévy measure μ is finite. Then (12) can be written as

$$X_t = x + \int_0^t b_0(X_s)ds + \sum_{s \le t} \gamma(X_{s-}, \Delta\Lambda_s)$$

with

$$b_0(x) = b(x) - \gamma_0(x) \int_{\{|\lambda|\leq 1\}} \lambda \, \mu(d\lambda).$$

Fix t, let J be the number of jumps of Λ before t, and let $(T_j;\ 1 \leq j \leq J)$ be the times of the jumps. Then $x \mapsto X_t(x)$ can be written as

$$X_t = \phi_{t-T_J} \circ \psi(\Delta\Lambda_{T_J}) \circ \phi_{T_J - T_{J-1}} \circ \ldots \circ \phi_{T_2 - T_1} \circ \psi(\Delta\Lambda_{T_1}) \circ \phi_{T_1}$$

where ϕ_t is the flow of the equation $\dot{x}_t = b_0(x_t)$, and $\psi(\lambda)$ is the map $x \mapsto x + \gamma(x, \lambda)$. Thus

$$X_t^{-1} = \phi_{T_1}^{-1} \circ \psi(\Delta\Lambda_{T_1})^{-1} \circ \phi_{T_2 - T_1}^{-1} \circ \ldots \circ \psi(\Delta\Lambda_{T_J})^{-1} \circ \phi_{t-T_J}^{-1}.$$

Notice that ϕ_t^{-1} is the flow of $\dot{x}_t = -b_0(x_t)$ and that $\psi(\lambda)^{-1}$ is the map $x \mapsto x + \overline{\gamma}(x, \lambda)$. Moreover

$$(J, T_1, \Delta\Lambda_{T_1}, T_2 - T_1, \Delta\Lambda_{T_2}, \ldots, T_J - T_{J-1}, \Delta\Lambda_{T_J}, t - T_J)$$

and

$$(J, t - T_J, \Delta\Lambda_{T_J}, \ldots, T_3 - T_2, \Delta\Lambda_{T_2}, T_2 - T_1, \Delta\Lambda_{T_1}, T_1)$$

have the same law. Thus X_t^{-1} has the law of a variable which looks like X_t, but with b_0 and γ replaced by $-b_0$ and $\overline{\gamma}$; this is exactly $X_t^\star$. The general case (when μ is infinite) is obtained by approximating the Lévy process Λ_t by

$$\sum_{s \leq t} 1_{\{|\Delta\Lambda_s| > \rho\}} \Delta\Lambda_s - t \int_{\{\rho < |\lambda| \leq 1\}} \lambda \, \mu(d\lambda)$$

which has finitely many jumps for $\rho > 0$. We obtain the solutions X_t^ρ and $(X_t^\rho)^\star$ of the corresponding equations, and we now know that $(X_t^\rho)^\star$ and $(X_t^\rho)^{-1}$ have the same law. In order to take the limit as $\rho \downarrow 0$ in this property, we apply the following deterministic result; if f_n is a sequence of homeomorphisms of $\mathbb{R}^d$, if f_n and f_n^{-1} converge respectively to limits f and g in $\mathcal{C}$, then f and g are homeomorphisms and $g = f^{-1}$. Here, the uniform estimates

$$\mathbb{E}\Big[|X_t^\rho(x) - X_t^\rho(y)|^p\Big] \leq C_p |y - x|^p,$$

which can be deduced from the techniques of [2, 8], show that the law of $(X_t^\rho;\ \rho > 0)$ is tight in $\mathcal{C}$. Similarly, the law of $(X_t^\rho)^\star$, and therefore of $(X_t^\rho)^{-1}$, is tight. Thus the law of the couple $(X_t^\rho, (X_t^\rho)^{-1})$ is tight, and the above deterministic result shows that if (Ξ_1, Ξ_2) is any limit, then Ξ_1 and

Ξ_2 are almost surely homeomorphisms, and $\Xi_2 = \Xi_1^{-1}$. On the other hand, one can deduce from the results of Section 5 of [2] that $X_t^\rho(x)$ converges in probability to $X_t(x)$ as $\rho \downarrow 0$, for any x. Thus the limit Ξ_1 is X_t, and $(X_t^\rho)^{-1}$ converges in law to X_t^{-1}; similarly, the variable $(X_t^\rho)^\star$, which has the same law, converges to $X_t^\star$, so X_t^{-1} and $X_t^\star$ have the same law. $\square$

By applying Lemma 5 and the reduction of Lemma 4, a change of variables shows that

$$\mathbb{E}\int f_1(x) f_2(X_t(x))dx = \mathbb{E}\int f_2(x) f_1(X_t^\star(x))dx.$$

This means that the semigroups of X_t and $X_t^\star$ are in duality with respect to the Lebesgue measure. As a consequence ([7]), if we consider the law of the process (X_t) with marginal law the Lebesgue measure, then the right continuous modification of its time reversal has the law of the process $(X_t^\star)$. Similarly, the processes

$$Z_t(s,x) = (s-t, X_t(x)), \qquad Z_t^\star(s,x) = (s+t, X_t^\star(x))$$

are in duality with respect to the Lebesgue measure on $\mathbb{R}^{d+1}$. For $r > 0$ fixed, let $\sigma = \sigma(s,x)$ and $\sigma^\star = \sigma^\star(s,x)$ be the exit times of $(0,r)\times D$ for the processes Z_t and $Z_t^\star$. Denote by $\overline{Z}_t$ and $\overline{Z}_t^\star$ the corresponding killed processes; they are also in duality with respect to the Lebesgue measure on $(0,r)\times D$.

Let us go back to the solution $h(t,x)$ of the heat equation; it is supposed to be bounded, so by adding a constant, we can also suppose that it is non negative. We are going to use the duality between non negative excessive functions of $\overline{Z}_t$ and excessive measures of $\overline{Z}_t^\star$. The process $h(Z_t)$ is a non negative martingale up to σ, so the killed process $h(\overline{Z}_t)$ is a right continuous supermartingale; this means that the function h is excessive for the process $(\overline{Z}_t)$. Let ν be the measure

$$\nu(dt,dx) = 1_{(0,r)}(t) 1_D(x) h(t,x) dt\, dx. \tag{20}$$

The fact that h is excessive for $(\overline{Z}_t)$ implies that ν is excessive for $(\overline{Z}_t^\star)$ (see XII.71 of [6]); this can be viewed from

$$\begin{aligned}
\int \mathbb{E}[f(\overline{Z}_t^\star(z))]\nu(dz) &= \int \mathbb{E}[f(\overline{Z}_t^\star(z)) h(\overline{Z}_0^\star(z))]dz \\
&= \int \mathbb{E}[f(\overline{Z}_0(z)) h(\overline{Z}_t(z))]dz \\
&\le \int_{(0,r)\times D} f(z) h(z) dz = \int f(z)\nu(dz)
\end{aligned}$$

for f non negative. The process $\overline{Z}_t^\star$ has no non trivial invariant measure (the lifetime is bounded), so ν is purely excessive and is therefore the increasing limit of potentials (XII.38 of [6]). We now want to reduce the study of ν to the study of a potential; to this end, we need the following result.

Lemma 6 *Let $\mathcal{K}$ be a compact subset of $(0, r) \times D$. Then the function a defined by $a(z) = \mathbb{E}[\sigma^\star(z)]$ is bounded below by a positive constant on $\mathcal{K}$.*

Proof. For any $(s, x) \in \mathcal{K}$, we have

$$a(s, x) \geq \theta \mathbb{P}\Big[\sup_{0<t<\theta} |X_t^\star(x) - x| < \delta\Big]$$

where θ and δ are less than the distance between $\mathcal{K}$ and the complement of $(0, r) \times D$. We fix δ and use the Doob-Meyer decomposition $X_t^\star(x) - x = V_t^x + M_t^x$ into a predictable process with finite variation and a martingale, so

$$V_t^x = -\int_0^t b(X_s^\star)ds + \int_0^t \int \Big(\overline{\gamma}(X_s^\star, \lambda) + \gamma_0(X_s^\star)\lambda 1_{\{|\lambda| \leq 1\}}\Big)\mu(d\lambda)ds.$$

Then $\int_0^\theta |dV_t^x|$ is bounded by $\delta/2$ if θ is small enough, so

$$a(s, x) \geq \theta \mathbb{P}\Big[\sup_{0<t<\theta} |M_t^x| < \delta/2\Big] \geq \theta(1 - c' \mathbb{E}|M_\theta^x|^2)$$

for θ small, by applying the Doob inequality. Since the predictable compensator $\langle M^x, M^x \rangle_\theta$ of the quadratic variation of M^x is dominated by θ, we deduce

$$a(s, x) \geq \theta(1 - c''\theta)$$

so the lemma is proved by choosing θ small enough. □

Lemma 7 *Let $\mathcal{K}$ be a compact subset of $(0, r) \times D$. Then the measure ν of (20) coincides on $\mathcal{K}$ with the potential for $(\overline{Z}_t^\star)$ of a finite measure ξ on $\mathcal{K}$.*

Proof. The measure ν is the increasing limit of potentials ν_j of measures ξ_j on $(0, r) \times D$. We have

$$\int f(z)\nu_j(dz) = \mathbb{E} \int_0^\infty \int f(\overline{Z}_t^\star(z))\xi_j(dz)dt.$$

If f is non negative and if $\tau(\mathcal{K})$ is the entrance time in $\mathcal{K}$, then

$$\int f(z)\nu_j(dz) \geq \mathbb{E}\int_{\tau(\mathcal{K})}^{\infty}\int f(\overline{Z}_t^\star(z))\xi_j(dz)dt = \mathbb{E}\int_0^{\infty}\int f(\overline{Z}_t^\star(z))\overline{\xi}_j(dz)dt \tag{21}$$

where $\overline{\xi}_j$ is the law of $\overline{Z}^\star_{\tau(\mathcal{K})}$ when the initial law is ξ_j; in particular, $\overline{\xi}_j$ is supported by $\mathcal{K}$. Thus, if $\overline{\nu}_j$ is the potential of $\overline{\xi}_j$, then $\nu_j \geq \overline{\nu}_j$. If f is supported by $\mathcal{K}$, then the inequality (21) becomes an equality, so $\nu_j = \overline{\nu}_j$ on $\mathcal{K}$. Then

$$\int \nu(dz) \geq \int \overline{\nu}_j(dz) = \int a(z)\overline{\xi}_j(dz) \geq c\,\overline{\xi}_j(\mathcal{K})$$

from Lemma 6. Recall that the problem has been reduced to the case where D is bounded, so ν is finite. Thus $\overline{\xi}_j$ is bounded and has a converging subsequence for the weak topology. Its limit ξ is a finite measure on $\mathcal{K}$; if $\overline{\nu}$ is its potential, then $\overline{\nu}$ is the limit of $\overline{\nu}_j$, and $\nu = \overline{\nu}$ on $\mathcal{K}$. □

We now see that the measure ξ cannot have mass everywhere. We fix x_0 in D and consider an open neighbourhood B_0; let B_n be the set of y such that $(x, y) \in \mathcal{A}_n$ for some $x \in B_0$; they are open sets. The assumption of Theorem 2 saying that D^c is not accessible from x_0 in n jumps implies that B_n is relatively compact in D if B_0 is chosen small enough. Thus we can choose $\mathcal{K}$ of the form

$$\mathcal{K} = [r/3, 2r/3] \times K$$

where K is a compact subset of $\mathbb{R}^d$, K^0 is its interior, and

$$\overline{B}_n \subset K^0 \subset K \subset D.$$

Lemma 8 *Let $\mathcal{K}$ satisfy the above condition. Then the measure ν coincides on $\mathcal{K}$ with the potential for $(\overline{Z}_t^\star)$ of a finite measure ξ on $\mathcal{K}$ satisfying $\xi((r/3, 2r/3) \times B_{n-1}) = 0$.*

Proof. Let f be the indicator of $B'_{n-1} = (r/3, 2r/3) \times B_{n-1}$ and let $\mathcal{K}^0$ be the interior of $\mathcal{K}$; if the measure ξ of Lemma 7 satisfies $\xi(B'_{n-1}) > 0$, then

$$\mathbb{E}\int f(\overline{Z}_0^\star(z))\xi(dz) = \xi(B'_{n-1}) > 0,$$

so, by right lower semicontinuity (because B'_{n-1} is open),

$$\mathbb{E}\int f(\overline{Z}_s^\star(z))\xi(dz) \geq c$$

for s small. Thus

$$\begin{aligned}
\mathbb{E}\int_0^\infty \int f(\overline{Z}^\star_s(z))\xi(dz)ds &\geq \mathbb{E}\int_t^\infty \int f(\overline{Z}^\star_s(z))\xi(dz)ds + c\,t \\
&\geq \mathbb{E}\int_t^\infty \int 1_{\mathcal{K}^0}(\overline{Z}^\star_t(z))f(\overline{Z}^\star_s(z))\xi(dz)ds + c\,t
\end{aligned}$$

for t small. This can be written as

$$\begin{aligned}
\nu(B'_{n-1}) &\geq \int_{\mathcal{K}^0} \mathbb{E}[f(\overline{Z}^\star_t(z))]\nu(dz) + c\,t \\
&= \int \mathbb{E}\Big[f(\overline{Z}^\star_t(z))(h\,1_{\mathcal{K}^0})(\overline{Z}^\star_0(z))\Big]dz + c\,t \\
&= \int \mathbb{E}\Big[f(\overline{Z}_0(z))(h\,1_{\mathcal{K}^0})(\overline{Z}_t(z))\Big]dz + c\,t \\
&\geq \int_{B'_{n-1}} \mathbb{E}[h(\widetilde{Z}_t(z))]dz + c\,t \qquad (22)
\end{aligned}$$

where $\widetilde{Z}_t(z)$ is the process $Z_t(z)$ killed at the first exit time σ_0 of $\mathcal{K}^0$. On the other hand, h is solution of the heat equation, so $h(z)$ is the expectation of $h(Z_{t\wedge\sigma_0})$ for $Z_t = Z_t(z)$, and

$$\mathbb{E}[h(\widetilde{Z}_t)] = \mathbb{E}[h(Z_t)1_{\{t<\sigma_0\}}] = h(z) - \mathbb{E}[h(Z_{\sigma_0})1_{\{\sigma_0\leq t\}}]$$

for $z \in \mathcal{K}^0$. By using this equality in (22), we obtain the second inequality in

$$\int_{B'_{n-1}} \mathbb{P}[\sigma_0(z) \leq t]dz \geq c\int_{B'_{n-1}} \mathbb{E}[h(Z_{\sigma_0})1_{\{\sigma_0\leq t\}}]dz \geq c'\,t.$$

If $\tau_0 = \tau_0(x)$ is the first exit time of K^0 for the process $X_t(x)$, we deduce

$$\int_{B_{n-1}} \mathbb{P}[\tau_0(x) \leq t]dx \geq c\,t.$$

The process $X_t(x)$ cannot jump from $\overline{B}_{n-1}$ into $(K^0)^c$, so the left hand side is $O(t^2)$ from Lemma 1. Our assumption $\xi(B'_{n-1}) > 0$ is therefore false. □

Lemma 9 *The statement of Theorem 2 holds true for the smoothness with respect to x. In particular, Theorem 1 holds true.*

Proof. We are going to apply the result of Lemma 3 to $X^\star_t$ instead of X_t. Notice that the set $\mathcal{A}^\star_n$ of Definition 2 corresponding to this process is symmetric to $\mathcal{A}_n$, so that

$$\mathcal{A}^\star_n = \Big\{(x,y);\ (y,x) \in \mathcal{A}_n\Big\}.$$

The process $\overline{X}_t^\star$ has a transition density $x \mapsto \overline{p}^\star(t, y, x) = \overline{p}(t, x, y)$; let us extend this function by 0 for $t < 0$. Thus the potential for $\overline{Z}_t^\star$ of a mass at (s, y) has density $(t, x) \mapsto \overline{p}^\star(t - s, y, x)$, and Lemma 7 says that h is equal almost everywhere on $\mathcal{K}$ to

$$h_1(t, x) = \int_{\mathcal{K}} \overline{p}^\star(t - s, y, x)\xi(ds, dy)$$

In the integral, Lemma 8 says that (s, y) is outside $(r/3, 2r/3) \times B_{n-1}$; thus, if t is in a neighbourhood T_r of $r/2$, we can decompose this integral into an integral for which $t - s$ is bounded below, and an integral for which $y \notin B_{n-1}$. If x is in B_0, the process X_t cannot go from x to D^c in n jumps, so $X_t^\star$ cannot go from D^c to x in n jumps; similarly, if $y \notin B_{n-1}$, the process $X_t^\star$ cannot go from y to x in $n - 1$ jumps. Thus we can apply the two parts of Lemma 3 to the two parts of the integral, and deduce that if n is large enough, then h_1 satisfies the required smoothness on $T_r \times B_0$. We still have to prove that h satisfies the same smoothness. One has

$$h(t, x) = \lim_{s \downarrow 0} h(t - s, X_s) = \lim h_1(t - s_k, X_{s_k})$$

almost surely along some sequence $s_k \downarrow 0$; it follows from the smoothness of h_1 with respect to x that

$$h(t, x) = \lim h_1(t - s_k, x),$$

so h inherits the smoothness of h_1 on $T_r \times B_0$. The choice of B_0 does not depend on r, so h is smooth on $(0, \infty) \times B_0$. $\square$

Proof of Theorem 2. We have to study the smoothness of $h(t, x)$ with respect to t. Like previously, we consider a neighbourhood B_0 of x_0, the set B_n of points accessible in n jumps from B_0, and suppose that B_n is relatively compact in D. Lemma 9 says that h is C^j with respect to $x \in B_0$ if n is large enough. Actually, if B_N is relatively compact in D for some $N \geq n$, the method shows that the smoothness holds for $x \in B_{N-n}$. On the other hand, by applying the heat equation, we obtain on $[t_0, \infty) \times B_{N-n}$ the estimate

$$h(t, x) = \mathbb{E}[h(t - (s \wedge \tau), X_{s \wedge \tau})] = \mathbb{E}[h(t - s, X_s)] + O(s^{n+1})$$

as $s \downarrow 0$, because $\mathbb{P}[\tau < s]$ is $O(s^{n+1})$(Lemma 1). From the smoothness of h with respect to x and Ito's formula, we get

$$h(t, x) = h(t - s, x) + Lh(t - s, x)s + o(s),$$

and this estimate is uniform. We deduce that h is differentiable with respect to t and its derivative is Lh. We also notice that $h(t,x)-h(t-s,x)$ is, for s fixed, a solution of the heat equation; by taking the limit in the martingale property, we prove that $\partial h/\partial t = Lh$ is also a bounded solution of the heat equation on $[t_0,\infty)\times B_{N-n}$. Let $t_1 > t_0$, let B'_0 be an open neigbourhood of x_0 which is relatively compact in B_0, and let B'_n be the corresponding sets of accessible points. If $N \geq 2n$, we can iterate the procedure and prove that $\partial h/\partial t$ is C^j with respect to x and differentiable with respect to t on $[t_1,\infty)\times B'_{N-2n}$. Thus we can obtain any order of smoothness. □

5 Extensions

In this section, we derive two extensions of the main result. First, we consider a wider class of operators L. In the continuous case, it is well known that one can consider the heat equation with a potential (a term of order 0 in L); this is also possible here. Moreover, contrary to the continuous case, the class of processes X_t is not stable with respect to Girsanov transforms, so applying Girsanov transforms enables to obtain a richer class of operators L. The new class consists of operators

$$\begin{aligned} Lf(x) &= f'(x)b(x) + g(x)f(x) \\ &\quad + \int \Big(f(x+\gamma(x,\lambda)) - f(x) - f'(x)\gamma(x,\lambda) \Big) \psi(x,\lambda)\mu(d\lambda) \end{aligned} \tag{23}$$

when $\beta \geq 1$, and

$$Lf(x) = g(x)f(x) + \int \Big(f(x+\gamma(x,\lambda)) - f(x) \Big) \psi(x,\lambda)\mu(d\lambda) \tag{24}$$

when $\beta < 1$ ((24) is obtained from (23) by taking a particular b). The measure μ and the coefficients b and γ satisfy the assumptions of Section 2, and the new coefficients satisfy the following conditions.

Assumptions on g and ψ. We suppose that

$$\psi(x,\lambda) = 1 + \psi_0(x)\lambda + O(|\lambda|^\alpha) \tag{25}$$

for some $\alpha > 1 \vee \beta$, as $\lambda \to 0$, uniformly in x, that the coefficients g and ψ_0 are C_b^∞, that ψ is C_b^∞ with respect to x uniformly in (x,λ), and that the relation (25) also holds for the derivatives with respect to x. We also suppose that ψ is bounded below by a positive constant.

We now give the probabilistic interpretation of the semigroup of L. To this end, we write the operator L in the form

$$Lf(x) = L_0 f(x) + g(x)f(x) + \int \Big(f(x+\gamma(x,\lambda)) - f(x) \Big) \psi_1(x,\lambda)\mu(d\lambda)$$

with

$$\begin{aligned} L_0 f(x) &= f'(x)b_0(x) \\ &\quad + \int \Big(f(x+\gamma(x,\lambda)) - f(x) - f'(x)\gamma_0(x)\lambda 1_{\{|\lambda|\le 1\}} \Big) \mu(d\lambda), \end{aligned}$$

$$b_0(x) = b(x) + \int (\gamma_0(x)\lambda 1_{\{|\lambda|\le 1\}} - \gamma(x,\lambda)\psi(x,\lambda))\mu(d\lambda),$$

$$\psi_1(x,\lambda) = \psi(x,\lambda) - 1.$$

The operator L_0 is the generator of the process X_t solution of

$$dX_t = b_0(X_t)dt + \gamma(X_{t-}, d\Lambda_t)$$

to which we can apply previous results.

Lemma 10 *The semigroup generated by L can be expressed as*

$$P_t^L f(x) = \mathbb{E}_x[\Gamma_t f(X_t)],$$

where the process $\Gamma_t = \Gamma_t(x)$ is solution of

$$d\Gamma_t = \Gamma_{t-}\Big(g_1(X_t)dt + \psi_1(X_{t-}, d\Lambda_t) \Big), \qquad \Gamma_0 = 1 \tag{26}$$

with

$$g_1(x) = g(x) - \int (\psi_1(x,\lambda) - \psi_0(x)\lambda 1_{\{|\lambda|\le 1\}})\mu(d\lambda).$$

Proof. We are looking for the predictable finite variation part in the Doob-Meyer decomposition of $f(X_t)\Gamma_t$. We use the Ito formula

$$f(X_t)\Gamma_t = f(x) + \int_0^t f(X_{s-})d\Gamma_s + \int_0^t \Gamma_{s-} df(X_s) + \sum_{s\le t} \Delta\Gamma_s \, \Delta f(X_s),$$

where

$$\Delta\Gamma_s\,\Delta f(X_s) = \Gamma_{s-}\psi_1(X_{s-},\Delta\Lambda_s)\Big(f(X_{s-}+\gamma(X_{s-},\Delta\Lambda_s)) - f(X_{s-})\Big).$$

The predictable finite variation parts of the processes $f(X_t)$ and Γ_t are respectively $\int_0^t L_0 f(X_s)ds$ and $\int_0^t \Gamma_s g(X_s)ds$. Thus the predictable finite variation part of $f(X_t)\Gamma_t$ is

$$\begin{aligned} & \int_0^t \Gamma_s L_0 f(X_s)ds + \int_0^t \Gamma_s g(X_s) f(X_s)ds \\ & + \int_0^t \Gamma_s \int \psi_1(X_s,\lambda)(f(X_s+\gamma(X_s,\lambda)) - f(X_s))\mu(d\lambda) \\ = & \int_0^t \Gamma_s Lf(X_s)ds. \end{aligned}$$

We can verify that the local martingale is a martingale for $f \in C_b^2$, so

$$\mathbb{E}[f(X_t)\Gamma_t] = f(x) + \int_0^t \mathbb{E}[Lf(X_s)\Gamma_s]ds.$$

This proves the lemma. $\square$

From Lemma 10, we can say that the function $h(t,x)$ is a solution of the heat equation for the operator L on $\mathbb{R}_+ \times D$ if for any $r > 0$, the process $h(r-t, X_t)\Gamma_t$, $0 \le t \le r$, is a local martingale up to the exit time of D.

Corollary 1 *Under the above assumptions, the result of Theorem 2 holds true for the operator L of (23) or (24).*

Proof. Let H_t be an independent truncated β-stable process as in the proof of Lemma 4, choose η so that $\exp(H_t - \eta t)$ is a martingale, and define

$$V_t(x,v) = v + H_t - \eta t + \log\Gamma_t(x).$$

Then from the equation (26) of Γ_t,

$$V_t = v + H_t - \eta t + \int_0^t g_1(X_s)ds + \int_0^t \log\psi(X_{s-}, d\Lambda_s).$$

The process $\widetilde{X}_t(x,v) = (X_t(x), V_t(x,v))$ is the solution of an equation driven by the Lévy process (Λ_t, H_t); moreover, it satisfies the previous assumptions. If $\overline{h}(t,x,v) = h(t,x)e^v$, then

$$\overline{h}(r-t, X_t, V_t) = h(r-t, X_t)\Gamma_t \exp(v + H_t - \eta t).$$

We deduce that the function $\overline{h}$ is a solution of the heat equation for the generator of $\widetilde{X}_t$; we can apply Theorem 2, and obtain the smoothness of $\overline{h}$, and therefore of h. $\square$

In the second extension, we consider locally bounded solutions of

$$(L - \partial/\partial t)h = \ell \tag{27}$$

on $[0, r] \times D$ for smooth functions ℓ and the operator L of (23) or (24). This means that the process

$$M_t^h(s,x) = h(s-t, X_t(x))\Gamma_t(x) - \int_0^t \ell(s-u, X_u(x))\Gamma_u(x)du, \quad 0 \le t \le s$$

is for any (s,x) a local martingale up to the first exit time of D for the process $X_t(x)$.

Corollary 2 *If ℓ is a C^∞ function, the smoothness result of Theorem 2 holds for solutions of (27).*

Proof. Since the problem is linear, solutions of (27) can be deduced from solutions of the homogenous equation (the heat equation), and from one particular solution of the non homogenous PDE; such a particular solution is given by the resolvent applied to ℓ. The probabilistic interpretation of this method can be described as follows. As in Corollary 1, we reduce the problem to the case where L is the operator of Section 2 (so that $\Gamma_t = 1$); we localize the problem and suppose that ℓ is C_b^∞. If $\mathcal{F}_t$ is the filtration of Λ_t, we define

$$\overline{M}_t^h(s,x) = M_t^h(s,x) + \mathbb{E}\Big[\int_0^s \ell(s-u, X_u(x))du \;\Big|\; \mathcal{F}_t\Big].$$

The martingale properties for M_t^h and $\overline{M}_t^h$ are equivalent. On the other hand, from the smoothness $x \mapsto X_t(x)$, the function

$$\phi(s,x) = \mathbb{E}\int_0^s \ell(s-t, X_t(x))dt$$

can be proved to be smooth, and one has

$$\overline{M}_t^h(s,x) = (h+\phi)(s-t, X_t(x)).$$

This is a martingale, so $h+\phi$ is a solution of the heat equation and Theorem 2 implies that $h+\phi$ is smooth; thus h is smooth. $\square$

6 The case of processes with smooth jumps

Up to now, we have proved the C^∞ smoothness of h only at points from which the process cannot quit D with a finite number of jumps. Here, we prove that this condition can be removed under additional smoothness assumptions on μ and γ, but only in the framework of Theorem 1 (a counterexample concerning the heat equation of Theorem 2 has been given in (13)). We will denote by γ'_λ the Jacobian matrix of $\gamma(x,\lambda)$ with respect to λ; in particular

$$\gamma_0(x) = \gamma'_\lambda(x,0).$$

Theorem 3 *Consider the operator L with the conditions on μ and the coefficients as in Section 2 or 5; suppose moreover that μ has a bounded support and has a density which is C^∞ on $\mathbb{R}^m \setminus \{0\}$, that $\gamma(x,\lambda)$ is C^∞ with respect to $(x,\lambda) \in \mathbb{R}^d \times \mathbb{R}^m$, and that $\gamma'_\lambda \gamma'^\star_\lambda$ is elliptic. If $\ell(x)$ is C^∞, then any solution $h(x)$ of $Lh = \ell$ in D is C^∞ on D.*

Example. This theorem can for instance be applied when Λ_t is the rotation-invariant β-stable Lévy process with Lévy measure defined in (7). Notice that the assumption of smoothness of μ excludes the example (8), for instance the case where the components of the Lévy process are independent.

Proof. Fix some integer j; we want to prove that h is C^j on D. For $\varepsilon > 0$ fixed (which will be chosen small enough), one writes the Lévy measure as $\mu = \mu_1 + \mu_2$, where μ_2 is supported by $\{|\lambda| \le \varepsilon\}$, and $\mu_1(d\lambda) = \rho_1(\lambda)d\lambda$ for a density ρ_1 which is C^∞ on $\mathbb{R}^m$; this is associated to a decomposition $\Lambda_t = \Lambda_t^1 + \Lambda_t^2$ where Λ_t^i has Lévy measure μ_i, and Λ_t^1 is of pure jump type (it is the sum of its jumps) and has finitely many jumps. By proceeding as in Section 5, we reduce the problem to the case where L is the operator of Section 2 and $\ell = 0$; we can choose D bounded and the function h is a bounded harmonic function. Let τ_1 be the first jump of Λ_t^1, let X_t^2 be the solution of the equation (12) with Λ_t replaced by Λ_t^2 and let τ_2 be the first exit time of D for X_t^2. Then τ_1 is an exponential variable with parameter $c = \mu_1(\mathbb{R}^m)$, the variable $\Delta\Lambda_{\tau_1}$ has law $\mu_1(d\lambda)/\mu_1(\mathbb{R}^m)$ and $(\tau_1, \Delta\Lambda_{\tau_1}, X^2)$ are independent. For $X_t = X_t(x)$, one has

$$\begin{aligned} h(x) &= \mathbb{E}[h(X_{\tau_1 \wedge \tau_2})] \\ &= \mathbb{E}[h(X_{\tau_2}) 1_{\{\tau_2 < \tau_1\}}] + \mathbb{E}[h(X_{\tau_1}) 1_{\{\tau_1 \le \tau_2\}}] \end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}[h(X^2_{\tau_2})1_{\{\tau_2<\tau_1\}}] + \mathbb{E}[h(X^2_{\tau_1} + \gamma(X^2_{\tau_1}, \Delta\Lambda_{\tau_1}))1_{\{\tau_1\le\tau_2\}}] \\
&= \mathbb{E}[h(X^2_{\tau_2})e^{-c\tau_2}] + \int\!\!\int \mathbb{E}[h(X^2_t + \gamma(X^2_t, \lambda))1_{\{t\le\tau_2\}}]e^{-ct}\rho_1(\lambda)dt\,d\lambda \\
&= h_1(x) - h_2(x) + h_3(x)
\end{aligned}
$$

with

$$
h_1(x) = \mathbb{E}[h(X^2_{\tau_2})e^{-c\tau_2}],
$$

$$
h_2(x) = \int\!\!\int \mathbb{E}[h(X^2_t + \gamma(X^2_t, \lambda))1_{\{\tau_2<t\}}]e^{-ct}\rho_1(\lambda)dt\,d\lambda,
$$

$$
h_3(x) = \int\!\!\int \mathbb{E}[h(X^2_t + \gamma(X^2_t, \lambda))]e^{-ct}\rho_1(\lambda)dt\,d\lambda.
$$

The process Λ^2 has jumps bounded by ε, so if one fixes x in D and by choosing ε small enough, X^2 needs an arbitrarily large number of jumps to reach D^c from x; if L_2 denotes the generator of X^2, then h_1 is solution of $L_2h_1 - ch_1 = 0$ in D, so from Corollary 1, the function h_1 is C^j in a neighbourhood of x for ε small enough. The function h_2 can be written in the form

$$
h_2(x) = \mathbb{E}[\phi(\tau_2, X^2_{\tau_2})]
$$

with $\phi(t,x) = h_3(x)\exp(-ct)$; thus $h_2(x) = h_2(0,x)$ with

$$
h_2(t,x) = \mathbb{E}[\phi(\tau_2 - t, X^2_{\tau_2})].
$$

This is a solution of the heat equation $\partial h_2/\partial t = L_2h_2$ on $\mathbb{R}\times D$, so from Theorem 2, h_2 is also C^j in a neighbourhood of x for ε small enough. For h_3, we fix some $T > 0$ and decompose the integral with respect to t into integrals on $[0,T]$ and $[T,\infty)$; we obtain $h_3 = h_4 + h_5$ with

$$
h_4(x) = e^{-cT}\mathbb{E}[h_3(X^2_T)]
$$

and h_5 consists of the integral on $[0,T]$. The smoothness of h_4 is again obtained by applying Theorem 2. For h_5, we write it as $h_5 = \mathcal{T}h$ with

$$
\mathcal{T}f(x) = \int\!\!\int_0^T \mathbb{E}[f(X^2_t + \gamma(X^2_t, \lambda))]e^{-ct}\rho_1(\lambda)dt\,d\lambda.
$$

The stochastic flow $x \mapsto X^2_t(x)$ is smooth, so if f is C^∞ with compact support, then we can differentiate with respect to the ith component of x

and obtain

$$\begin{aligned}
\nabla_i(\mathcal{T}f)(x) &= \int\!\!\int_0^T \mathbb{E}[f'(X_t^2+\gamma(X_t^2,\lambda))(I+\gamma'_x(X_t^2,\lambda))\nabla_i X_t^2] \\
&\qquad e^{-ct}\rho_1(\lambda)d\lambda\,dt \\
&= \int\!\!\int_0^T \mathbb{E}\Big[\frac{\partial}{\partial\lambda}\big[f(X_t^2+\gamma(X_t^2,\lambda))\big]\Big(\gamma_\lambda'^{\star}(\gamma'_\lambda\gamma_\lambda'^{\star})^{-1}\Big)(X_t^2,\lambda) \\
&\qquad (I+\gamma'_x(X_t^2,\lambda))\nabla_i X_t^2\Big]e^{-ct}\rho_1(\lambda)d\lambda\,dt.
\end{aligned}$$

We can apply an integration by parts formula for the integral with respect to λ, and we deduce from our assumptions that $\nabla_i(\mathcal{T}f)$ is dominated by the supremum norm of f. Derivatives of higher order are dealt with similarly, so we obtain that $(\mathcal{T}f)^{(k)}$ is dominated by $\|f\|_\infty$ for any order k. Then a usual approximation procedure enables to prove that $\mathcal{T}f$ is C_b^∞ for any bounded Borel function f; in particular, for $f=h$, we obtain that h_5 is C^∞. Thus we obtain the C^j smoothness of $h=h_1-h_2+h_4+h_5$ in a neighbourhood of x. □

References

[1] J. Bertoin, *Lévy processes*, Cambridge Tracts in Math. **121**, Cambridge Univ. Press, 1996.

[2] K. Bichteler, J.B. Gravereaux and J. Jacod, *Malliavin calculus for processes with jumps*, Stochastics Monographs **2**, Gordon and Breach, 1987.

[3] J.M. Bismut, Calcul des variations stochastique et processus de sauts, *Z. Wahrscheinlichkeitstheorie Verw. Geb.* **63** (1983), 147–235.

[4] P. Cattiaux, Calcul stochastique et opérateurs dégénérés du second ordre, I. Résolvantes, théorème de Hörmander et applications, *Bull. Sc. Math., 2ème série* **114** (1990), 421–462.

[5] Z.Q. Chen, First passage distributions and reflected Dirichlet spaces, preprint.

[6] C. Dellacherie and P.A. Meyer, *Probabilités et potentiel, Chapitres XII à XVI*, Hermann, 1987.

[7] C. Dellacherie, B. Maisonneuve and P.A. Meyer, *Probabilités et potentiel, Chapitres XVII à XXIV*, Hermann, 1992.

[8] T. Fujiwara and H. Kunita, Stochastic differential equations of jump type and Lévy processes in diffeomorphisms group, *J. Math. Kyoto Univ.* **25** (1985), 71–106.

[9] W. Hoh and N. Jacob, On the Dirichlet problem for pseudodifferential operators generating Feller semigroups, *J. Functional Analysis* **137** (1996), 19–48.

[10] N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes*, Kodansha and North Holland, 1981.

[11] Y. Ishikawa, Asymptotic behavior of the transition density for jump type processes in small time, *Tôhoku Math. J.* **46** (1994), 443–456.

[12] S. Kusuoka and D. Stroock, Applications of the Malliavin calculus, part II, *J. Fac. Sci. Univ. Tokyo, Sect. IA, Math.* **32** (1985), 1–76.

[13] R. Léandre, Régularité de processus de sauts dégénérés, *Ann. Inst. Henri Poincaré, Prob. Stat.* **21** (1985), 125–146.

[14] R. Léandre, Densité en temps petit d'un processus de sauts, in: *Séminaire de Probabilités XXI*, Lect. Notes in Math. **1247**, Springer, 1987.

[15] R. Léandre, Régularité de processus de sauts dégénérés (II), *Ann. Inst. Henri Poincaré, Probab. Stat.* **24** (1988), 209–236.

[16] J. Picard, On the existence of smooth densities for jump processes, *Probab. Theory Relat. Fields* **105** (1996), 481–511.

[17] J. Picard, Density in small time for Lévy processes, *ESAIM: Probability and Statistics* **1** (1997), 357–389.

[18] J. Picard, Density in small time at accessible points for jump processes, *Stochastic Proc. Applic.* **67** (1997), 251–279.

[19] D. Stroock and S.R.S. Varadhan, On degenerate elliptic-parabolic operators of second order and their associated diffusions, *Comm. Pure Applied Math.* **XXV** (1972), 651–713.