SUPPORT THEOREM FOR LÉVY DRIVEN SDES

ALEXEI KULIK

ABSTRACT. We provide a support theorem for the law of the solution to an SDE with jump noise. This theorem applies to general SDEs with jumps and is illustrated by examples of SDEs with quite degenerate jump noises, where the theorem leads to an informative description of the support.

1. INTRODUCTION

The aim of this note is to describe the topological support of the law of the solution to a general SDE with jump noise. For diffusions such a description is provided by the classical Stroock-Varadhan support theorem, see [15]. The natural question of an extension of this theorem to SDEs with jumps was studied intensively at the end of 90-ies – early 2000-ies by H.Kunita, Y.Ishikawa, and T.Simon, see [8], [10], [4]. It appears that, in the jump noise setting, the two cases should be naturally separated, named in [10] the ‘Type I’ and ‘Type II’ SDEs. In plain words, an SDE is of ‘Type I’ if the small jumps are absolutely integrable and of ‘Type II’ otherwise. A ‘Type I’ SDE admits a simple and intuitively clear description of the support; this case was studied completely in [10]. ‘Type II’ SDEs had been studied under some additional limitations only, which can be understood as certain technical convexity and scaling-type assumptions on the Lévy measure of the noise. The latter ‘scaling’ assumption (see H2 in [10] or [3] below) is quite restrictive and requires the jump noise to be close to an α-stable one, in a sense. This limitation is of a technical nature and is caused by the method used in [10], [3] rather than the problem itself, hence a natural question arises how to remove it in order to get a general support theorem for SDEs with jumps, free from any technical limitations. This question was solved in the one-dimensional setting in [12] and for canonical (Markus) equations in [13]. For general multidimensional Itô equations with jumps it apparently requires alternative methods, and remained open since early 2000ies.

In this note, we propose a method for proving a support theorem for general ‘Type II’ SDEs, based on a change of measure and free from any non-natural technical limitations. The description we obtain for the topological support for the law of the solutions to jumping SDEs is of a considerable importance because of its natural applications, in particular, to the study of the ergodic properties of the solution to SDE, considered as a Markov process and the strong maximum principle for the generator of this process; see more discussion in Section 2.3 below.

The structure of the paper is as follows. In Section 2 the main statement is formulated and provided by a discussion and examples. The proof of the main statement is given in Section 3 and to improve readability the proof of the key lemma (Lemma 3.1) is postponed to a separate Section 4. For the same reason, proof of a technical estimate (18) is given in Appendix A.
Acknowledgements. This research was supported by the Polish National Science Center grant 2019/33/B/ST1/02923. The author is grateful to Tomas Simon for inspiring discussion and bibliography information.

2. Main Result

2.1. Preliminaries. Let $N(du,dt)$ be a Poisson point measure (PPM) on $\mathbb{R}^d \times [0, \infty)$ with the intensity measure $\mu(du)dt$, where $\mu(du)$ is a Lévy measure, i.e.

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

We consider an SDE in $\mathbb{R}^m$

$$dX_t = b(X_t) \, dt + \int_{|u|\leq 1} c(X_{t-}, u) \tilde{N}(du, dt) + \int_{|u|> 1} c(X_{t-}, u) N(du, dt), \quad X_0 = x_0,$$

where $\tilde{N}(du, dt) = N(du, dt) - \mu(du)dt$ is the corresponding compensated PPM.

The coefficient $c(x, u)$ is assumed to have the form

$$c(x, u) = \sigma(x)u + r(x, u),$$

where $r(x, u)$ is negligible, in a sense, w.r.t. $|u|$ for small values of $|u|$ (see below). In the case $r(x, u) \equiv 0$ equation (1) transforms to

$$dX_t = b(X_t) \, dt + \sigma(X_{t-})dZ_t, \quad X_0 = x_0,$$

where the Lévy process $Z$ is given by its Itô-Lévy decomposition

$$dZ_t = \int_{|u|\leq 1} u \tilde{N}(du, dt) + \int_{|u|> 1} u N(du, dt).$$

We assume the following.

$\textbf{H}_1$. The functions $b : \mathbb{R}^m \to \mathbb{R}^m$ and $\sigma : \mathbb{R}^m \to \mathbb{R}^{m \times d}$ are Lipschitz continuous.

$\textbf{H}_2$. There exist constants $C > 0, \beta > 0$ such that

$$|r(x, u) - r(y, u)| \leq C|x - y||u|^\beta, \quad |r(x_0, u)| \leq C|u|^\beta,$$

and

$$\int_{|u|\leq 1} |u|^\beta \mu(du) < \infty. \quad (4)$$

If $\beta \leq 1$, then $\textbf{H}_2$ yields that the entire function $c(x, u)$ is absolutely integrable w.r.t. $\mu(du)$ on the set $\{|u| \leq 1\}$; that is, equation (1) has Type I in the terminology of [10]. This (comparatively simple) case is already studied completely, hence in what follows we consider the case $\beta > 1$, only. Note that, in this case $r(x, u) \leq C(|x - x_0| + 1)|u|^\beta \ll |u|$ for small $|u|$, and the linear part $\sigma(x)u$ is the principal one in the decomposition (2) for $c(x, u)$.

Recall that the Skorokhod space $D([0, T], \mathbb{R}^m)$ is the set of càdlàg functions on $[0, T]$ with the metric

$$d(f, g) = \sup_{\lambda \in \Lambda_0,T} \left( \sup_{0 \leq s < t \leq T} \log \frac{\lambda_t - \lambda_s}{t - s} + \sup_{t \in [0, T]} |f(\lambda_t) - g(t)| \right),$$

where $T > 0$ and $\Lambda_0, T$ denotes the set of strictly increasing continuous functions $\lambda : [0, T] \to [0, T]$ such that $\lambda(0) = 0, \lambda(T) = T$. It is known that $D([0, T], \mathbb{R}^m)$ with $d(\cdot, \cdot)$ is a Polish space, e.g. [11] Section 14. 
Under the assumptions $H_1, H_2$, SDE (1) has unique strong solution $X$, see [3, Theorem IV.9.1]. We fix a time horizon $T > 0$, and consider this solution on the time interval $[0, T]$. The law of this solution in the Skorokhod space $D([0, T], \mathbb{R}^m)$ will be denoted by $\text{Law}_{x_0,T}(X)$. We aim to describe the support of this law; recall that the (topological) support of a measure $\kappa$ on a metric space $S$ with Borel $\sigma$-algebra is the minimal closed subset $F$ such that $\kappa(S \setminus F) = 0$, this set is denoted by $\text{supp}(\kappa)$. Alternatively, $x \in \text{supp}(\kappa)$ if, and only if, for any open ball $B$ centered at $x$ one has $\kappa(B) > 0$.

2.2. The main statement. To describe the support of $\text{Law}_{x_0,T}(X)$ we introduce some constructions. First, we introduce a kernel $J(x, dy)$ on $\mathbb{R}^m$ by

$$J(x, A) = \mu \left( \{ u : x + c(x, u) \in A \} \right), \quad A \in \mathcal{B}(\mathbb{R}^m),$$

and define the set $\mathcal{A} \subset \mathbb{R}^m \times \mathbb{R}^m$ of ‘admissible jumps’ by

$$(x, y) \in \mathcal{A} \iff y \in \text{supp}(J(x, \cdot)).$$

Next, denote by $L$ the set of $\ell \in \mathbb{R}^d$ such that

$$\int_{|u| \leq 1} |\ell| \mu(du) < \infty;$$

here and below we use notation $a \cdot b$ for the scalar product in $\mathbb{R}^d$. It is easy to check that $L$ is a vector subspace in $\mathbb{R}^d$; we call it the ‘integrability subspace’ for $\mu$ furthermore. Denote by $u_L$ the orthogonal projection of $u$ on $L$, then

$$\int_{|u| \leq 1} |u_L| \mu(du) < \infty$$

and under the assumptions $H_1, H_2$ the following function is well defined and is Lipschitz continuous:

$$\bar{b}(x) = b(x) - \int_{|u| \leq 1} \sigma(x) u_L \mu(du) - \int_{|u| \leq 1} r(x, u) \mu(du).$$

Denote by $F_{0,T}^{\text{const}}$ and $F_{0,T}^{\text{step}}$ the following classes of functions $f : [0, T] \to \mathbb{R}^d$:

- any $f \in F_{0,T}^{\text{const}}$ takes a constant value in the orthogonal complement $L^\perp$ of the subspace $L$ in $\mathbb{R}^d$;
- for any $f \in F_{0,T}^{\text{step}}$ there exists a partition $[0, T] = [0, \tau_1] \cup [\tau_1, \tau_2] \cup \cdots \cup [\tau_j, T]$ such that, on each interval of the partition, $f$ takes a constant value in $L^\perp$.

Finally, denote by $S_{0,T,x_0}^{\text{const}}$ and $S_{0,T,x_0}^{\text{step}}$ the classes of the solutions to the Cauchy problem for the following piece-wise ordinary differential equations

$$\phi_t = x_0 + \int_0^t \left( \bar{b}(\phi_s) + \sigma(\phi_s) f_s \right) ds + \sum_{t_k \leq t} \Delta t_k \phi,$$

where $\{t_k\} \subset (0, T)$ is an arbitrary finite set, $f$ is an arbitrary function from the class $F_{0,T}^{\text{const}}$ or $F_{0,T}^{\text{step}}$, respectively, and the jumps of the function $\phi$ at the time moments $\{t_k\}$ satisfy

$$(\phi_{t_k^-}, \phi_{t_k}) \in \mathcal{A} \text{ for any } k.$$  

Denote by $S_{0,T,x_0}^{\text{const}}$ and $S_{0,T,x_0}^{\text{step}}$ the closures of these classes in $D([0, T], \mathbb{R}^m)$. 
Theorem 2.1. Assume $H_1, H_2$. Then for any $T > 0$
\[ \text{supp} \left( \text{Law}_{x_0,T}(X) \right) = S_{0,T,x_0}^{\text{const}} = S_{0,T,x_0}^{\text{step}}. \]

In plain words the above description of the support can be explained as follows. The SDE (1) contains two parts, the ‘deterministic flow’ part which corresponds to the drift, and the ‘stochastic jump part’ which corresponds to the PPM. These two parts are still related through the compensator term, which is involved in the stochastic integral w.r.t. $\tilde{N}(du,dT)$. In the simple case where $L = \mathbb{R}^d$, and thus the SDE is of the Type I, these parts can be completely separated, and (1) can be written in the form
\[ dX_t = \tilde{b}(X_t) dt + \int_{\mathbb{R}^d} c(X_{t^-}, u) N(du, dt), \]
where the effective drift coefficient is defined by $\tilde{b}(x) = b(x) - \int_{|u|\leq 1} c(x, u) \mu(du)$. In this case (5) does not contain the part with $f$ because $L^\perp = \{0\}$, and the description of the support of the law of the solution $X$ is intuitively clear: the solution follows the deterministic flow defined by the effective drift, then makes a jump, admissible for the stochastic jump part, then again follows the deterministic flow, etc. This description was thoroughly proved in [10, Theorem I]. The general (Type II) case is more sophisticated, since the compensator term cannot be separated from the stochastic integral. Theorem 2.1 actually tells us that the above description of the support remains essentially correct, with the following two important changes:

- the effective drift involves only the parts of $c(x, u)$ which are absolute integrable, i.e. $r(x, u)$ and $\sigma(x)u_L$;
- the ‘non-integrability subspace’ $L^\perp$ induces an extra drift part, which may act at arbitrary direction of the form $\sigma(x)h, h \in L^\perp$.

Such a description of the support was conjectured by T. Simon, see [10, Remark 3.3(c)], by analogy with a similar result proved for Lévy processes in [11]. Theorem 2.1 proves this conjecture in a wide generality, i.e. without any specific assumptions on the SDE except the natural conditions $H_1, H_2$ which yield strong existence of the solution.

Remark 2.2. Our description of admissible jumps differs from the one adopted in [10], which requires that
\[ y = x + c(x, u), \quad u \in \text{supp}(\mu). \]
These two descriptions coincide for $c(x, u)$ continuous in $u$; for discontinuous $c(x, u)$ the latter one is no longer applicable. To see this, one can consider a simple example where $m = d = 1$, $\mu$ is a finite discrete measure $\mu$ which has a full support in $\mathbb{R}$, $m = 1$, and $c(x, u) = 1_K(u) + 21_{\mathbb{R}\setminus K}(u)$, where $K$ is a countable dense subset in $\mathbb{R}$ such that $\mu(\mathbb{R}\setminus K) = 0$.

2.3. Discussion: applications and examples.
Support theorems are involved naturally in the study of ergodic properties of the Markov process associated to the SDE; namely, they provide a natural tool for proving that the process is topologically irreducible, see [2, Proof of Proposition 5.3] or [4, Theorem 1.3 and Section 4.3]. This gives a natural application field for Theorem 2.1. Namely, denote by $S_T(x_0)$ the closure of the set of the values $\phi_T$, where $\phi$ runs through the solutions to (5) with arbitrary finite set $\{t_k\} \subset (0, T)$, jumps of $\phi$
satisfying (6), and arbitrary function $f$ from the class $\mathbf{F}^{\text{step}}_{0,T}$. By Theorem 2.1, the solution to (1) satisfies

$$\text{supp} \left( \text{Law} \left( X_T \right) \right) = S_T(x_0).$$

Thus, in order to provide the topological irreducibility for the Markov process associated to (1), it is sufficient to show that

$$S_T(x) = \mathbb{R}^m$$

for all $x \in \mathbb{R}^m$. (8)

Another application of the support theorems dates back to the original paper by Stroock and Varadhan [15], and concerns the strong maximal principle for the generator $L$ of the Markov process $X$, which states that any sub-harmonic function for $L$ which reaches its maximum on the given set is constant on this set. We refer an interested reader for a detailed discussion to [3, Theorem IV.8.3] or [10, Remark 3.3(b)]. Here, we mention briefly the following simple corollary: if $\bigcup_{T > 0} S_T(x)$ is dense in $\mathbb{R}^m$ for any $x \in \mathbb{R}^m$, then the strong maximal principle for $L$ holds true on the entire $\mathbb{R}^m$. For this property to hold, it is clearly sufficient that (8) holds true for some $T > 0$.

Below, we give two simple sufficient conditions where (8) holds true for any $T > 0$.

Example 2.3. (Jump noise satisfying the ‘cone condition’). Let there exist $\theta \in (0, 1)$ such that, for any $\ell \in \mathbb{R}^d$, $|\ell| = 1$ and $\varepsilon > 0$, the intersection of the cone $\{u : u \cdot \ell \geq \theta |u|\}$ with the ball $\{|u| < \varepsilon\}$ has positive measure $\mu$. Assume also that $\sigma(x)$ is point-wise degenerate, i.e. $\text{rank} \sigma(x) = m \leq d, x \in \mathbb{R}^m$. Then $S_T(x) = \mathbb{R}^m$ for all $x \in \mathbb{R}^m, T > 0$. To prove this assertion, one can take $f \equiv 0$ and organize for given $x, y \in \mathbb{R}^m, \varepsilon > 0$ and $T > 0$ sequences of (frequent) jump times $\{t_k\}$ and (small) jumps amplitudes $u_k$ which would force the solution to (5) with $x_0 = x$ to take the final value $y$ with $|\phi_T - y| < \varepsilon$. This construction is essentially the same as in the proof of [7, Proposition 4.17], thus we omit the details here.

Example 2.4. (Jump noise of the ‘strong Type II’). Let there be no integrability directions for $\mu(du)$; that is, $L = \{0\}$. Assume also that $\sigma(x)$ is point-wise degenerate. Then $S_T(x) = \mathbb{R}^m$ for all $x \in \mathbb{R}^m, T > 0$. To prove this assertion, one can ignore the jump part in (5) and consider the solution to the SDE

$$d\phi_t = \tilde{b}(\phi_t) \, dt + \sigma(\phi_t) \, f_t \, dt, \quad \phi_0 = x.$$  

Since $f$ can be arbitrary piece-wise constant function $[0, T] \rightarrow \mathbb{R}^d$ and $\sigma(x)$ is point-wise degenerate, by a proper choice of $f$ respective solution $\phi$ can be made arbitrarily close to

$$\phi_T^{x,y,T} = x + \frac{t}{T}(y - x).$$  

Let us give two more particular examples illustrating the range of applicability of Theorem 2.1 and sufficient conditions from Example 2.3 and Example 2.4. For that purpose, we recall the auxiliary ‘scaling’ assumption imposed in [10], [4]: for some $\alpha \in (0, 2)$,

$$\int_{|u| \leq \varepsilon} (u \cdot \ell)^2 \mu(du) \asymp \varepsilon^{2-\alpha}, \quad |\ell| = 1.$$  

(9)

The following example concerns the cylindrical Lévy processes, which have been studied extensively in the last decades, e.g. [5], [6], [9].
Example 2.5. Let $\mu(du)$ be the Lévy measure of the $Z = (Z^1, \ldots, Z^d)$ with the independent components $Z^i, i = 1, \ldots, d$. Let the components be symmetric $\alpha_i$-stable processes on $\mathbb{R}$, then the scaling assumption (9) fails unless all the stability indices $\alpha_i, i = 1, \ldots, d$ are the same; that is, [10, Theorem II] is not applicable. If at least one $\alpha_i > 1$, then this is the Type II Lévy noise and [10, Theorem I] is not applicable, as well. On the other hand, Theorem 2.1 can be applied assuming the coefficients satisfy $H_1, H_2$; note that $\beta$ can be taken arbitrary $> \max_i \alpha_i$. The Lévy measure $\mu(du)$ is supported by the collection of the coordinate axes in $\mathbb{R}^d$ and thus satisfies the ‘cone condition’ from Example 2.3. Hence, if $\sigma(x)$ is point-wise degenerate, then identity (8) holds true for any $T > 0$.

Our second example shows that Lévy measure can be of a ‘strong Type II’ even if even the (rather mild) ‘cone condition’ fails.

Example 2.6. Let the Lévy measure $\mu(du)$ on $\mathbb{R}^2$ be the image of the symmetric $\alpha$-stable Lévy measure $\frac{dz}{|z|^{1+\alpha}}$ on $\mathbb{R}$ under the mapping $z \mapsto (z, |z|^{\gamma} \text{sgn } z)$ with some $\gamma > 1$. Then the scaling condition (9) fails, and for $\alpha > 1$ (i.e. when the noise is of Type II) the previous results are not applicable. On the other hand, for $\gamma \leq \alpha$ the noise has strong Type II. Hence, if $\sigma(x)$ is point-wise degenerate, then identity (8) holds true for any $T > 0$.

3. Proof of Theorem 2.1

To prove the required statement, it is sufficient to prove the following two inclusions:

$$\text{supp} \left( \text{Law}_{x_0,T}(X) \right) \subset S_{\text{const}}^{0,T,x_0}, \quad S_{\text{step}}^{0,T,x_0} \subset \text{supp} \left( \text{Law}_{x_0,T}(X) \right).$$

The proof of the first inclusion is simple and standard, corresponding argument was explained in [10]; for the reader’s convenience we outline the argument here. Consider a family of SDEs

$$dX_t^\eta = b_\eta(X_t^\eta) \, dt + \int_{|u| \leq 1} c(X_t^\eta, u) \tilde{N}(du, dt) + \int_{|u| > 1} c(X_t^\eta, u) N(du, dt),$$

with $X_0^\eta = x_0$, where

$$b_\eta(x) = b(x) - \int_{|u| < \eta} \sigma(x) u L \mu(du) - \int_{|u| < \eta} r(x, u) \mu(du).$$

It is easy to check that $b_\eta \to b, \eta \to 0$ uniformly on compact subsets of $\mathbb{R}^m$. Then by the usual stochastic calculus technique one can show that

$$\sup_{t \in [0,T]} |X_t^\eta - X_t| \to 0, \quad \eta \to 0$$

in probability; see the proof of a similar statement in Lemma 4.2 below. On the other hand, equation (11) can be written in the form

$$dX_t^\eta = \tilde{b}(X_t^\eta) \, dt + \sigma(X_t^\eta) \nu_\eta \, dt + \int_{|u| \geq \eta} c(X_t^\eta, u) N(du, dt),$$

with $X_0^\eta = x_0$, where

$$\tilde{b}(x) = b(x) - \int_{|u| < \eta} \sigma(x) u L \mu(du) - \int_{|u| < \eta} r(x, u) \mu(du).$$

It is easy to check that $\tilde{b} \to b, \eta \to 0$ uniformly on compact subsets of $\mathbb{R}^m$. Then by the usual stochastic calculus technique one can show that

$$\sup_{t \in [0,T]} |X_t^\eta - X_t| \to 0, \quad \eta \to 0$$

in probability; see the proof of a similar statement in Lemma 4.2 below.
where

\[ u_\eta = \int_{\eta \leq |u| \leq 1} (u - u_L) \mu(du) \in L^1. \]  

(13)

The PPM \( N \), restricted to \( \{|u| \geq \eta, t \in [0,T]\} \), has a finite set of atoms ("jumps"). Then the solution to the latter equation can be represented path-wise as a collection of solutions to ODEs of the form (5), where \( f \) belongs to \( F_{0,T}^{\text{const}} \), \( \{t_k\} \) are equal to the time instants of the jumps, and \( \Delta t_k \phi = c(\phi_{t_{k-}}, u_k) \), where \( \{u_k\} \) are equal to the amplitudes of the jumps for \( N \). Note that, for any \( t_k \), the value \( \phi_{t_k} \) is an image of \( x = \phi_{t_{k-}}, u = u_k \) under the mapping

\[(x, u) \mapsto x + c(x, u), \]

and thus the pair \( (\phi_{t_{k-}}, \phi_{t_k}) \) is admissible. Therefore, for any \( \eta \),

\[ \text{supp} \left( \text{Law}_{x_0,T}(X^\eta) \right) \subset S^{\text{const}}_{0,T;x_0} \iff \mathbb{P}(X^\eta \in [0,T] \in S^{\text{const}}_{0,T;x_0}) = 1. \]

By (12), the laws of \( X^\eta \) in \( D([0,T], \mathbb{R}^m) \) weakly converge to the law of \( X \), which gives for a closed set \( S^{\text{const}}_{0,T;x_0} \)

\[ \mathbb{P}(X \in [0,T] \in S^{\text{const}}_{0,T;x_0}) \geq \lim_{\eta \to 0} \mathbb{P}(X^\eta \in [0,T] \in S^{\text{const}}_{0,T;x_0}) = 1, \]

and completes the first inclusion in (10).

The second inclusion is the main part of the theorem. In order to proceed with its proof, we re-write the original SDE (1) to the form

\[ dX_t = \tilde{b}_\eta(X_t) \, dt + \sigma(X_t)u_\eta \, dt + \int_{|u| < \eta} c(X_{t-}, u) \tilde{N}(du, dt) \]

\[ + \int_{|u| \geq \eta} c(X_{t-}, u) N(du, dt), \]

where \( \eta > 0 \) is a (small) parameter which is yet to be chosen and

\[ \tilde{b}_\eta(x) = b(x) - \int_{\eta \leq |u| \leq 1} \sigma(x)u_L \mu(du) - \int_{\eta \leq |u| \leq 1} r(x, u) \mu(du), \]

recall that \( u_\eta \) is given by (13). Consider SDE

\[ dX^\eta_{t} = \tilde{b}_\eta(X^\eta_{t}) \, dt + \sigma(X^\eta_{t})v_\eta \, dt + \int_{|u| < \eta} c(X^\eta_{t-}, u) \tilde{N}(du, dt), \]

(14)

which can be seen as a modification of (14) with 'large jumps' being truncated, i.e. with the PPM \( N(du, dt) \) changed to its restriction to \( \{|u| < \eta\} \times [0,T] \). We will denote by \( X^\eta_{x,S;\eta,\text{trunc}} \) the solution of (15) with \( X_S = x \).

Let \( f \in F_{0,T}^{\text{step}}, x \in \mathbb{R}^m, S \in [0,T] \) be fixed and \( \phi^{x,S,f} \) be the solution of the ODE

\[ \phi_t = x + \int_{S}^{t} \left( \tilde{b}(\phi_s) + \sigma(\phi_s)f_s \right) ds, \quad t \in [S,T] \]

(16)

The following statement is the cornerstone of the entire proof. In what follows we denote by \( B(x, r) \) the open ball in \( \mathbb{R}^m \) with the center \( x \) and radius \( r \).
Lemma 3.1. (The Key Lemma) There exists $\rho \in (0,1)$ such that, for any given $f \in F_{0,T}^{\text{step}}, x \in \mathbb{R}^m,$ and $\gamma > 0$, there exists $\eta^{f,x,\gamma} > 0$ such that

$$p^{\text{trunc}}(\eta, f, x, \gamma) := \inf_{x' \in B(x,\rho \eta)} \mathbb{P} \left( \sup_{t \in [S,Q]} |X_t^{x',S,\eta,\text{trunc}} - \phi_t^{x,S,f}| \leq \gamma \right) > 0$$

for any $\eta \in (0, \eta^{f,x,\gamma})$. We postpone the proof of Lemma 3.1 to a separate Section 4; here we explain the argument (quite standard, though a bit cumbersome) which provides the second inclusion in (10) once this key lemma is proved.

Fix $\eta > 0$ and decompose

$$N(du,dt) = 1_{|u| < \eta} N(du,dt) + 1_{|u| \geq \eta} N(du,dt) =: N_\eta(du,dt) + N^\eta(du,dt).$$

The PPM $N^\eta(du,dt)$ a.s. has a finite number of atoms with $t \in [0,T]$; say, $\{(\xi_j, \tau_j)\}_j^{\tau_f}$. It is well known that the PPMs $N_\eta(du,dt)$ and $N^\eta(du,dt)$ are independent. In addition, the random variable

$$J = N(\{|u| \geq \eta\} \times [0,T])$$

has the Poisson distribution with the intensity $T\mu(|u| \geq \eta)$, and conditioned by the event $\{J = K\}$, the laws of the random vectors $\{\xi_j\}_{j=1}^K, \{\tau_j\}_{j=1}^K$ are independent. These laws are the $K$-fold product of

$$\mathbb{P}^\eta(du) := \frac{1}{\mu(\{|v| \geq \eta\})} \mu(du)$$

for $\{\xi_j\}_{j=1}^K$, and the uniform distribution on the simplex

$$\Delta_K(0,T) := \{(s_1, \ldots, s_K): 0 \leq s_1 \leq \cdots \leq s_K \leq T\}$$

for $\{\tau_j\}_{j=1}^K$. With this information in mind, it is easy to show that $X$ has a positive probability to stay in any neighbourhood of a given function $\phi \in S_{\text{step}}^{0,T,x_0}$.

Namely, let $\phi \in S_{\text{step}}^{0,T,x_0}$ be fixed with the corresponding function $f \in F_{0,T}^{\text{step}}$ and points $\{t_k, k = 1, \ldots, K\}$ from the equation (5) for $\phi$. Denote $x_k = \phi_{t_k}$, $y_k = \phi_{t_k}$; we can and will assume that $x_k \neq y_k$ because otherwise we can simply exclude the point $t_k$ from the equation (5). Denote also $t_0 = 0, t_{K+1} = T$, then for any positive

$$\delta < \delta_{\{t_k\}} := \frac{1}{2} \min_{k=0,\ldots,K} |t_{k+1} - t_k|$$

we have

$$p(\eta, \delta, \{t_k\}) := \mathbb{P}(J = K, |t_k - \tau_k| < \delta, k = 1, \ldots, K) = e^{-T\mu(|u| \geq \eta)} \left(2\delta \mu(|u| \geq \eta)\right)^K,$$

which is positive.

Process $X$ follows the truncated SDE (15) between the ‘big jumps instants’ $\tau_j$, and at these instants satisfies

$$\Delta_{\tau_j} X = c(X_{\tau_j}, \xi_j), \quad j = 1, \ldots, J.$$ 

Denote by $X^{s_1, \ldots, s_K}$ the similar process with $J = K$ and $\tau_k = s_k$; it follows the truncated SDE (15) between $s_k$, and at these time instants satisfies

$$\Delta_{s_k} X = c(X_{s_k}, \xi_k), \quad k = 1, \ldots, K;$$
the random vector $\{\xi_j\}_{j=1}^K$ has the law $(P^\eta)^\otimes K$. Then
\[
\mathbb{P}(X \in A) \geq \mathbb{P}(X \in A, J = K)
= \int_{\Delta_K(0,T)} \mathbb{P}(X^{s_1,\ldots,s_K} \in A) \mathbb{P}(J = K, (\tau_1, \ldots, \tau_k) \in ds_1 \ldots ds_K) \geq p(\eta, \delta, \{t_k\}) \inf_{|s_k - s_| < \delta, k = 1, \ldots, K} \mathbb{P}(X^{s_1,\ldots,s_K} \in A).
\]
We will apply this inequality with
\[
A = \{x(\cdot) : d(x(\cdot), \phi) \leq \varepsilon\},
\]
in this case $\mathbb{P}(X^{s_1,\ldots,s_K} \in A)$ can be bounded from below as follows. Denote $s_0 = 0$, $s_{K+1} = T$, and $y_0 = x_0$. Define a function $\phi^{s_1,\ldots,s_K}$ as follows: on each of the intervals $[s_k, s_{k+1})$, $k = 0, \ldots, K$ it satisfies (16) with $x = y_k$, $S = s_k$; it is also continuous at the end point $T = s_{K+1}$. Recall that the initial function $\phi_0$ follows similar description with $\{t_k\}$ instead of $\{s_k\}$. Using this observation, it is easy to show that
\[
\sup_{|s_k - s| < \delta, k = 1, \ldots, K} d(\phi^{s_1,\ldots,s_K}, \phi) \to 0, \quad \delta \to 0;
\]
for the reader’s convenience we prove this relation in Appendix A. Hence we can fix $\delta(\varepsilon) > 0$ such that
\[
\sup_{|s_k - s| < \delta, k = 1, \ldots, K} d(\phi^{s_1,\ldots,s_K}, \phi) \leq \frac{\varepsilon}{2}, \quad \delta \in (0, \delta(\varepsilon)),
\]
and thus for any $\{s_k\}$ with $|t_k - s_k| < \delta, k = 1, \ldots, K$
\[
\mathbb{P}(d(X^{s_1,\ldots,s_K}, \phi) \leq \varepsilon) \geq \mathbb{P}\left(d(X^{s_1,\ldots,s_K}, \phi^{s_1,\ldots,s_K}) \leq \frac{\varepsilon}{2}\right).
\]
Next, denote
\[
\mathcal{F}_t = \sigma(X^s_{s_1,\ldots,s_K}, s \leq t) = \sigma(N_\eta([0, s] \times du), s \leq t, \{\xi_k, s_k \leq t\}),
\]
and
\[
\mathcal{G}_t = \mathcal{F}_{t-} = \sigma(N_\eta([0, s] \times du), s \leq t, \{\xi_k, s_k < t\}),
\]
note that $\xi_k$ is independent on $\mathcal{G}_{s_k}$ for any $k$. The following holds:

(1) For any $k = 0, \ldots, K$ and $\gamma > 0$, the conditional probability w.r.t. $\mathcal{F}_{s_k}$ for the event

\[
\{|X^s_{s_1,\ldots,s_K} - \phi^{s_1,\ldots,s_K}_t| \leq \gamma, t \in [s_k, s_{k+1})\}
\]
equals
\[
\mathbb{P}\left(\sup_{t \in [s_k, s_{k+1})} |X^s_{t, s_k, \eta, t, \text{trunc}} - \phi^{s_1,\ldots,s_K}_t| \leq \gamma\right)\bigg|_{x=X^s_{s_k}}
\]
and by Lemma 3.1, is bounded from below by $p^{\text{trunc}}(\eta, f, y_k, \gamma)$ on the set
\[
\left\{|X^s_{s_k} - y_k| < \frac{\gamma}{2}\right\} \in \mathcal{F}_{s_k},
\]
provided that $\eta \leq \eta^{f, y_k, \gamma}$.
(II) for any \( k = 1, \ldots, K \) and \( \gamma > 0 \), the conditional probability w.r.t. \( \mathcal{G}_{s_k} \) for
the event
\[
\left\{ |X_{s_k}^{s_1, \ldots, s_K} - y_k| < \frac{\gamma}{2} \right\}
\]
equals
\[
P^n \left( |x + c(x, \xi) - y_k| < \frac{\gamma}{2} \right| x = X_{s_k}^{s_1, \ldots, s_K}
\]
equal
\[
\frac{1}{\mu(\{u : |v| \geq \eta\})} \mu \left( \left\{ u : |u| \geq \eta, |x + c(x, u) - y_k| < \frac{\gamma}{2} \right\} \right| x = X_{s_k}^{s_1, \ldots, s_K}.
\]

Recall that each pair \((x_k, y_k) = (\phi_{t_k}, \phi_{t_k})\) is admissible, hence for any \( \gamma > 0 \) and \( k \),
\[
J(x_k, B(y_k, \gamma)) = \mu(\{u : x_k + c(x, u) \in B(y_k, \gamma)\}) > 0.
\]
Take
\[
\gamma_* = \frac{1}{3} \min_k |x_k - y_k|,
\]
then by the assumptions \( \mathbf{H}_1, \mathbf{H}_2 \) there exists \( \eta^* > 0 \) such that, for all \( k \),
\[
|u| \leq \eta^* \implies |c(x_k, u)| < |x_k - y_k| - \gamma_* \implies x_k + c(x, u) \notin B(y_k, \gamma_*).
\]
This yields for any \( \gamma \in (0, \gamma_*] \)
\[
\mu(\{u : |u| > \eta^*, x_k + c(x, u) \in B(y_k, \gamma)\}) > 0, \quad k = 1, \ldots, K.
\]
Moreover, because \( c(x, u) \) is continuous w.r.t. \( x \) we have by the usual weak continuity arguments that, for each \( \gamma \in (0, \gamma_*] \) and \( k \), there exists \( \gamma' > 0 \) such that
\[
\inf_{x \in B(x_k, 2\gamma')} \mu(\{u : |u| > \eta^*, x + c(x, u) \in B(y_k, \gamma)\}) > 0.
\]
Now we can finalize the construction and the estimate. Define iteratively \( \gamma_k, \gamma'_k \) for \( k = K, \ldots, 1 \) as follows: Take
\[
\gamma_K = \min \left( \frac{\rho \varepsilon}{2}, \gamma_* \right)
\]
(with \( \rho \) given by Lemma 3.1) and define \( \gamma'_k \) such that (21) holds true with \( \gamma = \gamma_K \). Once \( \gamma_{k+1}, \gamma'_k \) are defined, take
\[
\gamma_k = \min \left( \rho \gamma_{k+1}, \gamma_{k+1} \right)
\]
and define \( \gamma'_k \) such that (21) holds true with \( \gamma = \gamma_{k+1} \).

It follows from the calculations in Appendix A that \( \delta \in (0, \delta(\varepsilon)) \) can be taken small enough such that, for any \( s_1, \ldots, s_K \) with \( |t_k - s_k| < \delta, k = 1, \ldots, K \),
\[
|\phi_{s_k}^{s_1, \ldots, s_K} - x_k| < \gamma'_k.
\]
We fix such \( \delta > 0 \) and the truncation level
\[
\eta = \min\{\eta_*, \eta'_k, \gamma'_k, \gamma_{k+1}, k = 0, \ldots, K\}.
\]
Denote
\[
A_k = \left\{ \sup_{t \in [s_k, s_{k+1}]} |X_{t}^{s_1, \ldots, s_K} - \phi_t^{s_1, \ldots, s_K}| \leq \gamma'_k \right\}, \quad k = 0, \ldots, K - 1,
\]
\[
A_K = \left\{ \sup_{t \in [s_K, T]} |X_{t}^{s_1, \ldots, s_K} - \phi_t^{s_1, \ldots, s_K}| \leq \frac{\varepsilon}{2} \right\},
\]
and

\[ B_k = \left\{ |X_{s_k}^{s_1, \ldots, s_K} - y_k| < \gamma_k \right\}, \quad k = 1, \ldots, K. \]

Then \( A_k \in \mathcal{G}_{s_k+1}, B_k \in \mathcal{F}_{s_k}, k = 1, \ldots, K, \) and we have the following:

- Since \( \gamma_k \leq \rho \gamma_k' \) for any \( k = 1, \ldots, K - 1, \) by Lemma \ref{lem:3.1} we have

\[ \mathbb{P}(A_k | \mathcal{F}_{s_k}) \geq p^\text{trunc}(\eta, f, y_k, \gamma_k') \text{ a.s. on the set } B_k. \]

We also have by Lemma \ref{lem:3.1}

\[ \mathbb{P}(A_1) \geq p^\text{trunc}(\eta, f, y_0, \gamma_1') \]

and

\[ \mathbb{P}(A_K | \mathcal{F}_{s_K}) \geq p^\text{trunc} \left( \eta, f, y_K, \frac{\varepsilon}{2} \right) \text{ a.s. on the set } B_K. \]

- By \eqref{eq:22}, for any \( k = 1, \ldots, K \) we have

\[ |X_{s_k}^{s_1, \ldots, s_K} - x_k| < 2\gamma_k' \text{ on the set } A_k - 1. \]

Then

\[ \mathbb{P}(B_k | \mathcal{G}_{s_k}) \geq p^\text{jump}_k(\eta) := \frac{1}{\mu(\{u : |u| \geq \eta\})} \inf_{x \in B(x_k, 2\gamma_k')} \mu(\{u : |u| > \eta, x + c(x, u) \in B(y_k, \gamma_k)\}), \]

and since \( \eta \leq \eta_k, k = 1, \ldots, K \) we have by \eqref{eq:21}

\[ p^\text{jump}_k(\eta) > 0, \quad k = 1, \ldots, K. \]

By the construction, we have \( \gamma_k, \gamma_k' \leq \frac{\varepsilon}{2} \) for all \( k = 1, \ldots, K. \) Then by the telescopic property of the conditional expectations we have for any \( s_1, \ldots, s_K \) with \( |t_k - s_k| < \delta, k = 1, \ldots, K \)

\[ \mathbb{P} \left( \sup_{t \in [0,T]} |X_{t}^{s_1, \ldots, s_K} - \phi_{t}^{s_1, \ldots, s_K}| \leq \frac{\varepsilon}{2} \right) \geq \mathbb{P}(A_0 \cap B_1 \cap A_1 \cap \cdots \cap B_K \cap B_K) \]

\[ \geq p^\text{trunc} \left( \eta, f, y_K, \frac{\varepsilon}{2} \right) \prod_{k=1}^{K} \left( p^\text{trunc}(\eta, f, y_{k-1}, \gamma_k') p^\text{jump}_k(\eta) \right) =: q(\eta, \delta) > 0. \]

Then by \eqref{eq:19}

\[ \mathbb{P}(d(X^{s_1, \ldots, s_K}, \phi) \leq \varepsilon) \geq \mathbb{P} \left( d(X^{s_1, \ldots, s_K}, \phi^{s_1, \ldots, s_K}) \leq \frac{\varepsilon}{2} \right) \]

\[ \geq \mathbb{P} \left( \sup_{t \in [0,T]} |X_{t}^{s_1, \ldots, s_K} - \phi_{t}^{s_1, \ldots, s_K}| \leq \frac{\varepsilon}{2} \right) \]

\[ \geq q(\eta, \delta). \]

and by \eqref{eq:17}

\[ \mathbb{P}(d(X, \phi) \leq \varepsilon) \geq p(\eta, \delta, \{t_k\}) q(\eta, \delta). \]

In these estimates, the choice of \( \delta, \eta \) depends on \( \phi \in S^{\text{step}}_{0,T,x_0} \) and \( \varepsilon > 0, \) only, hence the proof of \eqref{eq:10} is complete.
4. PROOF OF THE KEY LEMMA

We begin the proof of Lemma 3.1 with the following auxiliary result.

**Lemma 4.1.** For any \( w \in L^\perp \) and \( \eta > 0 \) there exist \( \zeta \in (0, \eta) \) and a function \( g : \mathbb{R}^d \rightarrow [-\frac{1}{2}, \frac{1}{2}] \) such that \( g(u) = 0 \) whenever either \( |u| \leq \zeta \) or \( |u| \geq \eta \) and
\[
\int_{\mathbb{R}^d} (u - u_L) g(u) \, \mu(du) = w.
\]

**Proof.** For a given \( 0 < \zeta < \eta \), denote by \( G^\eta_\zeta \) the set of all functions \( g : \mathbb{R}^d \rightarrow [-\frac{1}{2}, \frac{1}{2}] \) such that \( g(u) = 0 \) whenever either \( |u| \leq \zeta \) or \( |u| \geq \eta \). This set is convex and symmetric; hence
\[
V^\eta_\zeta = \left\{ \int_{\mathbb{R}^d} (u - u_L) g(u) \, \mu(du), g \in G^\eta_\zeta \right\}
\]
is a symmetric convex subset of \( L^\perp \), and so is the set
\[
V^\eta_0 = \bigcup_{\zeta \in (0, \eta)} V^\eta_\zeta.
\]
The statement of the lemma is equivalent to
\[
V^\eta_0 = L^\perp. \tag{23}
\]
Assuming (23) to fail, we have that \( V^\eta_0 \) is a proper symmetric convex subset of \( L^\perp \), and thus there exist \( \ell \in L^\perp \setminus \{0\} \) and \( c > 0 \) such that
\[
-c \leq v \cdot \ell \leq c, \quad v \in V^\eta_0; \tag{24}
\]
recall that \( v \cdot \ell \) denotes the scalar product in \( \mathbb{R}^d \). It follows from (24) that, for every \( \zeta \in (0, \eta) \) and \( g \in G^\eta_\zeta \),
\[
\left| \int_{\mathbb{R}^d} \ell \cdot u g(u) \, \mu(du) \right| = \left| \int_{\mathbb{R}^d} \ell \cdot (u - u_L) g(u) \, \mu(du) \right| = \left| \ell \cdot \int_{\mathbb{R}^d} (u - u_L) g(u) \, \mu(du) \right| \leq c,
\]
in the second identity we have used that \( u_L \in L \) is orthogonal to \( \ell \in L^\perp \). Taking
\[
g^\eta_\zeta(u) = \frac{1}{2} \text{sign} (\ell \cdot u) 1_{\zeta < |u| < \eta},
\]
we get from the previous inequality that
\[
\int_{\zeta < |u| < \eta} |\ell \cdot u| \, \mu(du) \leq 2c, \quad \zeta \in (0, \eta),
\]
and passing to the limit as \( \zeta \to 0 \) we obtain
\[
\int_{|u| < \eta} |\ell \cdot u| \, \mu(du) \leq 2c < +\infty.
\]
This means that \( \ell \in L \), which contradicts to the fact that \( \ell \in L^\perp \setminus \{0\} \). This contradiction shows that \( V^\eta_0 \) should coincide with entire \( L^\perp \), which completes the proof. \( \square \)
Similarly to the notation used in Lemma 3.1, we denote by $Y_t$ hence taking (25) into account we can write this equation in the form

$$
\int_{\mathbb{R}^+} (u - u_L) g_{t,n}(u) \mu(du) = v_\eta - f_t, \quad t \in [0, T]
$$

(25)

with $v_\eta \in L^1$ given by (13), and, for some $\zeta_{t,n} > 0$, one has $g_{t,n}(u) = 0$ whenever either $|u| \geq \eta$ or $|u| \leq \zeta_{t,n}$.

We write the compensated PPM in (15) in the form $N(du, dt) - \mu(du)dt$ and consider the same SDE with another PPM $Q_{t,n}(du, dt)$ which has the intensity measure $(1 + g_{t,n}(u))\mu(du)dt$:

$$
dY_t^n = \tilde{b}_n(Y_t^n) dt + \sigma(Y_t^n) v_\eta dt + \int_{|u| < \eta} c(Y_{t-}, u)(Q_{t,n}(du, dt) - \mu(du)dt).
$$

(26)

Similarly to the notation used in Lemma 3.1 we denote by $Y^{x,S,\eta}_{t} , t \geq S$ the solution to (26) with $Y^\eta_{S} = x$. The following lemma is based on quite standard stochastic calculus estimates.

**Lemma 4.2.** There exists $\rho \in (0, 1)$ such that, for arbitrary $\gamma \in (0, 1], x \in \mathbb{R}^d$,

$$
\inf_{x' \in B(x,\rho\gamma), 0 \leq S \leq T} P\left( \sup_{t \in [S, T]} |X^{x',S,\eta}_{t} - \phi^{x',S,\eta}_{t}| \leq \gamma \right) \to 1, \quad \eta \to 0.
$$

**Proof.** For simplicity of notation we take $S = 0, Q = T$ and omit the index $S$. We also consider the scalar case $d = 1$; for $d > 1$ similar estimates should be performed coordinate-wise.

The stochastic integral part in the equation (26) can be written as

$$
\int_{|u| < \eta} c(Y_{t-}, u) \tilde{Q}_{t,n}(du, dt) - \int_{|u| < \eta} c(Y_{t-}, u) g_{t,n}(u) \mu(du)dt,
$$

hence taking (25) into account we can write this equation in the form

$$
dY_t^n = \tilde{b}_{t,n}(Y^n_t) dt + \sigma(Y^n_t) f_t dt + \int_{|u| < \eta} c(Y_{t-}, u) \tilde{Q}_{t,n}(du, dt),
$$

(27)

where

$$
\tilde{b}_{t,n}(x) = \tilde{b}_n(x) - \int_{|u| < \eta} \sigma(x) u_L g_{t,n}(u) \mu(du) - \int_{|u| < \eta} r(x, u) g_{t,n}(u) \mu(du)
$$

$$
= \tilde{b}(x) + \int_{|u| < \eta} \sigma(x) u_L (1 - g_{t,n}(u)) \mu(du) + \int_{|u| < \eta} r(x, u) (1 - g_{t,n}(u)) \mu(du).
$$

Since the functions $u_L, |u|^\beta$ are integrable w.r.t. $\mu(du)$ on $\{|u| \leq 1\}$,

$$
|1 - g_{t,n}(u)| \leq \frac{3}{2},
$$

and assumptions $H_1, H_2$ hold, we have that

$$
\Delta^K_\eta := \sup_{x \in K, t \in [0, T]} |\tilde{b}_{t,n}(x) - \tilde{b}(x)| \to 0, \quad \eta \to 0
$$

for any compact subset $K \subset \mathbb{R}^d$. Denote $\phi_t = \phi^{x,f}$ and take

$$
K = \text{closure}\left( \{\phi_s, s \in [0, T]\} \right), \quad K' = \{y : \text{dist}(y, K) \leq 1\}.
$$
By the assumptions $H_1, H_2$, there exists a constant $C_{K'}$ such that
$$|c(y, u)| \leq C_{K'}|u|, \quad |u| \leq 1, \quad y \in K'.$$

Denote
$$\tau_{K'} = \inf\{t : Y^{x', \eta}_t \notin K'\}$$
with the usual convention $\inf \emptyset = T$. We have
$$Y^{x', \eta}_t - \phi_t = x' - x + \int_0^t \left( \tilde{b}_{t, \eta}^{x, \eta}(\phi_s) - \tilde{b}(\phi_s) \right) ds + \int_0^t \left( \tilde{b}_{t, \eta}^{x, \eta}(Y^{x', \eta}_s) - \tilde{b}_{t, \eta}^{x, \eta}(\phi_s) \right) ds + \int_0^t \int_{|u| < \eta} c(Y^{x', \eta}_s, u) \tilde{Q}_f^{x, \eta}(du, ds).
$$

By the Doob maximal inequality and the Itô isometry, for any $\varepsilon > 0$
$$\mathbb{P} \left( \sup_{t \in [0, \tau_{K'}]} \int_0^t \int_{|u| < \eta} c(Y^{x, \eta}_{s-}, u) \tilde{Q}_f^{x, \eta}(du, ds) \geq \varepsilon \right)$$
$$\leq \frac{1}{\varepsilon^2} E \left( \int_0^{\tau_{K'}} \int_{|u| < \eta} c(Y^{x, \eta}_{s-}, u) \tilde{Q}_f^{x, \eta}(du, ds) \right)^2$$
$$= \frac{1}{\varepsilon^2} E \int_0^{\tau_{K'}} \int_{|u| < \eta} c(Y^{x, \eta}_{s-}, u)^2 (1 + g_{y u}^{x, \eta}(u)) \mu(du) ds$$
$$\leq C' \frac{3T}{2} \int_{|u| < \eta} |u|^2 \mu(du) \to 0, \quad \eta \to 0;$$
in the last inequality we have used that $1 + g_{y u}^{x, \eta}(u) \leq \frac{3}{2}$. It is easy to check that the functions $\tilde{b}_{t, \eta}^{x, \eta}$ are uniformly Lipschitz, i.e. there exists $L$ such that, for any $t \in [0, T], \eta \in (0, 1]$
$$|\tilde{b}_{t, \eta}^{x, \eta}(x) - \tilde{b}_{t, \eta}^{x, \eta}(y)| \leq L|x - y|.$$

Then on the set
$$A_{\varepsilon, \eta} := \left\{ \sup_{t \in [0, \tau_{K'}]} \int_0^t \int_{|u| < \eta} c(Y^{x, \eta}_{s-}, u) \tilde{Q}(du, ds) < \varepsilon \right\}$$
we have
$$|Y^{x', \eta}_t - \phi_t| \leq |x' - x| + T\Delta_{\eta}^K + \varepsilon + L \int_0^t |Y^{x', \eta}_s(\phi_s) - \phi_s| ds, \quad t \in [0, \tau_{K'}],$$
which by the Gronwall inequality yields
$$\sup_{t \in [0, \tau_{K'}]} |Y^{x', \eta}_t - \phi_t| \leq \left( |x' - x| + T\Delta_{\eta}^K + \varepsilon \right) e^{LT}. \quad (28)$$

Take
$$\rho = \frac{1}{2} e^{LT},$$
and $\varepsilon, \eta$ small enough for
$$\left( T\Delta_{\eta}^K + \varepsilon \right) e^{LT} < 1.$$
Then, for any $\gamma \in (0, 1]$ and $x'$ with $|x' - x| < \rho \gamma$ we have on the set $A_{\varepsilon, \eta}$
$$\sup_{t \in [0, \tau_{K'}]} |Y^{x', \eta}_t - \phi_t| < 1.$$
Because trajectories of \( Y^{x',\eta} \) are right continuous this yields that, on this set, \( \tau_{K'} = T \) and actually
\[
\sup_{t \in [0,T]} |Y_t^{x',\eta} - \phi_t| \leq \left( |x' - x'| + T\Delta^K_\eta + \varepsilon \right)e^{LT}.
\] (29)

Now, for a given \( \gamma \), we take \( \varepsilon < \frac{1}{4} \gamma e^{-LT} \) and we get that for \( \eta \) small enough
\[
\inf_{x' \in B(x,\gamma)} \mathbb{P}(\sup_{t \in [0,T]} |Y_t^{x',\eta} - \phi_t| \leq \gamma) \geq \mathbb{P}(A_{\varepsilon,\eta}),
\]
and because
\[
\mathbb{P}(A_{\varepsilon,\eta}) \to 1, \ \eta \to 0
\]
this completes the proof. \( \square \)

Now we are ready to complete the proof of the Key Lemma. Denote
\[
\mathcal{E}^{f,\eta} := \exp \left( -\int_{\mathbb{R}^d \times [0,T]} \log(1 + g^f_\eta(u))\widetilde{Q}^{f,\eta}(du, dt) \right.
\]
\[
+ \int_{\mathbb{R}^d \times [0,T]} \left( g^f_\eta(u) - \log(1 + g^f_\eta(u)) \right) \mu(du)dt \right)
\]
recall that by the construction \( 1 + g^f_\eta(u) \in [\frac{1}{2}, \frac{3}{2}] \) and \( 1 + g^f_\eta(u) = 1 \) whenever either \( |u| \geq \eta \) or \( |u| \leq \zeta f,\eta \). Hence the stochastic integral under the exponent is well defined in the standard (quadratic) Itô sense. The second (deterministic) integral is easily bounded; hence it is easy to see that
\[
\mathbb{P}\left( \mathcal{E}^{f,\eta} < \frac{1}{N} \right) = \mathbb{P}(\log \mathcal{E}^{f,\eta} < -\log N) \to 0, \ \ N \to \infty
\] (30)
for any \( \eta \in (0, 1] \).

On the other hand, the classical result by Skorokhod [13] tells us that the laws of the PPMs \( N(du, dt) \) and \( Q^{f,\eta}(du, dt) \) are equivalent, and \( \mathcal{E}^{f,\eta} \) equals to the Radon-Nikodym derivative of \( \text{Law}(N) \) w.r.t. \( \text{Law}(Q^{f,\eta}) \) evaluated at \( Q^{f,\eta} \). Recall that \( X^{x',S,\eta,\text{trunc}} \) and \( Y^{x',S,\eta} \) are defined as the strong solutions to the same SDE with the noises \( N \) and \( Q^{f,\eta} \), respectively. Hence we can treat them as images of \( N, Q^{f,\eta} \) under some measurable mapping, which yields the identity
\[
\mathbb{P}\left( \sup_{t \in [S,Q]} |X_t^{x',S,\eta,\text{trunc}} - \phi^f_t,S| \leq \gamma \right) = \mathbb{E}1_{\sup_{t \in [S,Q]} |Y_t^{x',S,\eta} - \phi^f_t,S| < \varepsilon} \mathcal{E}^{f,\eta}.
\]

By Lemma 4.2 for a given \( x \in \mathbb{R}^d, \gamma > 0 \), and \( f \in \mathbb{F}^{\text{step}}_{0,T} \), there exists \( \eta^{f,x,\gamma} > 0 \) such that
\[
\inf_{x' \in B(x,\gamma \rho), 0 \leq S \leq Q \leq T} \mathbb{P}\left( \sup_{t \in [S,Q]} |Y_t^{x',S,\eta} - \phi^f_t,S| \leq \gamma \right) \geq \frac{2}{3}
\] (31)
for any \( \eta \in (0, \eta^{f,x,\gamma}] \). By (30), for any such \( \eta \) there exists \( N^{f,\eta} \) large enough for
\[
\mathbb{P}(\mathcal{E}^{f,\eta} \geq \frac{1}{N^{f,\eta}}) \geq \frac{2}{3}.
\]
Then for any \( x' \in B(x,\gamma \rho), 0 \leq S \leq Q \leq T \) we have
\[
\mathbb{P}\left( \sup_{t \in [S,Q]} |Y_t^{x',S,\eta} - \phi^f_t,S| \leq \gamma, \mathcal{E}^{f,\eta} \geq \frac{1}{N^{f,\eta}} \right) \geq \frac{1}{3}
\]
and
\[
P \left( \sup_{t \in [S,Q]} |X_t^{x',S,\eta,\text{trunc}} - \phi_t^{x,S,f}| \leq \gamma \right) \\
\geq \frac{1}{N_f,\eta} P \left( \sup_{t \in [S,Q]} |X_t^{x',\eta} - \phi_t^{x,S,f}| \leq \gamma, \phi^{f,\eta} \geq \frac{1}{N_f,\eta} \right) \\
\geq \frac{1}{3N_f,\eta} > 0,
\]
which completes the proof. \(\square\)

**Appendix A. Proof of (18)**

For given \(s_1, \ldots, s_K\) define \(\lambda^{s_1,\ldots,s_K}_j \in \Lambda_0,T\) as a function, linear on each of the intervals \([0, t_1), \ldots, [t_K, T]\), and satisfying
\[
\lambda^{s_1,\ldots,s_K}_j = s_j, \quad j = 1, \ldots, K.
\]

Denote
\[
\tilde{\phi}^{s_1,\ldots,s_K}_t = \phi^{s_1,\ldots,s_K}_t, \quad f^{s_1,\ldots,s_K}_t = \phi^{s_1,\ldots,s_K}_t, \quad t \in [0, T].
\]

It is clear that, for a fixed set \(\{t_k\}\),
\[
\sup_{|t_k - s_k| < \delta, k = 1, \ldots, K} \frac{d\lambda^{s_1,\ldots,s_K}_t}{dt} \to 0, \quad \delta \to 0, \quad (32)
\]
and, for a given \(f \in F_{0,T}^{\text{step}}\),
\[
\sup_{|t_k - s_k| < \delta, k = 1, \ldots, K} \|f^{s_1,\ldots,s_K} - f\|_{L_1(0,T)} \to 0, \quad \delta \to 0. \quad (33)
\]

By (32) we have
\[
\sup_{|t_k - s_k| < \delta, k = 1, \ldots, K} \sup_{0 \leq s \leq t \leq T} \left| \log \frac{\lambda^{s_1,\ldots,s_K}_t - \lambda^{s_1,\ldots,s_K}_s}{t - s} \right| \to 0, \quad \delta \to 0,
\]
hence in order to prove (18) it is sufficient to prove that
\[
\sup_{|t_k - s_k| < \delta, k = 1, \ldots, K} \sup_{t \in [0,T]} |\tilde{\phi}^{s_1,\ldots,s_K}_t - \phi_t| \to 0, \quad \delta \to 0. \quad (34)
\]

The functions \(\tilde{\phi}^{s_1,\ldots,s_K}, \phi\) have the same set \(\{t_k\}\) of discontinuity points. Moreover,
\[
\tilde{\phi}^{s_1,\ldots,s_K}_k = \phi_k = y_k, \quad k = 0, \ldots, K,
\]
and on each of the time intervals \([0, t_1), [t_1, t_2), \ldots, [t_K, T]\) function \(\tilde{\phi}^{s_1,\ldots,s_K}\) satisfies the ODE
\[
d\tilde{\phi}^{s_1,\ldots,s_K}_t = \left( \tilde{b}(\tilde{\phi}^{s_1,\ldots,s_K}_t) + \tilde{\sigma}(\tilde{\phi}^{s_1,\ldots,s_K}_t) f^{s_1,\ldots,s_K}_t \right) \frac{d\lambda^{s_1,\ldots,s_K}_t}{dt} dt.
\]

Function \(\phi\) at each of these intervals satisfies a similar ODE (10) with the same initial value; in addition, it is easy to verify that the values of the functions \(f^{s_1,\ldots,s_K}\),
are uniformly bounded for \( |t_k - s_k| < \delta, k = 1, \ldots, K \). This yields, by the Lipschitz continuity of \( \tilde{b}, \sigma \),

\[
\sup_{t \in [0,T]} |\phi_{t_1, \ldots, t_K} - \phi_t| \leq C \sup_{s \in [0,T]} \left| \frac{d\lambda_{s_1, \ldots, s_K}}{ds} - 1 \right| + C \| f_{s_1, \ldots, s_K} - f \|_{L_1(0,T)}
\]

\[
+ C \int_{t_k}^{t_{k+1}} |\phi_{s_1, \ldots, s_K} - \phi_s| ds, \quad t \in [t_k, t_{k+1}), \quad k = 0, \ldots, K
\]

with a constant \( C \) which does not depend on the choice of \( \{s_k\} \). By the Gronwall inequality, this gives

\[
\sup_{t \in [0,T]} |\phi_{t_1, \ldots, t_K} - \phi_t| \leq C e^{CT} \left( \sup_{s \in [0,T]} \left| \frac{d\lambda_{s_1, \ldots, s_K}}{ds} - 1 \right| + \| f_{s_1, \ldots, s_K} - f \|_{L_1(0,T)} \right).
\]

Combined with (32), (33) this proves (34) and completes the proof of (18). \( \square \)

REFERENCES

[1] P. Billingsley, Convergence of probability measures. John Wiley, New York, 1968.
[2] M. Hairer, J.C. Mattingly, M. Scheutzow (2011). Asymptotic coupling and a general form of Harris’ theorem with applications to stochastic delay equations. *Probab. Theory Related Fields*. 149 223–259.
[3] N. Ikeda, S. Watanabe, Stochastic Differential Equations and Diffusion Processes. North-Holland, Amsterdam, 1989.
[4] Ya. Ishikawa (2001) Support theorem for jump processes of canonical type. Proc. Japan Acad., 77, 79 – 83.
[5] T. Kulczycki, M. Ryznar, Transition density estimates for diagonal systems of SDEs driven by cylindrical \( \alpha \)-stable process, *ALEA Lat. Am. J. Probab. Math. Stat.* 15 (2018) 1335–1375.
[6] T. Kulczycki, M. Ryznar, Semigroup properties of solutions of SDEs driven by Lévy processes with independent coordinates, *Stochastic Process. Appl.* 130 (2020) 7185–7217.
[7] A. Kulik (2009) Exponential ergodicity of the solutions to SDEs with a jump noise. *Stochastic Processes and Applications*, 119, 602 — 632.
[8] H. Kunita, Canonical stochastic differential equations based on Lévy processes and their supports. In: Stochastic Dynamics. Springer, Berlin, 1999, 283 – 304.
[9] E. Priola, J. Zabczyk (2011) Structural properties of semilinear SPDEs driven by cylindrical stable processes, *Probab. Theory Related Fields* 149, 97–137.
[10] T. Simon (2000) Support theorem for jump processes. *Stochastic Processes and their Applications* 89 1–30.
[11] T. Simon (2001) Sur les petites déviations d’un processus de Lévy. *Potential Analysis* 14 (2), 155–173.
[12] T. Simon (2002) Support d’une équation d’Itô avec sauts en dimension 1. *Séminaire de probabilités de Strasbourg*, Tome 36, pp. 314–330.
[13] T. Simon (2003) Small deviations in p-variation for multidimensional Lévy processes. *J. Math. Kyoto Univ.* 43 (3), 523–565
[14] A.V. Skorokhod (1957) On the Differentiability of Measures Which Correspond to Stochastic Processes. I. Processes with Independent Increments, *Theory of Probability and its Applications*, 2, 407–432.
[15] Stroock, D.W., Varadhan, S.R.S. (1972) On the support of diffusion processes with applications to the strong maximum principle. In: Proceedings of Sixth Berkeley Symposium on Math. Stat. Probability, Vol. III. Univ. California Press, Berkeley, pp. 333–359.

Faculty of Pure and Applied Mathematics, Wroclaw University of Science and Technology, Wyb. Wyspianskiego 27, 50-370 Wroclaw, Poland.