The University of Manchester

MANCHESTER
1824

Delay differential equations driven by Lévy processes: Stationarity and Feller properties

Reiß, M. and Riedle, M. and Van Gaans, O.

2006

MIMS EPrint: **2007.168**

Manchester Institute for Mathematical Sciences

School of Mathematics

The University of Manchester

ELSEVIER

Delay differential equations driven by Lévy processes: Stationarity and Feller properties

M. Reiß[a], M. Riedle[b,*], O. van Gaans[c]

[a] *Institute of Applied Mathematics, University of Heidelberg, Germany*
[b] *Institute of Mathematics, Humboldt-University of Berlin, Germany*
[c] *Mathematical Institute, Leiden University, Leiden, The Netherlands*

Abstract

We consider a stochastic delay differential equation driven by a general Lévy process. Both the drift and the noise term may depend on the past, but only the drift term is assumed to be linear. We show that the segment process is eventually Feller, but in general not eventually strong Feller on the Skorokhod space. The existence of an invariant measure is shown by proving tightness of the segments using semimartingale characteristics and the Krylov–Bogoliubov method. A counterexample shows that the stationary solution in completely general situations may not be unique, but in more specific cases uniqueness is established.
ⓒ 2006 Elsevier B.V. All rights reserved.

1. Introduction

Stochastic delay differential equations, also known as stochastic functional differential equations, are a natural generalisation of stochastic ordinary differential equations by allowing the coefficients to depend on values in the past. When only the drift coefficient depends on the past, main stochastic tools and results for stochastic ordinary differential equations can be applied, for example by removing the drift via a change of measure. If the stochastic perturbation depends on the past, however, surprising new phenomena emerge, see Mohammed and Scheutzow [20] for a discussion on flow and stability properties.

---

* Corresponding author. Tel.: +49 30 20935874; fax: +49 30 20935848.
  E-mail address: riedle@mathematik.hu-berlin.de (M. Riedle).

The main purpose of the present work is to investigate the stationarity of delay differential equations driven by Lévy processes of the form

$$dX(t) = \left(\int_{[-\alpha,0]} X(t+s) \, \mu(ds) \right) dt + F(X)(t-) \, dL(t). \tag{1.1}$$

L denotes a general Lévy process, the drift term is obtained by integrating past values with respect to a signed measure μ and the nonlinear coefficient $F(X)$ depends on $(X(s) : s \in [t-\alpha, t])$ at time t, see Section 2 for details. We do not consider a nonlinearity in the drift in order to concentrate on the effects of the nonlinear noise term, which is facilitated by a variation of constants formula. While the solution processes are not Markovian any longer, one can retrieve the Markov property by regarding segments of the trajectories as processes in a function space. The delayed noise term causes a fundamental degeneration of the segment process: we show that the Markov semigroup is not Feller and not eventually strong Feller, but eventually Feller. Consequently the uniqueness of an invariant measure cannot be derived by the strong Feller property.

Stationarity results for Lévy-driven stochastic differential equations, even in the non-delay case, are not so widespread. Only for non-Gaussian Ornstein–Uhlenbeck processes are necessary and sufficient conditions guaranteeing stationarity well known, cf. Sato [26, Thm. 17.5] and Wolfe [28]. Non-Gaussian stationary Ornstein–Uhlenbeck processes have been attracting increasing attention recently due to their use in financial modelling and the relationship with self-decomposable distributions, cf. Barndorff-Nielsen and Shephard [4]. Invariant measures for ordinary differential equations with a nonlinear drift term and additive stable noise have been studied analytically by Albeverio et al. [1], but our general results, even when specified to the non-delayed case, seem to be new.

The question of the existence of stationary solutions of stochastic equations with delay goes back to the 1960s in the work of Itô and Nisio [12]. They have proved the existence, but not the uniqueness of a stationary solution for Wiener-driven delay differential equations under the condition that the drift is obtained by a delayed perturbation of a stable instantaneous feedback. For a more general non-linear drift functional and additive white noise, Scheutzow [27] derived sufficient conditions for the existence of an invariant probability measure in terms of Lyapunov functionals. For a similar approach and connections to stochastic partial differential equations see Bakhtin and Mattingly [3]. For the Lévy-driven equation (1.1) with constant F, that is additive noise, Gushchin and Küchler [11] have established necessary and sufficient conditions for the existence and uniqueness of stationary solutions.

Our work is based on analyzing the segment process in a function space and it is therefore closely related to results for stochastic evolution equations in infinite-dimensional spaces. In the case of additive noise an extensive literature for the stationarity of solutions of stochastic evolution equations exists, see Da Prato and Zabczyk [8]. Much less is known for non-additive noise, see for example Chow and Khasminskii [6] for some general results. An infinite-dimensional analogue of Eq. (1.1) driven by a Wiener process is considered by Bonaccorsi and Tessitore [5]. They obtain a stationarity result for small Lipschitz constants by a fixed point argument.

To prove the existence of a stationary solution of (1.1) under rather general conditions, we consider the segment process with values in the Skorokhod space $D([-\alpha, 0])$. First, we establish the Feller property for the Markov semigroup after time $t = \alpha$. Under the main assumption of a stable drift, we establish the tightness of the solution segments using semimartingale characteristics and apply the Krylov–Bogoliubov method to obtain an invariant measure on

the Skorokhod space. Due to the absence of the strong Feller property Doob's method fails to prove uniqueness of the invariant measure. From an abstract point of view the loss of the strong Feller property is due to the degeneracy of the diffusion term when the equation is lifted to the segment space: the driving process is only one-dimensional, cf. Gatarek and Goldys [10] for the abstract non-degenerate case. The question of uniqueness of the stationary solution turns out to be subtle and the degeneracy of the noise process prevents a straightforward analytical treatment. While for certain cases uniqueness will be shown to hold, a counterexample leads us to suspect that uniqueness fails in greater generality. Nevertheless, the correlation structure of the solution process, if it exists, is uniquely determined and analytically tractable.

In the next section we briefly review some basic facts about stochastic delay differential equations. Section 3 is devoted to the variation of constants formula and properties of the Markov semigroup. The existence and uniqueness of stationary solutions are discussed in Sections 4 and 5, respectively.

2. Preliminaries

We follow standard notation, in particular we write $C[a, b]$ for the space of real-valued continuous functions on $[a, b]$. The Skorokhod space of all real-valued functions on $[a, b]$ that are right-continuous and have left limits at every point (*càdlàg* for short) is denoted by $D[a, b]$. It is endowed with the *Skorokhod metric d_S* given by

$$d_S(\varphi, \psi) := \inf_{\lambda \in \Lambda[a,b]} \left(\|\varphi \circ \lambda - \psi\|_\infty + \|\lambda - \mathrm{Id}\|_\infty \right),$$

where $\Lambda[a, b] := \{\lambda : [a, b] \to [a, b] : \lambda$ is an increasing homeomorphism$\}$. Note that $d_S(\varphi_n, \varphi) \to 0$ implies the convergences $\varphi_n(a) \to \varphi(a)$, $\varphi_n(b) \to \varphi(b)$, but not the pointwise convergence in the interior (a, b). The space $D[a, \infty)$ of all real-valued càdlàg functions on $[a, \infty)$ can be similarly equipped with the Skorokhod metric. Here a sequence (φ_n) converges to φ if and only if

$$d_S^N(\varphi_n, \varphi) := \inf_{\lambda \in \Lambda[a,\infty)} \left(\sup_{t \in [a,N]} |(\varphi_n \circ \lambda)(u) - \varphi(u)| + \|\lambda - \mathrm{Id}\|_\infty \right) \to 0,$$

for $n \to \infty$ and all $N \in \mathbf{N}$ where $\Lambda[a, \infty) := \{\lambda : [a, \infty) \to [a, \infty) : \lambda$ is an increasing homeomorphism$\}$.

The space $(D[a, b], d_S)$ is a separable metric space. Moreover, there exists an equivalent metric d on $D[a, b]$ such that $(D[a, b], d)$ is a complete separable metric space, see for instance Jacod and Shiryaev [13]. We endow $D[a, b]$ with the corresponding Borel σ-algebra $\mathcal{B}(D[a, b])$. For $\varphi \in D[a, b]$ we denote by $\varphi(t-)$ its left-hand limit at t and we define $\Delta\varphi(t) := \varphi(t) - \varphi(t-)$, $t \in (a, b]$, and $\Delta\varphi(a) = 0$. For $\alpha > 0$ and a function $\varphi \in D[-\alpha, \infty)$ we introduce the *segment* of φ at time $t \geqslant 0$ as the function

$$\varphi_t : [-\alpha, 0] \to \mathbf{R}, \qquad \varphi_t(u) := \varphi(t + u).$$

Let us first turn our attention to the deterministic delay equation underlying the stochastic equation (1.1):

$$x(t) = \varphi(0) + \int_0^t \left(\int_{[-\alpha,0]} x(s + u) \, \mu(\mathrm{d}u) \right) \mathrm{d}s \quad \text{for } t \geqslant 0,$$

$$x(u) = \varphi(u) \quad \text{for } u \in [-\alpha, 0],$$

(2.1)

where μ is a signed finite Borel measure and the initial function φ is in $D[-\alpha, 0]$. Note that the inner integral exists because φ and a fortiori also x are measurable and locally bounded.

As the fundamental system in linear ordinary differential equations and the Green function in partial differential equations, the so-called *fundamental solution* or *resolvent* plays a major role in the analysis of (2.1). It is the function $r : \mathbf{R} \to \mathbf{R}$ which satisfies (2.1) with the initial condition $r(0) = 1$ and $r(u) = 0$ for $u \in [-\infty, 0)$. The solution $x(\cdot, \varphi)$ of (2.1) for an arbitrary initial segment $\varphi \in D[-\alpha, 0]$ exists, is unique, and can be represented as

$$x(t, \varphi) = \varphi(0)r(t) + \int_{[-\alpha, 0]} \int_s^0 r(t + s - u)\varphi(u) \, \mathrm{d}u \, \mu(\mathrm{d}s) \quad \text{for } t \geqslant 0, \tag{2.2}$$

cf. Chapter I in Diekmann et al. [9]. The fundamental solution converges for $t \to \infty$ to zero if and only if

$$v_0(\mu) := \sup \left\{ \mathrm{Re}(\lambda) : \lambda \in \mathbf{C}, \ \lambda - \int_{[-\alpha, 0]} \mathrm{e}^{\lambda s} \, \mu(\mathrm{d}s) = 0 \right\} < 0, \tag{2.3}$$

where $\mathrm{Re}(z)$ denotes the real part of a complex number z. In this case the decay is exponentially fast (see Theorem 5.4 in [9]) and the zero solution of (2.1) is uniformly asymptotically stable.

Subsequently we shall always work on a fixed complete probability space $(\Omega, \mathcal{F}, P)$ with a filtration $(\mathcal{F}_t)_{t \geqslant 0}$ satisfying the usual conditions. We study the following stochastic differential equation with time delay:

$$\mathrm{d}X(t) = \left(\int_{[-\alpha, 0]} X(t + s) \, \mu(\mathrm{d}s) \right) \mathrm{d}t + F(X)(t-) \, \mathrm{d}L(t) \quad \text{for } t \geqslant 0, \tag{2.4}$$
$$X(u) = \Phi(u) \quad \text{for } u \in [-\alpha, 0],$$

where μ is a signed finite Borel measure and the initial process $(\Phi(u) : u \in [-\alpha, 0])$ is assumed to have trajectories in $D[-\alpha, 0]$ and to be $\mathcal{F}_0$-measurable. The driving process $L = (L(t) : t \geqslant 0)$ is a Lévy process. We denote its Lévy–Khintchine characteristic by (b, σ^2, ν) with respect to the truncation function $x \mapsto x\mathbf{1}_{[-1,1]}(x)$.

Turning to the specification of the nonlinear mapping F, we remark that results for the existence and uniqueness of strong or weak solutions of stochastic delay differential equations driven by Brownian motion appear in different generalities: Mohammed [19] provides a result under random functional Lipschitz conditions, Mao [16] discusses in addition the *method of steps*, which provides a unique solution without a regular dependence of the coefficients on values in the past, Liptser and Shiryaev [15] give general results for weak solutions and Itô and Nisio [12] consider the existence of weak solutions for equations with finite and infinite delay. Since our equations are driven by Lévy processes and the most general conditions are not our concern here, we follow Protter [22] and merely assume that the deterministic functional $F : D[-\alpha, \infty) \to D[-\alpha, \infty)$ is *functional Lipschitz and autonomous*, i.e. it is continuous with respect to the Skorokhod topology and it satisfies for all $\varphi_1, \varphi_2 \in D[-\alpha, \infty)$:

(a) there exists a constant $K > 0$, independent of φ_1, φ_2 and t, such that

$$|F(\varphi_1)(t) - F(\varphi_2)(t)| \leqslant K \sup_{t-\alpha \leqslant s \leqslant t} |\varphi_1(s) - \varphi_2(s)| \quad \text{for all } t \geqslant 0; \tag{2.5}$$

(b) $F(\varphi_1(s + \cdot))(t) = F(\varphi_1)(t + s) \quad$ for all $t, s \geqslant 0$.

Equivalently, setting $\tilde{F}(\varphi|_{[-\alpha,0]}) := F(\varphi)(0)$ the two conditions can be stated as $F(\varphi)(t) = \tilde{F}(\varphi_t)$ with a functional $\tilde{F}$ which is Lipschitz continuous on $D[-\alpha, 0]$ equipped with the supremum norm.

We can rewrite the differential equation (2.4) as the integral equation

$$X(t) = \Phi(0) + \int_0^t G_\Phi(X)(s)\,\mathrm{d}s + \int_0^t H_\Phi(X)(s-)\,\mathrm{d}L(s) \quad \text{for } t \geqslant 0, \tag{2.6}$$

when introducing G_φ, $H_\varphi : D[0,\infty) \to D[0,\infty)$ for $s \geqslant 0$ and $\psi \in D[0,\infty)$ by (abusing notation slightly)

$$G_\varphi(\psi)(s) = \int_{[-\alpha,0]} \big(\varphi_s(u)\mathbf{1}_{[-\alpha,-s)}(u) + \psi_s(u)\mathbf{1}_{[-s,0]}(u)\big)\,\mu(\mathrm{d}u), \tag{2.7}$$

$$H_\varphi(\psi)(s) = F\big(\varphi\mathbf{1}_{[-\alpha,0)} + \psi\mathbf{1}_{[0,\infty)}\big)(s). \tag{2.8}$$

For $\mathcal{F}_0$-measurable initial segments Φ the mappings G_Φ and H_Φ are functional Lipschitz in the definition of Protter [22] and we can invoke Theorem V.7 in Protter [22] which ensures a unique strong solution of (2.4). Recall that a strong solution of (2.4) is an adapted, stochastic process X with càdlàg paths satisfying (2.6). The solution is called unique if all solutions are indistinguishable. We denote the solution by $(X(t) : t \geqslant -\alpha)$ or $(X(t, \Phi) : t \geqslant -\alpha)$.

Examples 2.1. (a) The no-delay case: if $\mu = b\delta_0$, a point mass at zero, and $F(\varphi)(t) = f(\varphi(t))$, $t \geqslant -\alpha$, then the equation reads

$$\mathrm{d}X(t) = bX(t)\,\mathrm{d}t + f(X(t-))\,\mathrm{d}L(t) \quad \text{for } t \geqslant 0.$$

If f is Lipschitz continuous, then F is easily seen to be functional Lipschitz and autonomous.

(b) The point-delay case: suppose $\mu = \sum_{i=1}^n b_i \delta_{\alpha_i}$ and $F(\varphi)(t) = f(\varphi(t-\alpha_1), \ldots, \varphi(t-\alpha_n))$, $t \geqslant 0$, and $F(\varphi)(u) = F(\varphi)(0)$, $u \in [-\alpha, 0]$, with $\alpha_i \in [-\alpha, 0]$. Then the equation reads

$$\mathrm{d}X(t) = \sum_{i=1}^n b_i X(t-\alpha_i)\,\mathrm{d}t + f(X(t-\alpha_1-), \ldots, X(t-\alpha_n-))\,\mathrm{d}L(t) \quad \text{for } t \geqslant 0$$

and F is again autonomous and functional Lipschitz if f is Lipschitz in all its arguments.

(c) The distributed-delay case: for $\mu(\mathrm{d}s) = b(s)\,\mathrm{d}s$ and $F(\varphi)(t) = f(\int_{[-\alpha,0]} \varphi(t+s)c(s)\,\mathrm{d}s)$, $t \geqslant 0$, with $b, c \in L^1[-\alpha, 0]$ and $F(\varphi)(u) = F(\varphi)(0)$, $u \in [-\alpha, 0]$, we obtain for $t \geqslant 0$

$$\mathrm{d}X(t) = \left(\int_{[-\alpha,0]} X(t+s)b(s)\,\mathrm{d}s\right)\mathrm{d}t + f\left(\int_{[-\alpha,0]} X(t+s)c(s)\,\mathrm{d}s\right)\mathrm{d}L(t).$$

Again, we need f to be Lipschitz in order to have F functional Lipschitz and autonomous.

(d) Further examples and counterexamples: other useful path-dependent mappings like $F(\varphi)(t) = \sup_{u\in[t-\alpha,t]} \varphi(u)$ and their combinations with Lipschitz functions are functional Lipschitz and autonomous. Beware, however, that not all Lipschitz continuous functionals $\tilde{F}$ on $D[-\alpha, 0]$ give rise to a functional $F : D[-\alpha, \infty) \to D[-\alpha, \infty)$, for instance the jump size functional $F(\varphi)(t) = \Delta\varphi(t)$ is not càdlàg for càdlàg functions φ with jumps. It is interesting to note that all admissible linear functionals are given by $\tilde{F}(\varphi) = \int_{[-\alpha,0]} \varphi(u)\,\rho(\mathrm{d}u)$ with ρ ranging through the space of finite Borel measures, which follows from the result given by Pestman [21] when excluding the part based on jump sizes.

3. Properties of the solution

3.1. The variation of constants formula

Many of our considerations will be based on a stochastic convolution equation, the *variation of constants formula*. This formula is easily derived if the driving process has bounded second moments, but no longer for processes where an Itô isometry or inequality fails. We provide a proof separately in Reiß et al. [24].

Theorem 3.1. *Let F be functional Lipschitz. Then for a stochastic process $X = (X(t) : t \geqslant -\alpha)$ and initial condition Φ the following are equivalent:*

(1) *X is the unique solution of* (2.4) *with $X_0 = \Phi$;*
(2) *X obeys the variation of constants formula:*

$$X(t) = \begin{cases} x(t, \Phi) + \displaystyle\int_0^t r(t-s)F(X)(s-)\,\mathrm{d}L(s), & t \geqslant 0, \\ \Phi(t), & t \in [-\alpha, 0], \end{cases} \tag{3.1}$$

where r is the fundamental solution of Eq. (2.1).

3.2. Measurability of the segment process

Our further work will be strongly based on considering the segment process $(X_t : t \geqslant 0)$ in $D[-\alpha, 0]$ instead of the real-valued process $(X(t) : t \geqslant -\alpha)$. This approach is natural because the segment process is Markovian and turns out to be eventually Feller. These properties will pave the way for our further analysis.

In the case where L is a Brownian motion the segment process is immediately a Feller process on the path space $C([-\alpha, 0])$, see Theorem III.3.1 in Mohammed [19], which is not true in our setting because of the discontinuity of the shift semigroup on $D[-\alpha, 0]$.

The following lemma establishes certain measurability and continuity properties of the segment process. For a continuous path space similar properties have been studied in Chapter 3.7 of Da Prato and Zabczyk [7] and Lemma II.2.1 in Mohammed [19] in Chapter 3.7.

Lemma 3.2. *Let $(Y(t) : t \geqslant 0)$ be a stochastically continuous process with càdlàg paths. Then the segment process $(Y_t : t \geqslant \alpha)$ in $D[-\alpha, 0]$ is stochastically continuous as well. Moreover, there exists a jointly measurable modification of $(Y_t : t \geqslant \alpha)$.*

Proof. Since the Borel-σ-algebra of $D[-\alpha, 0]$ coincides with the cylindrical σ-algebra generated by all point evaluations $\varphi \mapsto \varphi(c)$ for $\varphi \in D[-\alpha, 0]$ and arbitrary $c \in [-\alpha, 0]$ the segment Y_t is a $D[-\alpha, 0]$-valued random variable.

For $h > 0$ we define the homeomorphism $\lambda_h : [-\alpha, 0] \to [-\alpha, 0]$ by $\lambda_h(s) := s - h$ for $s \in [-\alpha + 2h, -h]$ and affine respectively on $[-\alpha, -\alpha + 2h]$ such that $\lambda_h(-\alpha) = -\alpha$ and on $[-h, 0]$ such that $\lambda_h(0) = 0$. Then $\|\lambda_h - \mathrm{Id}\|_\infty \leqslant h$ and

$$Y_{t+h}(\lambda_h(s)) = \begin{cases} Y(s+t), & s \in [-\alpha + 2h, -h], \\ Y(t + h + \lambda_h(s)), & s \in [-\alpha, -\alpha + 2h) \cup (-h, 0]. \end{cases}$$

Therefore, we obtain

$$d_S(Y_{t+h}, Y_t) \leqslant \|Y_{t+h} \circ \lambda_h - Y_t\|_\infty + \|\lambda_h - \mathrm{Id}\|_\infty \to |\Delta Y(t)| \quad \text{as } h \downarrow 0.$$

Hence, $t \mapsto Y_t$ is right-continuous at t_0 if Y is continuous at t_0. Similarly, one can establish $\overline{\lim}_{h\downarrow 0} d_S(Y_{t-h}, Y_t) \leqslant |\Delta Y(t - \alpha)|$. We conclude that $t \mapsto Y_t$ is stochastically continuous at t_0 if $P(\Delta Y(t_0 - \alpha) \neq 0) = P(\Delta Y(t_0) \neq 0) = 0$, which follows from the stochastic continuity of Y.

Any stochastically continuous process with values in a Polish space has a jointly measurable modification, which is proved following [7, Prop. 3.2], but measuring the distance with the metric of this space. This gives the final assertion.   □

3.3. The Feller property

Basic tools for deriving the existence and uniqueness of invariant measures are the Feller and strong Feller property of the Markov semigroup defined by the segment process. We establish here the Markov property, the Feller property after time α and give examples that the immediate Feller and the eventually strong Feller property fail in general. For our purposes the ordinary Markov property of the segment process will be sufficient, but the strong Markov property can also be derived following the lines of Chapter 9.2 in Da Prato and Zabczyk [7].

Proposition 3.3. *Let X be the unique solution of* (2.4)*. Then the segment process* $(X_t : t \geqslant 0)$ *is a Markov process on* $D[-\alpha, 0]$*:*

$$P(X_t \in B \mid \mathcal{F}_s) = P(X_t \in B \mid X_s) \quad P\text{-a.s.}$$

for all $t \geqslant s \geqslant 0$ *and Borel sets* $B \in \mathcal{B}(D[-\alpha, 0])$.

Proof. We fix $u \geqslant 0$ and consider for $t \geqslant u$ the equation

$$X^{u,\varphi}(t) = \varphi(0) + \int_u^t \int_{[-\alpha,0]} X^{u,\varphi}(s + v)\, \mu(\mathrm{d}v)\, \mathrm{d}s + \int_u^t F(X^{u,\varphi})(s-)\, \mathrm{d}L(s),$$

$$X^{u,\varphi}(m) = \varphi(m - u) \quad \text{for } m \in [u - \alpha, u] \text{ and } \varphi \in D[-\alpha, 0].$$

We denote the unique strong solution by $(X^{u,\varphi}(t) : t \geqslant u - \alpha)$ and the segment process by $(X_t^{u,\varphi} : t \geqslant u)$.

We define $\mathcal{G}_u := \sigma(L(s) - L(u) : s \geqslant u)$ which is independent of the σ-algebra $\mathcal{F}_u$ from the given filtration. The solution $X^{u,\varphi}(t)$ is $\mathcal{G}_u$-measurable for every $t \geqslant u$ and, as in the proof of Lemma 3.2, it follows that the segment $X_t^{u,\varphi}$ is $\mathcal{G}_u$-measurable as well. The uniqueness of the solution implies $X(s) = X^{u,X_u}(s)$ for every $s \geqslant u - \alpha$ and thus $X_t = X_t^{u,X_u}$ for every $t \geqslant u$ with probability one. By construction (cf. [22]) the solution process depends in a measurable way on the initial condition so that the function

$$A : D[-\alpha, 0] \times \Omega \to \mathbf{R}, \qquad A(\varphi, \omega) := \mathbf{1}_B(X_t^{u,\varphi}(\omega))$$

is measurable for every $B \in \mathcal{B}(D[-\alpha, 0])$ and independent of $\mathcal{F}_u$ for fixed φ. An application of the factorisation lemma [8, Prop. 1.12] yields P-almost surely

$$P(X_t \in B \mid \mathcal{F}_u) = \mathbf{E}[\mathbf{1}_B(X_t^{u,X_u}) \mid \mathcal{F}_u] = \mathbf{E}[A(X_u, \cdot) \mid \mathcal{F}_u] = \mathbf{E}[A(\varphi, \cdot)]|_{\varphi = X_u},$$

which ends the proof because the right-hand side is $\sigma(X_u)$-measurable.   □

Let us consider the Skorokhod topology on $D[-\alpha, 0]$ and let $B_b(D[-\alpha, 0])$ denote the space of all real-valued functions which are Borel with respect to the Skorokhod topology and bounded, i.e. $\sup\{|f(\varphi)| : \varphi \in D[-\alpha, 0]\} < \infty$. Let $C_b(D[-\alpha, 0])$ denote its subspace of continuous

functions with respect to the Skorokhod topology on $D[-\alpha, 0]$. Due to Proposition 3.3 the operators

$$P_{s,t} : B_b(D[-\alpha, 0]) \to B_b(D[-\alpha, 0]), \qquad P_{s,t} f(\varphi) := \mathbf{E}[f(X_t^{s,\varphi})]$$

have the property that $P_{u,s} P_{s,t} = P_{u,t}$ for $0 \leqslant u \leqslant s \leqslant t$. By homogeneity we have $P_{s,t} = P_{0,t-s}$ for $0 \leqslant s \leqslant t$, cf. Thm. V.32 in Protter [22], and the operators $P_t := P_{0,t}$, $t \geqslant 0$, form a *Markovian semigroup*. The Markovian semigroup will be called *eventually Feller* if there exists a $t_0 \geqslant 0$ such that for any $f \in C_b(D[-\alpha, 0])$ the following two conditions are satisfied:

$$P_t f \in C_b(D[-\alpha, 0]) \quad \text{for every } t \geqslant t_0, \tag{3.2}$$

$$\lim_{s \downarrow t} P_s f(\varphi) = P_t f(\varphi) \quad \text{for every } \varphi \in D[-\alpha, 0], \ t \geqslant t_0. \tag{3.3}$$

By definition the solution $(X(t) : t \geqslant 0)$ has càdlàg paths and it is easily observed from (2.6) that it has no fixed times of discontinuity: $P(\Delta X(t) \neq 0) \leqslant P(\Delta L(t) \neq 0) = 0$. Hence, the process $(X(t) : t \geqslant 0)$ is stochastically continuous. By Lemma 3.2 so is the segment process $(X_t : t \geqslant \alpha)$ and thus condition (3.3) is fulfilled for $t_0 = \alpha$. The semigroup is not stochastically continuous for $t_0 < \alpha$ and condition (3.2) fails for $t_0 < \alpha$ due to the discontinuity of the shift semigroup, as the following example demonstrates.

Consider $0 < \beta < \alpha$ and the initial functions $\varphi^n := \mathbf{1}_{[-\beta(1-n^{-1}),0]}$ which for $n \to \infty$ converge in $D[-\alpha, 0]$ to $\varphi^\infty := \mathbf{1}_{[-\beta,0]}$. The corresponding solution segments X_t^n, for an arbitrary specification of F and L in the differential equation, satisfy $X_{\alpha-\beta}^n(-\alpha) = \varphi^n(-\beta) = 0$, while $X_{\alpha-\beta}^\infty(-\alpha) = 1$ holds. Hence, $d_S(X_{\alpha-\beta}^n, X_{\alpha-\beta}^\infty) \geqslant 1$, which implies that $\varphi \mapsto P_t f(\varphi)$ is not continuous for $f(\psi) := |\psi(-\alpha)| \wedge 1 \in C_b(D[-\alpha, 0])$ and any time $t \in (0, \alpha)$. Similarly, $t \mapsto P_t f(\varphi^\infty)$ is seen to be discontinuous at $t = \alpha - \beta$.

We now establish condition (3.2) for $t_0 = \alpha$ by showing even more, namely that $\varphi \mapsto (X(t, \varphi) : t \in [0, T])$ is continuous from $D[-\alpha, 0]$ to the space of càdlàg processes with the uniform convergence on $[0, T]$ in probability, which is stronger than convergence in the Skorokhod topology in law. We start with a norm estimate in the spirit of Émery's inequality before proving the main result. In accordance with Section V.2 in Protter [22] we employ the following norms for semimartingales $(Z(t) : t \geqslant 0)$ and adapted càdlàg processes $(Y(t) : t \geqslant 0)$:

$$\|Y\|_{S^2[0,T]}^2 := \mathbf{E}\left[\sup_{0 \leqslant t \leqslant T} Y(t)^2\right],$$

$$\|Z\|_{H^2[0,T]}^2 := \inf\left\{\mathbf{E}[M, M]_T + \mathbf{E}[\mathrm{TV}(A)(T)^2]\right\},$$

where the infimum is taken over all possible decompositions $Z = M + A$ where M is a local martingale and A a bounded variation process with $M(0) = A(0) = 0$. The total variation of A on $[0, T]$ is denoted by $\mathrm{TV}(A)(T)$.

The quadratic variation process is defined by $[Z, Z] := Z^2 - \int Z(s-) \, \mathrm{d}Z(s)$. On the basis of these norms the spaces $H^2[0, T]$ and $S^2[0, T]$ are constructed canonically. Moreover, they are Banach spaces, and $H^2[0, T]$ is continuously embedded in $S^2[0, T]$.

Lemma 3.4. *Suppose the Lévy process L has a finite second moment and $(H(t) : 0 \leqslant t \leqslant T)$ is an adapted càdlàg process with $\int_0^T \mathbf{E}[H(t-)^2] \, \mathrm{d}t < \infty$. Then*

$$\left\|\int_0^{\cdot} H(s-) \, \mathrm{d}L(s)\right\|_{H^2[0,T]}^2 \leqslant \left(\sigma^2 + \int x^2 \, \nu(\mathrm{d}x) + (\mathbf{E}L(1))^2 T\right) \int_0^T \mathbf{E}[H(t-)^2] \, \mathrm{d}t.$$

Proof. This follows from the decomposition $L(t) = M(t) + t\mathbf{E}L(1)$ with M a square integrable martingale. $\square$

The surprising result of the next proposition, which says that convergence of the initial conditions in Skorokhod metric implies uniform convergence of the solution processes, is essentially due to the fact that the driving Lévy process is a semimartingale without fixed time of discontinuity.

Proposition 3.5. *Assume $F : D[-\alpha, \infty) \to D[-\alpha, \infty)$ is continuous with respect to the Skorokhod metric and satisfies (2.5). Let X^n be the solution of Eq. (2.4) with deterministic initial segment φ^n, let $\varphi^n \to \varphi$ in $D[-\alpha, 0]$ and let X be the solution with initial segment φ. Then $(X^n(t) : t \geqslant 0)$ converges to $(X(t) : t \geqslant 0)$ uniformly on compact sets in probability.*

Proof. We consider first a stopping time R such that the process L^{R-} is α-sliceable for some suitably small $\alpha > 0$ in the sense of [22].

In analogy to [22, Thm. V.10] we use the representation (2.6) and put

$$Y^n(t) := \int_0^t (G_\varphi(X) - G_{\varphi^n}(X))(s)\,\mathrm{d}s + \int_0^t (H_\varphi(X) - H_{\varphi^n}(X))(s-)\,\mathrm{d}L^{R-}(s),$$
$$\mathcal{G}^n(U)(t) := G_{\varphi^n}(X)(t) - G_{\varphi^n}(X - U)(t),$$
$$\mathcal{H}^n(U)(t) := H_{\varphi^n}(X)(t) - H_{\varphi^n}(X - U)(t),$$

$t \geqslant 0$, to obtain for $U^n := X - X^n$ the equation

$$U^n(t) = \varphi(0) - \varphi^n(0) + Y^n(t) + \int_0^t \mathcal{G}^n(U^n)(s)\,\mathrm{d}s + \int_0^t \mathcal{H}^n(U^n)(s-)\,\mathrm{d}L^{R-}(s).$$

Since L^{R-} is α-sliceable, [22, Lemma V.3.2], extended to two driving semimartingales, yields that the solution U^n of this equation satisfies $\|U^n\|_{S^2[0,T]} \leqslant C \|\varphi(0) - \varphi^n(0) + Y^n\|_{S^2[0,T]}$ for any $T > 0$ with a constant $C > 0$ depending on the process L^{R-} and a uniform bound for the Lipschitz constants of G_{φ^n} and H_{φ^n}. The Skorokhod metric ensures $\varphi^n(0) \to \varphi(0)$, so that $\|U^n\|_{S^2[0,T]} \to 0$ follows if Y^n tends to zero in $S^2[0,T]$. The latter is fulfilled if

$$\mathbf{E}\left[\int_0^T (G_\varphi(X)(t) - G_{\varphi^n}(X)(t))^2 + (H_\varphi(X)(t) - H_{\varphi^n}(X)(t))^2\mathrm{d}t\right] \tag{3.4}$$

tends to 0 as $n \to \infty$, due to the continuous embedding $H^2 \hookrightarrow S^2$ from [22, Thm. V.2] and Lemma 3.4 with the additional observation that $\|\int_0^\cdot J(s-)\mathrm{d}L^{R-}(s)\|_{H^2[0,T]} \leqslant \|\int_0^\cdot J(s-)\mathrm{d}L(s)\|_{H^2[0,T]}$ for any process $J \in S^2[0,T]$. Let ω be fixed for the moment. The functions $H_{\varphi^n}(X(\omega))$ converge in the Skorokhod topology to $H_\varphi(X(\omega))$, which implies convergence in $L^2[0,T]$. Concerning the sequence G_{φ_n} we have

$$\int_0^T (G_{\varphi_n}(X(\omega))(t) - G_\varphi(X(\omega))(t))^2\,\mathrm{d}t = \int_{[-\alpha,0)} \int_{[-\alpha,0)} \int_0^T (\varphi_n(t+u)$$
$$- \varphi(t+u))(\varphi_n(t+v) - \varphi(t+v))\mathbf{1}_{[-\alpha,-t)}(u)\mathbf{1}_{[-\alpha,-t)}(v)\,\mathrm{d}t\,\mu(\mathrm{d}u)\,\mu(\mathrm{d}v).$$

This expression converges to zero as $n \to \infty$, since the Skorokhod convergence of φ_n to φ implies $\varphi_n \to \varphi$ Lebesgue a.e. and the sequence $(\varphi_n)_n$ is uniformly bounded. Again by [22, Lemma V.3.2] and Lemma 3.4 the solution process X is an element of $S^2[0,T]$, whence by the uniform linear growth of (G_{φ^n}) and (H_{φ^n}) the argument inside the expectation in (3.4)

is dominated by a P-integrable function. The Dominated Convergence Theorem thus gives the convergence in (3.4) such that $\|X - X^n\|_{S^2[0,T]} = \|U^n\|_{S^2[0,T]} \to 0$ for any $T > 0$.

Next, let L be an arbitrary Lévy process. According to [22, Theorem V.5, p. 192] there exist stopping times $0 = T_0 \leqslant T_1 \leqslant T_2 \leqslant \cdots$ such that $\sup_\ell T_\ell = \infty$ a.s. and $L^{T_\ell-}$ is α-sliceable for each ℓ. Consider Eq. (2.4) with L replaced by $L^{T_\ell-}$ and let $X^{n,\ell}$ denote the solution with initial segment φ_n and let $X^{\infty,\ell}$ denote the solution with initial segment φ, for $n, \ell \in \mathbf{N}$. We have shown above that $X^{n,\ell} \to X^{\infty,\ell}$ uniformly on compact sets in probability for every ℓ. Further, it is clear from the equation that $X^{n,\ell} = (X^n)^{T_\ell-}$. Let now $t > 0$, $r > 0$, and $\varepsilon > 0$ be arbitrary. Choose an ℓ such that $P(T_\ell < t) < \varepsilon/2$. Then

$$
P\left(\sup_{0 \leqslant s \leqslant t} |X^n(s) - X(s)| \geqslant r \right)
$$

$$
\leqslant P\left(\sup_{0 \leqslant s \leqslant t} |X^n(s) - X(s)| \geqslant r \text{ and } T_\ell \geqslant t \right) + P(T_\ell < t)
$$

$$
\leqslant P\left(\sup_{0 \leqslant s \leqslant t} |X^{n,\ell}(s) - X^{\infty,\ell}(s)| \geqslant r \right) + \varepsilon/2 < \varepsilon
$$

for n large. Hence $X^n \to X$ uniformly on compact sets in probability. $\qquad\square$

Let us finally show that in general we cannot expect the solution to be *eventually strongly Feller*, which is characterised by the existence of a $t_0 > 0$ such that for all $f \in B_b(D[-\alpha, 0])$

$$
P_t f \in C_b(D[-\alpha, 0]) \quad \text{for every } t \geqslant t_0.
$$

Using indicator functions for f, this implies

$$
\varphi \mapsto P(X_t(\varphi) \in B) \in C_b(D[-\alpha, 0]) \quad \text{for every } B \in \mathcal{B}(D[-\alpha, 0]), \ t \geqslant t_0.
$$

Suppose the functional F in the Eq. (2.4) is of the form $F(\psi)(t) = f(\psi(t - \alpha))$ for $t \geqslant 0$ and $F(\psi)(t) = 0$ for $t < 0$, with a Lipschitz-continuous homeomorphism $f : \mathbf{R} \to (a, b)$, $b > a > 0$, and consider the case where L is standard Brownian motion. Then the quadratic variation $\langle X_t \rangle$ of the solution segment X_t, $t \geqslant \alpha$, satisfies

$$
\langle X_t \rangle_u = \int_0^{t+u} f^2(X(s - \alpha)) \, \mathrm{d}s \quad P\text{-a.s. for } u \in [-\alpha, 0].
$$

Since both sides of the equation are continuous in u for continuous X, there is one P-null exception set for all $u \in [-\alpha, 0]$. Consider the map

$$
V(\varphi)(u) := f^{-1}\left(\left(\frac{\mathrm{d}\langle\varphi\rangle_u}{\mathrm{d}u} \right)^{1/2} \right), \quad u \in [-\alpha, 0],
$$

defined on the functions φ with finite quadratic variation such that $\frac{\mathrm{d}\langle\varphi\rangle_u}{\mathrm{d}u} \in (a^2, b^2)$ for Lebesgue-almost every $u \in [-\alpha, 0]$. We have $P(V(X_t)(u) = X(t+u-\alpha), \ u \in [-\alpha, 0]) = 1$ for all $t \geqslant \alpha$. Iterating this map, we can recover with probability one the initial segment X_0 from observing $X_{m\alpha}$ since $V^m(X_{m\alpha}) = X_0$ for every integer m. This identifiability property shows that the laws of the segments $X_{m\alpha}(\varphi_1)$ and $X_{m\alpha}(\varphi_2)$ for different initial segments φ_1 and φ_2 must be singular. Hence, there is a contradiction to the strong Feller property at $t_0 = m\alpha$, which asserts the continuous dependence of the laws on the initial condition.

In fact, this example even shows that the Markov semigroup is not eventually regular in the sense of Da Prato and Zabczyk [8]. We shall see in Section 5.2 that this counterexample is due to the delay in the diffusion coefficient.

4. Existence of a stationary solution

4.1. Tightness

We establish the tightness of the laws $\{\mathcal{L}(X_t)\}_{t \geqslant 0}$ in $D[-\alpha, 0]$ by considering the semimartingale characteristics. Recall that (b, σ^2, ν) denotes the Lévy–Khintchine characteristic of the Lévy process L.

Assumption 4.1.

(a) The delay measure μ in the drift satisfies $v_0(\mu) < 0$ with v_0 from Eq. (2.3).
(b) The jump measure ν satisfies $\int_{|x|>1} \log |x| \, \nu(\mathrm{d}x) < \infty$.
(c) The coefficient F in Eq. (2.4) is functional Lipschitz, uniformly bounded and autonomous.

Condition (a) yields the exponential decay of the fundamental solution, while condition (b) ensures that $\int_0^t f(s) \, \mathrm{d}L(s)$, for exponentially decaying functions f of locally bounded variation, converges in law as $t \to \infty$ and is already for constant F necessary for the existence of a stationary solution, as was shown by Gushchin and Küchler [11], cf. also Thm. 4.3.17 in Applebaum [2]. In condition (c) restrictions on F are imposed such that the differential equation is autonomous, has a unique solution and the impact of the driving process cannot become too large. For the latter the imposed boundedness of F can certainly be relaxed considerably, but will then depend on the large jumps of L, that is, on fine properties of ν.

Proposition 4.2. *Grant Assumption 4.1. Then the solution process $(X(t) : t \geqslant -\alpha)$ of (2.4) with initial condition $X_0 = 0$ has one-dimensional marginal laws $\{\mathcal{L}(X(t))\}_{t \geqslant 0}$ that are tight.*

Proof. Let us split the Lévy process L into two parts, one of them consisting of jumps of size larger than one:

$$L(t) = N(t) + R(t) \quad \text{with } N(t) = \sum_{s \leqslant t} \Delta L(s) \mathbf{1}_{\{|\Delta L(s)| > 1\}}.$$

Then the variation of constants formula (3.1) yields $X = Y + Z$ with

$$Y(t) := \int_0^t r(t-s) F(X)(s-) \, \mathrm{d}N(s), \quad t \geqslant 0,$$

and

$$Z(t) := \int_0^t r(t-s) F(X)(s-) \, \mathrm{d}R(s), \quad t \geqslant 0.$$

Tightness of $(X(t) : t \geqslant 0)$ will follow from tightness of Y and Z.

The fundamental solution r decays exponentially with $|r(t)| \leqslant c e^{-\beta t}$ for some constants $c, \beta > 0$ due to Assumption 4.1(a). Considering Y first, we obtain for any $K > 0$ with $m := \sup_\psi |F(\psi)(0)|$ by time reversal for the compound Poisson process N the estimate

$$P(|Y(t)| > K) \leqslant P\left(\sum_{s \leqslant t} |r(t-s) F(X)(s-) \Delta N(s)| > K\right)$$

$$\leqslant P\left(\sum_{s\leqslant t} c e^{-\beta(t-s)} m\, |\Delta N(s)| > K\right)$$

$$= P\left(\sum_{s\leqslant t} e^{-\beta s}\, |\Delta N(s)| > \frac{K}{cm}\right).$$

The tightness for Y follows from the tightness of $\sum_{s\leqslant t} e^{-\beta s}\, |\Delta N(s)|$ which has been established in [11, Lemma 4.3] under Assumption 4.1(b).

Since R is a Lévy process with bounded jumps its canonical decomposition is, by means of [2, p. 103], of the simple form $R(t) = R_0(t) + t\mathbf{E}R(1)$ where $(R_0(t) : t \geqslant 0)$ is a square-integrable martingale. We split Z into the sum $Z = Z_0 + Z_1$ with

$$Z_0(t) := \int_0^t r(t-s)F(X)(s-)\,dR_0(s), \quad t \geqslant 0,$$

and

$$Z_1(t) := \mathbf{E}[R(1)]\int_0^t r(t-s)F(X)(s-)\,ds, \quad t \geqslant 0.$$

For Z_1 we easily obtain $P(|Z_1(t)| > K) \leqslant P(cm\beta^{-1}|\mathbf{E}[R(1)]| > K) = 0$ for K sufficiently large, implying tightness of Z_1. As in Lemma 3.4 we obtain

$$\mathbf{E}[Z_0(t)]^2 \leqslant \left(\sigma^2 + \int_{|x|\leqslant 1} x^2\, \nu(dx)\right) m^2 \int_0^t r^2(t-s)\,ds, \quad t \geqslant 0.$$

Hence by the exponential decay of r, the sequence $(Z_0(t))_{t\geqslant 0}$ is bounded in $L^2_P(\Omega)$ and thus tight. $\quad\square$

Proposition 4.3. *In the setting of Proposition 4.2 we have that the laws $\{\mathcal{L}(X(t+s) - X(t), s \in [0,\alpha])\}_{t\geqslant 0}$ are tight in $D[0,\alpha]$.*

Proof. We are led to consider for $t \geqslant 0$ and $s \in [0,\alpha]$

$$Y_t(s) := X(t+s) - X(t)$$
$$= \int_t^{t+s}\left(\int_{[-\alpha,0]} X(u+v)\,\mu(dv)\right) du + \int_t^{t+s} F(X)(u-)\,dL(u).$$

Let us introduce for $t \geqslant 0$ the semimartingale $(I_t(s) : s \in [0,\alpha])$ by letting

$$I_t(s) := \int_t^{t+s} F(X)(u-)\,dL(u) \quad \text{for } s \in [0,\alpha].$$

Now, either by following the lines in [13, III.2.c] and using the Lévy–Itô decomposition or by applying [23, Prop. 7.6] the semimartingale characteristic $(B_{I_t}, C_{I_t}, \nu_{I_t})$ of $(I_t(s) : s \in [0,\alpha])$ is found to be

$$B_{I_t}(s) = \int_t^{t+s}\left(bF(X)(u-)\right.$$
$$\left. + \int xF(X)(u-)\left(\mathbf{1}_{(-1,1)}(xF(X)(u-)) - \mathbf{1}_{(-1,1)}(x)\right)\nu(dx)\right) du,$$

$$C_{I_t}(s) = \sigma^2 \int_t^{t+s} F^2(X)(u-)\,du,$$

$$\nu_{I_t}(ds, dx) = ds \times K_{I_t}(X, t+s, dx),$$

where for $y \in D[-\alpha, \infty)$, $u \geqslant t$ and a Borel set $A \in \mathcal{B}(\mathbf{R})$

$$K_{I_t}(y, u, A) := \int \mathbf{1}_{A \setminus \{0\}} \left(F(y)(u-)x\right) \nu(\mathrm{d}x).$$

Hence, the semimartingale $(Y_t(s) : s \in [0, \alpha])$ has the characteristic $(B_{Y_t}, C_{Y_t}, \nu_{Y_t})$ with $C_{Y_t} = C_{I_t}$, $\nu_{Y_t} = \nu_{I_t}$ and

$$B_{Y_t}(s) = B_{I_t}(s) + \int_t^{t+s} \left(\int_{[-\alpha, 0]} X(u + v) \, \mu(\mathrm{d}v) \right) \mathrm{d}u.$$

We prove the tightness of $(Y_t)_{t \geqslant 0}$ by means of [13, Thm. VI.4.18] and [13, VI.4.20]. For that we have to verify that

$$a_{Y_t}(s) := \mathrm{TV}(B_{Y_t})(s) + C_{Y_t}(s) + \int_{[0,s] \times \mathbf{R}} (|x|^2 \wedge 1) \, \nu_{Y_t}(\mathrm{d}u, \mathrm{d}x), \quad s \in [0, \alpha],$$

forms a tight sequence $(a_{Y_t})_{t \geqslant 0}$ of processes and all limit points of the sequence $\{\mathcal{L}(a_{Y_t})\}_{t \geqslant 0}$ as $t \to \infty$ are laws of continuous processes. According to [13, Prop. VI.3.33 and VI.3.35] this will follow if there exist some increasing processes $(A_{Y_t})_{t \geqslant 0}$ satisfying these conditions and in addition $A_{Y_t} - a_{Y_t}$ defines for every $t \geqslant 0$ an increasing process, since $a_{Y_t}(s) \geqslant 0$ a.s. for all $s \in [0, \alpha]$, $t \geqslant 0$. To obtain such processes $(A_{Y_t}(s) : s \in [0, \alpha])$, we estimate

$$\mathrm{TV}(B_{Y_t})(s) = \int_t^{t+s} \left| \int_{[-\alpha, 0]} X(u + v) \, \mu(\mathrm{d}v) + bF(X)(u-) \right.$$
$$\left. + \int_{\mathbf{R}} x F(X)(u-) \left(\mathbf{1}_{(-1,1)} \left(x F(X)(u-) \right) - \mathbf{1}_{(-1,1)}(x) \right) \nu(\mathrm{d}x) \right| \mathrm{d}u$$
$$\leqslant \int_t^{t+s} \left(\left| \int_{[-\alpha, 0]} X(u + v) \, \mu(\mathrm{d}v) \right| + |b| \, m + c \right) \mathrm{d}u,$$

where $m := \sup_\psi |F(\psi)(0)|$ and the finite constant c is defined by

$$c := \int_{\frac{1}{m} \leqslant |x| < 1} m \, |x| \, \nu(\mathrm{d}x) + \nu(\mathbf{R} \setminus (-1, 1)).$$

Therefore, the process a_{Y_t} is majorised by

$$A_{Y_t}(s) := \int_t^{t+s} \left| \int_{[-\alpha, 0]} X(u + v) \, \mu(\mathrm{d}v) \right| \mathrm{d}u$$
$$+ s \left(|b| \, m + c + \sigma^2 m^2 + \int_{\mathbf{R}} ((m^2 x^2) \wedge 1) \, \nu(\mathrm{d}x) \right) \quad \text{for } s \in [0, \alpha]$$

and $A_{Y_t} - a_{Y_t}$ is increasing, which can be seen by some similar estimates. Since A_{Y_t} is continuous and only depends on t in the first term, it suffices to prove tightness in $C[0, \alpha]$ of the first term:

$$J_{Y_t}(s) := \int_t^{t+s} \left| \int_{[-\alpha, 0]} X(u + v) \, \mu(\mathrm{d}v) \right| \mathrm{d}u \quad \text{for } s \in [0, \alpha].$$

Recalling $r(u) = 0$ for $u < 0$, we obtain by the variation of constants formula

$$I(u) := \int_{[-\alpha, 0]} X(u + v) \, \mu(\mathrm{d}v)$$

$$= \int_{[-\alpha,0]} \left(\int_0^u r(u+v-s)F(X)(s-)\, L(ds) \right) \mu(dv)$$

$$= \int_0^u \dot{r}(u-s)F(X)(s-)\, dL(s).$$

To prove tightness of the absolutely continuous processes $(J_{Y_t})_{t\geqslant 0}$, it suffices to show that the process $(I(t+s) : s \in [0,\alpha])_{t\geqslant 0}$ is bounded in probability in $C[0,\alpha]$, that is,

$$\lim_{K\to\infty} \sup_{t\geqslant 0} P\left(\sup_{t\leqslant u\leqslant t+\alpha} |I(u)| > K \right) = 0. \tag{4.1}$$

Note that we have the exponential decay estimate $|\dot{r}(t)| \leqslant c' e^{-\beta t}$. Decomposing L into its drift, diffusion, and large and small jump parts, it is clear that only integration with respect to the large jump part N may pose problems. As in the proof of Proposition 4.2, however, the restriction on the large jumps in L and the finite intensity of N yield in a similar manner

$$\sup_{t\leqslant u\leqslant t+\alpha} \left| \int_0^u \dot{r}(u-s)F(X)(s-)\, dN(s) \right| \leqslant \sum_{s\leqslant t+\alpha} c' m e^{-\beta(t-s)} |\Delta N(s)|$$

and the tightness of the right-hand side by [11, Lemma 4.3]. Thus we infer (4.1). $\quad\square$

Theorem 4.4. *Grant Assumption 4.1. Then for the solution process $(X(t) : t \geqslant -\alpha)$ of (2.4) with initial condition $X_0 = 0$ the laws of the segments $\{\mathcal{L}(X_t)\}_{t\geqslant\alpha}$ are tight in $D[-\alpha,0]$.*

Proof. If we let $Z_t(s) := X(t-\alpha)$ for $s \in [-\alpha,0]$, then the processes $(Z_t)_{t\geqslant 0}$ of constant functions are tight in $C[-\alpha,0]$ by Proposition 4.2. On the other hand, $\{\mathcal{L}(X_t - Z_t)\}_{t\geqslant\alpha}$ are tight in $D[-\alpha,0]$ by Proposition 4.3 applying the time shift $t \mapsto t-\alpha$. Therefore the sum $(X_t - Z_t) + Z_t$ is tight in $D[-\alpha,0]$ using the result in [13, VI.3.33(a)]. $\quad\square$

4.2. From tight solutions to stationary solutions

We use the construction due to Krylov and Bogoliubov, see for example Da Prato and Zabczyk [8]. For the reader's convenience we include a proof, which is tailored for our purposes. Consider Eq. (2.4) and its Markovian semigroup $(P_t)_{t\geqslant 0}$ as defined below Proposition 3.3. Denote by $\mathcal{P} = \mathcal{P}(D[-\alpha,0])$ the set of Borel probability measures on $D[-\alpha,0]$, endowed with the topology of weak convergence of measures. Let $\langle\cdot,\cdot\rangle$ denote the duality pairing of $\mathcal{P}$ and $B_b := B_b(D[-\alpha,0])$ given by $\langle\zeta, f\rangle = \int f d\zeta$, $\zeta \in \mathcal{P}$, $f \in B_b$. Define for $t\geqslant 0$ and $\zeta \in \mathcal{P}$ the functional $P_t^*\zeta$ by

$$(P_t^*\zeta)f := \langle\zeta, P_t f\rangle, \quad f \in B_b.$$

If ζ is the distribution of an initial segment Φ, then $P_t^*\zeta$ is the distribution of $X_t(\Phi)$, since

$$\langle P_t^*\zeta, f\rangle = \int \mathbf{E}[f(X_t(\varphi))]\,\zeta(d\varphi) = \mathbf{E}\left[\mathbf{E}[f(X_t(\Phi))|\mathcal{F}_0]\right] = \mathbf{E}[f(X_t(\Phi))],$$

for $f \in B_b$. A measure $\zeta \in \mathcal{P}$ is called an *invariant measure* or *stationary distribution* of (2.4) if $P_t^*\zeta = \zeta$ for all $t\geqslant 0$, that is, $\langle\zeta, P_t f\rangle = \langle\zeta, f\rangle$ for all $f \in B_b$ and all $t\geqslant 0$.

It follows from Lemma 3.2 that $t \mapsto P_t^*\zeta$ is a continuous map from $[\alpha,\infty)$ to $\mathcal{P}$ and moreover $P_{s+t}^*\zeta = P_s^* P_t^*\zeta$ for $s,t\geqslant 0$. Further, P_t maps $C_b := C_b(D[-\alpha,0])$ into C_b for all $t\geqslant\alpha$, by Proposition 3.5.

Because of Theorem 4.4, the next theorem follows from Theorem 4.6 below.

Theorem 4.5. *Grant Assumption* 4.1. *Then there exists a stationary distribution for* (2.4).

Theorem 4.6. *If for some $\zeta \in \mathcal{P}$ the set $\{P_t^* \zeta : t \geqslant \alpha\}$ is tight, then there exists an $\eta \in \mathcal{P}$ such that $P_t^* \eta = \eta$ for all $t \geqslant 0$. Moreover, η is an element of the closed convex hull of $\{P_t^* \zeta : t \geqslant \alpha\}$ in $\mathcal{P}$.*

Proof. Define for convenience $T_t := P_{t+\alpha}$ and $\zeta(t) := T_t^* \zeta$, $t \geqslant 0$. By standard arguments it follows that there exists a unique ϑ_t in the closed convex hull of $\{P_t^* \zeta : t \geqslant \alpha\}$ in $\mathcal{P}$ such that

$$\langle \vartheta_t, f \rangle = \frac{1}{t} \int_0^t \langle \zeta(s), f \rangle \mathrm{d}s \quad \text{for all } f \in C_b.$$

Since $\{\zeta(s) : s \geqslant 0\}$ is tight, its convex hull is tight and hence relatively compact in $\mathcal{P}$ by Prohorov's Theorem. Thus the set $\{\vartheta_t : t \geqslant 0\}$ is contained in a compact set and therefore there exist a sequence $t_n \uparrow \infty$ and a measure $\eta \in \mathcal{P}$ such that $\vartheta_{t_n} \to \eta$.

Finally, for $t \geqslant \alpha$ and $f \in C_b$ we have

$$\lim_{n \to \infty} \frac{1}{t_n} \int_0^{t_n} \langle \zeta(t+s), f \rangle \mathrm{d}s = \lim_{n \to \infty} \frac{1}{t_n} \int_0^{t_n} \langle T_s^*, P_t f \rangle \mathrm{d}s = \langle \eta, P_t f \rangle$$

and on the other hand

$$\frac{1}{t_n} \int_0^{t_n} \langle \zeta(t+s), f \rangle \mathrm{d}s = \frac{1}{t_n} \int_0^{t_n} \langle \zeta(s), f \rangle \mathrm{d}s - \frac{1}{t_n} \int_0^t \langle \zeta(s), f \rangle \mathrm{d}s + \frac{1}{t_n} \int_{t_n}^{t_n+t} \langle \zeta(s), f \rangle \mathrm{d}s,$$

which converges to $\langle \eta, f \rangle$ as $n \to \infty$. Hence, $P_t^* \eta = \eta$ and it follows that $P_t^* \eta = P_t^*(P_\alpha^* \eta) = P_{t+\alpha}^* \eta = \eta$, for every $t \geqslant 0$. $\square$

We remark that the proof given above remains true in a more general setting. Indeed, we need only to replace $D[-\alpha, 0]$ by an arbitrary separable metric space E and assume that $(T_t)_{t \geqslant 0}$ is a family of bounded linear operators on $C_b(E)$ such that $T_{s+t+\alpha} = T_s T_t$ for all $s, t \geqslant 0$, that for some $\zeta \in \mathcal{P}(E)$ one has that $T_t^* \zeta \in \mathcal{P}(E)$ for all $t \geqslant 0$, that the map $t \mapsto \langle T_t^* \zeta, f \rangle$ is measurable from $[0, \infty)$ to $\mathbf{R}$ for all $f \in C_b(E)$, and that the set $\{T_t^* \zeta : t \geqslant 0\}$ is tight.

5. Uniqueness of the stationary solution

As we have seen, the Markovian semigroup is in general not eventually strongly Feller so that a main tool for establishing uniqueness of the invariant measure is not available. Moreover, when considered as a stochastic evolution equation, the generator of the deterministic equation (2.1) is only eventually compact (see [9]) and the Markov semigroup is only weakly continuous with a generator which is analytically not easily tractable (see [19]). Hence, typical analytical methods for proving uniqueness (see [17] for a survey) cannot be easily applied either.

We therefore consider several specific cases where uniqueness can be proved nevertheless: for small Lipschitz constants by a contraction argument, in the Wiener case for non-delayed diffusion coefficients by establishing the strong Feller property via Girsanov's Theorem and for compound Poisson driving processes and non-delayed drift terms by studying the deterministic behaviour between the jumps. After that, we relax the requirements and show that in full generality second-order uniqueness holds up to a constant factor. We conclude by giving an example where the invariant measures are not unique.

5.1. Small Lipschitz constants

If the function F is not too far from being constant, as measured by the Lipschitz constant, then uniqueness holds. The upper bound for the Lipschitz constant below can be reconstructed by our proof, but it is certainly not the best possible.

Theorem 5.1. *Grant Assumption* 4.1 *and suppose that the Lévy process has finite second moments. If the Lipschitz constant K of F in* (2.5) *is sufficiently small then the laws of all stationary solutions X of* (2.4) *coincide.*

Proof. Let X and Y be two stationary solutions with corresponding initial conditions X_0 and Y_0. As mentioned above Proposition 5.7 the moments $\mathbf{E}\|X_0\|_\infty^2$ and $\mathbf{E}\|Y_0\|_\infty^2$ are finite.

As $v_0(\mu) < 0$ the fundamental solution r decays exponentially with $|r(t)| \leqslant c\mathrm{e}^{-\beta t}$, $\int_0^\infty r^2(s)\exp(2\beta s)\,\mathrm{d}s < \infty$, and $\int_0^\infty \dot{r}^2(s)\exp(2\beta s)\,\mathrm{d}s < \infty$ for some constants c, $\beta > 0$. We choose an arbitrary constant $\gamma \in (0, \beta)$ and we let $Z(u) := F(X)(u) - F(Y)(u)$ for convenience. By use of the decomposition $L(t) = M(t) + \mathbf{E}[L(1)]\,t$ with a martingale M, the variation of constants formula implies for $t \geqslant \alpha$:

$$
\mathbf{E}\left[\sup_{t-\alpha\leqslant s\leqslant t}\left|\mathrm{e}^{\gamma s}(X(s) - Y(s))\right|^2\right] \leqslant 3\mathbf{E}\left[\sup_{t-\alpha\leqslant s\leqslant t}\left|\mathrm{e}^{\gamma s}x(s, X_0 - Y_0)\right|^2\right]
$$

$$
+ 3\mathbf{E}\left[\sup_{t-\alpha\leqslant s\leqslant t}\left|\int_0^s \mathrm{e}^{\gamma s}r(s - u)Z(u-)\,\mathrm{d}M(u)\right|^2\right]
$$

$$
+ 3(\mathbf{E}[L(1)])^2\mathbf{E}\left[\sup_{t-\alpha\leqslant s\leqslant t}\left|\int_0^s \mathrm{e}^{\gamma s}r(s - u)Z(u-)\,\mathrm{d}u\right|^2\right]. \tag{5.1}
$$

An application of representation (2.2) yields

$$
\mathbf{E}\left[\sup_{t-\alpha\leqslant s\leqslant t}\left|\mathrm{e}^{\gamma s}x(s, X_0 - Y_0)\right|^2\right] \leqslant d\mathbf{E}\|X_0 - Y_0\|_\infty^2 \tag{5.2}
$$

for a finite constant d depending only on the measure μ. Let r_1 be the function defined by $r_1(s) := r(s)\exp(\gamma s)$. Then we obtain for the second term in (5.1)

$$
\mathbf{E}\left[\sup_{t-\alpha\leqslant s\leqslant t}\left|\int_0^s \mathrm{e}^{\gamma s}r(s - u)Z(u-)\,\mathrm{d}M(u)\right|^2\right]
$$

$$
= \mathbf{E}\left[\sup_{t-\alpha\leqslant s\leqslant t}\left|\int_0^s \left(r_1(0) + \int_0^{s-u}\dot{r}_1(m)\,\mathrm{d}m\right)\mathrm{e}^{\gamma u}Z(u-)\,\mathrm{d}M(u)\right|^2\right]
$$

$$
\leqslant 2\mathbf{E}\left[\sup_{t-\alpha\leqslant s\leqslant t}\left|\int_0^s \mathrm{e}^{\gamma u}Z(u-)\,\mathrm{d}M(u)\right|^2\right]
$$

$$
+ 2\mathbf{E}\left[\sup_{t-\alpha\leqslant s\leqslant t}\left|\int_0^s \left(\int_0^{s-m}\mathrm{e}^{\gamma u}Z(u-)\,\mathrm{d}M(u)\right)\dot{r}_1(m)\,\mathrm{d}m\right|^2\right]. \tag{5.3}
$$

By Lemma 3.4 the first term in (5.3) can be estimated by

$$
\mathbf{E}\left[\sup_{t-\alpha\leqslant s\leqslant t}\left|\int_0^s \mathrm{e}^{\gamma u}Z(u-)\,\mathrm{d}M(u)\right|^2\right]
$$

$$\leqslant \left(\sigma^2 + \int x^2 \, \nu(\mathrm{d}x) \right) \int_0^t \mathrm{e}^{2\gamma u} \mathbf{E} \, |Z(u-)|^2 \, \mathrm{d}u. \tag{5.4}$$

Note that $\dot{r}_1$ has essentially the same asymptotic as $s \mapsto \exp(\gamma s)r(s)$. Hence if we choose a constant $\delta > 0$ such that $\gamma + \delta \leqslant \beta$ we obtain $d_1 := \int_0^\infty \exp(2\delta m) \, |\dot{r}_1(m)|^2 \, \mathrm{d}m < \infty$. Applying Hölder's inequality and a parameter transformation to the second term in (5.3) result in

$$\mathbf{E} \left[\sup_{t-\alpha \leqslant s \leqslant t} \left| \int_0^s \left(\int_0^{s-m} \mathrm{e}^{\gamma u} Z(u-) \, \mathrm{d}M(u) \right) \dot{r}_1(m) \, \mathrm{d}m \right|^2 \right]$$

$$\leqslant \mathbf{E} \left[\sup_{t-\alpha \leqslant s \leqslant t} \int_0^s \mathrm{e}^{2\delta m} \, |\dot{r}_1(m)|^2 \, \mathrm{d}m \int_0^s \left| \int_0^{s-m} \mathrm{e}^{\gamma u} Z(u-) \, \mathrm{d}M(u) \right|^2 \mathrm{e}^{-2\delta m} \, \mathrm{d}m \right]$$

$$\leqslant d_1 \mathrm{e}^{-2\delta(t-\alpha)} \mathbf{E} \left[\sup_{t-\alpha \leqslant s \leqslant t} \int_0^s \left| \int_0^m \mathrm{e}^{\gamma u} Z(u-) \, \mathrm{d}M(u) \right|^2 \mathrm{e}^{2\delta m} \, \mathrm{d}m \right]$$

$$\leqslant d_1 \mathrm{e}^{2\delta\alpha} \left(\sigma^2 + \int x^2 \, \nu(\mathrm{d}x) \right) \left(\int_0^\infty \mathrm{e}^{-2\delta m} \, \mathrm{d}m \right) \int_0^t \mathrm{e}^{2\gamma u} \mathbf{E} \, |Z(u-)|^2 \, \mathrm{d}u, \tag{5.5}$$

where we obtained the last line by Lemma 3.4. The last term in (5.1) can be estimated similarly by Hölder's inequality

$$\mathbf{E} \left[\sup_{t-\alpha \leqslant s \leqslant t} \left| \int_0^s \mathrm{e}^{\gamma s} r(s-u) Z(u-) \, \mathrm{d}u \right|^2 \right]$$

$$\leqslant \left(\int_0^\infty \mathrm{e}^{2\gamma u} \, |r(u)|^2 \, \mathrm{d}u \right) \int_0^t \mathrm{e}^{2\gamma u} \mathbf{E} \, |Z(u-)|^2 \, \mathrm{d}u. \tag{5.6}$$

By collecting the inequalities (5.2)–(5.6), using the Lipschitz condition (2.5) and applying Gronwall's Lemma we conclude that $\mathbf{E}\|X_s - Y_s\|_\infty^2 \to 0$ for $s \to \infty$ if the Lipschitz constant K is sufficiently small. Consequently, the laws of X_0 and Y_0 coincide, which completes the proof. $\quad\square$

5.2. Non-delayed diffusion coefficient

We have seen in Section 3.3 that the Markov semigroup $(P_t)_{t \geqslant 0}$ of the solution segments is in general not eventually strong Feller. This is only an effect due to the delay in the diffusion term and cannot be caused by a delayed drift for the Wiener-driven case, as we shall see now. Let us consider as special case of Eq. (2.4)

$$\mathrm{d}X(t) = \left(\int_{[-\alpha,0]} X(t+s) \, \mu(\mathrm{d}s) \right) \mathrm{d}t + f(X(t)) \, \mathrm{d}W(t) \quad \text{for } t \geqslant 0, \tag{5.7}$$

with initial segment Φ as in (2.4), a Wiener process W and a Lipschitz function $f : \mathbf{R} \to \mathbf{R}$. By a simple argument based on Girsanov's Theorem we obtain the following result.

Proposition 5.2. *If f satisfies the ellipticity condition $\inf_{x \in \mathbf{R}} f(x) > 0$, then the solution segments $(X_t : t \geqslant 0)$ of (5.7) generate a Markov semigroup on $C([-\alpha, 0])$ that is strongly Feller after time α.*

Proof. First note that the continuous functions form a closed subspace of the Skorokhod space $D[-\alpha, 0]$ such that the formerly obtained results are in the Wiener-driven case also valid on

$C[-\alpha, 0]$. Referring to Theorem 7.19 for diffusion-type processes in [15], we infer from the Lipschitz continuity of the coefficients and from the ellipticity of f that the laws Q_1 and Q_2 of the solution processes of (5.7) on $C[0, T]$, $T > 0$ arbitrary, are equivalent for different delay measures μ_1 and μ_2. The corresponding Radon–Nikodym derivative is given by

$$
\frac{\mathrm{d}Q_2}{\mathrm{d}Q_1}(X) = \exp\left(\int_0^T \left(\int_{[-\alpha,0]} X(t+s)\,(\mu_1 - \mu_2)(\mathrm{d}s) \right) f(X(t))^{-2}\,\mathrm{d}X(t) \right.
$$
$$
\left. - \frac{1}{2} \int_0^T \left(\int_{[-\alpha,0]} X(t+s)\,(\mu_1 - \mu_2)(\mathrm{d}s) \right)^2 f(X(t))^{-2}\,\mathrm{d}t \right).
$$

As in [18, Thm. 2.1] one can show that the validity of the strong Feller property at each time is invariant under the change of measure. According to that result we need to check that the semigroup is Feller and that

$$
\lim_{n\to\infty} \mathbf{E}\left| \frac{\mathrm{d}Q_2}{\mathrm{d}Q_1}(X^1(\cdot; \varphi^n)) - \frac{\mathrm{d}Q_2}{\mathrm{d}Q_1}(X^1(\cdot; \varphi)) \right| = 0
$$

for initial segments $\varphi^n \to \varphi$ in $C[-\alpha, 0]$ and for the corresponding solution process X^1 with the choice μ_1. The Feller property has been established in Section 3.3. By Scheffé's Lemma it suffices for the second condition to prove convergence in probability. This is accomplished by the continuity of the map $\varphi \mapsto X^1(\cdot, \varphi)$ from $C[-\alpha, 0]$ to $L^2([0, T] \times \Omega)$ for any T, which follows from [19, Thm. 3.1].

We have thus reduced the problem to proving the strong Feller property of the Markov semigroup generated by the solution segments $(\tilde{X}_t)_{t \geqslant 0}$ of

$$
\mathrm{d}\tilde{X}(t) = f(\tilde{X}(t))\,\mathrm{d}W(t) \quad \text{for } t \geqslant 0, \tag{5.8}
$$

as special case of (5.7) with $\mu = 0$. It is well known that this diffusion equation generates a strongly Feller semigroup on $\mathbf{R}$ under our assumptions on f, see e.g. [8, Thm. 7.1.1]. We claim that this property is inherited by the segment process. For this consider a bounded measurable functional Ψ on $C[-\alpha, 0]$ and remark that $\tilde{X}(\cdot; \varphi) = \tilde{X}(\cdot; \varphi(0))$ only depends on the initial value, not the whole segment. By the scalar Markov and weak uniqueness property we obtain for $t \geqslant \alpha$ and any initial segment φ with obvious notation

$$
\mathbf{E}[\Psi(\tilde{X}_t(\varphi))] = \mathbf{E}[\mathbf{E}[\Psi(\tilde{X}_t(\varphi)) \mid \mathcal{F}_{t-\alpha}]] = \mathbf{E}_\omega[\mathbf{E}_{\omega'}[\Psi(\tilde{X}_\alpha(\tilde{X}(t-\alpha; \varphi, \omega), \omega'))]].
$$

Setting $H(\xi) := \mathbf{E}[\Psi(\tilde{X}_\alpha(\xi))]$, $\xi \in \mathbf{R}$, the scalar strong Feller property implies the continuity of

$$
\eta \mapsto \mathbf{E}[H(\tilde{X}(t-\alpha; \eta))] = \mathbf{E}_\omega[\mathbf{E}_{\omega'}[\Psi(\tilde{X}_\alpha(\tilde{X}(t-\alpha; \eta, \omega), \omega'))]]
$$

for $\eta \in \mathbf{R}$. Since $\varphi^n \to \varphi$ in $C[-\alpha, 0]$ yields $\varphi^n(0) \to \varphi(0)$, we thus infer the continuity of

$$
\varphi \mapsto \mathbf{E}_\omega[\mathbf{E}_{\omega'}[\Psi(\tilde{X}_\alpha(\tilde{X}(t-\alpha; \varphi, \omega)), \omega')]] = \mathbf{E}[\Psi(\tilde{X}_t(\varphi))]
$$

on $C[-\alpha, 0]$, which is the asserted strong Feller property at $t \geqslant \alpha$. $\quad\square$

Corollary 5.3. *The Markov semigroup $(P_t)_{t \geqslant 0}$ is regular after time 2α. Thus, any stationary solution of (5.7) is unique and strongly mixing.*

Proof. Recall that we have regularity at t_0 if all transition probabilities $P(X_{t_0}(\varphi) \in \cdot)$ are equivalent for $\varphi \in C[-\alpha, 0]$. By Doob's Theorem [8, Thm. 4.2.1] this property yields the uniqueness and strong mixing result.

The regularity property at $t_0 > 2\alpha$ is implied by the strong Feller property at α together with the irreducibility at $t_0 - \alpha$ [8, Prop. 4.1.1], which means that all transition probabilities at time $t_0 - \alpha$ have support in the entire space. To prove the latter, we may again restrict to the case $\mu = 0$ and consider $\tilde{X}$ as in (5.8) due to the equivalence of the laws. As in [25, Cor. VIII.2.3] it follows from Girsanov's Theorem that the support of the (regular) conditional law $\mathcal{L}(\tilde{X}_{t_0-\alpha} \mid \tilde{X}(t_0 - 2\alpha) = x)$ is given by $S_x := \{f \in C[-\alpha, 0] : f(-\alpha) = x\}$. Since the law of $\tilde{X}(t_0 - 2\alpha; \varphi)$ has for the same reasons the full support $\mathbf{R}$ for any initial segment φ, we conclude by composition that $\mathcal{L}(X_{t_0-\alpha}(\varphi))$ has full support $C[-\alpha, 0]$ independent of φ, which yields the required irreducibility. $\quad\square$

The result is proved by a reduction to the non-delay case via a change-of-measure argument. Once uniqueness of the invariant measure of Eq. (5.8) has been established for certain Lévy-driven cases without delay, the same method of proof can be used to extend the result to corresponding delay equations via the general Girsanov Theorem [22].

5.3. Uniqueness in the compound Poisson case

Let us consider here the case of a Lévy triplet (b, σ^2, ν) with $\sigma = 0$, $b = 0$, and the total variation $\lambda := \|\nu\|_{\mathrm{TV}}$ finite, that is L is a compound Poisson process. If there is no delay in the drift, then we can reduce the question of uniqueness of the invariant law on the Skorokhod space $D[-\alpha, 0]$ to a property of the one-dimensional invariant law.

Proposition 5.4. *Suppose L is a compound Poisson process and consider for $a > 0$ a differential equation of the form*

$$\mathrm{d}X(t) = -aX(t)\,\mathrm{d}t + F(X)(t-)\,\mathrm{d}L(t) \quad \text{for } t \geqslant 0, \tag{5.9}$$

admitting a strong solution for any initial segment. If an invariant solution measure on $D[-\alpha, 0]$ exists and the one-dimensional marginal distributions of any two invariant measures are non-singular, then the invariant measure is unique.

Proof. Let ρ_1 and ρ_2 be two invariant measures. By coupling methods we can construct a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ carrying the process L, and the $\mathcal{F}_0$-measurable random variables $Y, Z \in D[-\alpha, 0]$ with $P^Y = \rho_1$, $P^Z = \rho_2$ and $P(Y(0) = Z(0)) > 0$. Denote by X^1 and X^2 the corresponding strong solution processes with initial conditions Y and Z, respectively.

Since with probability $\mathrm{e}^{-\lambda\alpha} > 0$ the process L does not jump on the interval $[0, \alpha]$, we have

$$P(X_\alpha^1 = X_\alpha^2, Y(0) = Z(0)) \geqslant P\left(\sum_{t \leqslant \alpha} |\Delta L(t)| = 0, Y(0) = Z(0)\right)$$

$$= P\left(\sum_{t \leqslant \alpha} |\Delta L(t)| = 0\right) P(Y(0) = Z(0)) > 0.$$

Hence, introducing the set

$$S := \{\varphi \mid \exists \omega \in \Omega : X_\alpha^1(\omega) = X_\alpha^2(\omega) = \varphi\} \subseteq D[-\alpha, 0],$$

we find for any Borel set B in $D[-\alpha, 0]$

$$P(X_\alpha^1 \in B \cap S) \geqslant P(\{\omega \mid X_\alpha^2(\omega) = X_\alpha^1(\omega), X_\alpha^1(\omega) \in B\})$$

$$= P(\{\omega \mid X_\alpha^2(\omega) = X_\alpha^1(\omega), X_\alpha^2(\omega) \in B\})$$

$$=: \tau_S(B)$$

and equivalently $P(X_\alpha^2 \in B \cap S) \geqslant \tau_S(B)$. By invariance, we conclude

$$\min\{\rho_1(B \cap S), \rho_2(B \cap S)\} \geqslant \tau_S(B) \quad \text{for all } B \in \mathcal{B}(D[-\alpha, 0])$$

with a non-negative measure τ_S satisfying $\tau_S(S) > 0$. Hence ρ_1 and ρ_2 are non-singular; for if $\rho_1(A) = 0$ and $\rho_2(A^C) = 0$ for some Borel set A, then

$$\tau_S(S) = \tau_S(S \cap A) + \tau_S(S \cap A^C) \leqslant \rho_1(S \cap A) + \rho_2(S \cap A^C) = 0.$$

As extremal points of the set of invariant measures are singular (see [8, Prop. 3.2.7]), uniqueness follows. $\square$

Theorem 5.5. *Grant Assumption* 4.1. *Suppose L is a compound Poisson process and consider Eq.* (5.9) *with $a > 0$ and $F(\varphi)(0) > 0$ for all $\varphi \in D[-\alpha, \infty)$. Then there exists a unique invariant measure for* (5.9).

Proof. If the jump measure ν is zero, then the compound Poisson process vanishes and the only invariant measure is clearly the point measure in zero. Let us now first consider the case of possible positive jumps: $\nu((0, \infty)) > 0$. By Proposition 5.4 it suffices to show that any two invariant one-dimensional distributions are non-singular. We first show that they are absolutely continuous with respect to Lebesgue measure.

Let $B \subseteq \mathbf{R}$ denote any Borel set. For the solution process X of (5.9) we find

$$P(X(t) \in e^{-at} B) \geqslant P\left(\sum_{s \leqslant t} |\Delta L(s)| = 0, \; X(0)e^{-at} \in e^{-at} B\right)$$

$$= e^{-\lambda t} P(X(0) \in B). \tag{5.10}$$

Now assuming that X is stationary with one-dimensional marginal law ρ_0, we obtain by Fubini's Theorem for any Lebesgue null set B and $T > 0$

$$\int_0^T \rho_0(e^{-at} B)\, dt = \int_0^T \int_{\mathbf{R}} \mathbf{1}_{e^{-at} B}(x)\, \rho_0(dx)\, dt$$

$$= \int_{\mathbf{R}\setminus\{0\}} \int_x^{xe^{aT}} \frac{a}{t} \mathbf{1}_B(t)\, dt\, \rho_0(dx) + \rho_0(\{0\})\mathbf{1}_B(0)$$

$$= \rho_0(\{0\})\mathbf{1}_B(0).$$

By estimate (5.10), however, the left-hand side is bounded from below by $\rho_0(B)\frac{1-e^{-\lambda T}}{\lambda}$. Hence, we infer $\rho_0(B) = 0$ for all Lebesgue null sets B with $0 \notin B$. Since F is positive and $\nu \neq 0$, we can exclude a point mass in zero because the state $\{0\}$ will be eventually left by the process P-a.s. and the probability of jumping back exactly to this state is zero. We conclude that ρ_0 is absolutely continuous with respect to the Lebesgue measure.

Let S denote the support of ρ_0. Since F is positive and bounded away from zero and L has positive jumps, there will occur with positive probability sufficiently many positive jumps of L in short time that the trajectory X will take arbitrarily high values. This means for the support S of the marginal invariant measure ρ_0 that $\sup S = +\infty$.

For a Borel set $B \subseteq (0, \infty)$ we have

$$\int_0^\infty e^{-at} \rho_0(e^{at} B)\, dt = \int_0^\infty \int_{(0,\infty)} e^{-at} \mathbf{1}_{e^{at} B}(x)\, \rho_0(dx)\, dt$$

$$= \int_{(0,\infty)} \int_0^x \frac{1}{ax} \mathbf{1}_B(s)\, ds\, \rho_0(dx).$$

If $\rho_0(B) = 0$, then (5.10) with B replaced by $e^{at}B$ yields that $\rho_0(e^{at}B) \leqslant e^{\lambda t}\rho_0(B) = 0$ for all $t > 0$, and we obtain that

$$\int_{(0,x)} \mathbf{1}_B(s)\,\mathrm{d}s = 0 \quad \text{for } \rho_0\text{-a.e. } x > 0.$$

Since $\sup S = +\infty$, we infer that the Lebesgue measure of B equals 0. Thus the Lebesgue measure of $(0, \infty)$ is absolutely continuous with respect to ρ_0.

If $\nu((-\infty, 0))$ is also positive, then the symmetric argument yields that the Lebesgue measure on $\mathbf{R}$ is equivalent to ρ_0. In any case, we know that two invariant measures are both equivalent to the appropriate Lebesgue measure and hence to each other. An application of Proposition 5.4 completes the proof. $\quad\square$

Remark 5.6. We have derived the regularity property that, unless the jump measure is zero, the one-dimensional marginals are absolutely continuous with respect to the Lebesgue measure.

In some cases one can easily derive the density of the invariant measure. For example, if we assume L to have only positive jumps of size at least $J > 0$ and $F(\varphi)(0) \in [\sigma_0, \sigma_1]$ for all $\varphi \in D[-\alpha, \infty)$ and some $\sigma_0, \sigma_1 > 0$ then the density of the marginal of the invariant measure of (5.9) near zero is given by

$$f(x) = C\frac{\lambda}{\alpha}x^{(\lambda-a)/a}, \quad x \in [0, J\sigma_0),$$

with a suitable constant C.

5.4. Second-order uniqueness

A real-valued stochastic process $(X(t) : t \geqslant -\alpha)$ will be called *second-order stationary*, if $0 < \mathbf{E}[X(t)]^2 < \infty$, the values $\mathbf{E}[X(t)]$ are constant for all $t \geqslant -\alpha$, and the function $(s, t) \mapsto \mathbf{E}[X(s)X(t)]$ depends only on the difference $s - t$. Obviously, any stationary solution of (2.4) with finite second moments is second-order stationary. If the Lévy process is a square-integrable martingale, we establish second-order uniqueness for Eq. (2.4) up to a constant factor, more precisely the expectation and the correlation function are uniquely determined and can be calculated analytically.

Note that the invariant measure exhibited in Section 4 will have finite second moments for its one-dimensional marginal whenever the Lévy process has finite second moments. This follows from the fact that the constructed tight sequence of segments (X_t) will be uniformly bounded in $L_P^2(\Omega)$ by the variation of constants formula (3.1) and Lemma 3.4.

Proposition 5.7. *Grant Assumption* 4.1. *Suppose the Lévy process is a square-integrable martingale with characteristics* (b, σ^2, ν). *Then any stationary solution* $(X(t) : t \geqslant -\alpha)$ *of* (2.4) *with finite second moments is a centered random process with auto-covariance function*

$$c(h) := \mathbf{E}[X(0)X(h)] = \frac{\mathrm{Var}[X(0)]}{\|r\|_{L^2(\mathbf{R}_+)}^2}\int_0^\infty r(s)r(s+h)\,\mathrm{d}s, \quad h \geqslant 0.$$

The spectral density is given by

$$\xi \mapsto \mathbf{E}[X(0)^2]\left(\|r\|_{L^2(\mathbf{R}_+)}\left|\chi_\mu(i\xi)\right|\right)^{-2}, \quad \xi \in \mathbf{R},$$

where $\chi_\mu(z) := z - \int_{[-\alpha,0]} e^{zu}\,\mu(\mathrm{d}u)$ *is the characteristic function of the deterministic equation* (2.1).

Proof. By the variation of constants formula (3.1) and the martingale property of L we have for $t \geqslant 0$

$$\mathbf{E}X(t) = \mathbf{E}x(t, X_0) = r(t)\mathbf{E}[X(0)] + \int_{[-\alpha, 0]} \int_s^0 r(t + s - u)\mathbf{E}[X(u)] \, \mathrm{d}u \, \mu(\mathrm{d}s).$$

Due to $\lim_{t \to \infty} r(t) = 0$ and stationarity we conclude that X is centered. Again using the variation of constants formula, we find for $h, t \geqslant 0$

$$\mathbf{E}[X(t)X(t + h)] = \mathbf{E}\left[x(t + h, X_0) \int_0^t r(t - u)F(X)(u-) \, \mathrm{d}L(u)\right]$$

$$+ \mathbf{E}\left[x(t, X_0) \int_0^{t+h} r(t + h - u)F(X)(u-) \, \mathrm{d}L(u)\right]$$

$$+ \mathbf{E}\left[x(t, X_0)\, x(t + h, X_0)\right] + \mathbf{E}\left[\int_0^t r(t - u)F(X)(u-) \, \mathrm{d}L(u)\right.$$

$$\left. \times \int_0^{t+h} r(t + h - u)F(X)(u-) \, \mathrm{d}L(u)\right].$$

As in Lemma 3.4 we obtain

$$\mathbf{E}\left[\int_0^t r(t - u)F(X)(u-) \, \mathrm{d}L(u) \int_0^{t+h} r(t + h - u)F(X)(u-) \, \mathrm{d}L(u)\right]$$

$$= \left(\sigma^2 + \int x^2 \nu(\mathrm{d}x)\right) \int_0^t r(t - u)r(t + h - u)\mathbf{E}[F(X)(u-)]^2 \, \mathrm{d}u.$$

The variance is estimated as the expectation before:

$$\mathrm{Var}[x(t, X_0)]$$

$$\leqslant 2\left(r(t)^2 \mathrm{Var}[X(0)] + \left(\int_{[-\alpha, 0]} \int_{-s}^0 |r(t + s - u)| \, \mathbf{E}|X(u)| \, \mathrm{d}u \, |\mu| \, (\mathrm{d}s)\right)^2\right),$$

which converges to 0 as $t \to \infty$. Applying the Cauchy–Schwarz inequality to the first three terms in the equation above results in

$$\mathrm{Cov}(X(0), X(h)) = \lim_{t \to \infty} \mathbf{E}[X(t)X(t + h)]$$

$$= \mathbf{E}[F(X)(0)]^2 \left(\sigma^2 + \int x^2 \nu(\mathrm{d}x)\right) \int_0^\infty r(u)r(u + h) \, \mathrm{d}u.$$

This yields the expression for the covariance function. The formula for the spectral density follows from the fact that r is the inverse Fourier transform of $\chi_\mu(-i\cdot)^{-1}$, as obtained for affine stochastic delay differential equations driven by a Wiener process in [14]. $\square$

Remark 5.8. It is seen from the proof that

$$\mathrm{Var}[X(0)] = \mathbf{E}[F(X)(0)]^2 \left(\sigma^2 + \int x^2 \nu(\mathrm{d}x)\right) \|r\|_{L^2(\mathbf{R}_+)}^2,$$

which gives some information about the size of the variance depending on bounds for the functional F. We shall see in the counterexample of Section 5.5 that this variance term need not be uniquely determined, at least for measurable functionals F.

5.5. Non-uniqueness

In the Wiener-driven case we construct an elliptic diffusion functional F which remains constant in time for certain initial segments, but with different values for different initial segments. By doing so, we can recover, for instance, Ornstein–Uhlenbeck processes with different diffusion coefficients as solutions. Suppose F is of the form

$$F(\varphi)(t) := \sqrt{\max\left(1, \min\left(\frac{2}{\alpha}\langle\varphi\rangle_{t-\alpha}^{t-\alpha/2}, 2\right)\right)}\mathbf{1}_{\mathbf{R}_+}(t) \quad \text{for } t \geqslant -\alpha,$$

where $\langle\varphi\rangle_a^b$ denotes the quadratic variation of $\varphi \in D[-\alpha, \infty)$ on the interval $[a, b]$ which might be infinite. Then F is bounded away from zero and infinity and is measurable (as a limit of measurable functionals), but obviously not continuous. Leaving our framework slightly, let us consider for a Wiener process W the equation

$$\begin{aligned}
&\mathrm{d}X(t) = -X(t)\,\mathrm{d}t + F(X)(t-)\,\mathrm{d}W(t) \quad \text{for } t \geqslant 0, \\
&X(u) = \Phi(u) \quad \text{for } u \in [-\alpha, 0].
\end{aligned} \tag{5.11}$$

Due to the positive minimal delay $\alpha/2$ there exists a strong unique solution to this equation for any $\mathcal{F}_0$-measurable initial segment by the method of steps, cf. Mao [16]. On the other hand, there exists for every $\sigma \in [1, 2]$ a stationary Ornstein–Uhlenbeck process X^σ which solves the equation (we suppose that W is a two-sided Wiener process)

$$\mathrm{d}X^\sigma(t) = -X^\sigma(t)\,\mathrm{d}t + \sigma\,\mathrm{d}W(t) \quad \text{for } t \in \mathbf{R}.$$

Then choosing $\Phi^\sigma = X_0^\sigma$, we obtain that each X^σ is also a stationary solution of (5.11). This is due to the fact that $\langle X^\sigma\rangle_{t-\alpha}^{t-\alpha/2} = \frac{\alpha}{2}\sigma^2$ and thus $F(X^\sigma)(t) = \sigma$ hold for all $t \geqslant 0$ and $\sigma \in [1, 2]$.

This example shows that some kind of regularity of F has to be imposed to guarantee uniqueness, but we do not know whether for functionals F with large, but finite Lipschitz constants uniqueness already breaks down. It is interesting to note that a similar dichotomy has been described by Mohammed and Scheutzow [20] for the long time behaviour depending on the diffusion functional.

Acknowledgement

References

[1] S. Albeverio, B. Rüdiger, J.-L. Wu, Invariant measures and symmetry property of Lévy type operators, Potential Anal. 13 (2) (2000) 147–168.

[2] D. Applebaum, Lévy Processes and Stochastic Calculus, Cambridge University Press, Cambridge, 2004.

[3] Y. Bakhtin, J.C. Mattingly, Stationary solutions of stochastic differential equation with memory and stochastic partial differential equation, Commun. Contemp. Math. 7 (5) (2005) 553–582.

[4] O.E. Barndorff-Nielsen, N. Shephard, Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics (With discussion), J. R. Stat. Soc. Ser. B Stat. Methodol. 63 (2) (2001) 167–241.

[5] S. Bonaccorsi, G. Tessitore, Asymptotic behavior of infinite dimensional stochastic differential equations by anticipative variation of constants formula, Appl. Math. Optim. 44 (3) (2001) 203–225.

[6] P.-L. Chow, R.Z. Khasminskii, Stationary solutions of nonlinear stochastic evolution equations, Stoch. Anal. Appl. 15 (5) (1997) 671–699.

[7] G. Da Prato, J. Zabczyk, Stochastic Equations in Infinite Dimensions, Cambridge University Press, Cambridge, 1992.

[8] G. Da Prato, J. Zabczyk, Ergodicity for Infinite Dimensional Systems, Cambridge University Press, Cambridge, 1996.

[9] O. Diekmann, S.A. van Gils, S.M. Verduyn Lunel, H.-O. Walther, Delay Equations. Functional-, Complex-, and Nonlinear Analysis, Springer-Verlag, New York, 1995.

[10] D. Gatarek, B. Goldys, On invariant measures for diffusions on Banach spaces, Potential Anal. 7 (2) (1997) 539–553.

[11] A. Gushchin, U. Küchler, On stationary solutions of delay differential equations driven by a Lévy process, Stochastic Process. Appl. 88 (2000) 195–211.

[12] K. Itô, M. Nisio, On stationary solutions of a stochastic differential equation, J. Math. Kyoto Univ. 4 (1964) 1–75.

[13] J. Jacod, A. N. Shiryaev, Limit Theorems for Stochastic Processes, Springer, Berlin, 2003.

[14] U. Küchler, B. Mensch, Langevin's stochastic differential equation extended by a time-delayed term, Stoch. Stoch. Rep. 40 (1–2) (1992) 23–42.

[15] R.S. Liptser, A.N. Shiryaev, Statistics of Random Processes. 1, Springer, Berlin, 2001.

[16] X. Mao, Stochastic Differential Equations and Their Applications, Horwood Publishing, Chichester, 1997.

[17] B. Maslowski, J. Seidler, Invariant measures for nonlinear SPDE's: Uniqueness and stability, Arch. Math. (Brno) 34 (1) (1998) 153–172.

[18] B. Maslowski, J. Seidler, Probabilistic approach to the strong Feller property, Probab. Theory Related Fields 118 (2) (2000) 187–210.

[19] S.-E.A. Mohammed, Stochastic Functional Differential Equations, Pitman Advanced Publishing Program, Melbourne, 1984.

[20] S.-E.A. Mohammed, M.K.R. Scheutzow, Lyapunov exponents of linear stochastic functional differential equations. II: Examples and case studies, Ann. Probab. 25 (3) (1997) 1210–1240.

[21] W.R. Pestman, Measurability of linear operators in the Skorokhod topology, Bull. Belg. Math. Soc. 2 (4) (1995) 381–388.

[22] P. Protter, Stochastic Integration and Differential Equations. A New Approach, Springer-Verlag, Berlin, 1990.

[23] S. Raible, Lévy processes in finance: Theory, numerics, and empirical facts, Ph.D. Thesis, Freiburg i.Br.: Univ. Freiburg, Mathematische Fakultät. http://www.freidok.uni-freiburg.de/volltexte/51.

[24] M. Reiß, M. Riedle, O. van Gaans, On Émery's inequality and a variation-of-constants formula, Preprint 05-1, Institute of Mathematics, Humboldt-University of Berlin, 2005. http://www.mathematik.hu-berlin.de/publ/publ01.html.

[25] D. Revuz, M. Yor, Continuous Martingales and Brownian Motion, Springer, Berlin, 1999.

[26] K.-I. Sato, Lévy Processes and Infinitely Divisible Distributions, Cambridge University Press, Cambridge, 1999.

[27] M.K.R. Scheutzow, Qualitative behaviour of stochastic delay equations with a bounded memory, Stochastics 12 (1984) 41–80.

[28] S. J. Wolfe, On a continuous analogue of the stochastic difference equation $X_n = \rho X_{n-1} + B_n$, Stochastic Process. Appl. 12 (1982) 301–312.