Support theorem for jump processes

Thomas Simon
Équipe d’Analyse et Probabilités, Université d’Évry-Val d’Essonne Boulevard des Coquibus, 91025 Évry Cedex, France

Abstract

Let $X$ be the solution of an Itô differential equation with jumps over $\mathbb{R}^d$. Under some auxiliary assumptions on the parameters of the equation, we characterize the support of the law of $X$ in the Skorohod space $D$ as the closure of the set of solutions to piecewise ordinary differential equations. This gives an analogue in the Poisson space to the classical Stroock–Varadhan support theorem. © 2000 Published by Elsevier Science B.V.

MSC: 60H10; 60J25

Keywords: Lévy measure; Markov property; Support theorem

1. Introduction

In a celebrated paper, Stroock and Varadhan (1972) characterized for the local uniform topology over $C(\mathbb{R}^+,\mathbb{R}^d)$ the support of the solution to the following Stratonovitch differential equation:

$$X_t = x + \int_0^t \sigma(X_s) \circ dW_s + \int_0^t b(X_s) ds,$$

as the closure of a set of deterministic “skeletons” obtained by substituting an element of the Cameron–Martin space for the Brownian noise. Recently several authors improved this result in different directions: extending the so-called approximative continuity property (Hargé, 1995), getting a support theorem for any $x$-Hölder norm with $x < 1/2$ (Ben-Arous et al., 1994; Millet and Sanz-Solé, 1994), considering stochastic differential equations driven by general semi-martingales (Gyöngy, 1994; Mackevičius, 1986).

However, all the processes involved in the preceding articles have continuous sample paths, and the support of jump processes in the Skorohod space does not seem to have been studied as yet.

In this paper, we consider in $\mathbb{R}^d$ the solution of a stochastic differential equation driven by a Lévy process without Gaussian part. This is also a strong Markov process, whose infinitesimal generator is an homogeneous integro-differential (non-local)
operator. Its sample paths live on the Skorohod space $\mathbb{D}(\mathbb{R}^+, \mathbb{R}^d)$. Under auxiliary assumptions on the coefficients of the equation and the Lévy measure, we show that their support for the (locally) Skorohod topology is the closure of a set of deterministic càdlàg skeletons.

Those skeletons are obtained in the most natural fashion when the Lévy process has finite variations a.s. In that case one can rewrite the equation with just a drift and a non-compensated integral, and to obtain a skeleton one must choose an ordered sequence of jumps of the carrying process and then consider the piecewise ordinary differential equation which is obtained. Such skeletons had already been used by Léandre (1990), to search the points of positive density related to a class of such Markov processes, in the framework of Bismut’s stochastic calculus of variations.

When the Lévy process has infinite variations a.s. the situation is more complicated since one cannot remove the whole compensator into the drift. One obtains similar skeletons but with a different drift which does not contain the small jumps. In some sense, they are analogous to those obtained in the Wiener space (see the remark after the statement of Theorem II).

As for the Wiener space, we must show a double inclusion whose direct part is easy, and the tools for the latter are roughly the same: polygonal approximation of the carrying process and Gronwall’s lemma. The reverse inclusion is not much more difficult in the finite variation case, since we deal with Stieltjes integrals. We just make a small discussion about the Skorohod topology and, except an obvious independence argument, the proof is entirely non-probabilistic.

In the finite variation case, the proof is of course more complicated, since we deal with stochastic integrals. But we appeal neither to a Girsanov transformation nor to approximative continuity properties (though it is possible to define such properties for a large class of Lévy processes as shown in Simon (2000), but then it seems more difficult to handle with them analytically as for Brownian motion). We rather make a repeated use of the strong Markov property for the couple obtained with the carrying process and the solution itself, via a constructive procedure involving the skeleton.

This procedure allows us to reduce the reverse inclusion to a control of the small deviations of the martingale part of the equation, together with the $\alpha$-variation of the Lévy process. This claim, which may be interesting by itself, is also proved via the Markov property and a similar constructive procedure involving the compensation. The latter procedure is quite elementary when the Lévy process is “quasi-symmetric”, that is when the compensation plays a trifling part.

In the general case some technical difficulties appear, probably due to the shape of the skeleton. Indeed when the Lévy measure is not well-distributed, the latter may be dragged away by its drift in a direction which is not recovered by the process. Therefore we must suppose that this measure verifies two conditions (see Assumptions H.1 and H.2 in Section 2). The first one is always true in dimension one, and could probably be improved in higher dimensions, but our method then promises to be very technical. The second one is lifted from other problems in the stochastic analysis of jump processes (see Picard and Savona, 1999 and the references therein) and is only useful in a technical (but crucial) lemma. We also notice that the second assumption is somewhat similar to Assumption H in Gyöngy (1994) and Mackevičius (1986).
However, we give as a final remark a rather general example where the support theorem holds without H.1. In Simon (2000), the support of a Lévy process over \( \mathbb{R}^d \) is characterized in full generality.

The paper is organized as follows: Section 2 presents the framework and the assumptions. Section 3 states the main results and also gives an analytical corollary which is a non-local version of Stroock–Varadhan’s original result. Section 4 contains the proof in the finite variation case, Section 5 treats the infinite variation case.

The main results of this paper were announced in Simon (1999b).

2. Preliminaries

2.1. Framework

Over \( \mathbb{R}^m \) Euclidean, consider a Lévy measure \( \nu \), i.e. a positive Borel measure such that \( \nu(\{0\}) = 0 \) and

\[
\int_{\mathbb{R}^m} \frac{|z|^2}{|z|^2 + 1} \nu(dz) < +\infty.
\]

Let \( \mu \) be the Poisson measure over \( \mathbb{R}^+ \times \mathbb{R}^m \) with intensity measure \( ds \otimes \nu(dz) \), and introduce \( \tilde{\mu} = \mu - ds \otimes \nu \) its compensated measure. For \( x \in \mathbb{R}^m \), consider the following Lévy process \( X \), written in its Lévy–Itô decomposition:

\[
X_t = x + \int_0^t \int_{|z| < 1} z \tilde{\mu}(ds,dz) + \int_0^t \int_{|z| > 1} z \mu(ds,dz)
\]

for every \( t \geq 0 \). Let \( \{\mathcal{F}_t, t \geq 0\} \) be its natural completed filtration and set \( \mathcal{F} = \mathcal{F}_\infty \).

Obviously \( \{\mathcal{F}_t, t \geq 0\} \) is also the completed filtration of \( \{p_t, t \geq 0\} \), the Poisson point process associated with the counting measure \( \mu \) (see, e.g. Section II.3 in Ikeda and Watanabe, 1989). For any \( \eta > 0 \), \( t \geq 0 \) we shall denote

\[
\tilde{X}^\eta_t = \int_0^t \int_{|z| \leq \eta} z \tilde{\mu}(ds,dz).
\]

Over some probability space, the sample paths of \( X \) have infinite variations a.s. if and only if

\[
\int_{|z| < 1} |z| \nu(dz) = +\infty.
\]

We will say that \( X \) is of type I (resp. of type II) when \( X \) has finite variations a.s. (resp. infinite variations a.s.).

Over \( \mathbb{R}^d \) Euclidean, let \( \alpha : \mathbb{R}^d \to \mathbb{R}^d \) and \( b : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d \) be global Lipschitz functions with \( b(x,0) = 0 \) for every \( x \in \mathbb{R}^d \). We suppose that \( \alpha \) is bounded and denote by \( |\alpha| \) its Sup norm. We also suppose that the following holds, for some positive
constant $K$:
\[
\int_{|z|<1} |b(x,z) - b(y,z)|^2 \nu(dz) \leq K|x - y|^2,
\]
for every $x, y \in \mathbb{R}^d$. Then it is well-known (see Theorem IV.9.1 in Ikeda and Watanabe, 1989) that for every $y \in \mathbb{R}^d$, there is a unique (strong) solution to the following stochastic differential equation:
\[
Y_t = y + \int_0^t a(Y_s) \, ds + \int_0^t \int_{|z|>1} b(Y_{r-}, z) \mu(ds, dz) + \int_0^t \int_{0<|z|<1} b(Y_{r-}, z) \tilde{\mu}(ds, dz). \tag{1}
\]
It is also well-known that $Y$ is then an $\mathcal{F}_t$-strong Markov process, whose infinitesimal generator acts on the smooth bounded functions from $\mathbb{R}^d$ into $\mathbb{R}$ in the following way:
\[
Lf(x) = \langle \nabla f(x), a(x) \rangle + \int_{|z|>1} (f(x + b(x,z)) - f(x)) \nu(dz)
+ \int_{0<|z|<1} (f(x + b(x,z)) - f(x) - \langle \nabla f(x), b(x,z) \rangle) \nu(dz).
\]
Besides, $(X,Y)$ is itself an $\mathcal{F}_t$-strong Markov process, since it satisfies
\[
(X_t, Y_t) = (x, y) + \int_0^t (0, a(Y_s)) \, ds + \int_0^t \int_{|z|>1} (z, b(Y_{r-}, z)) \mu(ds, dz)
+ \int_0^t \int_{0<|z|<1} (z, b(Y_{r-}, z)) \tilde{\mu}(ds, dz).
\]
The canonical space $\Omega$ associated with $(X,Y)$ is the canonical space associated with $(\mu, X_0, Y_0) : \Omega = \Omega_0 \times \mathbb{R}^m \times \mathbb{R}^d$, where $\Omega_0$ is the set of integer-valued measures over $\mathbb{R}^+ \times \mathbb{R}^m$ such that $\omega(t) \times \mathbb{R}^m) \leq 1$. We define over this space the following translation operator $\theta_t, t \geq 0, A$ compact in $\mathbb{R}^m - \{0\}$,
\[
\mu \circ \theta_t([0,t] \times A) = \mu([r,t + r] \times A),
\]
\[
X_0 \circ \theta_t = X_t,
Y_0 \circ \theta_t = Y_t.
\]
We also set $\mathbb{P}^{(x,y)}$ (resp. $\mathbb{P}^{(x)}$, $\mathbb{P}^{(y)}$) for the conditional law of $(X,Y)$ knowing $(X_0, Y_0) = (x, y)$ (resp. the conditional law of $X$ knowing $X_0 = x$, the conditional law of $Y$ knowing $Y_0 = y$). It is almost straightforward to see that outside of $\mathbb{P}^{(x,y)}$-negligible sets, $\{\theta_t, t \geq 0\}$ is the usual translation operator associated with the strong Markov process $(X,Y)$:
\[
(X_s, Y_s) \circ \theta_t = (X_{r+s}, Y_{r+s}), \quad \forall t, s \geq 0.
\]
By a monotone class argument, the strong Markov property entails that if $f : \mathbb{R}^d \times \Omega \to \mathbb{R}$ is bounded and $\mathcal{B}(\mathbb{R}^d) \otimes \mathcal{F}$ measurable, then for any $(x,y) \in \mathbb{R}^m \times \mathbb{R}^d$, $T$ $\mathcal{F}_T$-stopping time and $V : \Omega \to \mathbb{R}^d$ $\mathcal{F}_T$-measurable,
\[
\mathbb{E}^{(x,y)}[f(V(\omega), \theta_T)/\mathcal{F}_T] = \mathbb{E}^{(X_T(\omega), Y_T(\omega))}[f(V(\omega), .)] \mathbb{P}^{(x,y)} \quad \text{a.s.}
\]
Fix \((x, y) \in \mathbb{R}^m \times \mathbb{R}^d\). For any real \(\mathcal{F}_t\)-predictable process \(W\) depending on the parameter \(z\) such that for every \(t > 0\)
\[
E^{(x, y)} \left[ \int_0^t \int_{0 < |z| \leq 1} |W(s, z)|^2 \, ds \, d\nu(dz) \right] < +\infty,
\]
the stochastic integral
\[
\int_0^t \int_{0 < |z| \leq 1} W(s, z) \tilde{\mu}(ds, dz)
\]
is a square-integrable \(\mathcal{F}_t\)-martingale whose predictable quadratic variation is given by
\[
\int_0^t \int_{0 < |z| \leq 1} |W(s, z)|^2 \, ds \, d\nu(dz)
\]
(see, e.g. Section II.4 in Ikeda and Watanabe (1989)). In particular, the isometry property of the stochastic integral gives for every \(t > 0\)
\[
E^{(x, y)} \left[ \left( \int_0^t \int_{0 < |z| \leq 1} W(s, z) \tilde{\mu}(ds, dz) \right)^2 \right] = E^{(x, y)} \left[ \int_0^t \int_{0 < |z| \leq 1} |W(s, z)|^2 \, ds \, d\nu(dz) \right].
\]
The above equality will be used repeatedly and referred to as equality I. In the following the constants will be denoted by \(C, K\) and the infinitesimal quantities by \(\varepsilon\), even though they may change from one line to another.

Let \(\mathbb{D} = \mathbb{D}(\mathbb{R}^+, \mathbb{R}^d)\) the space of càdlàg functions from \(\mathbb{R}^+\) to \(\mathbb{R}^d\) endowed with the Skorohod–Prokhorov distance:
\[
d(f, g) = \sum_{n \geq 1} 2^{-n}(1 \wedge d_n(f, g))
\]
where \(d_n\) is defined by
\[
d_n(f, g) = \inf_{A \in \mathcal{A}} \left\{ \sup_{s \leq t} \log \frac{\lambda_s - \lambda_t}{t - s} + \sup_{s > 0} |k_n f(\lambda_s) - k_n g(s)| \right\},
\]
\(A\) designing the set of all continuous functions \(\lambda : \mathbb{R}^+ \to \mathbb{R}^+\) that are strictly increasing, with \(\lambda_0 = 0\) and \(\lambda_t \uparrow +\infty\) as \(t \uparrow +\infty\), and \(k_n\) being given by
\[
k_n(t) = \begin{cases} 
1 & \text{if } t \leq n, \\
n + 1 - t & \text{if } n < t < n + 1, \\
0 & \text{if } t \geq n + 1.
\end{cases}
\]
Such a distance makes \(\mathbb{D}\) into a polish space (see Chapter VI.1 in Jacod and Shiryaev (1987)). We still denote by \(\mathbb{P}^{(y)}\) the law of \(Y\) starting from \(y\) over \(\mathbb{D}\). We shall also make use of the Sup distance and for every càdlàg process \(Z\) we will set \(|Z|_T^2\) for the following quantity:
\[
\sup_{t \leq T} |Z_t|
\]
where \(T\) is possibly a random time. We are interested in the support of \(\mathbb{P}^{(y)}\) over \((\mathbb{D}, d)\). Recall that this is the set of \(\phi \in \mathbb{D}\) such that for every \(n \in \mathbb{N}^*, \varepsilon > 0\),
\[
\mathbb{P}^{(y)}[d_n(Y, \phi) < \varepsilon] > 0.
\]
2.2. Assumptions and notations

In addition to the general assumptions on \(a, b\) and \(v\), we need some auxiliary assumptions which depend on whether \(X\) is of type I or not.

2.2.1. The case of type I

In that case we just need the following:

**Assumption A.** There exists \(K\) such that for all \(x, y \in \mathbb{R}^d\)

\[
\int_{|z| \leq 1} |b(x, z) - b(y, z)| v(dz) \leq K|x - y|.
\]

Notice that this holds when \(b\) is \(C^1\) with respect to \(x\), and \(|b_x(x, z)| \leq K|z|\) for some constant \(K\). Assumption A enables us to rewrite (1) like this:

\[
Y_t = y + \int_0^t \tilde{a}(Y_s) \, ds + \int_0^t \int_{\mathbb{R}^d} b(Y_s, z) \mu(ds, dz),
\]

where for all \(x \in \mathbb{R}^d\)

\[
\tilde{a}(x) = a(x) - \int_{|z| \leq 1} b(x, z) v(dz)
\]

is a bounded and global Lipschitz function. For any \(\eta > 0\), \(t \geq 0\) we shall also denote

\[
\hat{Y}_t^\eta = \int_0^t \int_{|z| \leq \eta} b(Y_s, z) \mu(ds, dz).
\]

2.2.2. The case of type II

The situation is here more complicated and our conditions stronger. First, we need the following assumption on \(b\):

**Assumption B.** For every \((y, z) \in \mathbb{R}^d \times \{|z| \leq 1\}\), \(b(y, z)\) decomposes into

\[
b(y, z) = \tilde{b}(y, z) + b'(y, z),
\]

where \(\tilde{b}\) is a \((d, m)\) matrix-valued bounded and global Lipschitz function, and

\[
|b'(y, z)| \leq K|z|^\alpha,
\]

uniformly in \(y\), for some \(\alpha \in (1, 2]\) such that

\[
\int_{|z| \leq 1} |z|^\alpha v(dz) < \infty.
\]

Besides, there exists \(K\) such that for all \(x, y \in \mathbb{R}^d:\)

\[
\int_{|z| \leq 1} |b'(x, z) - b'(y, z)| v(dz) \leq K|x - y|.
\]

By Taylor’s formula, notice that this assumption holds with \(\alpha = 2\) when \(b\) is \(C^2\) with respect to \(z\), \(b_z\) and \(b_{zz}\) are bounded and uniformly Lipschitz with respect to \(y\).
Assumption B also enables us to rewrite (1), in this manner:

\[ Y_t = y + \int_0^t \tilde{a}(Y_s) \, ds + \int_0^t \int_{\mathbb{R}^m} b'(Y_{s-}, z) \, \mu(ds, dz) \]

\[ + \int_0^t \int_{0 < |z| < 1} \tilde{b}(Y_{s-}) \cdot z \, \tilde{\mu}(ds, dz) + \int_0^t \int_{|z| > 1} \tilde{b}(Y_{s-}) \cdot z \, \mu(ds, dz) \]

where for all \( x \in \mathbb{R}^d \)

\[ \tilde{a}(x) = a(x) - \int_{|z| < 1} b'(x, z) \, v(dz), \]

is a bounded and global Lipschitz function. For any \( \eta > 0, t \geq 0 \) we shall denote

\[ \tilde{Y}^\eta_t = \int_0^t \int_{|z| < \eta} \tilde{b}(Y_{s-}) \cdot z \, \tilde{\mu}(ds, dz) + \int_0^t \int_{|z| < \eta} b'(Y_{s-}, z) \, \mu(ds, dz). \]

We use the same notations \( \tilde{a}, \tilde{Y}^\eta \) for the sake of conciseness, even though the corresponding objects are different according as \( X \) is of type I or not. No confusion is possible since we shall treat the two cases separately.

Setting, for every \( 0 \leq \eta < \rho \),

\[ u_\rho^\eta = \int_{\eta \leq |z| \leq \rho} z \, v(dz), \]

we say that \( X \) is quasi-symmetric if for every \( \rho > 0 \), there exists a sequence \( \{\eta_k\} \) decreasing to 0 such that

\[ |u_{\eta_k}^\rho| \to 0 \]

as \( k \uparrow +\infty \) (we name this property (*)). This means that for every \( \rho \) the compensation involved in the martingale part of \( \tilde{Y}^\rho \) is somehow negligible, and of course this is true when \( X \) is really symmetric.

When \( X \) is not quasi-symmetric, we need to suppose that \( v \) is well-distributed in the following sense:

**Assumption H.1.** For every \( \rho > 0 \) such that (*) does not hold, there exists \( \gamma^\rho > 1 \) and a sequence \( \{\eta_k\} \) decreasing to 0 such that \( \text{Supp} \, v \) intersects the ball \( \{|z| = \gamma^\rho \eta_k\} \) and

\[ x_{\eta_k}^\rho = o(1/|u_{\eta_k}^\rho|), \]

where \( x_{\eta_k}^\rho \) denotes the angle between the direction \( u_{\eta_k}^\rho \) and \( \text{Supp} \, v \) on \( \{|z| = \gamma^\rho \eta\} \).

This assumption is the important one and lies at the heart of our method for proving the claim in Section 5 (see however the last subsection of this paper). Its statement is a bit technical but notice first that it always holds in dimension \( m = 1 \) (with \( x_{\eta_0}^\rho = 0 \)). Besides, it is verified in higher dimensions whenever \( \text{Supp} \, v \) contains a sequence of spheres whose radius tend to 0 (in particular, a whole neighbourhood of 0), or when the intersection of \( \text{Supp} \, v \) with the unit ball coincides with that of a convex cone. Assumption H.1 is roughly speaking a convexity assumption.
For technical reasons (see Lemma 2 in Section 6), we also need to suppose that $v$ satisfies the following non-degeneracy and scaling condition, which is however a bit restrictive:

**Assumption H.2.** There exists $\beta \in [1,2)$ and positive constants $k, K$ such that for any $\rho \leq 1$

$$k \rho^{2-\beta} I \leq \int_{|z| \leq \rho} z z^* v(dz) \leq K \rho^{2-\beta} I.$$  

Besides, if $\beta = 1$, then

$$\limsup_{\eta \to 0} \left| \int_{\eta \leq |z| \leq 1} z v(dz) \right| < \infty.$$  

This means that around the origin, the behaviour of the projections of the Lévy measure on any axis is roughly the same, and analogous to that of a real-valued stable measure. The above inequalities stand for symmetric positive-definite matrices, but they are indeed equivalent to

$$\int_{|z| \leq \rho} |v + z|^2 v(dz) \propto \rho^{2-\beta}$$

uniformly for unit vectors $v \in S^{m-1}$, where $v^*$ denotes the usual scalar product with $v$. In particular,

$$\beta = \inf \left\{ \alpha; \int_{|z| \leq 1} |z|^2 v(dz) < \infty \right\}$$

and the inf is not reached. Notice also that the measure $v$ may be very singular and have a countable support. See Picard and Savona (1999) for further properties and examples concerning Assumption H.2.

3. Results

3.1. The case of type I

Consider $U$, the set of sequences $u = \{u_n, n \geq 1\} = \{(t_n, z_n), n \geq 1\}$, where $\{t_n\}$ is a strictly increasing sequence of $\mathbb{R}^+$ with limit $+\infty$, and $\{z_n\}$ any sequence in the support of $v$. For any $u \in U$, introduce the following piecewise ordinary differential equation:

$$\phi_t = y + \int_0^t \tilde{a}(\phi_s) \, ds + \sum_{t_k \leq t} b(\phi_{t_k-}, z_{t_k}).$$  

(4)

By the Cauchy–Lipschitz Theorem and Assumption A, there exists a unique solution to (4). We shall denote by $\mathcal{S}^y$ the set of solutions to (4), $u$ varying over $U$, and by $\overline{\mathcal{S}}^y$ the closure of $\mathcal{S}^y$ in $(\mathcal{D}, d)$.  

Theorem I. Under Assumption A, 
\[ \text{Supp } \mathbb{P}^{(y)} = \mathcal{F}^y. \]

3.2. The case of type II

We define \( U \) as above and set, for any \( \eta > 0 \),
\[ U_\eta = \{ u \in U/\forall \alpha \geq 1, \ |z_\alpha| > \eta \}. \]

For \( p \in \mathbb{N}^* \), \( u \in U_\eta \), introduce the following piecewise ordinary differential equation:
\[ \phi_t = y + \int_0^t a_p(\phi_s) \, ds + \sum_{s \leq t, z_s = 0} b(\phi_{s-}, z_s) \] \tag{5}

where for every \( x \in \mathbb{R}^d \)
\[ a_p(x) = \tilde{a}(x) - \tilde{b}(x) \left( \int_{|z| \leq 1} z \nu(dz) \right). \]

By the Cauchy–Lipschitz Theorem and Assumption B, there exists a unique solution to (5) as well. Again we shall denote by \( \mathcal{F}^y \) the set of solutions to (5), \( \eta \) varying over \((0,1)\) and \( u \) over \( U_\eta \), and by \( \bar{\mathcal{F}}^y \) the closure of \( \mathcal{F}^y \) in \((D,d)\). Under the additional assumptions on \( v \), we get the same support theorem:

Theorem II. Under Assumptions B, H.1 and H.2,
\[ \text{Supp } \mathbb{P}^{(y)} = \bar{\mathcal{F}}^y. \]

3.3. Remarks

(a) To understand why the skeletons are different according as \( X \) is of type I or not, one should compare them with the shape that the initial equation takes in the two cases. In Eq. (2) one only deals with Stieltjes integrals, whereas in (3) there is also a true stochastic integral which cannot be decomposed. Thus one can stipulate that, even though its jumps must be truncated, the aim of the additional term in the second skeleton is to re-establish the laws of a calculus “with differences” in the stochastic equation. Indeed, the latter hold for Eq. (2) but fail for Eq. (3), as it can be seen from Itô’s formula with jumps (see, e.g. Section II.5 in Ikeda and Watanabe (1989)).

Actually, a Stratonovich integral of jump type (which re-establishes the above laws) was defined in the seventies by Marcus (1978) and studied thoroughly in a recent work of Kurtz et al. (1995) who consider SDEs driven by general semimartingales. However, this integral concerns only a specific class of integrands (the coefficient \( b \) must be chosen suitably to make the computations work). Notice also that these new SDEs of jump-type are included in the class of (non-linear) Itô SDEs we decided to study here.

(b) As in Stroock and Varadhan (1972) and Ikeda and Watanabe (1989), the above theorems allow us to define, for every \( y \in \mathbb{R}^d \), a maximal set \( D(y) \) upon which the integro-differential operator \( L \) satisfies the strong maximum principle with respect to \( y \). Let us briefly recall the matter of this question.
A function $u : \mathbb{R}^d \to \mathbb{R}$ is said to be $L$-subharmonic if it is upper semi-continuous and if for every $n \in \mathbb{N}^*$, $y \in \mathbb{R}^d$,

$$t \mapsto u(Y_t) - u(y)$$

is a $\mathbb{P}^y$-submartingale. Notice that if $u$ is $C^2$, then $u$ is $L$-subharmonic if

$$Lu \geq 0.$$ 

Of course one can also define subharmonicity with respect to a general domain $D$, but here this notion is fruitless since the operator $L$ is non-local. A strong maximum principle over $\mathbb{R}^d$ for $L$ is then the following:

Any subharmonic function reaching its maximum is constant.

Such a principle clearly does not always hold, for example in dimension $m = 2$ when the Lévy measure is supported by an axis. So for every $y \in \mathbb{R}^d$, we want to define a closed subset $D(y)$ of $\mathbb{R}^d$ satisfying the following properties:

(i) For any $L$-subharmonic function $u$ such that $u \leq u(y)$ on $D(y)$, then $u(x) = u(y)$ for every $x \in D(y)$.

(ii) If $z \notin D(y)$, then there exists an $L$-subharmonic function $u$ such that $u \leq u(y)$ on $D(y)$ and $u(z) < u(y)$.

When $Y$ is transient, reasoning as in Ikeda and Watanabe (1989, pp. 529–532) almost verbatim, we get the following corollary (of course under the conditions on $L$ which ensure Theorem I or II accordingly):

**Corollary.** For every $y \in \mathbb{R}^d$,

$$D(y) = \{ x \mid \exists \phi \in \mathcal{H}^y, \ t > 0 \mid x = \phi_t \}$$

satisfies the above properties (i) and (ii).

When $Y$ is recurrent, then the strong maximum principle is satisfied over $\mathbb{R}^d$ itself. Indeed, for every $x, y \in \mathbb{R}^d$, $Y$ started from $y$ visits a.s. any small neighbourhood of $x$ in a finite time, and so it is easy to see that property (i) holds with $D(y) = \mathbb{R}^d$ for every $y$. This is clearly also true for diffusions, but seems surprisingly unnoticed in the literature.

(c) Consider the following vector subspace of $\mathbb{R}^m$:

$$L = \left\{ x \in \mathbb{R}^m \left\lfloor \int_{|z| \leq 1} |\langle x, z \rangle| \nu(dz) < \infty \right\} \right.$$ 

and the following set of càdlàg functions:

$$\phi_t = x - \left( \int_{|z| \leq 1} z_L \nu(dz) \right) t + \psi_t + \sum_{t_n \leq t} z_n,$$

where $z_L$ denotes the orthogonal projection of $z$ onto $L$ (so in particular the integral is convergent in the second term of the right-hand side), $L^\perp$ is the orthogonal complementary to $L$, $\psi$ a continuous function from $\mathbb{R}^+$ to $L^\perp$ null at 0, $\{t_n\}$ a strictly increasing sequence of $\mathbb{R}^+$ with limit $+\infty$ and $\{z_n\}$ any sequence in $\text{Supp} \nu$. 
In Simon (2000, Corollaire 1), it is shown that the support of the carrying Lévy process $X$ over $\mathbb{D}, d$ is the closure of the above set. One can wonder if such a result could not be useful for our equations in the case of type II. It is indeed very natural to conjecture that the support should be defined as the closure of the set of solutions to the following ODEs:

$$
\phi_t = y + \int_0^t \tilde{a}(\phi_s) \, ds + \int_0^t \tilde{b}(\phi_s) \, \psi_s \, ds + \sum_{t \leq r} b(\phi_{t-}, z),
$$

with the same notations as above and where

$$
\tilde{a}(x) = a(x) - \int_{|z| \leq 1} b'(x, z) \nu(dz) - \int_{|z| \leq 1} \tilde{b}(x, z) \nu(dz).
$$

Of course $L = \{0\}$ under Assumption H.2, but we introduce this notation in spite of everything, because Assumptions H.1 and H.2 are probably unnecessary. Such a description of the support is in any case more satisfying since no $\eta$’s enter in the definition. As a matter of fact this result is true in dimension $m = 1$, but under Assumption H.2 and so we did not include it here since we would like to get rid of this assumption. This could be done in a quite different manner, viewing the solution of the equation as a certain (continuous) functional of the carrying process, in the spirit of the well-known papers of Doss (1977) and Sussmann (1978). Actually, when the coefficient $b$ is chosen such that the SDE is canonical (i.e. defined through a Marcus–Stratonovich integral), one can work neatly with the underlying vector fields and, even in higher dimensions, obtain a representation of the solution in terms of multiple integrals involving the carrying process, under the classical assumption that the Lie algebra generated by those vector fields is solvable. This is done in Kunita (1996), where the computations are very similar to the diffusion case (Kunita, 1980). However, the existence of such a functional for a general Itô SDE of jump type on the line (where there should be no auxiliary assumption) remains an open question, which probably requires tools different from the continuous case.

4. Proof of Theorem I

4.1. First inclusion

In this paragraph we briefly recall how the easy inclusion $\text{Supp}^{\mathbb{P}(y)} \subset \mathcal{F}^y$ can be handled. It suffices to consider, as for the Wiener space, a polygonal approximation of the carrying process, i.e. for any $\eta > 0$ to introduce the following stochastic differential equation:

$$
Y^\eta_t = y + \int_0^t \tilde{a}(Y^\eta_s) \, ds + \int_0^t \int_{|z| > \eta} b(Y^\eta_{s-}, z) \mu(ds, dz).
$$

Notice that the integral with respect to $\mu$ has a.s. finitely many jumps on every finite time interval. Hence, reasoning on every sample path, it is obvious that Eq. (6) admits a unique strong solution, whose support is in $\mathcal{F}^y$. 

Since, by definition, the Sup distance bounds from above that of Skorohod, our inclusion will be proved if for every $n \in \mathbb{N}^*$, $\epsilon > 0$,

$$
P^{(\eta)} \left[ \sup_{0 \leq t \leq n} |Y^n_t - Y_t| > \epsilon \right] \to 0$$

as $\eta \downarrow 0$. Indeed, this entails

$$
P^{(\eta, y)} \Rightarrow P^{(y)} \text{ weakly}$$

where $P^{(\eta, y)}$ stands for the law of $Y^n$ in $(D, d)$ and we thus have

$$
P^{(y)}[\mathcal{F}^Y] \geq \limsup_{\eta \downarrow 0} P^{(\eta, y)}[\mathcal{F}^Y] = 1.$$

The above convergence is easily proved with the help of a suitable decomposition of $(Y^n_t - Y_t)$, equality I, Cauchy–Schwarz inequality, and Gronwall’s lemma. We leave this to the reader, who should refer to the third chapter of Simon (1999a) if the task seems too boring to him. When $\gamma$ has compact support, one can even show that

$$
\lim_{\eta \downarrow 0} \mathbb{E}^{(y)} \left[ \sup_{0 \leq t \leq n} |Y^n_t - Y_t|^2 \right] = 0.
$$

This stronger convergence seems however untrue in the general case. Notice finally that the proof does not take the finite variations into account, and so the first inclusion of Theorem II follows almost exactly in the same way.

4.2. Second inclusion

4.2.1. Preliminary results

We are now concerned with the other inclusion $\mathcal{F}^Y \subset \text{Supp} P^{(y)}$, which is rather easy in this finite variation case. In this subsection we establish some results which will also hold when $X$ is of type II, via some straightforward adaptations.

Fix $y \in \mathbb{R}^d$, $u \in U$ and consider $\phi$ the solution to the differential equation (4). We need to show that for every $n \in \mathbb{N}^*$, $\epsilon > 0$,

$$
P^{(y)}[d_\epsilon(Y, \phi) < \epsilon] > 0. \quad (7)
$$

Consider

$$
t_0 = 0 < t_1 < \cdots < t_{N_n} \leq n + 1 < t_{N_n+1},
$$

the first ordered jumping times of $\phi$. Introduce

$$
\eta = \inf \{|z_i|, \; i = 1, \ldots, N_n\}/2.
$$

In the following, $n$ and $\eta$ will be fixed. To prove (7), we need to consider separately the jumps of $\mu$ greater than $\eta$ in modulus and those whose size is smaller than $\eta$. Roughly, we will show that on $[0, n + 1]$ one can make $Y^n$ arbitrarily close to $\phi$ (for the Skorohod distance) in the same time as $Y - Y^n$ is arbitrarily close to 0 (for the Sup distance). The second task is in the general case much more difficult than the first one, though it does not appeal to the Skorohod topology.
Set \( T_0 = 0, \{ T_i, i \geq 1 \} \) the ordered jumping times of \( X \) such that \( |\Delta X_{T_i}| > \eta \), and \( Z_i = \Delta X_{T_i} \) for every \( i \). For every \( i = 1, \ldots, N_n + 1, \rho > 0 \), the event
\[
A(i, \rho) = \{ 0 < (t_i - t_{i-1}) - (T_i - T_{i-1}) < \rho, |Z_i - z_i| < \rho \}
\]
has positive \( \mathbb{P}^{(\rho)} \)-probability: for every \( i \geq 1 \) (\( T_i - T_{i-1} \)) follows an exponential law, and \( z_i \) belongs to the support of \( \nu \). Hence, by independence, the same property holds for
\[
\Omega_{\rho} = \bigcap_{i=1}^{N_n+1} A(i, \rho).
\]
For the sake of conciseness, we will write in the sequel “positive probability” for “positive \( \mathbb{P}^{(\rho)} \)-probability”. Notice that on \( \Omega_{\rho} \), \( T_i < t_i \) for all \( i \) and that \( T_{N_n+1} > n + 1 \) if \( \rho \) is small enough.

Proposition 1. For every \( \varepsilon > 0 \), there exists \( \rho > 0 \) such that on the corresponding \( \Omega_{\rho} \),
\[
d_{\varepsilon}(\tilde{Y}, \phi) < \varepsilon.
\]

Proof. The proof is easy but a bit lengthy. It relies on the construction, for every \( \rho > 0 \), of a suitable change of time \( \lambda^\rho \) of \( \mathbb{R}^+ \) piecewise affine such that for all \( i = 0, \ldots, N_n \),
\[
\lambda^\rho_i = T_i
\]
and a repeated use of Gronwall’s lemma. We leave the verification to the reader.

Remark 2. Actually, the above proposition is just a stability result on SDEs with finitely many jumps. However, it seems difficult to deduce it from the general theorems on the stability of SDEs driven by semimartingales (Dellacherie and Meyer, 1980), because the different semimartingale topologies are too strong for our purposes. Besides, the result is false for a distance defined without change of time, since \( Y^n \) cannot jump at a prescribed time.

The second task is not much harder and we first reduce the problem to an estimate on \( \hat{Y}^\eta \), which is more adapted than \( (Y - Y^n) \) to be handled together with \( Y^n \). We recall that in the case of type I, for every \( t \geq 0 \),
\[
\hat{Y}^\eta_t = \int_0^t \int_{|z| \leq \eta} b(Y_{s-}, z) \mu(ds, dz).
\]

Proposition 3. The reverse inclusion is shown if the following assertion is true: for every \( \varepsilon > 0 \), \( \rho > 0 \) small enough,
\[
\mathbb{P}^{(\rho)}[|\hat{Y}^\eta|_{n+1} < \varepsilon; \Omega_{\rho}] > 0.
\]

Proof. Choose \( \rho \) corresponding to \( \varepsilon \) in the preceding proposition. In the latter, the change of time \( \lambda^\rho \) may be chosen such that \( \lambda^\rho_t \leq t \) for all \( t \); so recalling the definition of \( d_n \), it is easy to see that on \( \Omega_{\rho} \),
\[
d_n(Y, \phi) < d_n(Y^n, \phi) + |Y - Y^n|_{n+1}.
\]
Besides, since on $\Omega_\rho$ we also have
\[ d_\nu(Y^n, \phi) < \varepsilon, \]
it suffices to show that on
\[ \tilde{\Omega}_\rho = \{ |\tilde{Y}^n|_{n+1} < \varepsilon \} \cap \Omega_\rho \]
holds the following:
\[ |Y - Y^n|_{n+1} < K\varepsilon, \]
for some constant $K$ independent of $\varepsilon$. To this end, we decompose $Y_t - Y^n_t$ into
\[ Y_t - Y^n_t = \tilde{Y}^n_t + \int_0^t (\tilde{a}(Y_s) - a(Y^n_s)) \, ds \]
\[ + \int_0^t \int_{|z| > \eta} (b(Y^n_{s-}, z) - b(Y_{s-}, z)) \mu(ds, dz). \]
Hence on $\tilde{\Omega}_\rho$, for every $t \leq n + 1$, we get
\[ |Y_t - Y^n_t| < \varepsilon + K \int_0^t |Y_s - Y^n_s| \, ds + K \sum_{T_i \leq t} |Y_{T_i} - Y^n_{T_i}|. \]
Applying Gronwall’s lemma on $[0, T_1)$ yields
\[ \sup_{0 \leq t < T_1} |Y_t - Y^n_t| < e^{KT_1}. \]
Repeating the same argument $N_n$ times on each $[t_i, t_{i+1})$, we get
\[ |Y - Y^n|_{n+1} < K\varepsilon \]
where $K$ depends only on $a$, $b$, $n$. This completes the proof.

4.2.2. End of the proof

We now prove the reverse inclusion when $X$ is of type I. First, it is clear from the general assumption on $b$ that there exists $K$ such that
\[ |\tilde{Y}^n|_{n+1} < K \sum_{s \leq n+1} |\Delta \tilde{X}^n_s| \quad \text{a.s.} \]
But the left-hand expression depends only on the small jumps of $X$, and so we can control the Sup norm of $\tilde{Y}^n$ independently of $\Omega_\rho$ for any $\rho > 0$. Indeed, since
\[ \sum_{s \leq n+1} |\Delta \tilde{X}^0_s| \]
tends to 0 in probability as $\delta \downarrow 0$, and since for any $\delta > 0$,
\[ \mu([0, n + 1] \otimes \{ \delta \leq |z| \leq \eta \}) = 0 \]
with positive probability, it is clear that for every $\varepsilon > 0$, there exists $\Omega_\delta$ of positive probability and independent of $\Omega_\rho$ such that on $\Omega_\delta$,
\[ |\tilde{Y}^n|_{n+1} < \varepsilon. \]
Hence, for every $\varepsilon$, $\rho > 0$,
\[ \mathbb{P}^{(\delta)}[|\tilde{Y}^n|_{n+1} < \varepsilon; \Omega_\delta] > 0. \]
5. Proof of Theorem II

We recall that the first inclusion Supp $P^X \subset \tilde{\mathcal{Y}}^Y$ can be treated exactly in the same way as in the case of type I. So we are only concerned with the second inclusion, which however appears to be much more delicate in this case.

5.1. Second inclusion

This subsection is devoted to the reduction of this inclusion to a claim. Actually, this reduction does not use Assumptions B, H.1 or H.2; it also does not use the fact that $X$ is of type II.

Fix $\eta > 0$, $u \in U_\eta$, and consider the solution of the corresponding deterministic equation. Recalling that for every $t > 0$,

$\tilde{Y}_t^\eta = \int_0^t \int_{|z|<\eta} \tilde{b}(Y_{s-},z) \mu(ds,dz) + \int_0^t \int_{|z|<\eta} b'(Y_{s-},z) \mu(ds,dz),$

and using the same $Y^\eta$ as in the preceding subsection, it is easy to see that

$Y_t - Y_t^\eta = \tilde{Y}_t^\eta + \int_0^t (a_s(Y_s) - a_s(Y_s^\eta))ds$

$+ \int_0^t \int_{|z|<\eta} (b(Y_{s-},z) - b(Y_{s-}^\eta,z)) \mu(ds,dz).$

Hence, fixing now $n \in \mathbb{N}^*$, we can reason exactly as in the case of type I, and reduce the reverse inclusion to the assertion of Proposition 3 (of course with the corresponding $Y^\eta$. We leave the verification to the reader.

Proposition 4. The claimed assertion of Proposition 3 is true if for all $y \in \mathbb{R}^d$, $\varepsilon > 0$, $T > 0$,

$P^{X}[|Y_T| > \varepsilon] < \varepsilon > 0.$

Proof. Recall that $\{\theta_t, t \geq 0\}$ is the translation operator associated with the strong Markov process $(X, Y)$. For every (possibly random) time $T$, set

$\theta_T^n = \theta_T \circ \cdots \circ \theta_T.$

for $n \geq 1$. For $n = 0$ set $\theta_T^0 = \text{Id}_\mathbb{P}$. Fix $\rho > 0$ small enough and define as before $T = \inf\{t > 0 \mid |\Delta X_t| > \eta\}$, $Z = \Delta X_T$. Since $\theta_T^i = \theta_T^i, \text{ for each } i = 0, \ldots, N_n$ one can rewrite

$A(i, \rho) = B(i, \rho) \circ \theta_T^{i-1}$

for each $i = 1, \ldots, N_n + 1$, with

$B(i, \rho) = \{0 < (t_i - t_{i-1}) - T < \rho\} \cap \{|Z - z_i| < \rho\}.$

Besides, we easily see from its definition itself that $\{\theta_t, t \geq 0\}$ acts on $Y^\eta$ in the following way: for every $t, r \geq 0$,

$Y_t^\eta \circ \theta_r = Y^\eta_{t+r} - Y^\eta_r.$
So we can also rewrite, for $0 \leq t \leq n + 1$,

$$
\hat{Y}^n_t = \sum_{i=0}^{n-1} \hat{Y}^n_T \circ \theta^T_i + \hat{Y}^n_u \circ \theta^T_u
$$

with $T_p \leq t < T_{p+1}$ and $u = t - T_p < T \circ \theta^T_T$.

Notice that $\hat{Y}^n_t$ is a.s. continuous at every $T_i$, $i = 1 \ldots N_n + 1$. In particular we have, conditionally on $\Omega_p$,

$$
\{ |\hat{Y}^n_T| < \varepsilon \} \supset \{ |\hat{Y}^n_T| < k_n \varepsilon \} \cap \cdots \cap \{ |\hat{Y}^n_T \circ \theta^T_T| < k_n \varepsilon \}
$$

with $k_n = (N_n + 1)^{-1}$. We introduce, for $i = 1 \ldots N_n + 1$,

$$
C(i, \rho) = \{ |\hat{Y}^n_T| < k_n \varepsilon \}
$$

with $t_i^\rho = (t_i - t_{i-1}) + \rho$. In the following we will set $B(i, \rho) = B_i$ and $C(i, \rho) = C_i$ for the sake of conciseness. We have

$$
\mathbb{P}^{(x,y)}[|\hat{Y}^n_T| < \varepsilon; \Omega_p] \geq \mathbb{P}^{(x,y)}[B_1 \cap C_1; \cdots; B_{N_n + 1} \cap C_{N_n + 1}] \circ \theta^T_N.
$$

Applying now $N_n$ times the Markov property, the right-hand expression turns into

$$
\mathbb{P}^{(x,y)}[B_1 \cap C_1; \mathbb{P}^{(X_1, Y_1)}[B_2 \cap C_2; \cdots; \mathbb{P}^{(X_{N_n}, Y_{N_n})}[B_{N_n + 1} \cap C_{N_n + 1}] \cdots].
$$

We now need to introduce a more complicated notation for the events $C_i$. We define for every $i \geq 0$, $i = 1 \ldots N_n + 1$,

$$
C_i^\kappa = \{ |\hat{Y}^n_T|^\kappa < k_n \varepsilon \}.
$$

For every $i \geq 1$, $(x, y) \in \mathbb{R}^m \times \mathbb{R}^d$, the only stochastic dependence under $\mathbb{P}^{(x,y)}$ between $B_i$ and $C_i = C_i^{T \wedge t_i^\rho}$ comes from the subscript $T \wedge t_i^\rho$, since by the chaos property of the Poisson measure $(T, Z)$ is independent of the values taken by $\hat{Y}^n$ stopped at $T \wedge t_i^\rho$ (time at which $\hat{Y}^n$ does not jump a.s.). So we get, applying the translation invariance of $X$,

$$
\mathbb{P}^{(x,y)}[B_i \cap C_i] = \int_0^{t_i} \mathbb{P}^{(0)}[(t_i - t_{i-1}) - T \in du, |Z - z_i| < \rho] \mathbb{P}^{(x,y)}[C_i^{t_i \wedge t_i^\rho}]
$$

$$
\geq \int_0^{t_i} \mathbb{P}^{(0)}[(t_i - t_{i-1}) - T \in du, |Z - z_i| < \rho] \mathbb{P}^{(x,y)}[C_i^{t_i}] 
$$

$$
\geq \mathbb{P}^{(0)}[B_i] \mathbb{P}^{(x,y)}[C_i^{t_i}].
$$

Now since from the assumption

$$
\mathbb{P}^{(x,y)}[C_i^{t_i} > 0]
$$

for every $y \in \mathbb{R}^d$, and since obviously

$$
\mathbb{P}^{(0)}[B_i] > 0,
$$

we have

$$
\mathbb{P}^{(x,y)}[B_i \cap C_i] > 0
$$
for every \((x, y) \in \mathbb{R}^m \times \mathbb{R}^d\). Hence, in the above right-hand expression, we integrate each time measurable functions everywhere strictly positive over events of strictly positive measure. This entails that for every \(\varepsilon > 0, \rho\) small enough,

\[
P^{(x)}[|\hat{Y}^\rho|^*_n < \varepsilon, \Omega_\rho] > 0.
\]

The above proposition shows that the second inclusion is reduced to the following claim:

**Claim.** For every \(y \in \mathbb{R}^d, n \in \mathbb{N}^*, \eta, \varepsilon > 0\),

\[
P^{(y)}[|\hat{Y}^\eta|^*_n < \varepsilon] > 0.
\]

5.2. Proof of the claim

We fix \(n\) and \(\eta\) and it is clearly sufficient to suppose that \(\eta = 1\). So we denote \(\hat{Y} = \hat{Y}^1\), \(u_\eta = u_\eta^1\) and \(\gamma = \gamma^1\) in the setting of Assumption H.1. Again, we want to decompose \(\hat{Y}\) into its small jumps and its big jumps. We recall that for every \(\rho > 0, t \geq 0\),

\[
\hat{Y}^\rho_t = \int_0^t \int_{|s| \leq \rho} \hat{b}(Y_s-).z \mu(ds, dz) + \int_0^t \int_{|s| \leq \rho} b'(Y_s-, z) \mu(ds, dz)
\]

and we introduce

\[
\hat{Y}^\eta_t = \int_0^t \int_{|s| \leq \eta} \hat{b}(Y_s-).z \mu(ds, dz).
\]

By Assumption B, there exists a constant \(K\) such that for every \(0 \leq t \leq n\),

\[
\left| \int_0^t \int_{|s| \leq 1} b'(Y_s-, z) \mu(ds, dz) \right| \leq K \sum_{s \leq n} |\Delta \hat{X}_s|^\gamma.
\]

Thus, denoting \(\bar{Y} = \bar{Y}^1\), the claim will be shown as soon as for every \(y \in \mathbb{R}^d, \varepsilon > 0\),

\[
P^{(y)} \left[ \left\{ \sum_{s \leq n} |\Delta \hat{X}_s|^\gamma < \varepsilon \right\} \cap \{ |\hat{Y}^\eta|^*_n < \varepsilon \} \right] > 0.
\]

5.2.1. The quasi-symmetric case

Recall that in this case, there exists a sequence \(\{\eta_k\}\) decreasing to 0 such that

\[
|u_{\eta_k}| \to 0
\]

when \(k \uparrow +\infty\). Let \(\{\rho_k\}\) be extracted from \(\{\eta_k\}\) such that the following quantities tend a.s. to 0:

\[
\sum_{s \leq n} |\Delta \hat{X}^\rho_k|^\gamma \quad \text{and} \quad |\hat{Y}^\rho_k|^*_n.
\]

For any \(\varepsilon > 0\) one can choose \(k\) sufficiently large such that

\[
P^{(y)} \left[ \left\{ \sum_{s \leq n} |\Delta \hat{X}^\rho_k|^\gamma < \varepsilon \right\} \cap \{ |\hat{Y}^\rho_k|^*_n < \varepsilon/2 \} \right] > 1/2
\]
together with
\[ n|\hat{h}|\|u_\eta\| < \varepsilon/2. \]
If we denote
\[ T = \inf \{ t > 0 \mid |\Delta \hat{X}_t| > \rho_k \}, \]
then the event \( \{ T > n \} \) has positive probability and is independent of the process without the big jumps. So we get
\[ \mathbb{P}(\{ \sum_{s < T \wedge n} \big| \Delta \hat{X}_s \big|^2 < \varepsilon/2 \} \cap \{ T > n \}) > 0. \]
But on this event, since \( n|\hat{h}|\|u_\eta\| < \varepsilon/2 \),
\[ \sum_{s < n} |\Delta \hat{X}_s|^2 < \varepsilon \quad \text{and} \quad |\hat{Y}^{\star}_{n}| < \varepsilon. \]

5.2.2. The general case
If there exists a sequence \( \{ \eta_k \} \) decreasing to 0 such that
\[ |u_{\eta_k}| \to 0 \]
when \( k \uparrow +\infty \) (this might hold even though \( X \) is not quasi-symmetric), we reason of course exactly as above. In the other case, we use the notations of Assumption H.1 and define
\[ \tau_\eta = \gamma|\eta|/|u_\eta|. \]
Notice that there exists a constant \( K \) such that
\[ \tau_\eta \leq K \eta \]
in the neighbourhood of 0.

In the general case the proof of the claim is much longer than above. We notice first that it suffices to prove the latter for \( \nu \) having a compact support: there are no big jumps on the event whose probability must be positive (the latter are handled in the construction of Proposition 4).

We begin with the following continuity result, which is a bit obvious but we give a proof for the sake of completeness.

**Lemma 5.** For every \( y \in \mathbb{R}^d, k > 0, \)
\[ \mathbb{P}(\sup_{0 \leq t, s \leq \eta} \big| Y_t - Y_s \big| < k \sqrt{\eta}) \to 1 \]
as \( \eta \downarrow 0 \), uniformly in \( y \).

**Proof.** For \( \eta > 0 \), \( t > s \), we can decompose \( Y_t - Y_s \) into
\[ Y_t - Y_s = \int_s^t a_\eta(Y_u) \, du + \int_s^t \int_{\big|\cdot\big| < \eta^{1/4}} b(Y_{s \leftarrow \cdot}, z) \tilde{\mu}(ds, dz) \]
\[ + \int_s^t \int_{\big|\cdot\big| \geq \eta^{1/4}} b(Y_{s \leftarrow \cdot}, z) \mu(ds, dz), \]
where for every $y \in \mathbb{R}^d$

$$a_\eta(y) = a(y) - \int_{|z| \leq 1} b(y,z) \nu(dz).$$

This can be written as

$$Y_t - Y_s = I_1(t,s) + I_2(t,s) + I_3(t,s),$$

with

$$|a_\eta(y)| \leq K + K \int_{|z| \leq \eta^{1/4}} |z| \nu(dz) \leq K \eta^{-1/4},$$

and so

$$\sup_{0 \leq t,s \leq \eta} |I_1(t,s)| < K \eta^{3/4}.$$

On the other hand, by the equality I,

$$\mathbb{E}^{(y)} \left[ \sup_{0 \leq t,s \leq \eta} |I_2(t,s)|^2 \right] \leq K \eta \int_{|z| \leq \eta^{1/4}} |z|^2 \nu(dz).$$

Hence, by the Chebyshev inequality,

$$\mathbb{P}^{(y)} \left[ \sup_{0 \leq t,s \leq \eta} |I_2(t,s)| < k \sqrt{\eta} \right] \geq 1 - K \int_{|z| \leq \eta^{1/4}} |z|^2 \nu(dz) \to 1 \text{ as } \eta \downarrow 0.$$

Finally, since

$$\nu(|z| \geq \eta^{1/4}) \leq K \eta^{-1/2},$$

the parameter of the Poisson random variable

$$\mu([s,t] \otimes \{|z| \geq \eta^{1/4}\})$$

is bounded from above by $K \sqrt{\eta}$ when $0 \leq t,s \leq \eta$. Using Assumption B and the fact that $\nu$ is supposed to have a compact support, we get

$$\mathbb{P}^{(y)} \left[ \sup_{0 \leq t,s \leq \eta} |I_3(t,s)| = 0 \right] \geq e^{-K \sqrt{\eta}} \to 1$$

as $\eta \downarrow 0$. Putting the pieces together, we find

$$\mathbb{P}^{(y)} \left[ \sup_{0 \leq t,s \leq \eta} |Y_t - Y_s| < k \sqrt{\eta} \right] \to 1$$

as $\eta \downarrow 0$. It is easy to see that all the above estimates are uniform in $y$, and so the same holds for the convergence. \qed

**Remark 6.** Writing $(Y_t - Y_s)$ as in equation (1) and using the same estimates as above, one can show that for some $K > 0$:

$$\mathbb{E}^{(y)} \left[ \sup_{0 \leq t,s \leq \eta} |Y_t - Y_s|^2 \right] \leq K \eta.$$

Notice that one cannot expect a higher exponent on $\eta$ in the right-hand side: by Kolmogorov’s lemma, this would force $Y$ to be continuous!
Thanks to H.2 we get the following technical lemma, partly inspired from a result of Picard (Lemma 3.1 in Picard (1997)). We also give a detailed proof for the sake of completeness. We set $P$ for any $P^{(x)}$.

**Lemma 7.**

\[
\inf_{(\eta, v) \in (0, 1) \times \mathcal{S}^{m-1}} P[v \ast \hat{X}_{\tau_v}^\eta < 0] > 0.
\]

**Proof.** For $(\eta, v) \in (0, 1) \times \mathcal{S}^{m-1}$, set

\[
\mu_{\eta, v} = \int_{|z| \leq \eta} |v \ast z|^2 \nu(dz), \quad \rho_{\eta, v} = \sqrt{\tau_{\eta, v}} \text{ and } Y_{\eta, v} = \frac{v \ast \hat{X}_{\tau_v}^\eta}{\rho_{\eta, v}}.
\]

$Y_{\eta, v}$ is non-zero since by the non-degeneracy condition, $\hat{X}$ lives on the whole space $\mathbb{R}^m$. Hence by a martingale property, for every $(\eta, v) \in (0, 1) \times \mathcal{S}^{m-1}$,

\[
P[v \ast \hat{X}_{\tau_v}^\eta < 0] = P[Y_{\eta, v} < 0] > 0.
\]

Besides, $Y_{\eta, v}$ is an infinitely divisible real random variable whose Lévy measure $\nu_{\eta, v}$ acts on positive functions in the following way:

\[
\nu_{\eta, v}(f) = \tau_{\eta} \int_{|z| \leq \eta} f \left( \frac{v \ast z}{\rho_{\eta, v}} \right) \nu(dz).
\]

In particular,

\[
\nu_{\eta, v}(x^2) = \frac{\tau_{\eta}}{\rho_{\eta, v}^2} \int_{|z| \leq \eta} |v \ast z|^2 \nu(dz) = 1
\]

for every $(\eta, v) \in (0, 1) \times \mathcal{S}^{m-1}$. Besides, Supp $\nu_{\eta, v}$ is contained in the centred closed ball of radius $\eta/\rho_{\eta, v}$ and since by H.2, there exists $c$ such that

\[
\rho_{\eta, v} \geq c \eta
\]

uniformly in $v$, we see that Supp $\nu_{\eta, v}$ is contained in a compact set independent of $\eta$ and $v$. Hence, by the Lévy–Khintchine formula, we see that the family

\[
\{Y_{\eta, v}, (\eta, v) \in (0, 1) \times \mathcal{S}^{m-1}\}
\]

is relatively compact for the convergence in law. Now $x \mapsto 1_{\{x < 0\}}$ is lower semicontinuous, and so just have to show that any element in the closure of the above family verifies (8). Indeed, we show that there exists a constant $c > 0$ such that for every $(\eta, v) \in (0, 1) \times \mathcal{S}^{m-1}$, $\varepsilon \in (0, 1]$,

\[
\nu_{\eta, v}(|z|^2 1_{|z| < \varepsilon}) > c \varepsilon^{2-\beta}.
\]

(9)

Then, if $\bar{\nu}$ is the Lévy measure of any element in the closure of the above family, it follows by weak convergence that

\[
\bar{\nu}([-\varepsilon, \varepsilon]) > 0 \quad \text{for all } \varepsilon.
\]

Besides,

\[
\int_{-\varepsilon}^\varepsilon |x| \bar{\nu}(dx) \geq \bar{\nu}(1_{|z|^2 < \varepsilon}/\varepsilon) \geq c \varepsilon^{1-\beta} \geq c.
\]
for all $0 < \varepsilon \leq 1$. Since the measure $|x| \bar{\nu}(dx)$ does not load the singleton $\{0\}$, this yields

$$\bar{\nu}(|x| \wedge 1) = +\infty.$$  

It is then classical (see Lemma 1.4 in Picard (1997)) that the support of any element in the above closure is $\mathbb{R}$ itself.

Hence our lemma will be shown if the above condition (9) holds. But we have

$$\nu_{v,\varepsilon}(|z|^2 1_{|z| < \varepsilon}) = \frac{\tau_{\eta}}{\nu_{v,\varepsilon}} \int_{|z| \leq \varepsilon} |v \ast z|^2 1_{|v \ast z| < \varepsilon \nu_{v,\varepsilon}} v(dz)$$

$$\geq 1 \wedge \left( \frac{\tau_{\eta}}{\nu_{v,\varepsilon}} \int_{|z| < \varepsilon \nu_{v,\varepsilon}} |v \ast z|^2 v(dz) \right)$$

$$\geq 1 \wedge \left( \frac{\mu_{v,\varepsilon}}{\mu_{v,\varepsilon}} \right).$$

with $\eta' = \varepsilon \nu_{v,\varepsilon}$. Now from the definitions of $\mu_{v,\varepsilon}$, $\rho_{v,\varepsilon}$ and H.2,

$$\frac{\mu_{v,\varepsilon}}{\mu_{v,\varepsilon}} \geq c \varepsilon^{2 \beta} \left( \frac{\rho_{v,\varepsilon}}{\eta} \right)^{2 - \beta} \geq c \varepsilon^{2 \beta},$$

where we used again

$$\rho_{v,\varepsilon} \geq c \eta$$

for all $\varepsilon$. The proof is now complete.  

**Remark 8.** Assumption H.2 is only useful to show the above lemma. Yet the latter is crucial to our purposes (at least with a large equality inside $P$, see the proof of the following corollary), and seems untrue when the measure is degenerated or has too much irregular variations near 0.

Remark 6 and Lemma 7 entail the following:

**Corollary 9.** For every $\varepsilon > 0$, there exists $c > 0$ such that if $\eta$ is small enough, for every $y \in \mathbb{R}^d$, $v \in \mathcal{F}^{d-1}$,

$$P(y) \left[ v \ast \left( \int_0^{\tau_y} \int_{|z| \leq \varepsilon} \tilde{b}(Y_{s-}).z \tilde{\mu}(ds, dz) \right) \leq \varepsilon \tau_y \right] \geq c.$$  

**Proof.** It is straightforward from the preceding lemma that there exists $c > 0$ such that for every $B : \mathbb{R}^m \to \mathbb{R}^d$ linear, $v \in \mathcal{F}^{d-1}$, $\eta \in (0, 1],$

$$P[v \ast (B.\tilde{X}_t^\eta) \leq 0] \geq 2c.$$  

Hence, since we can write

$$\int_0^{\tau_y} \int_{|z| \leq \varepsilon} \tilde{b}(Y_{s-}).z \tilde{\mu}(ds, dz) = \tilde{b}(y) \cdot \tilde{X}_t^\eta + \int_0^{\tau_y} \int_{|z| \leq \varepsilon} (\tilde{b}(Y_{s-}) - \tilde{b}(y)) \cdot z \tilde{\mu}(ds, dz),$$

it is sufficient to show that if $\eta$ is small enough, for every $y \in \mathbb{R}^d$,

$$P(y) \left[ \left| \int_0^{\tau_y} \int_{|z| \leq \varepsilon} (\tilde{b}(Y_{s-}) - \tilde{b}(y)) \cdot z \tilde{\mu}(ds, dz) \right| \leq \varepsilon \tau_y \right] \geq 1 - c.$$
Applying equality I yields

\[ E(\gamma) \left[ \left( \int_0^{\tau} \int_{|z| \leq \eta} (\tilde{b}(Y_x) - \tilde{b}(y)) \cdot \tilde{\mu}(ds, dz) \right)^2 \right] \]

\[ \leq K\tau_{\eta} E(\gamma) \left[ \sup_{x \leq t} |Y_x - y|^2 \right] \int_{|z| \leq \eta} |z|^2 \nu(dz) \]

\[ \leq K\tau_{\eta}^2 \int_{|z| \leq \eta} |z|^2 \nu(dz). \]

The result follows now from the Chebyshev inequality.

We can now proceed to the proof of the claim and we appeal to Assumption H.1.

Fix \( \varepsilon > 0 \), and introduce the following event:

\[ A_{\eta}(t, v) = \left\{ \omega : \sup_{0 \leq u \leq t} |Y_u - Y_s| < \varepsilon \sqrt{\tau_{\eta}}, \sum_{s < t} |\Delta \tilde{X}_s|^x < \varepsilon \tau_{\eta}, \right. \]

\[ \left. \int_0^t \int_{|z| \leq \eta} \tilde{b}(Y_{s-}) \cdot z \tilde{\mu}(ds, dz) \right|_t^u < \varepsilon \sqrt{\tau_{\eta}}, \]

\[ v \ast \left( \int_0^t \int_{|z| \leq \eta} \tilde{b}(Y_{s-}) \cdot z \tilde{\mu}(ds, dz) \right) < \varepsilon \tau_{\eta} \right\}. \]

In general, \( t \) and \( v \) will also depend on \( \omega \).

**Proposition 10.** There exists \( c > 0 \) such that if \( \eta \) is small enough, for every \( y \in \mathbb{R}^d \), \( v \in \mathcal{S}^{d-1} \),

\[ \mathbb{P}(\gamma)[A_{\eta}(\tau_{\eta}, v)] \geq c. \]

**Proof.** By the preceding corollary, it remains to show that the \( \mathbb{P}(\gamma) \) probabilities of the three first events defining \( A_{\eta}(\tau_{\eta}, v) \) tend to 1 as \( \eta \downarrow 0 \), uniformly in \( y \). For the first one, this is given by Lemma 5. For the second one, this is immediate. For the last one, we first apply equality I:

\[ E(\gamma) \left[ \sup_{0 \leq u \leq \tau_{\eta}} \int_0^u \int_{|z| \leq \eta} \tilde{b}(Y_{s-}) \cdot z \tilde{\mu}(ds, dz) \right]^2 \leq K\tau_{\eta} \int_{|z| \leq \eta} |z|^2 \nu(dz) \]

and the result follows directly from the Chebyshev inequality.

Our method for proving the claim is indeed analogous to that of Proposition 4, but the framework is a bit heavier. First, the above proposition allows us to choose \( \eta > 0 \) small enough in the subsequence of H.1 such that

\[ \mathbb{P}(\gamma)[A_{\eta}(\tau_{\eta}, v)] > c \]
uniformly in $y,v$, in the same time as
\[
\eta^{-1} |u_\eta| < \varepsilon, \quad z_\eta < \varepsilon / |u_\eta|.
\]
Let
\[
0 < \tau^{A}_n < \cdots < \tau^{N_n+1}_n \leq n < \tau^{N_n+1}_n
\]
be a regular subdivision of $[0,n]$ with size $\tau_n$.

As in the preceding subsection, set $T = \inf \{ t > 0 \mid |\Delta X_t| > \eta \}$, $Z = \Delta X_T, T_0 = 0,$
and for $i = 1 \ldots N_n + 1$,
\[
T_i = T + \cdots + T \circ \theta_{T}^{-1},
\]
where we defined $\theta_T$ as before: $T_i$ is the $i$th jumping time on $\{|z| > \eta\}$. Consider the event
\[
B_\eta = \{ \tau_\eta - \rho_\eta < T < \tau_\eta \} \cap \{|Z - \tau_\eta u_\eta| < \varepsilon \tau_\eta \}
\]
where $\rho_\eta$ is chosen small enough such that on
\[
B_\eta \cap B_\eta \circ \theta_T \cap \cdots \cap B_\eta \circ \theta_{T}^{N_n},
\]
the $N_n + 1$ first jumping times stay in a left-neighbourhood of the $\tau_i$’s, and that
\[
N_n^2 |u_\eta| \rho_\eta < \varepsilon.
\]
A crucial remark, which is a direct consequence of Assumption H.1, is that
\[
P^{(0)}[B_\eta] > 0.
\]
As a matter of fact H.1 was introduced to get the above inequality. See however the following subsection, where we consider the case when $\tau_\eta u_\eta$ can be reached for any $\eta$ with a finite number (independent of $\eta$) of jumps whose size is approximately $\eta$.

Set $T_\eta = T \wedge \tau_\eta$, and introduce the following vectors of $\mathbb{R}^d$:
\[
v_\eta = \int_{0}^{T} \int_{|Z| \leq \eta} \tilde{b}(Y_{s-}) z \tilde{\mu}(ds, dz),
\]
\[
v_i^\eta = v_\eta + v_\eta \circ \theta_T + \cdots + v_\eta \circ \theta_{T}^{i-1}
\]
for every $i \geq 0$.

Consider the event
\[
\Omega_\eta = \{ B_\eta ; A_\eta(T_\eta, 0) \} \cap B_\eta \circ \theta_T \cap A_\eta(T_\eta \circ \theta_T, v_\eta^1) \circ \theta_T
\]
\[
\cap \cdots \cap \{ B_\eta \circ \theta_T^{N_n}; A_\eta(T_\eta \circ \theta_T^{N_n} , v_\eta^{N_n}) \circ \theta_T^{N_n} \}.
\]
First of all we show that on $\Omega_\eta$ we control simultaneously
\[
\sum_{s \leq \eta} |\Delta \hat{X}_s|^2 \quad \text{and} \quad |\hat{Y}_s|^2.
\]
We thus reason with $\omega$ fixed in $\Omega_\eta$, possibly out of a negligible set. Remark first that obviously
\[
\sum_{s \leq \eta} |\Delta \hat{X}_s|^2 < \varepsilon
\]
together with
\[ \sum_{s \leq n} |\Delta \tilde{X}_s|^2 1_{|\Delta \tilde{X}_s| > q} \leq N_\eta (\tau_\eta |u_\eta| + \varepsilon \tau_\eta)^q \leq K \eta^{-1/2} |u_\eta| < \varepsilon \]

if \( \eta \) was chosen small enough. Thus
\[ \sum_{s \leq n} |\Delta \tilde{X}_s|^2 < \varepsilon. \]

Next, for every \( t \leq n \) such that \( T_j \leq t < T_{j+1} \),
\[ \left| \int_0^t \int_{\eta < |z| < 1} \tilde{b}(Y_{t-}) z \tilde{\mu}(ds, dz) \right| \leq \left| \int_0^t \int_{\eta < |z| < 1} \tilde{b}(Y_{t-}) z \tilde{\mu}(ds, dz) \right| \\
+ \sum_{i=0}^{j-1} \left| \int_{T_i}^{T_{i+1}} \int_{\eta < |z| < 1} \tilde{b}(Y_{t-}) z \tilde{\mu}(ds, dz) \right| \\
\leq \left| \tilde{b} \right| (\tau_\eta + N_\eta \rho_\eta) |u_\eta| + \sum_{i=0}^{j-1} \int \int \ldots. \right|. \]

But \( \left| \tilde{b} \right| (\tau_\eta + N_\eta \rho_\eta) |u_\eta| \leq \varepsilon \) if \( \eta \) was chosen small enough, and for every \( i \)
\[ \int_{T_i}^{T_{i+1}} \int_{\eta < |z| < 1} \tilde{b}(Y_{t-}) z \tilde{\mu}(ds, dz) = \tilde{b}(Y_{T_{i+1}-}) \Delta X_{T_{i+1}} - \left( \int_{T_i}^{T_{i+1}} \tilde{b}(Y_s) ds \right) u_\eta \\
= \left( \int_{T_i}^{T_{i+1}} (\tilde{b}(Y_{T_{i+1}-}) - \tilde{b}(Y_s)) ds \right) u_\eta + w_i, \]
with \( |w_i| \leq K(N_\eta \rho_\eta |u_\eta| + \varepsilon \tau_\eta) \), since
\[ |\Delta X_{T_{i+1}} - (T_{i+1} - T_i) u_\eta| \leq (N_\eta \rho_\eta |u_\eta| + \varepsilon \tau_\eta). \]

Notice that for every \( i \),
\[ \sup_{T_i \leq t < T_{i+1}} |Y_t - Y_s| < \varepsilon \sqrt{\tau_\eta}, \]
which entails
\[ \left| \int_{T_i}^{T_{i+1}} (\tilde{b}(Y_{T_{i+1}-}) - \tilde{b}(Y_s)) ds \right| \leq K \varepsilon \sqrt{\tau_\eta}. \]

We get finally
\[ \left| \int_0^t \int_{\eta < |z| < 1} \tilde{b}(Y_{t-}) z \tilde{\mu}(ds, dz) \right| \leq \varepsilon + N_\eta K \left[ \tau_\eta \varepsilon \sqrt{\tau_\eta} |u_\eta| + (N_\eta \rho_\eta |u_\eta| + \varepsilon \tau_\eta) \right] \\
\leq \varepsilon + K \varepsilon \left( \sqrt{|u_\eta|} + 1 \right). \]

But,
\[ \sqrt{|u_\eta|} \to 0 \]
as \( \eta \downarrow 0 \) and we thus get, again if \( \eta \) had been chosen small enough,
\[ \left| \int_0^t \int_{\eta < |z| < 1} \tilde{b}(Y_{t-}) z \tilde{\mu}(ds, dz) \right| \leq \varepsilon. \]
Finally, for every $0 \leq t \leq n$, we have on $\Omega_\eta$
\[
\left| \int_0^t \int_{|z| \leq \eta} \tilde{b}(Y_{s-}).z \tilde{\mu}(ds,dz) \right|^2 \leq 2 \left( \varepsilon \tau_\eta + \left| \int_0^T \int_{|z| \leq \eta} \tilde{b}(Y_{s-}).z \tilde{\mu}(ds,dz) \right|^2 \right)
\]
for some $i$. Besides,
\[
\left| \int_0^T \int_{|z| \leq \eta} \tilde{b}(Y_{s-}).z \tilde{\mu}(ds,dz) \right|^2
\]
\[
= |v_\eta|^2 + \cdots + |v_\eta \circ \theta_T^{-1}|^2 + 2 \sum_{j=1}^{i-1} (v_\eta \circ \theta_T^j \ast v_\eta^j).
\]
Recalling the definition of $A_\eta$ and that on $\Omega_\eta$, $T_\eta \circ \theta_T^j = T \circ \theta_T^j$ for all $i$, we get
\[
|v_\eta|^2 + \cdots + |v_\eta \circ \theta_T^{-1}|^2 \leq N_\eta \tau_\eta \varepsilon \leq K \varepsilon,
\]
\[
2 \sum_{j=1}^{i-1} (v_\eta \circ \theta_T^j \ast v_\eta^j) \leq 2N_\eta \tau_\eta \varepsilon \leq K \varepsilon.
\]
This yields
\[
|\bar{Y}_s|^*_{\eta} \leq K \sqrt{\varepsilon} \leq \varepsilon.
\]
Putting the pieces together entails that on $\Omega_\eta$
\[
\sum_{s \leq n} |\Delta \hat{X}_s|^2 < \varepsilon \quad \text{and} \quad |\bar{Y}_s|^*_{\eta} < \varepsilon.
\]

It remains to show that $P^{(i)}[\Omega_\eta] > 0$. Indeed we reason almost exactly as in the proof of Proposition 4, except that here we must deal carefully with the $v_\eta^i$'s. Notice that the latter are respectively $F_{T_i}$-measurable (and even $F_{T_0}$-measurable). We obtain first, applying the (extended) Markov property knowing $F_{T_n}$,
\[
P^{(i)}[\Omega_\eta] = P^{(i)}\left[ \{B_\eta; A_\eta(T_\eta,0) \cap \{B_\eta \circ \theta_T; A_\eta(T_\eta \circ \theta_T, v_\eta^i) \circ \theta_T \} \right]
\]
\[
\cap \cdots \cap \{B_\eta \circ \theta_T^{N-1}; A_\eta(T_\eta \circ \theta_T^{N-1}, v_\eta^{N-1}) \circ \theta_T^{N-1} \}
\]
\[
P^{(i)}\left[ \{B_\eta; A_\eta(T_\eta, v_\eta^0) \} \right].
\]

The important point is that on $\Omega_\eta$, $v_\eta^N(\omega)$ is bounded by $K \varepsilon$. So we can apply Proposition 4 and get a constant $c$ such that for every $\omega$
\[
P^{(i)}[A_\eta(\tau_\eta, v_\eta^N(\omega))] > c > 0.
\]

Using the same notations as in the preceding subsection, we also get $c$ such that for every $\omega$
\[
P^{(i)}[B_\eta^i] = P^{(0)}[B_\eta^i] > c > 0.
\]

But on $\Omega_\eta$, $T_\eta \leq \tau_\eta$ a.s. So we can use for every $\omega$ exactly the same independence argument as in the proof of Proposition 4, between $B_\eta$ and $A_\eta(T_\eta, v_\eta^N(\omega))$ under
This entails
\[
P^{(y)}[\Omega] > c^2 P^{(y)} \left\{ B_{\eta}; A_{\eta}(T_{\eta}, 0) \right\} \cap \left\{ B_{\eta} \circ \theta_T; A_{\eta}(T_{\eta} \circ \theta_T, \nu_{\eta}^T) \circ \theta_T \right\} \\
\cap \cdots \cap \left\{ B_{\eta} \circ \theta_T^{N_{\eta} - 1}; A_{\eta}(T_{\eta} \circ \theta_T^{N_{\eta} - 1}, \nu_{\eta}^{N_{\eta} - 1} - 1) \circ \theta_T^{N_{\eta} - 1} \right\}.
\]

Repeating the same argument \( N_{\eta} \) times yields finally
\[
P^{(y)}[\Omega] > 0 \quad \text{for every } y \in \mathbb{R}^d.
\]

### 5.3. A final remark

Assumption H.1 is only useful to make \( \mu \) jump in a suitable direction, so that \( P^{(0)}[B_{\eta}] > 0 \). This is a crucial point but one can wonder if a similar proof would not work under a different assumption on the support of \( \nu \). In this subsection we treat an example which also yields the support theorem, even though H.1 does not hold. Such configurations of \( \text{Supp} \nu \) were already studied by Léandre (1985), in the framework of Malliavin’s calculus with jumps.

We suppose that \( \text{Supp} \nu \) is made up of \( m \) smooth parametered arcs \( \gamma_j : \mathbb{R}^+ \to \mathbb{R}^d \), such that
\[
\gamma_j(0) = 0, \quad \gamma_j'(0) \neq 0 \quad \text{for } j = 1 \ldots m,
\]
\[
\text{Vect}[\gamma_j'(0), j = 1 \ldots m] = \mathbb{R}^d.
\]

Besides, we suppose that those arcs are injective, disjoint, and that they all quit a neighbourhood of 0 after a finite parameter. The measure \( \nu \) decomposes itself along the arcs into \( m \) measures \( \nu_j(dt) \) whose support is \( \mathbb{R}^+ \). For simplicity we suppose that the \( \nu_j \)'s are identical, and denote again by \( \nu \) their common value. Assumption H.2 entails that
\[
\int_0^\eta z^2 \nu(dz) > \eta^{2-\beta}
\]
for some \( \beta \in [1, 2) \). If
\[
\sum_{j=1}^m \nu_j'(0) = 0,
\]
(this holds in particular when \( \beta = 1 \), because of H.2), we also make the following assumption:
\[
\text{Conv}[\gamma_j'(0), j = 1 \ldots m] = \mathbb{R}^d,
\]
where Conv stands for the generated convex cone.

Of course H.1 may hold, but this example also concerns some non-convex configurations of \( \text{Supp} \nu \) which are not recovered by H.1. For example
\[
\nu(dx, dy) = 1_{\{y = |x|\}} |x|^{-5/2} \, dx
\]
in \( \mathbb{R}^2 - \{0\} \), where the angle between \( \text{Supp} \nu \) and the direction \( u_{\eta} \) is \( \pi/4 \) on any sphere.

We claim that under H.2 and the above assumptions, the support theorem remains true.
The information on the size of a jump is an index $j$ and a parameter $z$. Since the arcs are disjoint and injective, the jumps with parameter lower than $\eta$ are independent of those with parameter not lower than $\eta$, for all $\eta > 0$. Hence, we can reason exactly as above but with a division of the space of jumps into small-parameter jumps and big-parameter jumps. The compensator of the big-parameter jumps is given by

$$u_\eta = \sum_{j=1}^{m} \int_0^\infty \tilde{\gamma}_j(z) \nu(dz),$$

where $\tilde{\gamma}_j$ stands for $\gamma_j$ intersected with the unit ball, and the integral converges since the arcs quit a neighbourhood of 0 after a finite parameter. Fixing a parameter $z$ small enough we can write, thanks to Taylor’s formula,

$$u_\eta = \sum_{j=1}^{m} \int_0^z \gamma_j(z) \nu(dz) + \sum_{j=1}^{m} \int_z^\infty \tilde{\gamma}_j(z) \nu(dz) = \left( \sum_{j=1}^{m} \gamma_j(0) \right) \int_0^z z \nu(dz) + v_\eta,$$

where $v_\eta$ has bounded norm. This entails

$$\eta \frac{u_\eta}{|u_\eta|} = \eta \left[ c_0 \sum_{j=1}^{m} \gamma_j(0) \right] + O(\eta^{\beta'})$$

where

$$\beta' = 1 \text{ if } \sum_{j=1}^{m} \gamma_j(0) = 0,$$

$$\beta' = \beta \text{ otherwise.}$$

We can rewrite

$$\eta \frac{u_\eta}{|u_\eta|} = \eta \left[ c_0 \sum_{j=1}^{m} \gamma_j(0) \right] + \eta^{\beta'} \left[ c_0 \sum_{j=1}^{m} \beta_j \gamma_j(0) \right] + o(\eta^{\beta'})$$

where, considering possibly a subsequence, each $|\beta_j|$ (each $\beta_j$ when $\beta = 1$ or $u_\eta = v_\eta$) is either bounded below by a positive constant $c_1$ as $\eta \downarrow 0$, or naught identically.

We now change $B_\eta$, making this time $\mu$ jump successively along the $\gamma_j(0)$; for $j = 1 \ldots m$, we set

$$B^j_\eta = \{ \tau_\eta | m - \rho_\eta < T < \tau_\eta/m \} \cap \{|Z - \gamma_j(0)| < c\eta \}$$

where $T, Z$ are defined as above, again $\rho_\eta$ is chosen small enough and where we defined:

$$\gamma = \inf_{j=1 \ldots m} \{|\gamma_j(0)|\}, \quad c = c_0 \wedge c_1, \quad \tau_\eta = \frac{2\eta}{c\gamma |u_\eta|}, \quad \gamma_j^\eta = \frac{2}{c\gamma}(c\eta + \beta_j \gamma_j(0)).$$

Using $\gamma_j(t) = t\gamma_j(0) + O(t^2)$ and $\text{Supp } \nu = \mathbb{R}^+$, we see that for $\eta$ small enough, for all $j$,

$$P^{(0)}[B^j_\eta] > 0.$$
replace $B_\eta, \ldots, B_\eta \circ \theta_T$ respectively by $B_\eta^1, \ldots, B_\eta^m \circ \theta_T$, and do the same starting from $B_\eta \circ \theta_T^m, \ldots, B_\eta \circ \theta_T^{m+1}, \ldots$, etc.

Since $\mathbb{P}(\mathcal{B}_\eta | B_\eta^j) > 0$ for all $j$, we get similarly

$$\mathbb{P}(\mathcal{B}_\eta | \Omega_\eta) > 0.$$ 

The estimates on $\Lambda_\eta$ are the same, except those concerning

$$\sup_{0 \leq t \leq \eta} \left| \int_0^t \int_{z \in O_\eta} \tilde{b}(Y_s - z) \tilde{\mu}(ds, dz) \right|,$$

where $O_\eta$ is the set of arcs initiated from the parameter $\eta$, and intersected with the unit ball.

For every $t \leq \eta$ such that $T_m \leq t < T_{m+1}$,

$$\left| \int_0^t \int_{z \in O_\eta} \tilde{b}(Y_s - z) \tilde{\mu}(ds, dz) \right| \leq \left| \int_{T_m}^t \int_{z \in O_\eta} \tilde{b}(Y_s - z) \tilde{\mu}(ds, dz) \right| + \sum_{i=0}^{j-1} \left| \int_{T_{m+i+1}}^{T_{m+i}} \int_{z \in O_\eta} \tilde{b}(Y_s - z) \tilde{\mu}(ds, dz) \right|$$

$$\leq \left| \tilde{b} \right| (d \tau_\eta + N_\eta \rho_\eta) |u_\eta| + K\eta + \sum_{i=0}^{N_\eta-1} \left| \int \int \cdots \right|.$$ 

Again $|\tilde{b}| (k \tau_\eta + N_\eta \rho_\eta) |u_\eta| + K\eta \leq \epsilon$ if $\eta$ was chosen small enough and from the definition of $B_\eta^i$, for every $i$

$$\left| \int_{T_{m+i}}^{T_{m+i+1}} \int_{z \in O_\eta} \tilde{b}(Y_s - z) \tilde{\mu}(ds, dz) \right|$$

is dominated by

$$\sum_{i=1}^{m} \left( \tilde{b}(Y_{T_{m+i}}) - \frac{1}{\tau_\eta} \int_{T_{m+i}}^{T_{m+i+1}} \tilde{b}(Y_s) ds \right) \gamma_{m+i}^{l'}(0) + \epsilon \eta^\beta,$$

which is again dominated by

$$K\eta \left( \sup_{T_{m+i} \leq t < T_{m+i+1}} |Y_t - Y_s| \right) + \epsilon \eta^\beta.$$ 

In the latter expression, one must take $(k-1)$ jumps in consideration, but their modulus are bounded by $K\eta$. Since $\tilde{b}$ is bounded, we are led finally to the estimate of:

$$K\eta (\sqrt{\tau_\eta} + K\eta) + \epsilon \eta^\beta,$$

which is bounded above by

$$K\eta (\sqrt{\tau_\eta} + \epsilon \eta^\beta$$

for $\eta$ small enough. Since

$$N_\eta \approx \eta^{-\beta},$$
we get, summing on $i$,
\[
\sup_{0 \leq t \leq \tau} \left| \int_0^t \int_{x \leq \|\bar{\nu}\|_N} \tilde{b} (Y_{s-}, x) \tilde{\mu} (dx, dz) \right| \leq \varepsilon + \sqrt{\eta |\bar{\mu}|} \leq \varepsilon
\]
for $\eta$ small enough. 

**Note added in proof**

After this paper had been accepted, Peter Imkeller drew the author’s attention to the following article by H. Kunita: Canonical stochastic differential equations based on Lévy processes and their supports in H. Crauel (ed.) et al., Stochastic dynamics. Conference on Random dynamical systems, Bremen, Germany, April 28–May 2, 1997. New York, Springer 283–304 (1997). In this paper the support of a Marcus-type SDE is characterized for the Skorohod topology on $[0,1]$, with the help of an approximative continuity argument.

**Acknowledgements**

I thank Francis Hirsch for constant support and sagacious comments, Jean Picard for the interest he took in this work and insightful remarks, and Jean Jacod for a very valuable discussion about the claim. I am also grateful to an anonymous referee whose help greatly improved the loose presentation of the first draft of this paper.

**References**


