DENSITY IN SMALL TIME AT ACCESSIBLE POINTS FOR JUMP PROCESSES

Jean Picard
Laboratoire de Mathématiques Appliquées
Université Blaise Pascal et CNRS
63177 Aubière Cedex, France

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

Abstract

We consider a process Y_t which is the solution of a stochastic differential equation driven by a Lévy process with an initial condition $Y_0 = y_0$. We assume conditions under which Y_t has a smooth density for any $t > 0$. We consider a point y that the process can reach with a finite number of jumps from y_0, and prove that, as t tends to 0, the density at this point is of order t^Γ for some $\Gamma = \Gamma(y_0, y)$. Some applications to the potential analysis of the process are given.

Résumé

Nous considérons un processus Y_t qui est solution d'une équation différentielle stochastique conduite par un processus de Lévy avec condition initiale $Y_0 = y_0$. Nous nous plaçons sous des hypothèses permettant d'assurer que Y_t a une densité régulière pour tout $t > 0$. Nous considérons un point y que le processus peut atteindre en un nombre fini de sauts depuis y_0, et démontrons que, lorsque t tend vers 0, la densité en ce point est d'ordre t^Γ pour un $\Gamma = \Gamma(y_0, y)$. Quelques applications à la théorie du potentiel sont données.

1 Introduction

Consider a Markov process Y_t with jumps, without continuous diffusion and with a deterministic initial condition $Y_0 = y_0$; some sets of conditions are known to be sufficient for the existence of a smooth density $y \mapsto p(t, y_0, y)$ for Y_t, see [2, 1, 8, 10]. These results are proved by extending the techniques which were initially worked out for continuous diffusions (Malliavin's calculus); they can also yield estimation of the density in small time $t \to 0$, but they require some regularity on the Lévy measure of the process. Up to now, the more precise results were only obtained when this measure has a smooth density; in [9], the density at points $y \neq y_0$ that the process can reach with one jump from y_0 is studied, and in [7], other points $y \neq y_0$ are studied, but only in the case of processes with finite variation. In [13], we have worked out a method for proving the existence of a smooth density without assuming smoothness of the Lévy measure; the basis of this method is the duality formula of [11, 12]. The aim of this article is to study, under the framework of [13], the behaviour in small time of the density at $y = y_0$, and at points y which can be reached with a finite number of jumps from y_0. In contrast with previous works [9, 7], the Lévy measure is allowed to be singular; for instance, it may have a countable support. We prove that, under some assumptions, the density $p(t, y_0, y)$ is of order t^Γ for some $\Gamma = \Gamma(y_0, y)$, but we do not obtain a precise estimation $p(t, y_0, y) \sim Ct^\Gamma$ as in [9, 7]; actually, this type of estimation probably fails to hold under our assumptions. This study is an improvement and an extension of some of the results which we have proved for real-valued Lévy processes in [14].

Let us suppose that Y_t is a d-dimensional Markov process which is the solution of a stochastic differential equation

$$dY_t = b(Y_t)dt + a(Y_{t-}, dX_t), \qquad Y_0 = y_0$$

driven by a Lévy process X_t (a process with stationary independent increments and with value 0 at time $t = 0$); we do not consider the case where Y_t contains a continuous diffusion part, so we suppose that X_t has no Brownian part. Under the assumptions of [13], we know that Y_t admits a smooth transition density $p(t, y_0, y)$ for $t > 0$. In the particular case $Y_t = y_0 + X_t$ where X_t is a non-degenerate symmetric β-stable process (for $0 < \beta < 2$), then

$$p(t, y_0, y) \leq p(t, y_0, y_0) = C\, t^{-d/\beta}.$$

In Section 3, we derive more general conditions under which these two relations hold approximately as $t \to 0$; more precisely, the scaling condition satisfied by β-stable processes is replaced by an approximate scaling condition of index β, and we prove that $p(t, y_0, y)$ is at most of order $t^{-d/\beta}$ as $t \to 0$, and is exactly of order $t^{-d/\beta}$ on the diagonal $\{y = y_0\}$. In particular, one can deduce from our result that points are polar if $d \geq 2$ or $\beta \leq 1$, and that they are regular for themselves otherwise.

In Section 4, we estimate more precisely the density $p(t, y_0, y)$ for y in the set of points which are Δ-accessible from y_0; this set is defined as follows. Consider the maps $A_0(y_0) = y_0$ and

$$A_{n+1}(y_0, x_1, \ldots, x_{n+1}) = A_n(y_0, x_1, \ldots, x_n) + a(A_n(y_0, x_1, \ldots, x_n), x_{n+1}),$$

let μ be the Lévy measure of X, and let $\mathcal{S}_n = \mathcal{S}_n(y_0)$ be the support of the image of $\mu^{\otimes n}$ by the map $x \mapsto A_n(y_0, x)$; this is the set of points which are accessible with n jumps from y_0, and the set of Δ-accessible points is defined as the union of these $\mathcal{S}_n$. If y is such a point, we prove that under some assumptions, the density $p(t, y_0, y)$ is of order t^Γ as $t \to 0$, where $\Gamma = \Gamma(y_0, y)$ depends on the jumps which drive Y_t from y_0 to y. If $d = 1$ and $\beta > 1$, we deduce an estimate for the hitting times of points.

The behaviour at points y which are not Δ-accessible was considered in [14] in the case of real-valued Lévy processes; these points were called asymptotically Δ-accessible when they are in the closure of $\bigcup_n \mathcal{S}_n$, and Δ-inaccessible otherwise; in particular, the study of Δ-inaccessible points requires large deviations techniques (as for continuous diffusions). However, we do not consider these points here; we only prove that $p(t, y_0, y) = o(t^n)$ for any n.

2 Assumptions

In this section, we list the conditions which will be assumed in our results. The constant numbers will be denoted by C or c, though they may vary from line to line; the dependence on some parameter will be emphasized by an index. The vectors will be identified to 1-column matrices, the transpose will be denoted by a star, and the scalar product by a dot. The identity matrix will be denoted by I.

We first introduce a $\mathbb{R}^m$ valued Lévy process X_t without Brownian part; its law can be characterized by a drift parameter $\chi \in \mathbb{R}^m$ and a measure μ on $\mathbb{R}^m \setminus \{0\}$ satisfying

$$\int (|x|^2 \wedge 1)\mu(dx) < \infty,$$

through the Lévy-Khinchin formula

$$\mathbb{E}[e^{iw.X_t}] = \exp\Big(itw.\chi + t\int (e^{iw.x} - 1 - iw.x1_{\{|x|\leq 1\}})\mu(dx)\Big). \tag{1}$$

The measure μ is called the Lévy measure of the process X_t. We suppose that the tail of μ at 0 satisfies the following approximate scaling and non-degeneracy condition; there exists some $0 < \beta < 2$ and positive c, C such that for any $\rho \leq 1$,

$$c\rho^{2-\beta} I \leq \int_{\{|x|\leq \rho\}} xx^\star \mu(dx) \leq C\rho^{2-\beta} I. \tag{2}$$

Equivalently, this means that

$$\int_{\{|x|\leq\rho\}} (x.u)^2 \mu(dx) \asymp \rho^{2-\beta}$$

uniformly for unit vectors u of $\mathbb{R}^m$, where the symbol $\asymp$ means that the quotient between the two sides is bounded below and above as $\rho \to 0$; this can also be written as

$$\int_{\{|x|\leq\rho\}} (x.w)^2 \mu(dx) \asymp \rho^{2-\beta} |w|^2 \tag{3}$$

uniformly for $w \in \mathbb{R}^m \setminus \{0\}$. This condition implies

$$\int_{\{|x|\leq\rho\}} |x|^2 \mu(dx) \asymp \rho^{2-\beta}, \tag{4}$$

which is equivalent to

$$\mu\{|x| > \rho\} \asymp \rho^{-\beta}. \tag{5}$$

Thus, if $\alpha > 0$,

$$\mathbb{P}\left[\sum_{s\leq t} |\Delta X_s|^\alpha < \infty\right] = 1 \Leftrightarrow \int (|x|^\alpha \wedge 1)\mu(dx) < \infty \Leftrightarrow \alpha > \beta. \tag{6}$$

In particular, X_t has finite variation if and only if $\beta < 1$. Actually, the approximate scaling condition (2) is not sufficient for our purpose in the case $\beta \leq 1$, so we now write the complete assumption about X.

Assumption $\mathcal{H}(X)$. There exists a $0 < \beta < 2$ such that the following conditions hold.

1. If $0 < \beta < 1$, we suppose that

$$\chi = \int_{\{|x|\leq 1\}} x\,\mu(dx) \tag{7}$$

   and

$$\int_{\{|x|\leq \rho\}} (x.u)^2 1_{\{x.u>0\}} \mu(dx) \asymp \rho^{2-\beta} \tag{8}$$

   as $\rho \to 0$ uniformly for unit vectors u.

2. If $\beta = 1$, we assume (2) and

$$\limsup_{\varepsilon \to 0} \Big| \int_{\{\varepsilon<|x|\leq 1\}} x\,\mu(dx) \Big| < \infty. \tag{9}$$

3. If $1 < \beta < 2$, we only assume (2).

Remarks. These conditions will imply that X_t is of order $t^{1/\beta}$ as $t \to 0$, and that the law of $X_t/t^{1/\beta}$ is not asymptotically supported by a strict closed subset of $\mathbb{R}^m$. In the case $\beta < 1$, the process X_t has finite variation; the assumption (7) means that X_t is a pure jump process (it is the sum of its jumps), and (8) is an enforcement of (2). Notice also that regularity is not assumed for μ; for instance, it can be supported by a countable set.

Example. Let X_t be a β-stable process ($0 < \beta < 2$), so that the variables X_1 and $X_t/t^{1/\beta}$ have the same law for any $t > 0$; from the Lévy-Khintchin formula (1), this means that

$$\int g(x)\mu(dx) = \int_0^\infty dr \int_{S^{d-1}} \sigma(dz) r^{-1-\beta} g(rz)$$

for any nonnegative function g and some finite measure σ on the sphere S^{d-1}, and that

$$(1-\beta)\chi = \int z\,\sigma(dz).$$

Then our conditions $\mathcal{H}(X)$ are fulfilled provided that σ is not supported by an hyperplane $\{z;\, z.u = 0\}$ if $\beta \geq 1$, and is not supported by a half space $\{z;\, z.u \geq 0\}$ if $\beta < 1$. By looking at the result stated in Theorem 1 below in the case $Y_t = X_t$, it is not difficult to verify that this condition is actually necessary; if σ is supported by an hyperplane, then X_t is supported by the same hyperplane so has no density, and if σ is supported by a half space with $\beta < 1$ (for instance if X_t is real-valued and non-decreasing), then X_t is supported by the same half space, so the density, when it exists and is continuous, is necessarily 0 at the initial point 0.

Then let $a(y,x)$ and $b(y)$ be $\mathbb{R}^d$ valued Borel functions defined respectively on $\mathbb{R}^d \times \mathbb{R}^m$ and $\mathbb{R}^d$ with

$$a(y,x) = a_1(y)x + a_2(y,x)$$

where a_1 and a_2 are respectively matrix-valued and vector-valued. We make the following assumptions on these coefficients; the derivatives are always taken with respect to y and are indexed by multiindices $k \in \mathbb{N}^d$; the 0th order derivative is by convention the function itself.

Assumption $\mathcal{H}(a,b)$. The coefficients a and b satisfy the following conditions.

1. The functions a_1 and b are C_b^∞, and $a_1 a_1^\star$ is uniformly elliptic.

2. The function $a_2(y,x)$ is infinitely differentiable with respect to y, and there exists some $\alpha > \beta \vee 1$ such that

$$|a_2^{(k)}(y,x)| \leq C_k |x|^\alpha \tag{10}$$

   for any multiindex $k \in \mathbb{N}^d$ and any $|x| \leq 1$.

3. The function a satisfies

$$\left|\det(I + a'(y,x))\right| \geq c, \qquad |a(y,x)| \leq C|x|, \qquad |a^{(k)}(y,x)| \leq C_k$$

   for any y, any $k \in \mathbb{N}^d \setminus \{0\}$, and μ-almost any x; moreover, the function $y \mapsto y + a(y,x)$ is for μ-almost every x a diffeomorphism.

4. In the case $\beta < 1$, we suppose that $b = 0$.

Then for any initial condition y_0, the equation

$$Y_t = y_0 + \int_0^t b(Y_s)ds + \int_0^t a(Y_{s-}, dX_s) \tag{11}$$

with coefficients (a, b) has a unique solution; the stochastic integral is defined as

$$\int_0^t a(Y_{s-}, dX_s) = \lim \sum_i a(Y_{t_i}, X_{t_{i+1}} - X_{t_i})$$

in probability as the step of the subdivision (t_i) of $[0, t]$ tends to 0, or equivalently as the sum

$$\int_0^t a(Y_{s-}, dX_s) = \int_0^t a_1(Y_{s-})dX_s + \sum_{s \le t} a_2(Y_{s-}, \Delta X_s)$$

of an Ito integral and of a sum which converges from (10) and (6) because $\alpha > \beta$. Notice that the equation (11) can be written as

$$dY_t = d\Xi_t(Y_{t-}), \qquad Y_0 = y_0,$$

with

$$\Xi_t(y) = b(y)t + \int_0^t a(y, dX_s).$$

With this notation we are in the framework of [6]; however, in contrast with [6], we have not required the functions $a(y, .)$ to be bounded in $L^p(\mu)$ for $p \ge 2$; they are only bounded when restricted to a bounded subset of $\mathbb{R}^m$; this difficulty can be passed over by noticing that there are only finitely many big jumps on a finite time interval, so that one can study separately the behaviour of the equation at big jumps and on intervals between these big jumps (on which one can apply the results of [6]). One can deduce that there exists a stochastic flow of diffeomorphisms ϕ_{st} such that $Y_t = \phi_{st}(Y_s)$; in particular, the inversibility of the flow follows from the inversibility of $y \mapsto y + a(y, x)$. In the case $\beta < 1$, the process X_t is a pure jump process, so Y_t has finite variation, and the condition $b = 0$ implies that Y_t is also a pure jump process.

Let us finish this section with a notation. If $r > 0$, we decompose the process X_s into the sum of

$$\widetilde{X}_s^r = \sum_{u \le s} \Delta X_u 1_{\{|\Delta X_u| > r\}}, \qquad X_s^r = X_s - \widetilde{X}_s^r. \tag{12}$$

Notice that X^r and $\widetilde{X}^r$ are independent. We also introduce the process Y^r as the solution of

$$Y_s^r = y_0 + \int_0^s b(Y_v^r)dv + \int_0^s a(Y_{v-}^r, dX_v^r), \tag{13}$$

and the pure jump process $\widetilde{Y}^r$ defined by

$$\Delta \widetilde{Y}_s^r = a(\widetilde{Y}_{s-}^r, \Delta \widetilde{X}_s^r), \qquad \widetilde{Y}_0^r = y_0. \tag{14}$$

Actually, we will consider a time interval $[0,t]$, and use the decomposition (12) on $[0,t]$ for an r depending on t, generally $r(t) = t^{1/\beta}$.

3 The density at the initial point

In this section, we prove that the density of Y_t at the initial point y_0 is of order $t^{-d/\beta}$ as $t \to 0$, and that it is also the order of the maximum of the density.

Theorem 1 *Assume $\mathcal{H}(X)$ and $\mathcal{H}(a,b)$. The Markov process Y_t has for any $t > 0$ a transition density*

$$p(t, y_0, y) = \mathbb{P}[Y_t \in dy \mid Y_0 = y_0] \Big/ dy$$

which is infinitely differentiable with respect to y, and which satisfies the following estimates as $t \to 0$.

1. *For t small and any $k \in \mathbb{N}^d$, if $p^{(k)}$ denotes the kth derivative with respect to y (with $p^{(0)} = p$), one has*

$$\sup_{y_0, y} |p^{(k)}(t, y_0, y)| \le C_k t^{-(|k|+d)/\beta} \tag{15}$$

with the notation $|k| = \sum k_j$.

2. *On the diagonal $y = y_0$,*

$$p(t, y_0, y_0) \asymp t^{-d/\beta} \tag{16}$$

as $t \to 0$ uniformly in y_0.

Before entering the proof, let us notice that with this result, one can decide whether the potential density

$$G_\lambda(y_0, y) = \int_0^\infty e^{-\lambda t} p(t, y_0, y) dt \tag{17}$$

is finite (for $\lambda > 0$). If $d = 1$ and $\beta > 1$, then G_λ is bounded, and otherwise $G_\lambda(y_0, y_0) = \infty$. As a consequence, we have the following result.

Corollary 1 *Assume $\mathcal{H}(X)$ and $\mathcal{H}(a, b)$.*

1. *If $d \geq 2$ or $\beta \leq 1$, then points are polar (with probability 1, the process $(Y_t)_{t>0}$ never hits a fixed point y).*
2. *If $d = 1$ and $\beta > 1$, then points are regular for themselves (with probability 1, the process starting at y_0 hits y_0 at arbitrarily small positive times). Moreover, if one considers a local time L_t at y_0 and the subordinator*
$$\tau_s = \inf\{t > 0;\ L_t > s\}$$
(which may jump to infinity), then the Lévy measure $\overline{\mu}$ of τ_s satisfies the approximate scaling condition
$$\overline{\mu}\{z;\ z > \rho\} \asymp \rho^{-\overline{\beta}} \tag{18}$$
as $\rho \to 0$ with index $\overline{\beta} = 1 - 1/\beta$ (this is an estimation on the number of small excursions).

Proof. One has

$$\int p(s, y_2, y_1) p(t, y_1, y_0) dy_1 = p(s + t, y_2, y_0) \leq C_\varepsilon$$

if $s \geq \varepsilon > 0$, so

$$\int G_\lambda(y_1, y_0) \Big(\int_\varepsilon^\infty e^{-\lambda s} p(s, y_2, y_1) ds \Big) dy_1 < \infty$$

for any y_0 and y_2. By letting $\varepsilon \to 0$, we deduce that for any y_0, the set

$$P(y_0) = \{y_1;\ G_\lambda(y_1, y_0) = \infty\}$$

has potential zero. On the other hand, it is classical and easily proved that the map $y_1 \mapsto G_\lambda(y_1, y_0)$ is λ-excessive, so the set $P(y_0)$ is actually polar (see [5, §XIV.85]). In particular, if $d/\beta \geq 1$, one has $G_\lambda(y_0, y_0) = \infty$ from (16), and $\{y_0\}$ is therefore polar. On the other hand, if $d/\beta < 1$ (so $d = 1$), then G_λ is uniformly bounded; let us explain briefly how it implies the regularity of points. From (16), the function $y_1 \mapsto G_1(y_1, y_0)$ is bounded and uniformly excessive, so, as it is explained in Theorem 3.7 of [3], one deduces the existence of a continuous additive functional L_t such that

$$G_1(y_1, y_0) = \mathbb{E}_{y_1} \int_0^\infty e^{-s} dL_s \tag{19}$$

(where the expectation is computed for the initial value $Y_0 = y_1$). This functional can be defined by

$$L_t = \lim_n \int_0^t f_n(Y_s) ds$$

where f_n are approximations of the Dirac distribution at y_0; in particular, the process L_t increases only when Y_t is at point y_0, so L_t is a local time at y_0, and y_0 is therefore regular for itself. Moreover, the functions G_λ when $\lambda > 0$ are related with one another by the resolvent equation, and one can deduce from (19) that

$$G_\lambda(y_1, y_0) = \mathbb{E}_{y_1} \int_0^\infty e^{-\lambda s} dL_s$$

for any $\lambda > 0$. We deduce from this equation written for $y_1 = y_0$ that if $g(\lambda)$ is the Laplace exponent for the subordinator τ_s, then

$$G_\lambda(y_0, y_0) = \mathbb{E} \int_0^\infty \exp(-\lambda \tau_s) ds = \int_0^\infty \exp(-s g(\lambda)) ds = 1/g(\lambda).$$

On the other hand, the estimate (16) implies that $G_\lambda(y_0, y_0) \asymp \lambda^{1/\beta - 1}$ as $\lambda \to \infty$, and therefore $g(\lambda) \asymp \lambda^{1-1/\beta}$. If $\overline{\mu}$ is the Lévy measure of τ_s, we have

$$g(\lambda) = \int_{\{z<\infty\}} (1 - e^{-\lambda z}) \overline{\mu}(dz) + \overline{\mu}\{z = \infty\}.$$

The fact that this integral is of order $\lambda^{1-1/\beta}$ as $\lambda \to \infty$ can be shown to be equivalent to (18). □

Remark. We have stated in Theorem 1 estimates for the density of Y_t; however, by looking at [13] and at the proofs below, similar estimates can be proved for the density of $\psi(Y_t)$, where ψ is an affine function from $\mathbb{R}^d$ onto some space of smaller dimension. In particular, one can prove as in Corollary 1 that smooth submanifolds of $\mathbb{R}^d$ of dimension at most $d-2$ are polar; hypersurfaces H are polar if $\beta \leq 1$, and if $\beta > 1$, any point of H is regular for H.

The proof of Theorem 1 will rely on several lemmas; the idea is to check that the variables $(Y_t - y_0)/t^{1/\beta}$ have densities which are bounded as well as their derivatives (to this end we will apply [13]); this will yield the first part of the theorem (the upper bound (15)). Then we will check that these variables are equivalent as $t \to 0$ to some infinitely divisible variables; a lower bound on the densities of these limit variables will give the lower bound in (16).

The first step is to study the regularity of the law. First notice that the Lévy process X_t is easily proved to have a smooth density; the Lévy-Khinchin formula (1) and the approximate scaling condition (3) indeed imply

$$\begin{aligned}\left|\mathbb{E}[e^{iw.X_t}]\right| &= \exp -t\int(1-\cos(w.x))\mu(dx) \\ &\leq \exp -ct\int_{\{|x|\leq 1/|w|\}}(w.x)^2\mu(dx) \leq \exp -c't|w|^\beta.\end{aligned}$$

The characteristic function is rapidly decreasing, so we deduce from the Fourier inversion formula that X_t has a smooth density $q(t,x)$ which satisfies (15).

Let us now consider the process Y_t; by taking our inspiration from Malliavin's calculus, we have proved in [13] the existence of a smooth density for systems of type (11). However, the assumptions were somewhat different; the diffeomorphism assumption of $\mathcal{H}(a,b)$ was not required, but the functions $a(y,.)$ were assumed to have at most linear growth in $L^p(\mu)$ for $p \geq 2$; this condition is not supposed here (anyway, assuming it does not simplify the following proof). Nevertheless, it is satisfied if the jumps are bounded because in this case $a(y,x)$ is dominated by $|x| \wedge 1$ which is in $L^p(\mu)$; in particular, the solution Y_s^r of (13) has a smooth density $p_r(s,y_0,y)$ for any $s > 0$. Moreover, by looking more precisely at the proofs of [13], one can check the following result.

Lemma 1 *Let $\overline{X}^r$ be a family of Lévy processes indexed by some parameter r, with jumps bounded by 1, with uniformly bounded drift parameter χ_r and with*

Lévy measures satisfying uniformly (2); let (a_r, b_r) be a family of coefficients satisfying uniformly the assumption $\mathcal{H}(a, b)$ (this means that the constant numbers which are involved in this assumption do not depend on r); let $\overline{Y}^r$ be the solution of the equation of type (11), with coefficients (a_r, b_r) and driven by $\overline{X}^r$. Then for any fixed $t > 0$, the variable $\overline{Y}^r_t$ has a C_b^∞ density $\overline{p}_r(t, y_0, y)$, and

$$\sup_{r,y_0,y} |\overline{p}_r^{(k)}(t, y_0, y)| < \infty$$

for any $k \in \mathbb{N}^d$.

We are going to use this result in order to estimate the density $p_r(t, y_0, y)$ of Y_t^r; we first introduce the family $\overline{X}^r$ to which we will apply this lemma.

Lemma 2 *Define*

$$\chi_r = \chi - \int_{\{r<|x|\leq 1\}} x\, \mu(dx).$$

Then $\chi_r = O(r^{1-\beta})$ as $r \to 0$. On the other hand, the family of Lévy processes $\overline{X}^r_t = X^r_{r^\beta t}/r$ indexed by the parameter $0 < r \leq 1$ satisfies the assumptions of Lemma 1; moreover, if $\beta < 1$, the family $\overline{X}^r$ satisfies uniformly (8).

Proof. In the case $\beta > 1$, we deduce from the scaling condition (5) that

$$\int_{\{r<|x|\leq 1\}} x\, \mu(dx) = O(r^{1-\beta}),$$

so χ_r is also $O(r^{1-\beta})$; this property is evident in the case $\beta = 1$ from (9), and if $\beta < 1$,

$$\chi_r = \int_{\{|x|\leq r\}} x\, \mu(dx) = O(r^{1-\beta})$$

by applying (7) and (5), so the first statement of the lemma is proved. For the second statement, one checks that the drift parameter of $\overline{X}^r$ is $r^{\beta-1}\chi_r$, so is bounded; its jumps are bounded by 1, and its Lévy measure is given by

$$\mu_r(A) = r^\beta \mu(r\, A)$$

for $A \subset \{|x| \leq 1\}$, so

$$\int_{\{|x|\leq\rho\}} xx^\star \mu_r(dx) = r^{\beta-2} \int_{\{|x|\leq r\rho\}} xx^\star \mu(dx)$$

for $\rho \leq 1$, and the approximate scaling condition (2) holds uniformly for μ_r. The proof of (8) in the case $\beta < 1$ is similar. □

Lemma 3 *For any fixed $h > 0$ and $k \in \mathbb{N}^d$, the density p_r of the process Y^r (defined by (13)) satisfies*

$$|p_r^{(k)}(r^\beta h, y_0, y)| \leq C_{kh} r^{-(|k|+d)}$$

for $0 < r \leq 1$.

Proof. Consider the family of processes $\overline{Y}^r_h = Y^r_{r^\beta h}/r$. Then $\overline{Y}^r_h$ is solution of

$$d\overline{Y}^r_h = b_r(\overline{Y}^r_h)dh + a_r(\overline{Y}^r_{h-}, d\overline{X}^r_h), \qquad \overline{Y}^r_0 = y_0/r$$

with the Lévy process $\overline{X}^r_h$ of Lemma 2, and the coefficients

$$b_r(y) = r^{\beta-1}b(ry), \qquad a_r(y,x) = r^{-1}a(ry, rx).$$

If $\beta < 1$, then $b_r = 0$, and otherwise, b_r is uniformly C_b^∞. The function a_r admits the decomposition

$$a_r(y,x) = a_1(ry)x + r^{-1}a_2(ry, rx),$$

and it is easy to verify that the assumptions $\mathcal{H}(a,b)$ are satisfied uniformly by the family (a_r, b_r) . Thus we deduce from Lemma 1 that for any fixed $h > 0$, the variables $(\overline{Y}^r_h; r > 0)$ have uniformly C_b^∞ transition densities $\overline{p}_r(h, \overline{y}_0, .)$. In order to conclude, it is now sufficient to notice that

$$p_r(r^\beta h, y_0, y) = r^{-d}\overline{p}_r(h, y_0/r, y/r).$$

□

Lemma 4 *Let ϕ_{st} be the stochastic flow of diffeomorphisms generated by the equation (11). Then for any $q \geq 1$ and any $k \in \mathbb{N}^d \setminus \{0\}$,*

$$\sup_{(t,y)\in[0,1]\times\mathbb{R}^d} \mathbb{E}\sup_{s\leq t}\left|(\phi_{st}^{-1})^{(k)}(y)\right|^q < \infty. \tag{20}$$

Proof. The reversed process $Z_s = Z_s(t,y) = \phi^{-1}_{(t-s)-,t}(y)$ is the solution of the equation

$$Z_s = y - \int_0^s b(Z_u)du + \int_0^s a_0(Z_{u-}, dV_u)$$

where $V_s = X_t - X_{(t-s)-}$, and a_0 is such that $y \mapsto y + a_0(y,x)$ is for μ-almost any x the inverse map of $y \mapsto y + a(y,x)$. This equation has the same type than the forward one, and the Lévy process V has the same law than X. Our assumptions $\mathcal{H}(a,b)$ imply estimates on a_0 and its y-derivatives, so thanks to [6], we deduce that the flow of this reversed equation has derivatives which satisfy (20); actually, the function $a_0(z,.)$ is not bounded in $L^p(\mu)$, but one can verify that only bounds on the derivatives of a_0 are needed for (20). $\square$

Proof of the first part of Theorem 1. Let γ be a positive constant which will be chosen later. Consider the sequence of intervals

$$I_n(t) = [t_n, t_{n+1}] = [t(1-2^{-n+1}), t(1-2^{-n})], \qquad n \geq 1,$$

and let N be the number of the first interval I_n on which X has no jump greater than $r = r(n,t) = \gamma t^{1/\beta} 2^{-n/\beta}$, that is

$$N = \min\Big\{n;\ \sup\{|\Delta X_u|;\ u \in I_n(t)\} \leq r(n,t)\Big\}.$$

Then

$$\begin{aligned}
\mathbb{P}[N > n] &= \prod_{k=1}^{n} \mathbb{P}\Big[\sup\{|\Delta X_u|;\ u \in I_k(t)\} > r(k,t)\Big] \\
&= \prod_{k=1}^{n} \Big(1 - \exp\Big[-|I_k|\mu\{|x| > r(k,t)\}\Big]\Big) \\
&\leq \prod_{k=1}^{n} \Big(|I_k|\mu\{|x| > r(k,t)\}\Big).
\end{aligned}$$

From (5), one has

$$\mu\{|x| > r(k,t)\} \leq C\gamma^{-\beta} 2^k / t,$$

and the length of I_k is $t2^{-k}$, so

$$\mathbb{P}[N > n] \leq C^n \gamma^{-n\beta}. \tag{21}$$

In particular, N is almost surely finite if γ is chosen large enough. Denote the random interval I_N by $[S', S]$ and let $R = r(N,t)$. Let $\psi_n : \mathbb{R}^d \to \mathbb{R}^d$ be the flow from time t_n to time t_{n+1} for the equation driven by the truncated process $X^{r(n,t)}$, so that

$$\mathbb{P}[\psi_n(y_1) \in dy] = p_{r(n,t)}(|I_n|, y_1, y)dy.$$

Since X has no jump greater than $R = r(N,t)$ on I_N, one has $Y_S = \psi_N(Y_{S'})$; for any n, ψ_n is independent from (N, Y_{t_n}), so

$$\begin{aligned} \mathbb{P}[Y_S \in dy \mid N = n, Y_{S'} = y_1] &= \mathbb{P}[\psi_n(y_1) \in dy \mid N = n, Y_{t_n} = y_1] \\ &= \mathbb{P}[\psi_n(y_1) \in dy], \end{aligned}$$

and therefore

$$\mathbb{P}[Y_S \in dy \mid N, Y_{S'}] \;/\; dy = p_R(|I_N|, Y_{S'}, y).$$

The length of I_N is $(R/\gamma)^\beta$, so

$$\mathbb{P}[Y_S \in dy \mid N] \;/\; dy = \mathbb{E}[p_R((R/\gamma)^\beta, Y_{S'}, y) \mid N].$$

This conditional density can be transported from time S to time t with the diffeomorphism ϕ_{St} generated by the equation; conditionally on N, ϕ_{St} is independent from Y_S and $Y_{S'}$, so we can deduce the conditional density of Y_t given N, and by taking the expectation, we get

$$p(t, y_0, y) = \mathbb{E}\Big[\Big|\det(\phi_{St}^{-1})'(y)\Big|\, p_R((R/\gamma)^\beta, Y_{S'}, \phi_{St}^{-1}(y))\Big]. \tag{22}$$

The determinant and p_R can be estimated from Lemmas 4 and 3 (use $h = \gamma^{-\beta}$), so

$$p(t, y_0, y) \le C_\gamma \|R^{-d}\|_2 = C'_\gamma t^{-d/\beta} \|2^{Nd/\beta}\|_2.$$

From (21), this L^2 norm is bounded if γ is chosen large enough, so (15) is proved for $k = 0$. The derivatives of p are dealt with by differentiating (22) and using again Lemmas 4 and 3; the absolute value in (22) does not cause any problem in the differentiation because the sign of the determinant does not depend on y (since ϕ_{St}^{-1} is a diffeomorphism). Thus

$$|p^{(k)}(t, y_0, y)| \le C_{k,\gamma} t^{-(d+|k|)/\beta} \|2^{N(d+|k|)/\beta}\|_2.$$

By choosing γ larger and larger, we check (15) for any k. $\Box$

Remark. Actually, with the above method, one can prove that the density of Y_t^r, for $r > 0$, satisfies

$$|p_r^{(k)}(t, y_0, y)| \le C_k t^{-(d+|k|)/\beta} \tag{23}$$

uniformly for $r \ge t^{1/\beta}$. This will be used in next section.

Any infinitely divisible variable X can be represented as the value at time $t = 1$ of a Lévy process X_t; in particular, this variable can be characterized by a drift parameter χ and a Lévy measure μ, and the characteristic function of X is given by the Lévy-Khintchin formula (1) for $t = 1$. With these definitions, we have the following result.

Lemma 5 *Let X be an $\mathbb{R}^d$-valued infinitely divisible variable, the Lévy measure μ of which satisfies (2), or (8) if $\beta < 1$; then its density is everywhere positive.*

The proof of this lemma is postponed until the appendix.

Lemma 6 *Consider a family of $\mathbb{R}^d$ valued infinitely divisible variables Ξ_t^i indexed by $t > 0$ and some parameter $i \in \mathcal{I}$, and a family Υ_t^i of $\mathbb{R}^d$ valued random variables. Assume the following conditions.*

1. *The variables Υ_t^i have densities which are C_b^1 uniformly in (t, i).*
2. *The variable $\Upsilon_t^i - \Xi_t^i$ converges in L^1 to 0 as $t \to 0$, uniformly in i.*
3. *The drift parameter of Ξ_t^i is uniformly bounded.*
4. *The Lévy measure μ_t^i of Ξ_t^i is supported by a bounded set, the measure $|x|^2 \mu_t^i(dx)$ is uniformly bounded, and μ_t^i satisfies uniformly (2), or (8) if $\beta < 1$.*

Then the density of Υ_t^i is bounded below on any compact set as $t \to 0$, uniformly in i.

Proof. The family of measures $|x|^2 \mu_t^i(dx)$ is relatively compact for the topology of convergence on bounded continuous functions; since the family of drift parameters is also relatively compact in $\mathbb{R}^d$, the family of laws of the variables Ξ_t^i is tight (because weak convergence of the measure $|x|^2 \mu_t^i(dx)$ and convergence of the drift parameter imply weak convergence of the infinitely divisible variable from (1)). The limits of Ξ_t^i are infinitely divisible variables which satisfy (2) or (8), so from Lemma 5, their densities are positive. On the other hand, the variables Υ_t^i have the same weak limits as $t \to 0$ and $i \in \mathcal{I}$; the densities of Υ_t^i are bounded in C_b^1, so they are relatively compact for the topology of uniform convergence on compact subsets, and we have just proved that their limits are positive, so we can conclude. □

Lemma 7 *For $r = r(t) = t^{1/\beta}$, the variable Y_t^r (solution of (13)) satisfies*

$$Y_t^r = y_0 + a_1(y_0)X_t^r + o(t^{1/\beta}) = y_0 + O(t^{1/\beta})$$

in L^2 if $\beta \neq 1$, uniformly in y_0, and

$$Y_t^r = y_0 + b(y_0)t + a_1(y_0)X_t^r + o(t) = y_0 + O(t)$$

if $\beta = 1$.

Proof. The Doob-Meyer decomposition of Y^r is

$$Y_s^r = y_0 + M_s + \int_0^s b(Y_u^r)du + \int_0^s a_1(Y_u^r)du\,\chi_r + \int_0^s \int_{\{|x|\leq r(t)\}} a_2(Y_u^r, x)\mu(dx)du \tag{24}$$

where $\chi_r = \chi_{r(t)}$ was defined in Lemma 2, and where M_s is a martingale. Let us estimate the three last terms of this sum for $s \leq t$. The integral involving b is uniformly of order t, and by assumption it is 0 in the case $\beta < 1$, so it is always at most of order $t^{1/\beta}$. Since χ_r is $O(t^{1/\beta-1})$ from Lemma 2, the term involving a_1 is also $O(t^{1/\beta})$. Finally, for the last integral, one applies the assumption (10) on a_2 and

$$\int_{\{|x|\leq r\}} |x|^\alpha \mu(dx) = O(r^{\alpha-\beta}),$$

so

$$\Big|\int_0^s \int_{\{|x|\leq r(t)\}} a_2(Y_u^r, x)\mu(dx)du\Big| \leq C\,t^{\alpha/\beta} = o(t^{1/\beta}). \tag{25}$$

Thus the predictable finite variation part of Y^r is at most of order $t^{1/\beta}$. For the martingale part, we denote by $[M, M]_t$ the quadratic variation of M_t and obtain

$$\mathbb{E}|M_s|^2 \leq \mathbb{E}[M, M]_t = \mathbb{E}\int_0^t \int_{\{|x|\leq r(t)\}} |a(Y_u^r, x)|^2 \mu(dx)du = O(t^{2/\beta})$$

from (4) since $|a(y, x)| \leq C|x|$. Thus

$$Y_s^r = y_0 + O(t^{1/\beta}) \tag{26}$$

in L^2, uniformly for $s \leq t$. If now we want to find a more precise estimate, we write the decomposition

$$X_s^r = s\chi_r + M_s^X$$

where M_s^X is a martingale, and deduce from (24) that

$$\begin{aligned} Y_s^r &= y_0 + b(y_0)s + a_1(y_0)X_s^r + M_s' + \int_0^s (b(Y_u^r) - b(y_0))du \\ &\quad + \int_0^s (a_1(Y_u^r) - a_1(y_0))du\,\chi_r + \int_0^s \int_{\{|x|\le r(t)\}} a_2(Y_u^r, x)\mu(dx)du \end{aligned}$$

for the martingale $M_s' = M_s - a_1(y_0)M_s^X$. From (25) and (26), the three last terms are negligible with respect to $t^{1/\beta}$ in L^2; since

$$\Delta M_s' = \Delta Y_s^r - a_1(y_0)\Delta X_s^r = a_2(Y_{s-}^r, \Delta X_s^r) + (a_1(Y_{s-}^r) - a_1(y_0))\Delta X_s^r,$$

we have

$$\begin{aligned} \mathbb{E}|M_t'|^2 &= \mathbb{E}[M', M']_t \\ &= \mathbb{E}\int_0^t \int_{\{|x|\le r(t)\}} \Big| a_2(Y_s^r, x) + (a_1(Y_s^r) - a_1(y_0))x \Big|^2 \mu(dx)ds \\ &\le C\int_0^t \int_{\{|x|\le r(t)\}} \Big(|x|^{2\alpha} + t^{2/\beta}|x|^2\Big)\mu(dx) \\ &= O(t^{2\alpha/\beta} + t^{4/\beta}). \end{aligned}$$

Thus M_t' is also $o(t^{1/\beta})$, and we deduce that

$$Y_t^r = y_0 + b(y_0)t + a_1(y_0)X_t^r + o(t^{1/\beta}).$$

If $\beta < 1$, one has $b = 0$, and if $\beta > 1$, t is $o(t^{1/\beta})$, so the estimate of the lemma follows. $\square$

Proof of the second part of Theorem 1. The upper bound was already proved. For $t > 0$ fixed, we are going to consider the truncated process X^r of (12) with $r = r(t) = t^{1/\beta}$, and the corresponding solution Y^r of (13). Since

$$t\,\mu\{|x| > r(t)\} \le C,$$

the probability of the event

$$A_t = \Big\{\sup_{s\le t} |\Delta X_s| \le r(t)\Big\}$$

is bounded below by a positive constant, and on this event, one has $X = X^r$ and $Y = Y^r$ on the time interval $[0, t]$. Since Y^r and A_t are independent, one has

$$p(t, y_0, y) \ge \mathbb{P}[A_t]\,\mathbb{P}[Y_t^r \in dy] \Big/ dy \ge c\,\mathbb{P}[Y_t^r \in dy] \Big/ dy.$$

Thus we only have to find a lower bound for the density of Y_t^r at y_0; more precisely, we have to prove that the density of $\Upsilon_t = (Y_t^r - y_0)/t^{1/\beta}$ is asymptotically bounded below at 0. Recall that the infinitely divisible variable $\overline{X}_1^r = X_t^r/t^{1/\beta}$ was proved in Lemma 2 to have bounded drift coefficient and to satisfy uniformly the approximate scaling condition (2) or (8); since a_1 is bounded and $a_1 a_1^\star$ is uniformly elliptic, one can check that $a_1(y_0)\overline{X}_1^r$ also satisfies these conditions. Moreover, from Lemma 7, the variable $\Upsilon_t - a_1(y_0)\overline{X}_1^r$ converges to 0 as $t \to 0$ (to $b(y_0)$ in the case $\beta = 1$), and the density of Υ_t has already been proved to be uniformly C_b^1 in (15); thus we can conclude from Lemma 6 (the index i of this lemma is not used here). $\square$

4 The density at accessible points

In this section, we obtain successively in Theorems 2 and 3 a lower and an upper bound for the density at points y which are accessible with a finite number of jumps from y_0. Then we emphasize in Corollary 2 the case where these two bounds are similar. When the points are regular for themselves ($d = 1$ and $\beta > 1$), we deduce in Corollary 3 an estimate on the hitting times of points. Finally, we give examples where one can apply our results.

Let ν be a probability measure on $\mathbb{R}^m$ such that ν and μ are mutually absolutely continuous; we suppose that $d\nu \,/\, d\mu$ is globally bounded above, and is locally bounded below on $\mathbb{R}^m \setminus \{0\}$; for instance, one can choose

$$\nu(dx) = (|x|^2 \wedge 1)\mu(dx) \Big/ \int (|z|^2 \wedge 1)\mu(dz).$$

We let U_n be the Markov chain given by

$$U_{n+1} = a(U_n, \xi_{n+1}), \qquad U_0 = y_0 \tag{27}$$

where (ξ_n) are independent variables with distribution law ν, which are chosen independent of the process X. The support of U_n consists of the points which are accessible with n jumps.

Theorem 2 *Assume $\mathcal{H}(X)$ and $\mathcal{H}(a,b)$, consider the Markov chain U_n, and let $y \neq y_0$ be a fixed point such that*

$$\mathbb{P}[|U_n - y| \le \varepsilon] \ge c\,\varepsilon^\gamma \tag{28}$$

for any $\varepsilon \le 1$ and some $\gamma \ge 0$, $n \ge 1$. Then, for t small,

$$p(t, y_0, y) \ge c\, t^{n+(\gamma-d)/\beta}.$$

For the proof of this theorem, if $1/c_0$ is an upper bound for $d\nu/d\mu$, we use the decomposition of X_s as the sum $\widetilde{X}_s + Z_s$ of two independent Lévy processes with respective Lévy measures $c_0\nu$ and $\mu - c_0\nu$, and such that $\widetilde{X}_s$ is a pure jump process. Let $\widetilde{Y}_s$ be the pure jump process satisfying

$$\Delta\widetilde{Y}_s = a(\widetilde{Y}_{s-}, \Delta\widetilde{X}_s), \qquad \widetilde{Y}_0 = y_0,$$

and let $N = N(t)$ be the number of jumps of $\widetilde{X}$ before time t. For t fixed, the process Z_s can also be decomposed as $Z_s^r + \widetilde{Z}_s^r$, where $\widetilde{Z}_s^r$ consists of the jumps greater than $r = r(t) = t^{1/\beta}$. Let $Y_s^{(r)}$ be the solution of the equation (11) driven by $\widetilde{X} + Z^r$. One proceeds as in the second part of Theorem 1; the processes $\widetilde{X}$, Z^r and $\widetilde{Z}^r$ are independent, the probability that $\widetilde{Z}^r \equiv 0$ on $[0, t]$ is bounded below by a positive constant, and on this event, one has $Y_t = Y_t^{(r)}$; thus it is sufficient to estimate the density of $Y_t^{(r)}$ at y. We are going to work conditionally on $\widetilde{X}$.

Lemma 8 *Use in this lemma the notation $H_t = o(t^{1/\beta})$ in order to say*

$$\mathbb{E}\Big[|H_t|^2 \mid \widetilde{X}\Big]^{1/2} \le \varepsilon_n(t) t^{1/\beta}$$

on the event $\{N = n\}$, where the function $\varepsilon_n(t)$ converges to 0 as $t \to 0$ for any n. Then there exists a matrix-valued function A such that

$$Y_t^{(r)} = \widetilde{Y}_t + \int_0^t A(s, \widetilde{X}) a_1(\widetilde{Y}_s) dZ_s^r + o(t^{1/\beta}) \tag{29}$$

in the case $\beta \neq 1$, and

$$Y_t^{(r)} = \widetilde{Y}_t + \int_0^t A(s, \widetilde{X})(a_1(\widetilde{Y}_s) dZ_s^r + b(\widetilde{Y}_s) ds) + o(t)$$

in the case $\beta = 1$; moreover $A(s, \widetilde{X})$ and its inverse are bounded by some $C(n)$ on $\{N = n\}$.

Proof. We only consider the case $\beta \neq 1$ (the generalization is easy). In order to understand the behaviour of $Y_t^{(r)} - \widetilde{Y}_t$ on $\{N = n\}$, let us consider the small values of n. For $n = 0$, one has $\widetilde{X}_s = 0$, $\widetilde{Y}_s = y_0$ and $Y_s^{(r)}$, $s \leq t$, reduces to the solution of the equation driven by Z^r; we can get estimates similar to those of Lemma 7, and get

$$Y_t^{(r)} = y_0 + a_1(y_0) Z_t^r + o(t^{1/\beta}), \tag{30}$$

so the result (29) holds with $A = I$. Let us now consider the case $n = 1$; denote by T the time of the jump of $\widetilde{X}$ in $[0, t]$; on the intervals where $\widetilde{X}$ does not jump, we can apply (30) to the semigroup of $Y^{(r)}$ conditioned on $\widetilde{X}$ which coincides with the semigroup of the Z^r driven process; in particular, on the interval $[0, T[$, we obtain

$$Y_{T-}^{(r)} = y_0 + a_1(y_0) Z_T^r + o(t^{1/\beta}) = y_0 + O(t^{1/\beta}).$$

At time T, we have

$$\widetilde{Y}_T = y_0 + a(y_0, \Delta \widetilde{X}_T)$$

and

$$Y_T^{(r)} = Y_{T-}^{(r)} + a(Y_{T-}^{(r)}, \Delta \widetilde{X}_T) = y_0 + a_1(y_0) Z_T^r + a(y_0 + a_1(y_0) Z_T^r, \Delta \widetilde{X}_T) + o(t^{1/\beta}),$$

so

$$Y_T^{(r)} = \widetilde{Y}_T + (I + a'(y_0, \Delta \widetilde{X}_T)) a_1(y_0) Z_T^r + o(t^{1/\beta}).$$

In particular, $Y_T^{(r)} - \widetilde{Y}_T$ is of order $t^{1/\beta}$. Moreover, on the interval $[T, t]$, one has, as on $[0, T[$,

$$Y_t^{(r)} = Y_T^{(r)} + a_1(Y_T^{(r)})(Z_t^r - Z_T^r) + o(t^{1/\beta}) = Y_T^{(r)} + a_1(\widetilde{Y}_T)(Z_t^r - Z_T^r) + o(t^{1/\beta})$$

and $\widetilde{Y}_t = \widetilde{Y}_T$, so that

$$Y_t^{(r)} = \widetilde{Y}_t + (I + a'(y_0, \Delta \widetilde{X}_T)) a_1(y_0) Z_T^r + a_1(\widetilde{Y}_T)(Z_t^r - Z_T^r) + o(t^{1/\beta}).$$

Therefore, the result (29) holds with

$$A(s, \widetilde{X}) = (I + a'(y_0, \Delta \widetilde{X}_T)) 1_{\{s \leq T\}} + I\, 1_{\{T < s\}}.$$

More generally, if T_i, $1 \leq i \leq N$ are the times of the jumps of $\widetilde{X}$ on $[0, t]$ with the convention $T_0 = 0$ and $T_{N+1} = t$, one can prove (29) with

$$A(s, \widetilde{X}) = (I + a'(\widetilde{Y}_{T_{N-1}}, \Delta \widetilde{X}_{T_N})) \dots (I + a'(\widetilde{Y}_{T_i}, \Delta \widetilde{X}_{T_{i+1}}))$$

on $\{T_i < s < T_{i+1}\}$, $0 \leq i < N$, and $A(s, \widetilde{X}) = I$ if $i = N$. $\square$

Proof of Theorem 2. Consider the event

$$A_n = \{N = n\} \cap \{T_n \leq t/2\}$$

on which $\widetilde{X}$ has n jumps on the time interval $[0, t/2]$, and no jump on $[t/2, t]$. The probability of A_n is bounded below by some $c\,t^n$. On the other hand, on A_n, the conditional law of $Y_t^{(r)}$ given $\widetilde{X}$ and $Y_{t/2}^{(r)}$ coincides with the transition kernel from time $t/2$ to time t for the equation driven by Z^r; in particular we can apply Lemma 3 to this transition kernel, and after taking the mean with respect to the law of $Y_{t/2}^{(r)}$, we deduce that the conditional density of $Y_t^{(r)}/t^{1/\beta}$ given $\widetilde{X}$ is uniformly C_b^1 on A_n. The variable $(Y_t^{(r)} - \widetilde{Y}_t)/t^{1/\beta}$ satisfies the same property, and from Lemma 8, it is equivalent (conditionally on $\widetilde{X}$) to some infinitely divisible variable satisfying the assumptions of bounded drift and approximate scaling; we deduce from Lemma 6 (the index i of this lemma is here the path of $\widetilde{X}$) that

$$\mathbb{P}[(Y_t^{(r)} - \widetilde{Y}_t)/t^{1/\beta} \in dz \mid \widetilde{X}] \geq c\,dz$$

on A_n and for z in a bounded set. Thus

$$\mathbb{P}[Y_t^{(r)} \in dy \mid \widetilde{X}] \geq ct^{-d/\beta}dy$$

on $\{|y - \widetilde{Y}_t| \leq t^{1/\beta}\} \cap A_n$, and therefore

$$\mathbb{P}[Y_t^{(r)} \in dy] \geq ct^{-d/\beta}\mathbb{P}[|\widetilde{Y}_t - y| \leq t^{1/\beta};\, A_n]dy. \tag{31}$$

On the other hand, the jumps $\Delta\widetilde{X}_{T_i}$ of $\widetilde{X}$ are independent with common law ν, and are independent from the sequence of their times (T_i); since the sequence (ξ_n) involved in the definition of U_n is also chosen independent of (T_i), we deduce that conditionally on A_n, the variables $\widetilde{Y}_t$ and U_n have the same law. Thus

$$\begin{aligned} \mathbb{P}[|\widetilde{Y}_t - y| \leq t^{1/\beta};\, A_n] &= \mathbb{P}[|U_n - y| \leq t^{1/\beta};\, A_n] \\ &= \mathbb{P}[|U_n - y| \leq t^{1/\beta}]\mathbb{P}[A_n] \\ &\geq c\,t^{n+\gamma/\beta}, \end{aligned} \tag{32}$$

and we can conclude from (31) and (32). $\square$

We have proved in Theorem 2 that a lower bound on the density of Y_t at y can be related to a lower bound on the distribution of the Markov chain U_n of (27) near y. For the upper bound, an estimate on the distribution of U_n is not sufficient, but more generally, we need an upper bound on a family of Markov chains V_n which are perturbations of U_n.

Theorem 3 *Assume $\mathcal{H}(X)$ and $\mathcal{H}(a,b)$; let (ξ_n) be a sequence of independent variables with law ν, as described in the beginning of this section; consider the Markov chain*

$$V_{n+1} = (f_{n+1} \circ a)(V_n, \xi_{n+1}), \qquad V_0 = f_0(y_0) \tag{33}$$

for some smooth functions f_n such that

$$\Phi_n = \sup_{k \le n, y} \Big(|f_k(y) - y| + |f_k'(y) - I| \Big) < \infty.$$

Fix some point $y \neq y_0$. Suppose that there exists a non decreasing sequence K_n and a sequence $\gamma_n \in [0, +\infty]$ such that for any n, for any sequence $(f_k)_{0 \le k \le n}$ satisfying $\Phi_n \le K_n$, one has

$$\mathbb{P}[|V_n - y| \le \varepsilon] \le C_n \varepsilon^{\gamma_n} \tag{34}$$

for $\varepsilon > 0$ if $\gamma_n < \infty$, and the probability is 0 for ε small enough if $\gamma_n = \infty$. Let

$$\Gamma = \min_n (n + (\gamma_n - d)/\beta). \tag{35}$$

If $\Gamma < \infty$, then $p(t, y_0, y) = O(t^\Gamma)$ as $t \to 0$, and if $\Gamma = \infty$, then $p(t, y_0, y) = o(t^n)$ for any n.

Notice that if y is not Δ-accessible, then y is not in the support of U_n, and by choosing the size of perturbations K_n small enough, it is not in the support of V_n; thus one can take $\gamma_n = \infty$ for any n, and therefore $\Gamma = \infty$.

We decompose the proof of this theorem into some lemmas; it is sufficient to prove it when γ_n is finite (by letting $\gamma_n \to \infty$ when it is infinite), so we make this assumption in the proof.

Lemma 9 *Consider a family of d-dimensional variables H admitting uniformly C_b^1 densities p. Then for any $q \ge 1$,*

$$p(y) \le C_q \Big(\frac{\mathbb{E}[|H|^q] + 1}{(1 + |y|)^q} \Big)^{1/(d+1)}. \tag{36}$$

Proof. Let p'_0 be an upper bound for the derivative of p. Then

$$p(z) \geq p(y) - p'_0|z-y| \geq p(y)/2$$

for $|z-y| \leq p(y)/(2p'_0)$, so

$$\begin{aligned}
\mathbb{E}[|H|^q] &\geq \frac{p(y)}{2}\int_{\{|z-y|\leq p(y)/(2p'_0)\}} |z|^q dz \\
&= c\,p(y)^{d+1}\int_{\{|z-y|\leq p(y)/(2p'_0)\}} |z|^q dz \Big/ \int_{\{|z-y|\leq p(y)/(2p'_0)\}} dz \\
&\geq c\,p(y)^{d+1}|y|^q.
\end{aligned}$$

We deduce

$$p(y) \leq C(\mathbb{E}[|H|^q] \;/\; |y|^q)^{1/(d+1)} \tag{37}$$

so the estimation (36) is proved for $|y| \geq 1/2$. For $|y| < 1/2$, we can choose a unit vector e, and since $p(y)$ is the density of $H + e$ at $y + e$, we obtain (37) with (H, y) replaced by $(H + e, y + e)$ in the right-hand side; we again deduce (36). □

Lemma 10 *Consider the pure jump processes $\widetilde{X}^r_s$ and $\widetilde{Y}^r_s$ of (12) and (14) for $r = r(t) = t^{1/\beta}$, and let $N = N(t)$ be the number of jumps of $\widetilde{X}^r$ before time t. Then, for any $\lambda > 0$, one has*

$$\mathbb{P}[|\widetilde{Y}^r_t - y| \leq \lambda t^{1/\beta} \mid N] \leq C(1 + \lambda^L + N^L)t^{\Gamma+d/\beta}$$

for some $L > 0$.

Proof. For $L \geq 4(\Gamma + d/\beta)$, the statement of the lemma holds on the event $\{N > t^{-1/4}\}$, so it is sufficient to prove it on $\{N \leq t^{-1/4}\}$. Conditionally on N, on $\{N > 0\}$, the number of jumps of size greater than $t^{1/(2\beta)}$ is a binomial variable with parameters N and

$$\kappa(t) = \frac{\mu\{x;\ |x| > t^{1/(2\beta)}\}}{\mu\{x;\ |x| > t^{1/\beta}\}} = O(\sqrt{t}).$$

For an integer ℓ fixed, consider the event

$$A = \Big\{\sum_{s\leq t} 1_{\{|\Delta X_s|>t^{1/(2\beta)}\}} < \ell\Big\}.$$

Then, on the event $\{\ell \leq N \leq t^{-1/4}\}$, one has

$$\mathbb{P}[A^c \mid N] = \sum_{j=\ell}^{N} \binom{N}{j} \kappa(t)^j (1-\kappa(t))^{N-j} \leq \sum_{j=\ell}^{N} N^j \kappa(t)^j = O(t^{\ell/4}) = O(t^{\Gamma+d/\beta}) \tag{38}$$

if we choose $\ell \geq 4(\Gamma + d/\beta)$, so it is sufficient to work on the event A. Now, in (12), consider both processes $\widetilde{X}^r$ for $r = r(t)$, and $\widetilde{X}^\rho$ for some $\rho > 0$ fixed. Let $J = J(t)$ be the number of jumps of $\widetilde{X}^\rho$ before time t, and let $(S_j)_{1 \leq j \leq J}$ the times of these jumps; put $S_0 = 0$ and $S_{J+1} = t$. For each j and y_1, consider the pure jump process

$$\Delta Z_s^{j,y_1} = a(Z_{s-}^{j,y_1}, \Delta \widetilde{X}_s^r - \Delta \widetilde{X}_s^\rho), \qquad Z_{S_j}^{j,y_1} = y_1 \tag{39}$$

for $s \geq S_j$, and let $f_j(y_1)$ be the value of this process at time S_{j+1}, so that

$$\widetilde{Y}^r_{S_{j+1}-} = f_j(\widetilde{Y}^r_{S_j}).$$

One can construct from (33) a chain V_j with these perturbations f_j, and a chain $\overline{V}_j$ with these functions f_j but with variables $(\overline{\xi}_n)$ with law $\mu(dx)$ conditioned on $\{|x| > \rho\}$; conditionally on $(\widetilde{X}^r - \widetilde{X}^\rho, (S_j))$, the variables $\widetilde{Y}_t^r$ and $\overline{V}_J$ have the same law, and the law of $\overline{V}_j$ is dominated by the law of V_j, so

$$\mathbb{P}[|\widetilde{Y}_t^r - y| \leq \lambda t^{1/\beta} \mid \widetilde{X}^r - \widetilde{X}^\rho; (S_i)] \leq C_j \mathbb{P}[|V_j - y| \leq \lambda t^{1/\beta} \mid \widetilde{X}^r - \widetilde{X}^\rho] \tag{40}$$

on $\{J = j\}$; since $J < \ell$ on A, the dependence of the constant C_j on j is not important. We now verify that if ρ is chosen small enough, then the perturbations f_j satisfy the assumptions of Theorem 3. Notice that on $A \cap \{N \leq t^{-1/4}\}$, one has

$$\begin{aligned}
&\sum_{s \leq t} \left| \Delta \widetilde{X}_s^r - \Delta \widetilde{X}_s^\rho \right| \\
&\quad = \sum_{s \leq t} |\Delta X_s| 1_{\{t^{1/\beta} < |\Delta X_s| \leq t^{1/(2\beta)}\}} + \sum_{s \leq t} |\Delta X_s| 1_{\{t^{1/(2\beta)} < |\Delta X_s| \leq \rho\}} \\
&\quad \leq N t^{1/(2\beta)} + \ell\rho \leq 2\ell\rho
\end{aligned}$$

for t small enough; since $a(y,x)$ and $a'(y,x)$ are of order $|x|$ as $x \to 0$, one deduces from the study of (39) that

$$|f_j(y) - y| \leq C \sum_{S_j < s \leq S_{j+1}} \left| \Delta \widetilde{X}_s^r - \Delta \widetilde{X}_s^\rho \right| \leq 2C\ell\rho$$

and

$$
\begin{aligned}
|f_j'(y) - I| &\leq C \sum_{S_j < s \leq S_{j+1}} \left|\Delta \widetilde{X}_s^r - \Delta \widetilde{X}_s^\rho\right| \prod_{S_j < s \leq S_{j+1}} \left(1 + C\left|\Delta \widetilde{X}_s^r - \Delta \widetilde{X}_s^\rho\right|\right) \\
&\leq 2C\ell\rho \exp(2C\ell\rho).
\end{aligned}
$$

These expressions are less than K_ℓ if ρ is chosen small enough, so we can apply (34) on the event $A \cap \{N \leq t^{-1/4}\}$ to the conditional law of V_j for $j \leq \ell$. Thus

$$
\begin{aligned}
&\mathbb{P}\Big[|\widetilde{Y}_t^r - y| \leq \lambda t^{1/\beta} \;\Big|\; N\Big] \\
&\leq \mathbb{P}[A^c \mid N] + \sum_{j=0}^{\ell-1} \mathbb{E}\Big[1_A 1_{\{J=j\}} \mathbb{P}\Big[|\widetilde{Y}_t^r - y| \leq \lambda t^{1/\beta} \;\Big|\; \widetilde{X}^r - \widetilde{X}^\rho; (S_i)\Big] \;\Big|\; N\Big] \\
&\leq \mathbb{P}[A^c \mid N] + C\sum_{j=0}^{\ell-1} \mathbb{E}\Big[1_{\{J=j\}} \mathbb{P}\Big[|V_j - y| \leq \lambda t^{1/\beta} \;\Big|\; \widetilde{X}^r - \widetilde{X}^\rho\Big] \;\Big|\; N\Big] \\
&\leq C' t^{\Gamma + d/\beta} + C' \sum_{j=0}^{\ell-1} \mathbb{P}[J = j \mid N] \lambda^{\gamma_j} t^{\gamma_j/\beta}
\end{aligned}
$$

where we have used (40) in the second inequality, (34) and (38) in the third one. Conditionally on N, the variable J has a binomial law with parameters N and $\mu\{|x| > \rho\} \;/\; \mu\{|x| > r\} = O(t)$, so the conditional probability of $\{J = j\}$ is dominated by $N^j t^j$; thus

$$
\begin{aligned}
\mathbb{P}\Big[|\widetilde{Y}_t^r - y| \leq \lambda t^{1/\beta} \;\Big|\; N\Big] &\leq C t^{\Gamma + d/\beta} + C \sum_{j=0}^{\ell-1} \lambda^{\gamma_j} N^j t^{j + \gamma_j/\beta} \\
&\leq C t^{\Gamma + d/\beta} \Big(1 + \sum_{j=0}^{\ell-1} \lambda^{\gamma_j} N^j\Big)
\end{aligned}
$$

from the definition (35) of Γ. The lemma follows from this estimation. □

Lemma 11 *With the notation of Lemma 10, for any $k \in \mathbb{N}^d$, the conditional density $\widetilde{p}$ of Y_t given $\widetilde{X}^r$ satisfies*

$$
|\widetilde{p}^{(k)}(y)| \leq C_k t^{-(d+|k|)/\beta} \exp(C_k N).
$$

Proof. Let $T_1, \ldots, T_N$ be the times of jumps of $\widetilde{X}^r$, let $T_0 = 0$, $T_{N+1} = t$; there is at least one time interval $[T_i, T_{i+1}]$ of length at least $t/(N{+}1)$, say $[T_I, T_{I+1}]$;

the conditional transition density from time T_I to time $S = T_I + t/(N+1)$ of Y coincides, on $\{N = n\}$, with the transition density of Y^r from time 0 to $t/(n+1)$. In particular, one deduces from (23) that this transition density and more generally its kth derivatives are dominated by $(t/(n+1))^{-(d+|k|)/\beta}$. By integrating with respect to the conditional law of Y_{T_I}, we deduce that the conditional density $\widetilde{p}_S$ of Y_S given $\widetilde{X}^r$ satisfies

$$|\widetilde{p}_S^{(k)}(y)| \leq C_k(t/(N+1))^{-(d+|k|)/\beta}. \tag{41}$$

On the other hand, this density $\widetilde{p}_S$ can be propagated until time t in order to get the conditional density of $Y_t = \phi_{St}(Y_S)$ given $\widetilde{X}^r$ and the increments of X between S and t; we deduce

$$\widetilde{p}(y) = \mathbb{E}\Big[|\det(\phi_{St}^{-1})'(y)|\widetilde{p}_S(\phi_{St}^{-1}(y)) \;\Big|\; \widetilde{X}^r\Big]. \tag{42}$$

In order to estimate this conditional expectation and its derivatives with respect to y, we decompose the diffeomorphism ϕ_{St}^{-1} by

$$\phi_{St}^{-1} = \phi_{S,T_{I+1}-}^{-1} \circ \phi_{T_{I+1}-,T_{I+1}}^{-1} \circ \ldots \circ \phi_{T_{N-1},T_N-}^{-1} \circ \phi_{T_N-,T_N}^{-1} \circ \phi_{T_N,t}^{-1}. \tag{43}$$

The maps ϕ_{T_j-,T_j}^{-1} are $\widetilde{X}^r$-measurable and have bounded derivatives, and conditionally on $\widetilde{X}^r$, the maps $\phi_{T_j,T_{j+1}-}^{-1}$ are independent and, as in Lemma 4, they satisfy

$$\mathbb{E}\Big[|(\phi_{T_j,T_{j+1}-}^{-1})^{(l)}(y)|^q \;\Big|\; \widetilde{X}^r\Big] \leq C_{l,q}$$

for $l \in \mathbb{N}^d \setminus \{0\}$. Thus, by differentiating (43), we deduce that

$$\mathbb{E}\Big[|(\phi_{St}^{-1})^{(l)}(y)| \;\Big|\; \widetilde{X}\Big] \leq \exp(C_l N),$$

so, from (42) and (41),

$$|\widetilde{p}^{(k)}(y)| \leq C_k(t/(N+1))^{-(d+|k|)/\beta} \exp(C_k N).$$

□

Lemma 12 *With the notation of Lemma 10, for $q \geq 1$, one has*

$$\mathbb{E}\Big[|Y_t - \widetilde{Y}_t^r|^q \;\Big|\; \widetilde{X}^r\Big]^{1/q} \leq C_q t^{1/\beta} \exp(C_q N).$$

Proof. This is done like Lemma 8; the difference is that we need all the moments and not only the second one. On $\{N=0\}$, we have $Y_t = Y_t^r$ and $\widetilde{Y}_t^r = y_0$, so we have to check that $Y_t^r - y_0 = O(t^{1/\beta})$ in L^q; from the Doob-Meyer decomposition (24), the only term which must be studied is M_t; since M_t is dominated by $[M,M]_t^{1/2}$ from the Burkholder-Davis-Gundy inequalities (see for instance [4]), it is sufficient to prove that $[M,M]_t$ is of order $t^{2/\beta}$. One can write the Doob-Meyer decomposition

$$[M,M]_t = \int_0^t \int_{\{|x|\le r\}} |a(Y_u^r, x)|^2 \mu(dx) du + \overline{M}_t \le Ct^{2/\beta} + \overline{M}_t \tag{44}$$

for a martingale $\overline{M}_t$. Since the jumps of M are bounded by some $Ct^{1/\beta}$, one has

$$[\overline{M}, \overline{M}]_t = \sum_{s\le t} |\Delta M_s|^4 \le Ct^{2/\beta}[M,M]_t,$$

so, for $q \ge 2$, by applying again the Burkholder-Davis-Gundy inequalities, we obtain from (44) that

$$\begin{aligned} \Big\| [M,M]_t \Big\|_q &\le Ct^{2/\beta} + \Big\| [\overline{M}, \overline{M}]_t^{1/2} \Big\|_q \\ &\le Ct^{2/\beta} + Ct^{1/\beta} \Big\| [M,M]_t \Big\|_{q/2}^{1/2}. \end{aligned} \tag{45}$$

We know from (44) that $[M,M]_t$ is $O(t^{2/\beta})$ in L^1, so (45) shows that it is also $O(t^{2/\beta})$ in L^2, and more generally, that is $O(t^{2/\beta})$ in L^q for any $q \ge 1$; thus our result is proved on $\{N=0\}$. More generally, for higher values of N, if T_i are the times of jumps of $\widetilde{X}^r$ with $T_0 = 0$ and $T_{N+1} = t$, one checks similarly that

$$\mathbb{E}\Big[|Y_{T_{i+1}-} - Y_{T_i}|^q \;\Big|\; \widetilde{X}^r \Big]^{1/q} \le C_q t^{1/\beta}.$$

One has $\widetilde{Y}^r_{T_{i+1}-} = \widetilde{Y}^r_{T_i}$, so

$$\mathbb{E}\Big[|Y_{T_{i+1}-} - \widetilde{Y}^r_{T_{i+1}-}|^q \;\Big|\; \widetilde{X}^r \Big]^{1/q} \le \mathbb{E}\Big[|Y_{T_i} - \widetilde{Y}^r_{T_i}|^q \;\Big|\; \widetilde{X}^r \Big]^{1/q} + C_q t^{1/\beta}.$$

From the Lipschitz continuity of $y \mapsto y + a(y,x)$, we have

$$|Y_{T_{i+1}} - \widetilde{Y}^r_{T_{i+1}}| \le C|Y_{T_{i+1}-} - \widetilde{Y}^r_{T_{i+1}-}|.$$

We deduce from the two above formulas an induction relation on the conditional moments of $Y_{T_i} - \widetilde{Y}^r_{T_i}$ from which we get

$$\mathbb{E}\Big[|Y_{T_n} - \widetilde{Y}^r_{T_n}|^q \;\Big|\; \widetilde{X}^r \Big]^{1/q} \le C_q t^{1/\beta} \sum_{i=1}^{n-1} C^i.$$

We conclude by taking $n = N + 1$. $\square$

Proof of Theorem 3. Consider the variable

$$H = \exp(CN)(Y_t - \widetilde{Y}_t^r)/t^{1/\beta}$$

where the constant C is chosen large enough, so that the conditional density of H given $\widetilde{X}^r$ is uniformly C_b^1 (this is possible from Lemma 11). The conditional moments of H have from Lemma 12 at most exponential growth with respect to N, so by applying Lemma 9, we obtain that for any q,

$$\mathbb{P}[H \in dh \mid \widetilde{X}^r] \Big/ dh \leq C_q(1 + |h|)^{-q} \exp(C_q N).$$

Thus the conditional density $\widetilde{p}$ of Y_t satisfies (with a modification of C_q)

$$\begin{aligned} \widetilde{p}(y) &\leq C_q t^{-d/\beta} \Big(\exp(C_q N) \Big(1 + t^{-1/\beta} |y - \widetilde{Y}_t^r| \exp(CN)\Big)^{-1} \Big)^q \\ &\leq C_q t^{-d/\beta} \Big(\exp(C_q N) \Big(1 + t^{-1/\beta} |y - \widetilde{Y}_t^r|\Big)^{-1} \Big)^q. \end{aligned}$$

The density $p(t, y_0, y)$ is the expectation of $\widetilde{p}(y)$; we apply the relation

$$\mathbb{E}[Z^q] = q \int_0^\infty u^{q-1} \mathbb{P}[Z \geq u] du$$

valid for positive variables Z and get

$$\begin{aligned} p(t, y_0, y) &\leq qC_q t^{-d/\beta} \int_0^\infty u^{q-1} \mathbb{P}\Big[1 + t^{-1/\beta}|y - \widetilde{Y}_t^r| \leq \exp(C_q N)/u\Big] du \\ &\leq C_q' t^\Gamma \int_0^\infty u^{q-1} \mathbb{E}\Big[\Big(1 + N^L + u^{-L} \exp(LC_q N)\Big) 1_{\{N \geq \log u / C_q\}}\Big] du \end{aligned}$$

from Lemma 10. The variable $N = N(t)$ is a Poisson variable with bounded mean, so it is not difficult to verify that the expectation is uniformly $o(u^{-k})$ for any k as $u \to \infty$, and is uniformly $O(u^{-L})$ as $u \to 0$; thus the integral is bounded if we choose $q > L$. $\square$

Here is an immediate consequence of Theorems 2 and 3.

Corollary 2 *Assume the conditions of Theorem 3 with $\Gamma < \infty$, and suppose that the lower bound (28) on the distribution of U_n holds for n and $\gamma = \gamma_n$ such that $\Gamma = n + (\gamma_n - d)/\beta$. Then $p(t, y_0, y) \asymp t^\Gamma$ as $t \to 0$.*

When the point y is regular for itself (see Corollary 1), we can also deduce an estimate on the hitting time of y.

Corollary 3 *Assume the conditions of Corollary 2, and suppose moreover $d = 1$ and $\beta > 1$. Define*

$$T_y = \inf\{t > 0;\ Y_t = y\}.$$

Then, as $t \to 0$,

$$\mathbb{P}[T_y \le t] \asymp t^{\Gamma+1/\beta}.$$

Proof. We deduce from Corollary 2 that the potential density $G_\lambda(y_0, y)$ defined in (17) is of order $\lambda^{-\Gamma-1}$ as $\lambda \to \infty$; from Theorem 1, $G_\lambda(y, y)$ is of order $\lambda^{1/\beta-1}$. Moreover, since $y_0 \mapsto G_\lambda(y_0, y)$ is λ-excessive, the process $e^{-\lambda t} G_\lambda(Y_t, y)$ is a supermartingale, and it is actually a martingale up to T_y, so

$$\mathbb{E}\Big[e^{-\lambda T_y}\Big] = \frac{G_\lambda(y_0, y)}{G_\lambda(y, y)} \asymp \lambda^{-\Gamma-1/\beta}.$$

This estimate can be shown to be equivalent to the statement of the corollary. □

Now let us explain how one can apply the above results. When one is given a point y, the first thing to do is to look at the minimal number of jumps n_0 with which y is accessible; then for $k < n_0$, the point y is at a positive distance of the support of U_k, and if the perturbations are small enough, it is also at a positive distance of the support of V_k, so we can take $\gamma_k = \infty$. Then one has to look for the values of γ_k, $k \ge n_0$, satisfying the assumptions of the theorems (more precisely the upper bound (34) for any k, and the lower bound (28) for a k for which the expression in the definition of Γ is minimal). However, since we are only interested in Γ, and since we know that the upper bound is always satisfied with $\gamma_k = 0$, it is not necessary to compute all the values of γ_k; we can stop at the first value n such that

$$n - d/\beta \ge \min_{k<n}(k + (\gamma_k - d)/\beta).$$

Moreover, one may sometimes know a better lower bound $\overline{\gamma} > 0$ on the values of γ_k, and in this case, this stopping criterion can be improved.

Example 1. Suppose that the support of μ is a countable set S such that $S \cap \{|x| > \rho\}$ is finite for any $\rho > 0$. Fix y_0 and y and let n be the minimal number of jumps which drive the process from y_0 to y; we suppose that n is finite (the point y is Δ-accessible). As it has been explained, one can take $\gamma_k = \infty$ for $k < n$. On the other hand we can take $\gamma_k = 0$ for $k \geq n$ in (34), and we obtain $\Gamma = n - d/\beta$. The point y is isolated in the support of U_n, so the lower bound (28) holds for n and $\gamma = 0$. Thus Corollary 2 can be applied. In this case, one can say that the process follows the path from y_0 to y which has the minimal number of jumps.

Example 2. Consider the example

$$\mu(dx) = g(x)|x|^{-d-\beta}dx$$

where g is bounded below and above by positive constant numbers. Suppose also that, if y is in some neighbourhood of y_0, the map $x \mapsto a(y,x)$ is a C^1 diffeomorphism and that its Jacobian and its inverse are bounded. We can choose ν so that it has a bounded density; from our assumption on $a(y_0,.)$, the variable U_1 has a bounded density, and by composing with the transition kernel, we deduce that the variables U_n have bounded densities; moreover, the perturbations $y \mapsto f_n(y)$ of Theorem 3 are also diffeomorphisms (if K_n is small enough), so the variables V_n also have bounded densities. Thus we can take $\gamma_n = d$ for any $n \geq 1$, so that $\Gamma = 1$. It is also clear that the lower bound (28) on the distribution of U_1 holds for $n = 1$, $\gamma = d$, so $p(t, y_0, y) \asymp t$ for any y; the process goes from y_0 to y with only one big jump. This is a particular case of the framework of [9]. If g has compact support, we can consider points y that the process can reach with n jumps and obtain as in [7] that the density is of order t^n.

Example 3. Suppose $\mu = \mu_1 + \mu_2$ where μ_1 and μ_2 satisfy respectively the assumptions of examples 1 and 2; suppose that the diffeomorphism assumption on $x \mapsto a(y,x)$ of previous example is satisfied for any y. Let y be a point that the chain U_n can reach with jumps in the support of μ_1, and let n be the minimal number of these jumps; we suppose $n \geq 2$. Then $\gamma_k = d$ for $1 \leq k < n$, and $\gamma_n = 0$, so $\Gamma = \min(1, n - d/\beta)$. Thus the process goes directly from y_0 to y if $n > 1 + d/\beta$, but it prefers to make n jumps if $n < 1 + d/\beta$. This means that the process does not always minimizes the number of jumps; it prefers singular jumps (coming from the singular part μ_1 of μ) to regular jumps (coming from μ_2).

Example 4. There may be some points where there is a gap between the lower and upper estimates of Theorems 2 and 3, so that one cannot conclude.

Suppose that $y_0 = 0$, $a(y,x) = x$ (so that $Y_t = X_t$) and that μ has a smooth density which is positive except at some point y_1. Then for $y \neq y_1$, one has $p(t,0,y) \asymp t$ as in example 2. On the other hand, let us consider $y = y_1$; the upper estimate (34) is satisfied with $\gamma_1 = d$ and it cannot be improved; however, the lower bound (28) does not hold for this value of γ because the density of U_1 is 0 at y. The problem appears because we have introduced a perturbed Markov chain V_n in the study, and here, the variable V_1 may have more mass near y than U_1. Thus we cannot conclude about the precise behaviour of the density at this point.

5 Appendix: Proof of Lemma 5

One can write $X = X_1$ as the value at time 1 of a Lévy process X_t with density $q(t,x)$. By writing the law of X_1 as the convolution of the law of $X_{1/2}$ with itself, one obtains

$$q(1,x) = \mathbb{E}q(1/2, x - X_{1/2}).$$

Since $q(1/2,z)$ is positive in the neighbourhood of a point z_0, it is sufficient to prove that $x - z_0$ is in the support of $X_{1/2}$. Thus the lemma will be proved if we check that the support of X_t is $\mathbb{R}^d$ for any $t > 0$. We need some notation. For $0 < \rho < 1$, let S_μ^ρ be the set of unit vectors z such that

$$\int_{\{|x|\le 1\}} 1_{\{x.z \ge (1-\rho)|x|\}} |x|^\beta \mu(dx) = \infty$$

and let S_μ be the intersection of these sets as $\rho \to 0$; in other words, if

$$\nu(A) = \int_{\{|x|\le 1\}} 1_A(x/|x|)|x|^\beta \mu(dx), \tag{46}$$

then S_μ is the set of unit vectors z such that the ν-measure of any neighbourhood of z in the unit sphere is infinite. Let Σ_μ be the closed additive semigroup generated by the support of μ, let Σ'_μ be the convex cone with vertex 0 generated by the directions $z \in S_\mu$, and let Σ''_μ be the linear subspace generated by the vectors of S_μ.

Lemma 13 *The closed support of X_t satisfies*

$$\operatorname{supp} X_t + \Sigma'_\mu \subset \operatorname{supp} X_t.$$

Proof. If one adds to X_t an independent variable with law

$$1_{\{|x|\geq\rho\}}\mu(dx) \;/\; \mu(\{|x|\geq\rho\}),$$

one obtains a law which is absolutely continuous with respect to the law of X_t, so by letting $\rho \to 0$, one deduces that

$$\operatorname{supp} X_t + \operatorname{supp} \mu \subset \operatorname{supp} X_t,$$

so

$$\operatorname{supp} X_t + \Sigma_\mu \subset \operatorname{supp} X_t. \tag{47}$$

On the other hand, let z be a point of S_μ^ρ; then there exists a point x in the support of μ satisfying

$$x.z \geq (1-\rho)|x|, \qquad |x| \leq \rho.$$

In particular, the first inequality implies

$$|z - x/|x|| \leq \sqrt{2\rho}.$$

For any $\lambda > 0$, if $[.]$ denotes the integer value, we have

$$\left|\left[\frac{\lambda}{|x|}\right]x - \lambda z\right| \leq \rho + \lambda\sqrt{2\rho}.$$

Thus the distance between λz and Σ_μ is dominated by $\rho + \lambda\sqrt{2\rho}$ for any λ, ρ and any $z \in S_\mu^\rho$. If now z is in S_μ, by letting $\rho \to 0$, we deduce that Σ_μ contains λz for any $\lambda > 0$; since Σ_μ is a semigroup, it contains Σ'_μ. Thus the lemma is proved from (47). □

Lemma 14 *If $\beta \geq 1$, then*

$$\operatorname{supp} X_t + \Sigma''_\mu \subset \operatorname{supp} X_t.$$

Proof. If $\{T_1, \ldots, T_K\}$ is a finite random subset of $[0,t]$, then the law of $X_t - \sum_{j=1}^K \Delta X_{T_j}$ is absolutely continuous with respect to the law of X_t, so

$$\Big(X_t - \sum_{j=1}^K \Delta X_{T_j}\Big) \in \operatorname{supp} X_t \qquad \text{a.s.} \tag{48}$$

Fix z in S_μ^ρ and consider the jumps ΔX_{T_j} of X in the set $\{x.z \geq (1-\rho)|x|\} \cap \{|x| \leq \rho\}$, and ordered in decreasing magnitude, so that $\Delta X_{T_j} \to 0$. From the definition of S_μ^ρ, one has

$$\sum_j |\Delta X_{T_j}|^\beta = \infty.$$

Since $\beta \geq 1$, one has $\sum |\Delta X_{T_j}| = \infty$, and since the jumps are in the cone $\{x.z \geq (1-\rho)|x|\}$, we deduce

$$\lim_k \Big| \sum_{j=1}^k \Delta X_{T_j} \Big| = \infty.$$

For $\lambda > 0$, define

$$K = \inf\Big\{k; \Big|\sum_{j \leq k} \Delta X_{T_j}\Big| \geq \lambda\Big\}.$$

Since the sum of these jumps is in the cone $\{x.z \geq (1-\rho)|x|\}$, we have

$$\Big| z - \sum_{j \leq K} \Delta X_{T_j} \Big/ \Big|\sum_{j \leq K} \Delta X_{T_j}\Big|\Big| \leq \sqrt{2\rho},$$

and by using the definition of K and $|\Delta X_{T_K}| \leq \rho$,

$$\Big|\lambda z - \sum_{j \leq K} \Delta X_{T_j}\Big| \leq \rho + \lambda\sqrt{2\rho}.$$

Thus, from (48), the distance between $X_t - \lambda z$ and the support of X_t is dominated by $\rho + \lambda\sqrt{2\rho}$. If now z is in S_μ, by letting $\rho \to 0$, we deduce that $X_t - \lambda z$ is in $\operatorname{supp} X_t$, and therefore

$$\operatorname{supp} X_t - \lambda z \subset \operatorname{supp} X_t.$$

We have proved in Lemma 13 that

$$\operatorname{supp} X_t + \lambda z \subset \operatorname{supp} X_t,$$

so the lemma follows. □

Proof of Lemma 5 in the case $\beta < 1$. From Lemma 13, it is sufficient to prove that $\Sigma'_\mu = \mathbb{R}^d$. We have

$$\int_{\{|x| \leq \rho\}} 1_{\{0 \leq x.u < \varepsilon|x|\}} (x.u)^2 \mu(dx) \leq \varepsilon^2 \int_{\{|x| \leq \rho\}} |x|^2 \mu(dx) \leq C\varepsilon^2 \rho^{2-\beta},$$

so it follows from the approximate scaling condition (8) that for $\varepsilon > 0$ small enough and for any unit vector u,

$$\int_{\{|x|\le\rho\}} 1_{\{x.u\ge\varepsilon|x|\}}(x.u)^2\mu(dx) \ge c\rho^{2-\beta}.$$

Let us fix such an ε. Then

$$\int_{\{|x|\le\rho\}} 1_{\{x.u\ge\varepsilon|x|\}}|x|^2\mu(dx) \asymp \rho^{2-\beta}$$

which is equivalent to

$$\mu\{x;\ x.u \ge \varepsilon|x|,\ |x| > \rho\} \asymp \rho^{-\beta}, \tag{49}$$

and therefore, with the definition (46),

$$\nu\Big(\{z;\ z.u \ge \varepsilon\}\Big) = \int_{\{|x|\le 1\}} 1_{\{x.u\ge\varepsilon|x|\}}|x|^{\beta}\mu(dx) = \infty.$$

This means that S_μ intersects $\{z;\ z.u \ge \varepsilon\}$ for any unit vector u; thus the convex cone Σ'_μ must be $\mathbb{R}^d$. $\square$

Proof of Lemma 5 in the case $\beta \ge 1$. From Lemma 14, it is sufficient to prove that $\Sigma''_\mu = \mathbb{R}^d$. By proceeding as in (49), there exists a positive ε such that for any unit vector u of $\mathbb{R}^d$,

$$\mu\Big\{x;\ |x.u| \ge \varepsilon|x|,\ |x| > \rho\Big\} \asymp \rho^{-\beta},$$

so the ν-measure of the set $\{z;\ |z.u| \ge \varepsilon\}$ is infinite; thus this set intersects S_μ, so Σ''_μ cannot be orthogonal to u; since u is arbitrary, Σ''_μ is $\mathbb{R}^d$. $\square$

References

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

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

[3] R.M. Blumenthal, *Excursions of Markov processes*, Probability and its Applications, Birkhäuser, 1992.

[4] C. Dellacherie and P.A. Meyer, *Probabilités et potentiel, Chapitres V à VIII*, Hermann, 1980.

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

[6] 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.

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

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

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

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

[11] D. Nualart and J. Vives, A duality formula and some applications, in: *Seminar on Stochastic Analysis, Random Fields and Applications*, Progress in Probab. **36**, Birkhäuser, 1995.

[12] J. Picard, Formules de dualité sur l'espace de Poisson, *Ann. Inst. Henri Poincaré, Probab. Stat.* **32** (1996), 509–548.

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

[14] J. Picard, Density in small time for Lévy processes, preprint.