A nonlinear Kolmogorov equation for stochastic functional delay differential equations with jumps

F. Cordoni, L. Di Persio[†] and I. Oliva

Abstract. We consider a stochastic functional delay differential equation, namely an equation whose evolution depends on its past history as well as on its present state, driven by a pure diffusive component plus a pure jump Poisson compensated measure. We lift the problem in the infinite dimensional space of square integrable Lebesgue functions in order to show that its solution is an L^2-valued Markov process whose uniqueness can be shown under standard assumptions of locally Lipschitzianity and linear growth for the coefficients. Coupling the aforementioned equation with a standard backward differential equation, and deriving some ad hoc results concerning the Malliavin derivative for systems with memory, we are able to derive a non-linear Feynman-Kac representation theorem under mild assumptions of differentiability.

Mathematics Subject Classification (2010). 34K50; 60H07; 60H10; 60H30; 60G51.

Keywords. stochastic delay differential equations , quadratic variation, Lévy processes, Feynman–Kac formula, mild solution.

Contents

[†]Corresponding author.

1. Introduction

During recent years, an increasing attention has been paid to stochastic equations whose evolution depends not only on the present state, but also on the past history. In particular, it has been shown that memory effects cannot be neglected when dealing with many natural phenomena. As examples, let us mention the coupled *atmosphere-ocean models*, see, e.g., [9], and their applications in describing climate changes in the environmental sciences setting, or the effect of time delay considering population dynamics, when suitable growth models are considered, see, e.g., [2]. Nevertheless, assumptions that will be made throughout the work are mainly taken into account having in mind concrete financial applications. For instance, in [38, 53] the authors pointed out how delay arises in commodity markets and energy markets, when it is necessary to take into account the impact of production and transportation, whereas in [5, 11] the authors provide applications to option pricing in markets with memory. Similarly, delay naturally arises when dealing with financial instruments as, e.g., *Asian options* or *lookback options*, as studied in, e.g. [16] and references therein.

For the mathematical foundations of the theory of stochastic functional delay differential equations (SFDDEs) we refer to [41], as well as to [42] to many motivating examples concerning the treatment of equations with delay. In particular the monograph [41] represents an early and deep treatment of SFDDE's, where several results concerning existence and uniqueness of solutions to SFDDE's as well as regularity results are derived. The theory of delay equations has seen a renewed attention recently, in particular in [13, 14] an ad hoc stochastic calculus, known as *functional Itô's calculus*, has been derived, based on a suitable Itô's formula for delay equations. Also, in past few years several different works have appeared deriving fundamental results on delay equations based on semigroup theory and infinite dimensional analysis, see, e.g. [30, 31], or based on the calculus via regularization, see, e.g. [22, 33]. Eventually, in [30, 22], it has been shown that SFDDE's, *path-dependent calculus* and delay equations via semigroup theory, are in fact closely related.

Having in mind possible financial applications, the aim of the present work is to extend some results concerning the *non-linear Feynman-Kac formula* for a forward-backward system with delay, where the driving noise is a non Gaussian Lévy process, using the theory of SFDDE's first introduced in [41]. It is worth to mention that, particularly during last decades, asset price dynamics and, more generally, financial instruments processes, have been widely characterized by trajectories showing sudden changes and ample jumps. It follows that the classical Black and Scholes picture has to be refined by allowing to consider random components constituted by both diffusive and jump components.

We thus consider the following $\mathbb{R}-$valued SFDDE with jumps

$$\begin{aligned} dX(t) =& \mu(t, X(t+\cdot), X(t))dt + \sigma(t, X(t+\cdot), X(t))dW(t) + \\ &+ \int_{\mathbb{R}_0} \gamma(t, X(t+\cdot), X(t), z)\tilde{N}(dt, dz)\,, \end{aligned} \tag{1.1}$$

where $W(t)$ is a standard Brownian motion, $\tilde{N}(dt, dz)$ is a compensated Poisson random measure with associated Lévy measure ν. Also the notation $X(t+\cdot)$ means that the *coefficients* μ, σ and γ, at time t, depend not only on the present state of the process X but also on its past values. Exploiting the concept of *segment* of a process X, see, e.g., [41, 42], we will lift the finite dimensional $\mathbb{R}$-valued process solution to (1.1) to an infinite dimensional stochastic process with values in a suitable path-space. More precisely, in what follows we will denote by $r > 0$ the maximum delay taken into account and $T < \infty$ a fixed finite time horizon. Thus, for an $\mathbb{R}-$valued stochastic process X, we indicate with $X(t)$ the value in $\mathbb{R}$ at time $t \in [0, T]$ and with X_t the corresponding segment, i.e. the trajectory in the time interval $[t-r, t]$, that is $X_t(\cdot) : [-r, 0] \to \mathbb{R}$ is such that $X_t(\theta) := X(t+\theta)$ for all $\theta \in [-r, 0]$.

Then equation (1.1) can be rewritten as s

$$\begin{cases} \mathrm{d}X(t) = & \mu(t, X_t, X(t))\mathrm{d}t + \sigma(t, X_t, X(t))\mathrm{d}W(t) + \\ & + \int_{\mathbb{R}_0} \gamma(t, X_t, X(t), z)\tilde{N}(\mathrm{d}t, \mathrm{d}z) \\ (X_0, X(0)) & = (\eta(\theta), x) \end{cases}, \tag{1.2}$$

for all $t \in [0, T]$, $\theta \in [-r, 0]$, $x \in \mathbb{R}$ and η a suitable $\mathbb{R}-$valued function on $[-r, 0]$.

Remark 1.1. In what follows we will only consider the $1-$dimensional case, the case of a $\mathbb{R}^d-$valued stochastic process, perturbed by a general $\mathbb{R}^m-$dimensional Wiener process and a $\mathbb{R}^n-$dimensional Poisson random measure, with $d > 1$, $m > 1$ and $n > 1$, can be easily obtained from the present one.

In order to take into account the delay component, we study the equation (1.2) in the *Delfour-Mitter space* defined as follows $M^2 := L^2([-r, 0]; \mathbb{R}) \times \mathbb{R}$, endowed with the scalar product

$$\langle (X_t, X(t)), (Y_t, Y(t)) \rangle_{M_2} = \langle X_t, Y_t \rangle_{L^2} + X(t) \cdot Y(t)\,,$$

and norm

$$\|(X_t, X(t))\|^2_{M_2} = \|X_t\|^2_{L^2} + |X(t)|^2\,, \quad (X_t, , X(t)) \in M_2\,, \tag{1.3}$$

where $\cdot$, resp. $|\cdot|$, stands for the scalar product in $\mathbb{R}$, resp. the absolute value, and $\langle \cdot, \cdot \rangle_{L^2}$, resp. $\|\cdot\|_{L^2}$, is the scalar product, resp. norm, in $L^2([-r, 0]; \mathbb{R}) =: L^2$. Note that the space M^2 is a separable Hilbert space, see, e.g., [6]. The *Delfour-Mitter space* can be generalized to be a separable Banach space if we consider $p \in (1, \infty)$, equipped with the appropriate norm. In this work we will consider the case $p = 2$.

Alternatively, we could have considered the space of *càdlàg functions*, i.e. right–continuous functions with finite left limit, on the interval $[-r, 0]$, $\mathbf{D} := D([-r, 0]; \mathbb{R})$ called *Skorokhod space*; in particular $\mathbf{D}$ is a non separable

Banach space if endowed with the sup norm $\|\cdot\|_{\mathbf{D}} = \sup_{t\in[-r,0]}|\cdot|$. We also have that $\mathbf{D} \subset M^2$ with the injection being continuous, see, e.g., [6]. Nevertheless, choosing M^2 as state space we cannot deal with the case of discrete delays, see, e.g., [6, pag. 3], or [41].

The choice of considering the Hilbert space M^2 instead of the *Skorokhod space* $\mathbf{D}$ has two main motivations. First, the separability of the Hilbert space M^2 allows us to prove a fundamental property for the SFDDE under investigation, that is we will show that although exhibiting delay, the SFDDE (1.2) is a M^2-Markov process. The same, primary due to the fact that the *Skorokhod space* $\mathbf{D}$ is not a separable Banach space, does not hold if one considers $\mathbf{D}$ as state space. One can nevertheless avoid this problem considering weaker topologies on $\mathbf{D}$, such as the so-called *Skorokhod topology*, under which $\mathbf{D}$ can be shown to be separable, see, e.g [46]. For the sake of simplicity we will address here the simpler case of an M^2-valued process, leaving the more technical case of $\mathbf{D}-$process to future investigations.

Second reason we are choosing here the Hilbert space M^2 is the extensive use we will do of *Malliavin calculus*. In fact Malliavin calculus provides a powerful tool to study general regularity properties of a process or, as in the present case, to obtain representation theorem under mild regularity assumptions for the process. Nevertheless its generalization to the infinite dimensional setting, mostly when the driving noise is a general Lévy process, is rather technical and the theory, even if promising results have been obtained, see [4] and references therein, is still not completely developed. For these reasons, in the present work, we will employ an approach similar to the one used in [25] for backward stochastic differential equations with time-delayed generator and in [33] for SFDDE with a Brownian noise. We will in fact exploit the fact that the original equation (1.2) has value in a finite dimensional space, so that one can use standard results in Malliavin calculus. This will imply that, exactly as in [33], we will not use a purely infinite dimensional formulation for our problem, such as for instance the one first formulated in [12] and subsequently used in [35]. In fact the M^2-setting will be mainly used to prove existence and uniqueness of a solution and most important, as mentioned above, we are able to prove that the SFDDE (1.2) is a M^2-Markov process.

We have already mentioned that, despite the fact that the process (1.2) exhibits memory effects, lifting the problem to consider a M^2-value solution leads to obtain a solution which is a *Markov process*. Taking in mind latter result and in order to derive the *Kolomogorov equation* associated to equation (1.2), we will consider, following [33, 34, 35], a classical $\mathbb{R}-$valued backward stochastic differential equation (BSDE), coupled with the forward equation equation (1.2), which evolves according to

$$\begin{cases} \mathrm{d}Y(t) &= \psi\left(t, X_t, X(t), Y(t), Z(t), \int_{\mathbb{R}_0} U(t,z)\delta(z)\nu(\mathrm{d}z)\right)\mathrm{d}t \\ &\quad + Z(t)\mathrm{d}W(t) + \int_{\mathbb{R}_0} U(t,z)\tilde{N}(\mathrm{d}t,\mathrm{d}z) \\ Y(T) &= \phi(X_T, X(T)) \end{cases}, \tag{1.4}$$

where ψ and ϕ are given suitable functions to be specified later on. We recall that a solution to equation (1.4) is a triplet (Y, Z, U), where Y is the *state process*, while Z and U are the *control processes*.

It is well known that, when the delay is not involved, there exists a Feynman-Kac representation theorem that connects the solution of the coupled forward-backward system (1.2) and (1.4), to a deterministic semi-linear partial integro-differential equation, see, e.g., [24, Chapter. 4] or [7] for further details. When the delay is taken into consideration, previous result has been recently proved in the Brownian case in [33, 35]. In the present paper we extend latter result taking into consideration a non Gaussian Lévy noise. In particular, exploiting notations already introduced, we will consider the following coupled forward-backward stochastic differential equation (FBSDE) with delay, for $t \in [\tau, T] \subset [0, T]$,

$$\begin{cases} \mathrm{d}X^{\tau,\eta,x}(t) & = \mu(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t))\mathrm{d}t + \sigma(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t))\mathrm{d}W(t) \\ & \quad + \int_{\mathbb{R}_0} \gamma(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t), z)\tilde{N}(\mathrm{d}t, \mathrm{d}z) \\ (X_\tau^{\tau,\eta,x}, X^{\tau,\eta,x}(\tau)) & = (\eta, x) \in M^2 \\ \mathrm{d}Y^{\tau,\eta,x}(t) & = \psi\left(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t), Y^{\tau,\eta,x}(t), Z^{\tau,\eta,x}(t), \tilde{U}^{\tau,\eta,x}(t)\right)\mathrm{d}t \\ & \quad + Z^{\tau,\eta,x}(t)\mathrm{d}W(t) + \int_{\mathbb{R}_0} U^{\tau,\eta,x}(t,z)\tilde{N}(\mathrm{d}t, \mathrm{d}z) \\ Y^{\tau,\eta,x}(T) & = \phi(X_T^{\tau,\eta,x}, X^{\tau,\eta,x}(T)) \end{cases}, \tag{1.5}$$

where we have denoted for short by

$$\tilde{U}^{\tau,\eta,x}(t) := \int_{\mathbb{R}_0} U^{\tau,\eta,x}(t,z)\delta(z)\nu(dz)\,.$$

Moreover we have denoted by $X^{\tau,\eta,x}$ the value of the process with starting time $\tau \in [0, T]$ and initial value $(\eta, x) \in M^2$. In what follows we will often omit the dependence on the initial value point (η, x) and we assume that the process starts at time $\tau = 0$, i.e. $X_t^{0,\eta,x} =: X_t$. Also, in order to simplify notation, most of the results will be proved for $\tau = 0$, the extension to the general case of $\tau \neq 0$ being straightforward.

We are going to connect the solution to the FBSDE (1.5) to the solution of the following partial integro-differential Hilbert–space valued equation

$$\begin{cases} \frac{\partial}{\partial t}u(t,\eta,x) + \mathcal{L}_t u(t,\eta,x) = \psi\left(t,\eta,x,u(t,\eta,x), \partial_x u(t,\eta,x)\sigma(t,\eta,x), \mathcal{J}u(t,\eta,x)\right) \\ u(T,\eta,x) = \phi(\eta,x)\,, \quad t \in [0,T]\ ,\ (\eta,x) \in M^2\,. \end{cases} \tag{1.6}$$

where $\mathcal{L}_t$ is the *infinitesimal generator* of the forward M^2-valued process in equation (4.1), ∂_x is the derivative with respect to the present state $X(t)$ and $\mathcal{J}$ is the operator

$$\mathcal{J}u(t,\eta,x) := \int_{\mathbb{R}_0} [u(t,\eta,x+\gamma(t,\eta,x,z)) - u(t,\eta,x)]\delta(z)\nu(dz)\,.$$

In particular, we will consider a mild notion of solution to equation (1.6), so that we say that a function $u : [0,T] \times M^2 \to \mathbb{R}$ is a *mild solution* to equation (1.6) if there exist $C > 0$ and $m \geq 0$, such that, for any $t \in [0,T]$ and any (η_1, x_1), $(\eta_2, x_2) \in M^2$, u satisfies

$$\begin{aligned} &|u(t,\eta_1,x_1) - u(t,\eta_2,x_2)| \leq C|(\eta_1,x_1)-(\eta_2,x_2)|_2(1+|(\eta_1,x_1)|_2+|(\eta_2,x_2)|_2)^m\,, \\ &|u(t,0,0)| \leq C\,, \end{aligned} \tag{1.7}$$

and the following equality holds true

$$u(t,\eta,x) = P_{t,T}\phi(\eta,x) + \int_t^T P_{t,s}[\psi(\cdot,u(s,\cdot),\partial_x u(s,\cdot)\sigma(s,\cdot),\mathcal{J}u(s,\cdot)](\eta,x)\mathrm{d}s\,, \tag{1.8}$$

for all $t \in [0,T]$, and $(\eta,x) \in M^2$, $P_{t,s}$ being the *Markov semigroup* related to the equation (1.2). In particular we would like to stress that we require the solution u to equation (3.5) to be locally Lipschitz continuous with respect to the second variable with at most polynomial growth, so that the derivative appearing in the right–hand–side of equation (1.8) is to be defined in a mild sense, to better specified later on.

We thus define

$$\begin{cases} Y^{\tau,\eta,x}(t) & := u(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t)) \\ Z^{\tau,\eta,x}(t) & := \partial_x u(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t))\ \sigma(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t)) \\ U^{\tau,\eta,x}(t,z) & := u\left(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t) + \gamma(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t), z)\right) \\ & \quad\ -u\left(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t)\right) \end{cases}$$

then the triplet $(Y^{\tau,\eta,x}, Z^{\tau,\eta,x}, U^{\tau,\eta,x})$ is the unique solution to the backward equation (1.4), where ∂_x is the derivative with respect to the $\mathbb{R}-$valued present state $X(s)$ of $(X_s, X(s))$, u being the *mild solution* to the *Kolmogorov equation*

$$\begin{cases} \frac{\partial}{\partial t}u(t,\eta,x) + \mathcal{L}_t u(t,\eta,x) = \psi\left(t,\eta,x,u(t,\eta,x),\partial_x u(t,\eta,x)\sigma(t,\eta,x),\mathcal{J}u(t,\eta,x)\right) \\ u(T,\eta,x) = \phi(\eta,x)\,, \quad t \in [0,T]\ ,\ (\eta,x) \in M^2\,. \end{cases}.$$

As regard the notion of mild solution for the *Kolmogorov equation* (1.6), we have to mention that different notions can be chosen. Our choice is due mainly to the fact that, since we do not require for differentiability assumptions, it seems to be the most suitable for financial applications. As an example, in option pricing one usually have that the terminal payoff of a given claim is Lipschitz continuous, without being differentiable. Moreover, mild differentiability assumptions and the use of delayed coefficients, allow the above notion to be particularly suited to price exotic options, as in the case of *Asian options*, see [16]. Furthermore, the notion of mild solution we have chosen well emphasize the intrinsic stochastic nature of the problem, also providing an immediate connection to BSDE theory, hence allowing to treat general semilinear PIDE. We refer to [24] for a comprehensive treatment of

BSDE's with general Lévy noise, see also [15] and references therein for a more financially oriented study of the topic.

We would also like to recall that different notions of mild solution for partial integro-differential already exist in literature, mostly considering *Volterra-type* equations, allowing also to exhibit delays, we refer the interested reader to [10, 36, 40]. Also, in a setting similar to the present one, a notion of mild solution for SPDE's driven by $\alpha-$stable noise can be found in [50, 51], where the authors study mild solutions of semilinear parabolic equations in an infinite dimensional Hilbert space, in order to obtain the associated Hamilton-Jacobi-Bellman equation with applications to the stochastic optimal control problems.

Last but not least, rather recently a further notion of mild solution to delay equations has appeared in literature. This is an ad hoc generalization of the standard notion of *viscosity solution*. In particular first in [27], and then in [28, 29], a new notion of viscosity solution to PDE with delays, called *path-dependent PDE*, has been formulated, based on the newly developed *functional Itô calculus* mentioned above. Latter notion has been also exploited to treat *path-dependent* PDE with delayed generator, see, e.g., [21], or [17] for an application to mathematical finance.

The paper is organized as follows: in Section 2 we introduce necessary notations and formalize the tools necessary to treat delay equations in the Hilbert space M^2. In particular Section 2 is devoted to the characterization of fundamental results on SFDDE, such as existence and uniqueness, as well as the Markov property of the forward process. Thus subsection 2.1 is devoted to results concerning Malliavin calculus for delay equations which will be needed in order to prove the main representation theorem. In Section 3 we prove the main result based on Malliavin calculus, which is related to the study of the joint quadratic variation of the forward equation and a suitable function; in Section 4 we give the *non-linear Feynman-Kac theorem* that is later used to derive a deterministic representation to the FBSDE. Finally in Section 5 we give an application of obtained result to optimal control.

2. Forward stochastic functional differential equation with delay

In this Section we introduce the notation used throughout the paper, also presenting basic definitions and main results related to the mathematical techniques involved in our approach. Some results are already established in literature, such as existence and uniqueness of solutions, whereas others are proved here for the first time.

Let us consider a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$, where $(\mathcal{F}_t)_{t\in[0,T]}$ is the *natural* filtration jointly generated by the random variables $W(s)$ and $N(ds, dz)$, for all $z \in \mathbb{R} \setminus \{0\} =: \mathbb{R}_0$ and for all $s \in [0, T]$, augmented by all $\mathbb{P}$-null sets, W being a 1-dimensional Brownian motion, while N is a 1-dimensional Poisson random measure, independent from W, with associated

Lévy measure $\nu(dz)$, satisfying

$$\int_{\mathbb{R}_0} \min\{1, z^2\}\nu(dz) < \infty\,; \tag{2.1}$$

also we define the compensated random measure $\tilde{N}(dt, dz) := N(dt, dz) - \nu(dz)dt$.

We will further assume in what follows that the Lévy measure ν satisfies

$$\int_{\mathbb{R}_0} |z|^2\nu(dz) < \infty\,. \tag{2.2}$$

We underline that condition (2.1) is a standard assumption in the definition of a Lévy measure ν, whereas assumption (2.2) implies that the process has a finite second moment, which is a natural assumption if one has in mind financial applications.

In the following, we fix a *delay* $r > 0$ and we will use the notation $X(t)$ to denote the present state at time t of the real valued process X, whereas we use X_t to denote the *segment* of the path described by X during the time interval $[t-r, t]$ with values in a suitable infinite dimensional path space. In particular, we refer to the couple

$$\Big((X(t+\theta))_{\theta\in[-r,0]}\,, X(t)\Big) =: (X_t, X(t))\ .$$

From now on, we define $M^2 := L^2 \times \mathbb{R} := L^2([-r, 0]; \mathbb{R}) \times \mathbb{R}$, endowed with the scalar product

$$\langle (X_t, X(t)), (Y_t, Y(t))\rangle_{M_2} = \langle X_t, Y_t\rangle_{L^2} + X(t)\cdot Y(t)\,,$$

and norm

$$\|\,(X_t, X(t))\,\|^2_{M^2} = \|X_t\|^2_{L^2} + |X(t)|^2\,, \tag{2.3}$$

namely the *Delfour-Mitter space*, which is a separable Hilbert space, see, e.g., [41] and reference therein for details.

Furthermore, for any $p \in [2, \infty)$, we denote by $\mathcal{S}^p(t) := \mathcal{S}^p([0, t]; M^2)$ and we say that a M^2-valued stochastic process $(X_s, X(s))_{s\in[0,t]}$ belongs to $\mathcal{S}^p(t)$ if

$$\|X\|^p_{\mathcal{S}^p(t)} := \mathbb{E}\left[\sup_{s\,\in\,[0,t]} \|(X_s, X(s))\|^p_{M^2}\right] < \infty\ .$$

We denote for short $\mathcal{S}^p := \mathcal{S}^p(T)$. For the sake of simplicity, the following notation is used throughout the paper: $|\cdot|_2$ denotes the norm in M^2 and $|\cdot|$ the absolute value in $\mathbb{R}$.

Remark 2.1. Let us stress that we will consider here a $\mathbb{R}-$valued stochastic process X, nevertheless any result that follows can be easily generalized to the case of an $\mathbb{R}^d-$ valued stochastic process. In particular we would have considered the *Delfour-Mitter space* $M^2([-r, 0]; \mathbb{R}^d) := L^2([-r, 0]; \mathbb{R}^d) \times \mathbb{R}^d$, see, e.g. [6].

As briefly said in Section 1, the main goal of this work is to study a stochastic functional delay differential equation (SFDDE) of the form

$$\begin{cases} \mathrm{d}X(t) = \mu(t, X_t, X(t))\mathrm{d}t + \sigma(t, X_t, X(t))\mathrm{d}W(t) + \int_{\mathbb{R}_0} \gamma(t, X_t, X(t), z)\tilde{N}(\mathrm{d}t, \mathrm{d}z) \\ (X_0, X(0)) = (\eta, x) \in M^2 \end{cases}, \tag{2.4}$$

for all $t \in [0, T]$. We will assume the functionals μ, σ and γ to fulfil the following assumptions.

Assumption 2.2. (A1) the coefficients

$$\mu : [0,T] \times M^2 \to \mathbb{R}\,, \quad \sigma : [0,T] \times M^2 \to \mathbb{R}\,, \quad \gamma : [0,T] \times M^2 \times \mathbb{R}_0 \to \mathbb{R}$$

are continuous.

(A2) There exists $K > 0$ such that for all $t \in [0, T]$ and for all (η_1, x_1), $(\eta_2, x_2) \in M^2$,

$$\begin{aligned} &|\mu(t, \eta_1, x_1) - \mu(t, \eta_2, x_2)|^2 + |\sigma(t, \eta_1, x_1) - \sigma(t, \eta_2, x_2)|^2 \\ &\qquad + \int_{\mathbb{R}_0} |\gamma(t, \eta_1, x_1, z) - \gamma(t, \eta_2, x_2, z)|^2 \nu(dz) \\ &\qquad \le K|(\eta_1, x_1) - (\eta_2, x_2)|_2^2 (1 + |(\eta_1, x_1)|_2^2 + |(\eta_2, x_2)|_2^2)\,. \end{aligned}$$

Throughout the paper, we will look for *strong solution* to equation (2.4) in the following sense.

Definition 2.3. We say that $X := (X_t, X(t))_{t\in[0,T]}$ is a *strong solution* to equation (2.4) if for any $t \in [0, T]$ X is indistinguishably unique and $(\mathcal{F}_t)_{t\in[0,T]}$-adapted and it holds $\mathbb{P}-$a.s.

$$\begin{aligned} X(t) &= x + \int_0^t \mu(s, X_s, X(s))ds + \int_0^t \sigma(s, X_s, X(s))dW(s) \\ &\quad + \int_0^t \int_{\mathbb{R}_0} \gamma(s, X_s, X(s), z)\tilde{N}(ds, dz)\,, \\ X_0 &= \eta\,. \end{aligned}$$

In what follows we will denote by $(X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t))$ the M^2-value of the process at time $t \in [\tau, T]$, with initial value $(\eta, x) \in M^2$ at initial time $\tau \in [0, T]$. However, for the sake of brevity, in most of the results, we will avoid to state the dependence on the initial value (τ, η, x) writing for short $(X_t, X(t))$ instead of $(X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t))$.

Now we provide an existence and uniqueness result for equation (2.4).

Theorem 2.4. *Suppose that μ, σ and γ satisfy conditions $(A1) - (A2)$ in Assumptions* 2.2. *Then, for all $t \in [0, T]$ and $(\eta, x) \in M^2$, there exists a unique strong solution to the SFDDE* (1.2) *in $\mathcal{S}^p$ and there exists $C_1 := C_1(K, L, T, p)$ such that*

$$\|X^{\eta,x}\|_{\mathcal{S}^p}^p \le C_1(1 + |(\eta, x)|_2^p)\,. \tag{2.5}$$

Moreover, the map $(\eta, x) \mapsto X^{\eta,x}$ *is Lipschitz continuous from* M^2 *to* $\mathcal{S}^p$ *and it exists* $C_2 := C_2(K, L, T)$ *such that*

$$\|X^{\eta_1,x_1} - X^{\eta_2,x_2}\|^p_{\mathcal{S}^p} \leq C_2|(\eta_1, x_1) - (\eta_2, x_2)|^p_2 \,. \tag{2.6}$$

Proof. Existence and uniqueness of the solution to equation (2.4), as well as the estimate in equation (2.5), are proved in [6, Th. 2.14].

As regards equation (2.6), exploiting the Burkholder-Davis-Gundy inequality, see, e.g. [3, Section 4.4.], we have that, for any $t \in [0, T]$, denoting for short by C several positive constants,

$$\begin{aligned}
&|X^{\eta_1,x_1} - X^{\eta_2,x_2}|^p_{\mathcal{S}^p} = \\
&= \mathbb{E} \sup_{t\in[0,T]} |(X^{\eta_1,x_1}_t, X^{\eta_1,x_1}(t)) - (X^{\eta_2,x_2}_t, X^{\eta_2,x_2}(t))|^p_2 \leq \\
&\leq C|(\eta_1, x_1) - (\eta_2, x_2)|^p_2 \\
&+ C\left[\int_0^t |\mu(s, X^{\eta_1,x_1}_s, X^{\eta_1,x_1}(s)) - \mu(s, X^{\eta_2,x_2}_s, X^{\eta_2,x_2}(s))|^p ds\right. \\
&+ \left(\int_0^t |\sigma(s, X^{\eta_1,x_1}_s, X^{\eta_1,x_1}(s)) - \sigma(s, X^{\eta_2,x_2}_s, X^{\eta_2,x_2}(s))|^2 ds\right)^{\frac{p}{2}} \\
&+ \left.\int_0^t \int_{\mathbb{R}_0} |\gamma(s, X^{\eta_1,x_1}_s, X^{\eta_1,x_1}(s), z) - \gamma(s, X^{\eta_2,x_2}_s, X^{\eta_2,x_2}(s), z)|^p \nu(dz) ds\right] ,
\end{aligned}$$

so that from the Lipschitz continuity in assumption 2.2 $(A2)$, it follows

$$\begin{aligned}
&\mathbb{E} \sup_{t\in[0,T]} |(X^{\eta_1,x_1}_t, X^{\eta_1,x_1}(t)) - (X^{\eta_2,x_2}_t, X^{\eta_2,x_2}(t))|^p_2 \leq \\
&\leq C|(\eta_1, x_1) - (\eta_2, x_2)|^p_2 + \\
&+ \int_0^T \sup_{s\in[0,q]} |(X^{\eta_1,x_1}_s, X^{\eta_1,x_1}(s)) - (X^{\eta_2,x_2}_s, X^{\eta_2,x_2}(s))|^p_2 ds \,,
\end{aligned}$$

and the claim follows from Grownall's inequality. □

Remark 2.5. We want to stress that a result analogous to Thm. 2.4 can be obtained by replacing the Delfour-Mitter space M^2 with the space $\mathbf{D}$ of càdlàg functions, with the corresponding sup norm $\|\cdot\|_{\mathbf{D}} = \sup_{t\in[-r,0]} |\cdot|$, see e.g. [6, 46].

One of the major results, when one is to lift the delay equation into an infinite dimensional setting exploiting the notion of *segment*, is that one is able to recover the Markov property of the driving equation, see, e.g [42, Theorem II.1]. Similarly also equation (2.4) results to be an M^2-valued Markov process.

Proposition 2.6. *Let* $X = ((X_t, X(t)))_{t\in[0,T]}$ *be the strong solution to equation* (2.4)*, then the process* X *is a Markov process in the sense that*

$$\mathbb{P}((X_t, X(t)) \in B|\mathcal{F}_s) = \mathbb{P}((X_t, X(t)) \in B|(X_s, X(s)) = (\eta, x)) \,, \quad \mathbb{P} - a.s. \,,$$

for all $0 \leq s \leq t \leq T$ *and for all Borel sets* $B \in \mathcal{B}(M^2)$.

Proof. See, e.g. [6, Th. 3.9], or also, see, e.g., [46, Prop. 3.3] or [44, Sec.9.6]. □

Having shown in Proposition 2.6 that X is a M^2-valued Markov process, we can therefore introduce the transition semigroup $P_{\tau,t}$, acting on the space of Borel bounded function on M^2, denoted by $B_b(M^2)$, namely, we define

$$P_{\tau,t} : B_b(M^2) \to B_b(M^2), \quad P_{\tau,t}[\varphi](x) := \mathbb{E}[\varphi(X_t^{\tau,\eta,x})], \qquad \varphi \in B_b(M^2). \tag{2.7}$$

Concerning the infinitesimal generator $\mathcal{L}_t$ of equation (2.4), following [33, 34, 35], we will not enter in further details concerning its explicit representation or the characterization of its domain, since this goes beyond the aim of the present work and it is not necessary in order to prove the main results. Nevertheless let us mentioned that its form can be derived from a direct application of Itô's formula, see, e.g. [6, Th. 3.6].

2.1. Malliavin calculus for jump processes with delay

In this subsection we recall some definitions and main results concerning Malliavin operator and Skorokhod integral for jump processes. We will give fundamental definition in order to fix the notation and to recall the most effective results, we refer to [26, 45] for further references and proofs of some results, or to [23, 25] for application of Mallavin calculus to delay equations.

In particular we stress that very few results concerning Malliavin calculus for jump processes in infinite dimension exist, where also the most simple case of jumps processes having values in an infinite dimensional Hilbert space is difficult to treat, we refer the interested reader to [4]. In order to avoid problems coming with the Hilbert space setting we will, in the present section, exploit the same ideas used in [25]. Using the fact that the original SDE has finite dimensional realizations. This will allow us to exploit standard results in Malliavin calculus for jumps processes with values in $\mathbb{R}^d$. Also, in order to be able to do so, as in [25], we must work with delay of integral type, which motivates the choice of the Hilbert space M^2.

In order to keep the present paper as much as self contained as possible, we will first recall definitions and fundamental results for Malliavin calculus for jumps processes mainly taken from [26]. Eventually we state the main result of the present subsection, that is, as done in [25] exploiting the finite dimensional nature of the SFDDE, we prove a Malliavin differentiability result for SFDDE. Also, for the sake of brevity, we will state the results just for the jump component and we refer to [35, 42] for the diffusive part.

Let us denote by $I_n(f)$ the n-fold *iterated stochastic integral* w.r.t. the random measure $\tilde{N}$, as

$$I_n(f_n) := \int_{([0,T]\times\mathbb{R}_0)^n} f((t_1,z_1),\dots,(t_n,z_n))\tilde{N}(\mathrm{d}t_1,\mathrm{d}z_1)\dots\tilde{N}(\mathrm{d}t_n,\mathrm{d}z_n) \in L^2(\Omega), \tag{2.8}$$

where

$$f \in L^2\left(([0,T]\times\mathbb{R}_0)^n\right) = L^2\left(([0,T]\times\mathbb{R}_0)^n), \otimes\nu(dz)dt\right) ,$$

is a deterministic function.

Thus, every random variable $F \in L^2(\Omega)$ can be represented as an infinite sum of iterated integrals of the form (2.8). This representation is known as *chaos expansion*, see, e.g.[26, Def. 12.1] or [45, Th. 1].

Theorem 2.7. *The stochastic Sobolev space $\mathbb{D}^{1,2}$ consists of $\mathcal{F}-$measurable random variable $F \in L^2(\Omega)$ such that, for $(f_n)_{n\geq 0}$, with $f_n \in L^2\left(([0,T]\times\mathbb{R}_0)^n\right)$, it holds*

$$F = \sum_{n=0}^{\infty} I_n(f_n)\,, \tag{2.9}$$

with the following norm

$$\|F\|^2 = \sum_{n=0}^{\infty} nn!\|I_n(f_n)\|^2_{L^2(([0,T]\times\mathbb{R}_0)^n)}\,.$$

Given the *chaos expansion* in equation (2.9), we can introduce the *Malliavin derivative* $D_{t,z}$ and its domain $\mathbb{D}^{1,2}$, see, e.g. [26, Def. 12.2].

Definition 2.8. Let us consider a random variable $F \in \mathbb{D}^{1,2}$, the *Malliavin derivative* is the operator $D : \mathbb{D}^{1,2} \subset L^2(\Omega) \to L^2(\Omega\times[0,T]\times\mathbb{R}_0)$ defined as

$$D_{t,z}F = \sum_{n=1}^{\infty} nI_{n-1}(f_n(\cdot,t,z)),\ F \in \mathbb{D}^{1,2},\ z \neq 0\,. \tag{2.10}$$

Since the operator D is closable, see, e.g., [26, Thm. 3.3 and Thm 12.6], we denote by $\mathbb{D}^{1,2}$ the domain of its closure.

The following result represents a *chain rule* for Malliavin derivative.

Theorem 2.9. *Let $F \in \mathbb{D}^{1,2}$ and let ϕ be a real continuous function on $\mathbb{R}$. Suppose $\phi(F) \in L^2(\Omega)$ and $\phi(F + D_{t,z}F) \in L^2(\Omega\times[0,T]\times\mathbb{R}_0)$. Then, $\phi \in \mathbb{D}^{1,2}$ and*

$$D_{t,z}\phi(F) = \phi(F + D_{t,z}F) - \phi(F)\,. \tag{2.11}$$

Proof. See, e.g. [26, Thm. 12.8]. □

Once the Malliavin derivative has been defined, we are able to introduce its adjoint operator, the *Skorokhod integral*, in particular next definition is taken from [26, Def. 11.1], see, also [45, Sec. 3] for details.

Definition 2.10. Let $\delta : L^2(\Omega\times[0,T]\times\mathbb{R}_0) \to L^2(\Omega)$ be the adjoint operator of the derivative D. The set of processes $h \in L^2(\Omega\times[0,T]\times\mathbb{R}_0)$ such that

$$\left|\mathbb{E}\int_0^T\int_{\mathbb{R}_0} D_{s,z}F\, h_t(z)\,\nu(dz)ds\right| \leq C\|F\|\,,$$

for all $F \in \mathbb{D}^{1,2}$, forms the domain of δ, denoted by $dom\,\delta$.

For every $h \in dom\,\delta$ we can define the *Skorokhod integral* as

$$\delta(h) := \int_0^T \int_{\mathbb{R}_0} h_t(z)\tilde{N}(\hat{dt}, dz)\,,$$

for any $F \in \mathbb{D}^{1,2}$.

Definition 2.11. We denote by $\mathbb{L}^{1,2}$ the space of $\mathcal{F}-$adapted processes $h : \Omega \times [0,T] \times \mathbb{R}_0 \to \mathbb{R}$ such that $h_t \in \mathbb{D}^{1,2}$ and

$$\mathbb{E}\int_0^T \int_{\mathbb{R}_0} |h_t(z)|\nu(dz)dt < \infty$$

$$\mathbb{E}\int_{([0,T]\times\mathbb{R}_0)^2} |D_{t,z}h_s(\zeta)|\nu(d\zeta)ds\nu(dz)dt < \infty\,.$$

From Definitions 2.10–2.11 above, we have that $\mathbb{L}^{1,2} \subset dom\,\delta$. If $h \in \mathbb{L}^{1,2}$ and $D_{t,z}h \in dom\,\delta$, then $\delta(h) \in \mathbb{D}^{1,2}$ and

$$D_{t,z}\delta(h) = h(z) + \delta(D_{t,z}h)\,, \tag{2.12}$$

see, e.g. [37]. Notice also that $\mathbb{L}^{1,2} \simeq L^2([0,T];\mathbb{D}^{1,2})$.

Proposition 2.12. *Let h_t be a predictable square integrable process. Then, if $h \in \mathbb{D}^{1,2}$, we have, for a.e. $(s,z) \in [0,t] \times \mathbb{R}_0$,*

$$D_{s,z}\int_0^t h_\tau d\tau = \int_s^t D_{\tau,z}h_\tau d\tau\,,$$

$$D_{s,z}\int_0^t h_\tau dW(\tau) = \int_s^t D_{\tau,z}h_\tau dW(\tau)\,,$$

$$D_{s,z}\int_0^t \int_{\mathbb{R}_0} h_\tau \tilde{N}(d\tau, dz) = h_s + \int_s^t \int_{\mathbb{R}_0} D_{\tau,\zeta}h_\tau \tilde{N}(d\tau, d\zeta)\,.$$

Proof. ee, e.g. [45, Prop. 6]. □

Next result is the chain rule for SFDDE, that is the generalization of Theorem 2.9 to the case of delay equations, that will be needed in the proof of the main result of the present Section as well as in subsequent sections.

Theorem 2.13. *Let F and $\psi \in \mathbb{D}^{1,2}$, let also ϕ be a real valued continuous function on M^2 Suppose $\phi(\psi, F) \in L^2(\Omega)$ and $\phi(\psi + D_{t,z}\psi, F + D_{t,z}F) \in L^2(\Omega \times [-r,T] \times \mathbb{R}_0)$. Then, $\phi \in \mathbb{D}^{1,2}$ and it holds*

$$D_{t,z}\phi(\psi, F) = \phi(\psi + D_{t,z}\psi, F + D_{t,z}F) - \phi(\psi, F)\,. \tag{2.13}$$

Proof. Following [43, Proposition 6.2], let us define a partition of $[-r, 0]$,

$$\Pi_k : -r \le s_1 < \ldots < s_k \le 0\,,$$

with

$$\|\Pi_k\| := \max_{2\le i\le k}(s_i - s_{i-1}) \to 0, \quad \text{as } k \to \infty\,.$$

Let $I_k : \mathbb{R}^k \to L^2([-r,0],\mathbb{R})$ be the continuous linear embedding associated to the partition Π_k as

$$I_k(x_1,\dots,x_k)(t) := \sum_{i=1}^{k} x_i I_{(s_{i-1},s_i]}(t)\ ,$$

and set $\underline{s}^k$ the tuple $(s_1,\dots,s_k)$. Let us also define

$$Q_{\underline{s}^k}(\psi) := \left(\frac{1}{s_1 - s_0}\int_{s_0}^{s_1} \psi(t)\mathrm{d}t, \dots, \frac{1}{s_k - s_{k-1}}\int_{s_{k-1}}^{s_k} \psi(t)\mathrm{d}t \right)\ ,$$

the L^2 projection for $\psi \in L^2([-r,0],\mathbb{R})$. Finally, we define a linear map $T^k : L^2([-r,0],\mathbb{R}) \to L^2([-r,0],\mathbb{R})$ as

$$T^k \,:\, \psi \mapsto \psi^k := T^k\psi := I_k \circ Q_{\underline{s}^k}(\psi)\,;$$

in particular it holds that $T^k\psi \to \psi$ in $L^2([-r,0],\mathbb{R})$ as $k \to \infty$, see, e.g. [43, Lemma 5.1].

We thus define the function $\phi^k : \mathbb{R}^k \times \mathbb{R} \to \mathbb{R}$, so that, from the classical chain rule Theorem 2.9 applied to ϕ^k we have

$$\begin{aligned} D_{t,z}\phi(\psi^k, F) &= D_{t,z}\phi^k\left(Q_{\underline{s}^k}(\psi), F\right) = \\ &= \phi^k(Q_{\underline{s}^k}\left(\psi + D_{t,z}\psi\right), F + D_{t,z}F) - \phi^k(Q_{\underline{s}^k}\left(\psi\right), F) = \\ &= \phi(I_k \circ Q_{\underline{s}^k}\left(\psi + D_{t,z}\psi\right), F + D_{t,z}F) - \phi(I_k \circ Q_{\underline{s}^k}\left(\psi\right), F)\,. \end{aligned}$$

Then the claim follows taking the limit as $k \to \infty$ together with [43, Lemma 5.1], the continuity of ϕ and the Dominated Convergence Theorem. □

We are finally able to prove next theorem, which is the main result of the current subsection concerning Malliavin differentiability of the SFDDE (2.4).

Theorem 2.14. *Let us suppose that Assumptions* 2.2 *(A1)-(A2) hold and* $X = (X(t))_{t\in[-r,T]}$ *is the solution to equation* (2.4)*. Then,* $X \in L^2\left([-r,T];\mathbb{D}^{1,2}\right)$ *and, for every* $s \in [0,T]$ *and* $z \in \mathbb{R}_0$*, the stochastic process* $\{D_{s,z}X(t) : t \in [s,T]\}$ *satisfies*

$$\mathbb{E}\left[\int_0^T \int_{\mathbb{R}_0} \sup_{t\in[s,T]} |D_{s,z}X(t)|^2 \nu(dz)ds\right] < \infty\ . \tag{2.14}$$

In particular, for any $t \in [0,T]$*,* $X(t) \in \mathbb{D}^{1,2}$ *and it holds*

$$\begin{cases} D_{s,z}X(t) & = \gamma(s, X_s, X(s), z) + \\ & + \int_s^t \left(\mu\left(X_u + D_{s,z}X_u, X(u) + D_{s,z}X(u)\right) - \mu(X_u, X(u))\right) du + \\ & + \int_s^t \left(\sigma\left(X_u + D_{s,z}X_u, X(u) + D_{s,z}X(u)\right) - \sigma(X_u, X(u))\right) dW(u) + \\ & + \int_s^t \int_{\mathbb{R}_0} \left(\gamma\left(X_u + D_{s,z}X_u, X(u) + D_{s,z}X(u)\right) - \gamma(X_u, X(u))\right) \tilde{N}(du, d\zeta)\,, \\ D_{s,z}X(t) & = 0\,, \quad t \in [-r,s)\,, \end{cases} \tag{2.15}$$

Moreover, for any $z \in \mathbb{R}_0$, there exists a measurable version of the two-parameter process

$$D_{s,z}X_t = \{D_{s,z}X_t(\theta) \,:\, s \in [0,T]\,, \theta \in [-r,0]\}\ .$$

Proof. We will use a standard Picard's approximation scheme, see, e.g. [26, Th. 17.2]. Let $X^0(t) = x$ and $X_t^0 = \eta$, then set, for $n > 0$,

$$\begin{aligned} X^{n+1}(t) &= x + \int_0^t \mu(s, X_s^n, X^n(s))ds + \int_0^t \sigma(s, X_s^n, X^n(s))dW(s) \\ &\quad + \int_0^t \int_{\mathbb{R}_0} \gamma(s, X_s^n, X^n(s), z)\tilde{N}(ds, dz)\,, \\ X_0^{n+1} &= \eta\,, \end{aligned}$$

where we use the notation $X_s^n := (X^n(s+\theta))_{\theta\in[-r,0]}$.

We are going to prove by induction over n that $X^n(t) \in \mathbb{D}^{1,2}$ for any $t \in [0,T]$, $D_{s,z}X(t)$ is a predictable process and that

$$\xi_{n+1}(t) \le C_1 + C_2 \int_{-r}^t \xi_n(s)ds\,,$$

where C_1, C_2 are some suitable constants and

$$\xi_n(s) := \sup_{0\le s\le t} \mathbb{E} \int_{\mathbb{R}_0} \sup_{s\le\tau\le t} |D_{s,z}X^n(\tau)|^2 \nu(dz) < \infty\,.$$

For $n = 0$ the above claim is trivially satisfied. Let us thus assume that the previous assumptions hold for n, we have to show that they hold also for $n+1$. Indeed we have that $\int_0^t \mu(s, X_s^n, X^n(s))ds$, $\int_0^t \sigma(s, X_s^n, X^n(s))dW(s)$ and $\int_0^t \gamma(s, X_s^n, X^n(s), z)\tilde{N}(ds, dz) \in \mathbb{D}^{1,2}$, and proposition 2.12 guarantees that

$$\begin{aligned} D_{s,z} \int_0^t \mu(\tau, X_\tau^n, X^n(\tau))d\tau &= \int_s^t D_{\tau,z}\mu(\tau, X_\tau^n, X^n(\tau))d\tau \\ D_{s,z} \int_0^t \sigma(\tau, X_\tau^n, X^n(\tau))dW(\tau) &= \int_s^t D_{\tau,z}\sigma(\tau, X_\tau^n, X^n(\tau))dW(\tau) \end{aligned}$$

and

$$\begin{aligned} D_{s,z} \int_0^t \gamma(\tau, X_\tau^n, X^n(\tau), z)\tilde{N}(d\tau, dz) &= \gamma(s, X_s^n, X^n(s), z) \\ &\quad + \int_s^t \int_{\mathbb{R}_0} D_{\tau,\zeta}\gamma(\tau, X_s^\tau, X^n(\tau), \zeta)\tilde{N}(d\tau, d\zeta) \end{aligned}$$

for $s \leq t$. Consequently, for any $t \in [0,T]$, $X^{n+1}(t) \in \mathbb{D}^{1,2}$ and

$$\begin{aligned} D_{s,z}X^{n+1}(t) = \gamma(s, X_s^n, X^n(s), z) &+ \int_s^t D_{\tau,z}\mu(\tau, X_\tau^n, X^n(\tau))d\tau \\ &+ \int_s^t D_{\tau,z}\sigma(\tau, X_\tau^n, X^n(\tau))dW(\tau) \\ &+ \int_s^t \int_{\mathbb{R}_0} D_{\tau,z}\gamma(\tau, X_\tau^\tau, X^n(\tau), \zeta)\tilde{N}(d\tau, d\zeta)\,, \end{aligned} \tag{2.16}$$

and the representation in equation (2.15) immediately follows from the chain rule Th. 2.13.

By squaring both sides of equation (2.16), we have

$$\begin{aligned} \left|D_{s,z}X^{n+1}(t)\right|^2 \leq 4\left|\gamma(s, X_s^n, X^n(s), z)\right|^2 &+ \left|\int_s^t \mu_{\tau,z}(\tau, X_\tau^n, X^n(\tau))d\tau\right|^2 \\ &+ 4\left|\int_s^t \sigma_{\tau,z}(\tau, X_\tau^n, X^n(\tau))dW(\tau)\right|^2 \\ &+ 4\left|\int_s^t \int_{\mathbb{R}_0} \gamma_{\tau,z}(\tau, X_s^\tau, X^n(\tau), \zeta)\tilde{N}(d\tau, d\zeta)\right|^2 . \end{aligned} \tag{2.17}$$

By exploiting Doob maximal inequality, stochastic Fubini's theorem and Itô isometry, we get

$$\begin{aligned} \mathbb{E}\int_{\mathbb{R}_0} \sup_{s\leq\tau\leq t} \left|D_{s,z}X^{n+1}(t)\right|^2 \nu(dz) &\leq C\left[\mathbb{E}\int_{\mathbb{R}_0} \left|\gamma(s, X_s^n, X^n(s), z)\right|^2 \nu(dz)\right. \\ &+ \mathbb{E}\left|\int_s^t \mu_{\tau,z}(\tau, X_\tau^n, X^n(\tau))d\tau\right|^2 + \mathbb{E}\left|\int_s^t \sigma_{\tau,z}(\tau, X_\tau^n, X^n(\tau))dW(\tau)\right|^2 \\ &+ \left.\mathbb{E}\left|\int_s^t \int_{\mathbb{R}_0} \gamma_{\tau,z}(\tau, X_s^\tau, X^n(\tau), \zeta)\tilde{N}(d\tau, d\zeta)\right|^2\right] \\ &\leq C\left[\mathbb{E}\int_{\mathbb{R}_0} \left|\gamma(s, X_s^n, X^n(s), z)\right|^2 \nu(dz)\right. \\ &+ \mathbb{E}\int_s^t \left|\mu_{\tau,z}(\tau, X_\tau^n, X^n(\tau))\right|^2 d\tau + \mathbb{E}\int_s^t \left|\sigma_{\tau,z}(\tau, X_\tau^n, X^n(\tau))\right|^2 d\tau \\ &+ \left.\mathbb{E}\int_s^t \int_{\mathbb{R}_0} \left|\gamma_{\tau,z}(\tau, X_s^\tau, X^n(\tau), \zeta)\right|^2 \nu(dz)d\tau\right]\,, \end{aligned} \tag{2.18}$$

where we denote for short by $C > 0$ a suitable constant.

Exploiting Assumptions 2.2 together with Theorem 2.9, we get

$$
\begin{aligned}
&\mathbb{E}\int_{\mathbb{R}_0}\sup_{s\le\tau\le t}|D_{s,z}X^{n+1}(\tau)|^2\nu(dz)\le\\
&\le C_1\int_s^t\mathbb{E}\int_{\mathbb{R}_0}|\left(D_{s,z}X^n_\tau,D_{s,z}X^n(\tau)\right)|_2^2\nu(dz)d\tau+C_2\left(1+\mathbb{E}|\left(X^n_\tau,X^n(\tau)\right)|_2^2\right)\\
&\le C_1\left(\mathbb{E}\int_s^t\int_{-r}^0\int_{\mathbb{R}_0}|D_{s,z}X^n(\tau+\theta)|^2\,\nu(dz)d\theta d\tau+\mathbb{E}\int_s^t\int_{\mathbb{R}_0}|D_{s,z}X^n(\tau)|^2\,d\nu(dz)\tau\right)+\\
&+C_3(1+\lambda)\\
&\le C_1\left(\mathbb{E}\int_{-r}^0\int_s^{t+\theta}\int_{\mathbb{R}_0}|D_{s,z}X^n(p)|^2\,\nu(dz)dpd\theta+\mathbb{E}\int_s^t\int_{\mathbb{R}_0}|D_{s,z}X^n(\tau)|^2\,\nu(dz)d\tau\right)+\\
&+C_3(1+\lambda)\\
&\le C_4\mathbb{E}\int_s^t\int_{\mathbb{R}_0}|D_{s,z}X^n(\tau)|^2\,\nu(dz)d\tau+C_3(1+\lambda)\,,
\end{aligned}
\tag{2.19}
$$

where C_1, C_2, C_3 and C_4 denote some suitable constants and λ is such that

$$\lambda=\sup_n\mathbb{E}\sup_{-r\le s\le T}|X^n(s)|_2^2<\infty\,.$$

Also, we obtain

$$X^{n+1}=\left(X^{n+1}(t)\right)_{t\in[-r,T]}\in L^2(\Omega\times[-r,T])\,,$$

and for any t, $X^{n+1}(t)\in\mathbb{D}^{1,2}$, so that $X^{n+1}\in L^2(\Omega\times[-r,T];\mathbb{D}^{1,2})$ and, for $p\le s$, $D_{s,z}X^{n+1}(p)=0$.

It follows that, for any $z\in\mathbb{R}_0$, it exists a measurable version of the two-parameter process

$$D_{s,z}X^{n+1}_t=\left\{D_{s,z}X^{n+1}_t(\theta)\,:\,s\in[0,T]\,,\theta\in[-r,0]\right\}\,,$$

such that $D_{s,z}X^{n+1}_t\in L^2(\Omega\times[0,T]\times[-r,0])$, see, e.g. [42, Sec. 4].

Therefore, the inductive hypothesis is fulfilled by X^{n+1} and

$$\mathbb{E}\sup_{s\le T}|X^n(s)-X(s)|^2\to0\quad\text{ as }\ n\to\infty\,.$$

Finally, thanks to a discrete version of Gronwall's lemma, see, e.g. [8, Lemma 4.1] or [26, Th. 17.2], and applying equation (2.19), we have

$$\sup_{n\ge0}\mathbb{E}\int_{-r}^T|D_{s,z}X^n(\tau)|^2d\tau<\infty\,,$$

so that $X(t)\in\mathbb{D}^{1,2}$.

By repeating the same reasoning as before, we have

$$X=(X(t))_{t\in[-r,T]}\in L^2(\Omega\times[-r,T]),\ X(t)\in\mathbb{D}^{1,2}\,,$$

for any t, so that $X\in L^2(\Omega\times[-r,T];\mathbb{D}^{1,2})$. The proof is complete by observing that, for any $z\in\mathbb{R}_0$, there exists a measurable version of the two-parameter process

$$D_{s,z}X_t=\{D_{s,z}X_t(\theta)\,:\,s\in[0,T]\,,\theta\in[-r,0]\}\,,$$

such that $D_{s,z}X_t\in L^2(\Omega\times[0,T]\times[-r,0])$. □

3. Joint quadratic variation

In order to prove the main result of this work, which consists in giving an explicit *Feynman-Kac* representation formula for a coupled forward-backward system with delay, we need first to prove a joint quadratic variation result. The main advantage of such an approach is to overcome difficulties that may arise in dealing with the Itô formula in infinite dimension, since, in general, the process X_t fails to be a semi-martingale, so we cannot rely on standard Itô calculus. Furthermore, with the present approach, we are able to relax hypothesis concerning the differentiability of the coefficients.

Following [33, 35], we introduce a *generalized covariation* process. The definition of joint generalized quadratic variation we consider in the present paper has been first introduced in [48], see also [47, 49], with the only difference that they consider the limit to hold uniformly on compacts sets in probability. We have chosen here, following [33, 35], to consider the limit in probability because the limiting procedure is easier with a stronger notion of convergence, such as the convergence in probability. Also, it is shown in [48, Prop. 1.1] that the standard definition of joint quadratic variation, see, e.g. [3, Section 4.4.3], coincides with the quadratic variation defined below.

Definition 3.1. Given a couple of $\mathbb{R}$-valued stochastic processes $(X(t), Y(t))$, $t \geq 0$, we define their *joint quadratic variation* on $[0, T]$, to be

$$\langle X(t), Y(t) \rangle_{[0,T']} := \mathbb{P} - \lim_{\epsilon \downarrow 0} C^{\epsilon}_{[0,T']}(X(t), Y(t)) \, ,$$

where $\mathbb{P} - \lim$ denotes the limit to be taken in probability and

$$C^{\epsilon}_{[0,T']}(X(t), Y(t)) := \frac{1}{\epsilon} \int_0^{T'} (X(t+\epsilon) - X(t))(Y(t+\epsilon) - Y(t)) \mathrm{d}t, \ \epsilon > 0 \, , \quad (3.1)$$

with $0 \leq T' + \epsilon < T$.

Before stating our main result we are to better introduce a mild notion of derivative we will use throughout the paper. In what follows we will consider a function $u : [0, T] \times M^2 \to \mathbb{R}$, such that there exist $C > 0$ and $m \geq 0$, such that, for any $t \in [0, T]$ and any (η_1, x_1), $(\eta_2, x_2) \in M^2$, u satisfies

$$\begin{aligned} &|u(t, \eta_1, x_1) - u(t, \eta_2, x_2)| \leq C|(\eta_1, x_1) - (\eta_2, x_2)|_2 (1 + |(\eta_1, x_1)|_2 + |(\eta_2, x_2)|_2)^m \, , \\ &|u(t, 0, 0)| \leq C \, . \end{aligned} \quad (3.2)$$

that is we require the function u to be Lipschitz continuous without requiring any further regularity concerning differentiability. Nevertheless, in what follows, we will use the notation of ∂^{σ}_x. In particular following [35] we will introduce a mild notion of derivative, called *generalized directional gradient* $\partial^{\sigma}_x u$. When u is sufficiently regular, it can be shown that the *generalized directional gradient*, in the direction $\sigma(t, \eta, x)$, of a function u, coincides with $\partial_x u(t, \eta, x)\sigma(t, \eta, x)$.

The definition, as well as the characterization of several properties, for the *generalized directional gradient* has been provided in [35]. We will only

state here the definition of *generalized directional gradient*, whereas we refer to [35] to a complete treatment of the topic.

In particular it has been shown in [35] that the following holds

$$\langle u(\cdot, X_\cdot, X(\cdot), W(\cdot)\rangle_{\tau,t} = \int_\tau^t \zeta(s, X_s, X(s))ds\,, \tag{3.3}$$

where $\langle\,\cdot\,,\,\cdot\,\rangle_{\tau,t}$ denotes the *joint quadratic variation* defined above and $\zeta : [0,T]\times M^2 \to \mathbb{R}$ is a suitable measurable map, see also [33, 34, 35] for details. Under suitable hypothesis of regularity, in [35] the authors show that

$$\langle u(\cdot, X_\cdot, X(\cdot), W(\cdot)\rangle_{\tau,t} = \int_\tau^t \partial_x u(t, X_s, X(s))\sigma(t, X_s, X(s))ds\,, \quad \mathbb{P}-a.s.\,, \tag{3.4}$$

where we denote by ∂_x the derivative w.r.t. the present state. Hence, equation (3.3) can be considered as the definition of the *generalized directional gradient* of the function u along the direction σ. We say that the map $\zeta : [0,T]\times M^2 \to \mathbb{R}$ belongs to the directional gradient of u, or equivalently that $\zeta \in \partial_x^\sigma u$, if equation (3.3) holds. Therefore, we use for short the notation $\partial_x^\sigma u$ to represent an element of the *generalized directional gradient*. Since this topic lies outside our goals, having been deeply studied in a more general setting in [35], we skip every technicality and invite the interested reader to [35].

The following result represents the core of this paper.

Theorem 3.2. *Let us assume that $u : [0,T]\times M^2 \to \mathbb{R}$ is locally Lipschitz w.r.t. the second variable and with at most polynomial growth, namely, there exist $C > 0$ and $m \geq 0$, such that, for any $t \in [0,T]$ and any (η_1, x_1), $(\eta_2, x_2) \in M^2$, u satisfies*

$$\begin{aligned} &|u(t,\eta_1,x_1) - u(t,\eta_2,x_2)| \leq C|(\eta_1,x_1)-(\eta_2,x_2)|_2(1+|(\eta_1,x_1)|_2+|(\eta_2,x_2)|_2)^m\,,\\ &|u(t,0,0)| \leq C\,. \end{aligned} \tag{3.5}$$

Then, for every $(\eta, x) \in M^2$ and $0 \leq \tau \leq T' \leq T$, the process

$$\{u(t, X_t^{(\tau,\eta,x)}, X^{\tau,\eta,x}(t)),\, t \in [\tau, T']\}$$

admits a joint quadratic variation on the interval $[\tau, T']$ with

$$J(t) := \int_0^t \int_{\mathbb{R}_0} z\tilde{N}(ds, dz),$$

given by

$$\begin{aligned} \langle u(\cdot, X_\cdot^{\tau,\eta,x}, X^{\tau,\eta,x}(\cdot)), J(\cdot)\rangle_{[\tau,T']} &= \int_\tau^{T'} \int_{\mathbb{R}_0} z\,[u(s, X_s^{\tau,\eta,x}, X^{\tau,\eta,x}(s)\\ &+\gamma(s, X_s^{\tau,\eta,x}, X^{\tau,\eta,x}(s), z) - u(s, X_s^{\tau,\eta,x}, X^{\tau,\eta,x}(s))]\, N(\mathrm{d}s, \mathrm{d}z)\,. \end{aligned} \tag{3.6}$$

Remark 3.3. An analogous of [35, Prop. 4.4] is valid in the present case, that is the following representation holds

$$\langle u(\cdot, X_\cdot^{\tau,\eta,x}, X^{\tau,\eta,x}(\cdot)), W(\cdot)\rangle_{[\tau,T']} = \int_\tau^{T'} \partial_x^\sigma u(s, X_s^{\tau,\eta,x}, X^{\tau,\eta,x}(s))ds\,,$$

where $\partial_x^\sigma u$ is the *generalized directional gradient*. The claim follows from [35] by observing that the Poisson random measure does not affect the result and the proof follows exactly the same steps as in [35].

Proof. Without loss of generality, we prove the result for $\tau = 0$, as the case of a general initial time $\tau \neq 0$ can be proved using the same techniques. Fix $(\eta, x) \in M^2$ and a time horizon $T' \in [0, T]$ and denote for brevity $X^{0,\eta,x}$ by X. In what follows we will denote with $\tilde{N}(\hat{dt}, dz)$ the Skorokhod integral.

In order to shorten the notation set

$$v_t := (u(t+\epsilon, X_{t+\epsilon}, X(t+\epsilon)) - u(t, X_t, X(t))\mathbb{1}_{[0,T']}(t)\,,$$

and

$$A^\epsilon := \{(t,s) \in [0,T'] \times [0,T'] \,:\, 0 \leq t \leq T'\,,\ t \leq s \leq t+\epsilon\}\,.$$

From equation (3.5) and theorem 2.14, we have $v_t \in \mathbb{L}^{1,2}$, so that, for any t, $v_t \in \mathbb{D}^{1,2}$ and then $v_t \mathbb{1}_{A^\epsilon}(t,\cdot) \in L^2(\Omega \times [0,T])$. Furthermore, equation (2.12) implies that v_t is Skorokhod integrable and from [26, Th. 12.11] we have

$$\begin{aligned}\int_0^{T'} \int_{\mathbb{R}_0} z v_t \mathbb{1}_{A^\epsilon}(t,s)\tilde{N}(\hat{dt}, dz) &= v_t \int_0^{T'} \int_{\mathbb{R}_0} z \mathbb{1}_{A^\epsilon}(t,s)\tilde{N}(\hat{dt}, dz) \\ &\quad - \int_0^{T'} \int_{\mathbb{R}_0} z D_{s,z} v_t \mathbb{1}_{A^\epsilon}(t,s) N(ds, dz) =: z_t\,,\end{aligned} \tag{3.7}$$

which holds since $z \in L^2(\Omega \times [0,T])$. Also, equation (3.7) implies, for a.a. $t \in [0,T']$,

$$\begin{aligned}&u(t+\epsilon, X_{t+\epsilon}, X(t+\epsilon)) - u(t, X_t, X(t))(J_{t+\epsilon} - J_t) \\ &= u(t+\epsilon, X_{t+\epsilon}, X(t+\epsilon)) - u(t, X_t, X(t)) \int_t^{t+\epsilon} \int_{\mathbb{R}_0} z\tilde{N}(ds, dz) \\ &= \int_t^{t+\epsilon} \int_{\mathbb{R}_0} z D_{s,z} \left(u(t+\epsilon, X_{t+\epsilon}, X(t+\epsilon)) - u(t, X_t, X(t))\right) N(ds, dz) \\ &+ \int_t^{t+\epsilon} \int_{\mathbb{R}_0} z \left(u(t+\epsilon, X_{t+\epsilon}, X(t+\epsilon)) - u(t, X_t, X(t))\right) \tilde{N}(ds, dz)\,.\end{aligned} \tag{3.8}$$

Let us integrate the right-hand side of equation (3.8) in $[0,T']$ w.r.t. t. By noticing that the left-hand side equals to ϵC^ϵ, we write the right-hand

side as follows

$$\begin{aligned}
&\int_0^{T'}\int_t^{t+\epsilon}\int_{\mathbb{R}_0} zD_{s,z}\left(u(t+\epsilon,X_{t+\epsilon},X(t+\epsilon))-u(t,X_t,X(t))\right)N(ds,dz)dt\\
&+\int_0^{T'}\int_t^{t+\epsilon}\int_{\mathbb{R}_0} z\left(u(t+\epsilon,X_{t+\epsilon},X(t+\epsilon))-u(t,X_t,X(t))\right)\tilde{N}(\hat{ds},dz)dt\\
&=\int_0^{T'}\int_t^{t+\epsilon}\int_{\mathbb{R}_0} zD_{s,z}\left(u(t+\epsilon,X_{t+\epsilon},X(t+\epsilon))-u(t,X_t,X(t))\right)N(ds,dz)dt\\
&+\int_0^{T'+\epsilon}\int_{\mathbb{R}_0}\int_{(s-\epsilon)^+}^{s\wedge T'} z\left(u(t+\epsilon,X_{t+\epsilon},X(t+\epsilon))-u(t,X_t,X(t))\right)dt\tilde{N}(\hat{ds},dz)\,.
\end{aligned}\tag{3.9}$$

It remains to verify that $\int_0^{T'} zv_t\mathbb{1}_{A^\epsilon}(t,\cdot)dt$ appearing in equation (3.9) is Skorokhod integrable. From the definition of Skorokhod integral, by using equation (3.7) for $G\in\mathbb{D}^{1,2}$ and the *duality formula*, see e.g. [26, equation (12.14)], we have

$$\begin{aligned}
&\mathbb{E}\left[\int_0^T\int_{\mathbb{R}_0}\int_0^T zv_t\mathbb{1}_{A^\epsilon}(t,s)dtD_{s,z}G\nu(dz)ds\right]\\
&=\int_0^T\mathbb{E}\left[\int_0^T\int_{\mathbb{R}_0} zv_t\mathbb{1}_{A^\epsilon}(t,s)D_{s,z}G\nu(dz)ds\right]dt\\
&=\int_0^T\mathbb{E}\left[G\int_0^T\int_{\mathbb{R}_0} zv_t\mathbb{1}_{A^\epsilon}(t,s)D_{s,z}\tilde{N}(\hat{ds},dz)\right]dt=\mathbb{E}\left[G\int_0^T z_tdt\right],
\end{aligned}$$

so that $\int_0^{T'} v_t\mathbb{1}_{A^\epsilon}(t,\cdot)dt$ is Skorokhod integrable. Hence,

$$\begin{aligned}
&\int_0^T\int_0^T\int_{\mathbb{R}_0} zv_t\mathbb{1}_{A^\epsilon}(t,s)dt\tilde{N}(\hat{ds},dz)\\
&\qquad=\int_0^T z_tdt=\int_0^T\int_0^T\int_{\mathbb{R}_0} zv_t\mathbb{1}_{A^\epsilon}(t,s)\tilde{N}(\hat{ds},dz)dt\,.
\end{aligned}$$

Exploiting again equation (3.7) we have

$$\begin{aligned}
\int_0^T\int_0^{T'}\int_{\mathbb{R}_0} zv_t\mathbb{1}_{A^\epsilon}(t,s)dt\tilde{N}(\hat{ds},dz)&=\int_0^T zv_t(J_{t+\epsilon}-J_t)\mathbb{1}_{[t,T]}(t)dt\\
&-\int_0^{T'}\int_0^T\int_{\mathbb{R}_0} zD_{s,z}v_t\mathbb{1}_{A^\epsilon}(t,s)N(ds,dz)dt\,,
\end{aligned}$$

and then equation (3.9) is proved.

On the other hand, thanks to the chain rule Theorem 2.13 and from theorem 2.14 together with the adeptness property of the Malliavin derivative,

i.e. $D_{s,z}X(t) = 0$ if $s > t$, we have that, for a.a. $s \in [t, t+\epsilon]$,

$$\begin{aligned} D_{s,z}v_t &= D_{s,z}[u(t+\epsilon, X_{t+\epsilon}, X(t+\epsilon)) - u(t, X_t, X(t))] \\ &= D_{s,z}[u(t+\epsilon, X_{t+\epsilon}, X(t+\epsilon))] \\ &= u(t+\epsilon, X_{t+\epsilon} + D_{s,z}X_{t+\epsilon}, X(t+\epsilon) + D_{s,z}X(t+\epsilon)) \\ &\quad - u(t+\epsilon, X_{t+\epsilon}, X(t+\epsilon)) \,. \end{aligned}$$

Now, we apply equation (3.9) to get

$$\begin{aligned} C^\epsilon &= \frac{1}{\epsilon}\int_0^{T'}\int_t^{t+\epsilon}\int_{\mathbb{R}_0} z\Big[u\left(t+\epsilon, X_{t+\epsilon} + D_{s,z}X_{t+\epsilon}, X(t+\epsilon) + D_{s,z}X(t+\epsilon)\right) N(\hat{\mathrm{d}}s, \mathrm{d}z)\mathrm{d}t \\ &- \frac{1}{\epsilon}\int_0^{T'}\int_t^{t+\epsilon} u\left(t+\epsilon, X_{t+\epsilon}, X(t+\epsilon)\right)] N(\hat{\mathrm{d}}s, \mathrm{d}z)\mathrm{d}t \\ &+ \frac{1}{\epsilon}\int_0^{T'+\epsilon}\int_{\mathbb{R}_0}\int_{(s-\epsilon)^+}^{s\wedge T'} z\left(u(t+\epsilon, X_{t+\epsilon}, X(t+\epsilon)) - u(t, X_t, X(t))\right) dt\tilde{N}(\hat{\mathrm{d}}s, \mathrm{d}z) \,. \end{aligned}$$

Let us consider separately the two terms

$$\begin{aligned} I_1^\epsilon &:= \frac{1}{\epsilon}\int_0^{T'}\int_t^{t+\epsilon}\int_{\mathbb{R}_0} z\Big[u\left(t+\epsilon, X_{t+\epsilon} + D_{s,z}X_{t+\epsilon}, X(t+\epsilon) + D_{s,z}X(t+\epsilon)\right) N(\hat{\mathrm{d}}s, \mathrm{d}z)\mathrm{d}t \\ &\quad - \frac{1}{\epsilon}\int_0^{T'}\int_t^{t+\epsilon} u\left(t+\epsilon, X_{t+\epsilon}, X(t+\epsilon)\right)] N(\hat{\mathrm{d}}s, \mathrm{d}z)\mathrm{d}t \,, \\ I_2^\epsilon &:= \frac{1}{\epsilon}\int_0^{T'+\epsilon}\int_{\mathbb{R}_0}\int_{(s-\epsilon)^+}^{s\wedge T'} z\left(u(t+\epsilon, X_{t+\epsilon}, X(t+\epsilon)) - u(t, X_t, X(t))\right) dt\tilde{N}(\hat{\mathrm{d}}s, \mathrm{d}z) \,. \end{aligned}$$

As regards I_2^ϵ, the proof proceed as in [35, Prop. 4.4.], see also [33, Th. 3.1]. We report in what follows its main steps for the sake of completeness. We have to show that

$$\frac{1}{\epsilon}\int_0^{T'} v_t \mathbb{1}_{A^\epsilon}(t,s)\mathrm{d}t \to 0 \,,$$

in $\mathbb{L}^{1,2}$, since this implies $I_2^\epsilon \to 0$ in $L^2(\Omega)$, together with the boundedness of the Skorokhod integral. Thus, for a general $y \in \mathbb{L}^{1,2}$, we have

$$T^\epsilon(y)_s = \frac{1}{\epsilon}\int_0^{T'} (y_{t+\epsilon} - y_t)\mathbb{1}_{A^\epsilon}(t,s)dt = \frac{1}{\epsilon}\int_{(s-\epsilon)\vee t}^{s\wedge T} (y_{t+\epsilon} - y_t)dt \,,$$

so that we have to show that $T^\epsilon(y) \to 0$ in $\mathbb{L}^{1,2}$.

Let us recall the isomorphism

$$L^2\left([0,T]; \mathbb{D}^{1,2}(\mathbb{R})\right) \simeq \mathbb{L}^{1,2} \,.$$

Following [35], we have to prove that $\|T^\epsilon\|_{\mathbb{L}^{1,2}}(\mathbb{R})$ is bounded uniformly w.r.t. ϵ. In fact, we have

$$\begin{aligned}
\|T^\epsilon(y)_s\|^2_{\mathbb{D}^{1,2}(\mathbb{R})} &\leq \frac{1}{\epsilon^2}\int_0^{T'} \mathbb{1}_{A^\epsilon}(t,s)dt \int_0^{T'} |y_{t+\epsilon}-y_t|^2_{\mathbb{D}^{1,2}(\mathbb{R})}\mathbb{1}_{A^\epsilon}(t,s)dt \\
&\leq \int_0^{T'} |y_{t+\epsilon}-y_t|^2_{\mathbb{D}^{1,2}(\mathbb{R})}\mathbb{1}_{A^\epsilon}(t,s)dt\,, \\
\|T^\epsilon(y)_s\|^2_{\mathbb{L}^{1,2}(\mathbb{R})} &= \int_0^{T'} \|T^\epsilon(y)_s\|^2_{\mathbb{D}^{1,2}(\mathbb{R})}ds \\
&\leq \int_0^{T'} |y_{t+\epsilon}-y_t|^2_{\mathbb{D}^{1,2}(\mathbb{R})} \int_0^{T'} \mathbb{1}_{A^\epsilon}(t,s)ds\,dt \\
&\leq \int_0^{T'} |y_{t+\epsilon}-y_t|^2_{\mathbb{D}^{1,2}(\mathbb{R})}dt \leq 2\|y\|^2_{\mathbb{L}^{1,2}(\mathbb{R})}\,,
\end{aligned}$$

and thus the claim follows by [35, Prop. 4.4.], or [33, Th. 3.1].

As regards I_1^ϵ, we have

$$\begin{aligned}
I_1^\epsilon = \frac{1}{\epsilon}\int_0^T \int_t^{t+\epsilon} \int_{\mathbb{R}_0} z\,u(t+\epsilon, X_{t+\epsilon}+D_{s,z}X_{t+\epsilon}, X(t+\epsilon)+D_{s,z}X(t+\epsilon))N(\hat{\mathrm{d}}s,\mathrm{d}z)\mathrm{d}t \\
-\frac{1}{\epsilon}\int_0^T \int_t^{t+\epsilon} \int_{\mathbb{R}_0} z\,u(t+\epsilon, X_{t+\epsilon}, X(t+\epsilon))]N(\hat{\mathrm{d}}s,\mathrm{d}z)\mathrm{d}t := K_1^\epsilon - K_2^\epsilon\,.
\end{aligned}$$

Let us first prove that

$$K_2^\epsilon \to \int_0^{T'}\int_{\mathbb{R}_0} zu(t,X_t,X(t))N(\hat{\mathrm{d}}t,\mathrm{d}z)\,, \quad \mathbb{P}-a.s. \tag{3.10}$$

as $\epsilon \to 0$.

From assumption (3.5) on the function u, the right-continuity of X, and exploiting the Lebesgue differentiation theorem together with the dominated convergence theorem, it follows that

$$\begin{aligned}
&\frac{1}{\epsilon}\int_0^{T'}\int_t^{t+\epsilon}\int_{\mathbb{R}_0} zu(t+\epsilon, X_{t+\epsilon}, X(t+\epsilon))N(\hat{\mathrm{d}}s,\mathrm{d}z)\mathrm{d}t \\
&= \int_0^{T'+\epsilon}\int_{\mathbb{R}_0} z\frac{1}{\epsilon}\int_{s\vee\epsilon}^{(s+\epsilon)\wedge T'} u(t,X_t,X(t))dtN(\hat{\mathrm{d}}s,dz) \\
&\to \int_0^{T'}\int_{\mathbb{R}_0} zu(s,X_s,X(s))N(\hat{\mathrm{d}}s,\mathrm{d}z)\,,
\end{aligned} \tag{3.11}$$

$\mathbb{P}-$a.s., as $\epsilon \to 0$.

Let us now prove that

$$K_1^\epsilon \to \int_0^{T'}\int_{\mathbb{R}_0} zu(t,X_t,X(t)+\gamma(t,X_t,X(t),z))N(\hat{\mathrm{d}}t,\mathrm{d}z)\,. \tag{3.12}$$

Theorem 2.14 assures that

$$\begin{aligned} D_{s,z}X(t+\epsilon) =& \gamma(s, X_s, X(s), z) + \int_s^{t+\epsilon} D_{s,z}[\mu(q, X_q, X(q))]\mathrm{d}q \\ &+ \int_s^{t+\epsilon} D_{s,z}[\sigma(q, X_q, X(q))]\mathrm{d}W(q) \\ &+ \int_s^{t+\epsilon}\int_{\mathbb{R}_0} D_{s,z}[\gamma(q, X_q, X(q), \zeta)]\tilde{N}(\mathrm{d}q, \mathrm{d}\zeta)\,. \end{aligned} \tag{3.13}$$

Proceeding as above, we get

$$\begin{aligned} &\frac{1}{\epsilon}\int_0^T \int_t^{t+\epsilon}\int_{\mathbb{R}_0} zu(t+\epsilon, X_{t+\epsilon} + D_{s,z}X_{t+\epsilon})N(\hat{\mathrm{d}}s, \mathrm{d}z)\mathrm{d}t \\ &= \int_0^{T'+\epsilon}\int_{\mathbb{R}_0} z\frac{1}{\epsilon}\int_{(s-\epsilon)^+}^{s\wedge T'} u\left(t+\epsilon, X_{t+\epsilon} + D_{s,z}X_{t+\epsilon}, X(t+\epsilon) + D_{s,z}X(t+\epsilon)\right) dtN(\hat{\mathrm{d}}s, dz)\,. \end{aligned} \tag{3.14}$$

The continuity of u, together with the right-continuity of X, and the Lebesgue differentiation theorem provide that

$$\begin{aligned} &\int_\epsilon^{T'+\epsilon}\int_{\mathbb{R}_0} z\frac{1}{\epsilon}\int_{s\vee\epsilon}^{(s+\epsilon)\wedge T'} u(t, X_t + D_{s,z}X_t, X(t) + D_{s,z}X(t))dtN(\hat{\mathrm{d}}s, dz) \\ &\qquad \to \int_0^{T'}\int_{\mathbb{R}_0} zu(t, X_t + D_{t,z}X_t, X(t) + D_{t,z}X(t))N(\hat{\mathrm{d}}t, \mathrm{d}z)\,, \end{aligned} \tag{3.15}$$

$\mathbb{P}-$a.s. as $\epsilon \to 0$.

Moreover, theorem 2.14 implies that

$$\begin{aligned} D_{s,z}X(t+\theta) &= \gamma(s, X_s, X(s), z) + \int_s^{t+\theta} D_{s,z}[\mu(q, X_q, X(q))]\mathrm{d}q \\ &\quad + \int_s^{t+\theta} D_{s,z}[\sigma(q, X_q, X(q))]\mathrm{d}W(q) \\ &\quad + \int_s^{t+\theta}\int_{\mathbb{R}_0} D_{s,z}[\gamma(q, X_q, X(q), \zeta)]\tilde{N}(\mathrm{d}q, \mathrm{d}\zeta)\,, \quad \theta \in [-r, 0]\,, \\ D_{s,z}X(t+\theta) &= 0\,, \quad s > t+\theta\,, \end{aligned}$$

and exploiting the adaptedness of the Malliavin derivative, namely

$$\begin{aligned} D_{t,z}X_t(\theta) = D_{t,z}X(t+\theta) = 0\,, \quad \text{for } \theta \in [-r, 0)\,, \\ D_{t,z}X(t) = \gamma(t, X_t, X(t), z)\,, \end{aligned} \tag{3.16}$$

and substituting (3.16) into eq. (3.15), we obtain the claim and (3.12) is proved. Equation (3.6) thus follows and the proof is then complete. □

4. Existence of mild solutions of Kolmogorov equation

The main goal of this section is to prove an existence and uniqueness result of a *mild* solution, in a sense to be specified later, of a non-linear path-dependent partial integro-differential equation. Such a solution is connected to a forward-backward system with delay of the form

$$\begin{cases} \mathrm{d}X^{\tau,\eta,x}(t) & = \mu(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t))\mathrm{d}t + \sigma(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t))\mathrm{d}W(s) \\ & \quad + \int_{\mathbb{R}_0} \gamma(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t), z)\tilde{N}(\mathrm{d}t, \mathrm{d}z) \\ (X_\tau^{\tau,\eta,x}, X^{\tau,\eta,x}(\tau)) & = (\eta, x) \in M^2 \\ \mathrm{d}Y^{\tau,\eta,x}(t) & = \psi\left(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t), Y^{\tau,\eta,x}(t), Z^{\tau,\eta,x}(t), \tilde{U}^{\tau,\eta,x}(t)\right)\mathrm{d}t \\ & \quad + Z^{\tau,\eta,x}(t)\mathrm{d}W(t) + \int_{\mathbb{R}_0} U^{\tau,\eta,x}(t,z)\tilde{N}(\mathrm{d}t, \mathrm{d}z) \\ Y^{\tau,\eta,x}(T) & = \phi(X_T^{\tau,\eta,x}, X^{\tau,\eta,x}(T)) \end{cases} , \tag{4.1}$$

where we have set for short

$$\tilde{U}^{\tau,\eta,x}(t) := \int_{\mathbb{R}_0} U^{\tau,\eta,x}(t,z)\delta(z)\nu(dz)\,.$$

In particular the solution to the forward–backward SFDDE (4.1) is the quadruple (X, Y, Z, U) taking values in $M^2 \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$. We refer to [24] for a detailed introduction to forward-backward system with jumps.

Let us assume the following assumptions to hold:

Assumption 4.1.

(B1) The map $\psi : [0,T] \times M^2 \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is continuous and there exists $K > 0$ and $m \geq 0$ such that

$$\begin{aligned} |\psi(t,\eta_1,x_1,y_1,z_1,u_1) - \psi(t,\eta_2,x_2,y_2,z_2,u_2)| &\leq K|(\eta_1,x_1) - (\eta_2,x_2)|_2 \\ &\quad + K(|y_1 - y_2| + |z_1 - z_2| + |u_1 - u_2|)\,; \\ |\psi(t,\eta_1,x_1,y,z,u) - \psi(t,\eta_2,x_2,y,z,u)| &\leq K(1 + |(\eta_1,x_1)|_2 + |(\eta_2,x_2)|_2 + |y|)^m \\ &\quad \cdot (1 + |z| + |u|)(|(\eta_1,x_1) - (\eta_2,x_2)|_2)\,; \\ |\psi(t,0,0,0,0,0)| &\leq K\,, \end{aligned}$$

for all $(t,\eta_1,x_1,y_1,z_1,u_1)$, $(t,\eta_2,x_2,y_2,z_2,u_2) \in [0,T] \times M^2 \times \mathbb{R}^3$;

(B2) the map $\phi : M^2 \to \mathbb{R}$ is measurable and there exist $K > 0$ and $m \geq 0$ such that

$$|\phi(\eta_1,x_1) - \phi(\eta_2,x_2)| \leq K(1 + |(\eta_1,x_1)|_2 + |(\eta_2,x_2)|_2)^m |(\eta_1,x_1) - (\eta_2,x_2)|_2\,,$$

for all (η_1,x_1), $(\eta_2,x_2) \in M^2$;

(B3) there exists $K > 0$ such that the function $\delta : \mathbb{R}_0 \to \mathbb{R}$ satisfies

$$|\delta(z)| \leq K|(1 \wedge |z|)\,, \quad \delta(z) \geq 0\,, \quad z \in \mathbb{R}_0\,.$$

Remark 4.2. Following [24], we have chosen this particular form for the generator ψ of the backward component in equation (4.1), due to the fact that it results to be convenient in many concrete applications.

Remark 4.3. We want to stress that assumptions 4.1 imply that there exists a suitable constant $C > 0$ such that

$$|\psi(t,\eta,x,y,z,u)| \leq C(1+|(\eta,x)|_2^{m+1}+|y|+|z|+|u|)\,,$$
$$|\phi(\eta,x)| \leq C(1+|(\eta,x)|_2^{m+1})\,.$$

In what follows we will denote by $\mathbb{K}([0,T])$ the space of all triplet (Y,Z,U) of predictable stochastic processes taking value in $\mathbb{R}\times\mathbb{R}\times\mathbb{R}$ and such that

$$\|(Y,Z,U)\|_{\mathbb{K}}^2 := \mathbb{E}\left[\sup_{t\in[0,T]}|Y(t)|^2\right]+\mathbb{E}\left[\int_0^T|Z(t)|^2\mathrm{d}\tau\right]$$
$$+\mathbb{E}\left[\int_0^T\int_{\mathbb{R}_0}|U(t,z)|^2\nu(dz)dt\right]<\infty\,. \tag{4.2}$$

The following Proposition ensures the existence and the uniqueness of the solution to the system (4.1), under suitable properties of the coefficients.

Proposition 4.4. *Let us consider the coupled forward-backward system* (4.1) *which satisfies Assumptions* 2.2 *and Assumptions* 4.1.

Then, the coupled forward-backward system admits a unique solution

$$(X^{\tau,\eta,x},Y^{\tau,\eta,x},Z^{\tau,\eta,x},U^{\tau,\eta,x})\in\mathcal{S}^p\times\mathbb{K}([0,T])\,.$$

Eventually we have that the map

$$(\tau,\eta,x)\mapsto(X^{\tau,\eta,x},Y^{\tau,\eta,x},Z^{\tau,\eta,x},U^{\tau,\eta,x})\,,$$

is continuous.

Proof. The existence and uniqueness of the solution to the forward component follows from theorem 2.4, since Assumptions 2.2 hold true by hypothesis, whereas the existence and uniqueness of the backward component under Assumptions 4.1 follows [7, Cor. 2.3] or [24, Thm. 4.1.3] .

The continuity of the map $(\tau,\eta,x)\mapsto X^{\tau,\eta,x}$ is guaranteed by theorem 2.4, whereas the continuity of $(\tau,\eta,x)\mapsto(Y^{\tau,\eta,x},Z^{\tau,\eta,x},U^{(\tau,\eta,x})$ follows from [7, Prop. 1.1].

□

Theorem 4.5. *Let us consider the coupled forward-backward system* (4.1) *which satisfies Assumptions* 2.2 *and* 4.1. *Let us define the function* $u:[0,T]\times M^2\to\mathbb{R}$,

$$u(t,\eta,x):=Y_t^{t,\eta,x}\,,$$

with $t\in[0,T]$ *and* $(\eta,x)\in M^2$,.

Then, there exist $C>0$ *and* $m\geq 0$, *such that, for any* $t\in[0,T]$ *and any* (η_1,x_1), $(\eta_2,x_2)\in M^2$, *the function* u *satisfies*

$$|u(t,\eta_1,x_1)-u(t,\eta_2,x_2)| \le C|(\eta_1,x_1)-(\eta_2,x_2)|_2(1+|(\eta_1,x_1)|_2+|(\eta_2,x_2)|_2)^m\,,$$
$$|u(t,0,0)| \le C\,. \tag{4.3}$$

Moreover, for every $t \in [0,T]$ *and* $(\eta,x) \in M^2$ *we have* $\mathbb{P}-a.s.$ *and for a.e.* $t \in [\tau,T]$

$$\begin{aligned} Y^{\tau,\eta,x}(t) &= u\left(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t)\right)\,,\\ Z^{\tau,\eta,x}(t) &= \partial_x^\sigma u\left(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t)\right)\,,\\ U^{\tau,\eta,x}(t,z) &= u(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t)+\gamma(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t),z))\\ &\quad - u(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t))\,, \end{aligned} \tag{4.4}$$

where ∂_x^σ *is the generalized directional gradient in the sense of equation* (3.3).

Remark 4.6. Let us recall that, if u is sufficiently regular, then

$$Z^{(\tau,\eta,x)}(t) = \partial_x u(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t))\sigma(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t))\,.$$

Proof. The fact that $u(t,\eta,x) := Y_t^{t,\eta,x}$ satisfies (4.3) immediately follows from the continuity of the map

$$(\tau,\eta,x) \mapsto (X^{\tau,\eta,x}, Y^{\tau,\eta,x}, Z^{\tau,\eta,x}, U^{\tau,\eta,x})\,,$$

proved in proposition 4.4 together with assumptions 2.2.

The representation of Y and Z follow from [33, Cor. 4.3].

As regards the process U, using the standard notion of joint variation we have

$$\langle Y^{\tau,\eta,x}(\cdot), J(\cdot)\rangle_{[\tau,T]} = \int_\tau^T \int_{\mathbb{R}_0} z\, U^{\tau,\eta,x}(s,z) N(ds,dz)\,. \tag{4.5}$$

On the other hand, Theorem 3.2 implies

$$\begin{aligned} &\langle u(\cdot, X_\cdot^{\tau,\eta,x}, X^{\tau,\eta,x}(\cdot)), J(\cdot)\rangle_{[\tau,T]}\\ &= \int_\tau^T \int_{\mathbb{R}_0} z\left[u(s, X_s^{\tau,\eta,x}, X^{\tau,\eta,x}(s)+\gamma(s, X_s^{\tau,\eta,x}, X^{\tau,\eta,x}(s),z))\right] N(ds,dz)\\ &- \int_\tau^T \int_{\mathbb{R}_0} z\left[u(s, X_s^{\tau,\eta,x}, X^{\tau,\eta,x}(s))\right] N(ds,dz)\,. \end{aligned} \tag{4.6}$$

Comparing now equation (4.5) and equation (4.6), the representation for U in equation (4.4) follows. □

4.1. The non-linear Kolmogorov equation

The present section is devoted to prove that the solution to the forward-backward system (4.1) can be connected to the solution of a path-dependent partial integro-differential equation with values in the Hilbert space M^2.

More precisely, let us consider the Markov process $(X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t))$ defined as the solution of equation (2.4), and the corresponding infinitesimal generator $\mathcal{L}_t$.

The path-dependent partial-integro differential equation we want to investigate has the following form

$$\begin{cases} \frac{\partial}{\partial t}u(t,\eta,x) + \mathcal{L}_t u(t,\eta,x) = \psi\left(t,\eta,x,u(t,\eta,x),\partial_x^\sigma u(t,\eta,x),\mathcal{J}u(t,\eta,x)\right), \\ u(T,\eta,x) = \phi(\eta,x), \end{cases} \tag{4.7}$$

for all $t \in [0,T]$, and $(\eta,x) \in M^2$, where $u : [0,T]\times M^2 \to, \mathbb{R}$ is an unknown function, ψ and ϕ are two given functions such that $\psi : [0,T] \times M^2 \times \mathbb{R} \times \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ and $\psi : M^2 \to \mathbb{R}$, $\partial_x^\sigma u$ is the *generalized directional gradient* and $\mathcal{J}$ is a functional acting as

$$\mathcal{J}u(t,\eta,x) = \int_{\mathbb{R}_0} \left(u(t,\eta,x+\gamma(t,\eta,x,z)) - u(t,\eta,x)\right)\delta(z)\nu(dz)\,.$$

In particular, we want to look for a *mild solution* of equation (4.7), according to the following definition.

Definition 4.7. A *mild solution* to equation (4.7) is a function $u : [0,T] \times M^2 \to \mathbb{R}$ such that there exist $C > 0$ and $m \geq 0$, such that, for any $t \in [0,T]$ and any (η_1,x_1), $(\eta_2,x_2) \in M^2$, u satisfies

$$\begin{aligned} &|u(t,\eta_1,x_1) - u(t,\eta_2,x_2)| \leq C|(\eta_1,x_1)-(\eta_2,x_2)|_2(1+|(\eta_1,x_1)|_2+|(\eta_2,x_2)|_2)^m \\ &|u(t,0,0)| \leq C \end{aligned} \tag{4.8}$$

and the following identity hold true

$$u(t,\eta,x) = P_{t,T}\phi(\eta,x) + \int_t^T P_{t,s}[\psi(\cdot,u(s,\cdot),\partial_x^\sigma u(s,\cdot),\mathcal{J}u(s,\cdot)](\eta,x)\mathrm{d}s\,, \tag{4.9}$$

for all $t \in [0,T]$, and $(\eta,x) \in M^2$ and where $P_{t,s}$ is the Markov semigroup for equation (1.2) introduced in equation (2.7).

Theorem 4.8. *Assume that Assumptions* 2.2 *and Assumptions* 4.1 *hold true. Then, the path-dependent partial integro-differential equation* (4.7) *admits a unique mild solution* u*, in the sense of definition* 4.7*. In particular, the mild solution* u *coincide with the function* u *introduced in theorem* 4.5*.*

Proof. In what follows, as above, we will denote for short

$$\tilde{U}^{\tau,\eta,x}(s) := \int_{\mathbb{R}_0} U^{\tau,\eta,x}(s,z)\delta(z)\nu(dz)\,,$$

Let us consider the backward stochastic differential equation in equation (4.1), namely,

$$
\begin{aligned}
Y^{t,\eta,x}(t) &= \phi(X_T^{t,\eta,x}, X^{t,\eta,x}(T)) + \\
&+ \int_t^T \psi\left(X_s^{t,\eta,x}, X^{t,\eta,x}(s), Y^{t,\eta,x}(s), Z^{t,\eta,x}(s), \tilde{U}^{t,\eta,x}(s)\right) \mathrm{d}s \\
&+ \int_t^T Z^{t,\eta,x}(s)\mathrm{d}W(s) + \int_t^T \int_{\mathbb{R}_0} U^{t,\eta,x}(s,z)\tilde{N}(\mathrm{d}s,\mathrm{d}z)\,.
\end{aligned}
$$

Taking the expectation and exploiting equation (4.4), then Y satisfies equation (4.9).

In order to show the uniqueness let $u(t,\eta,x)$, $0 \le \tau \le t \le T$, be a mild solution of equation (4.7), so that

$$
\begin{aligned}
u(t,&\eta,x) = \mathbb{E}\left[\phi(X_T^{t,\eta,x}, X^{t,\eta,x}(T))\right] \\
&+\mathbb{E}\left[\int_t^T \psi\left(X_s^{t,\eta,x}, X^{t,\eta,x}(s), Y^{t,\eta,x}(s), Z^{t,\eta,x}(s), \tilde{U}^{t,\eta,x}(s)\right) \mathrm{d}s\right]\,.
\end{aligned}
$$

By recalling that $(X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t))_{t\in[0,T]}$ is a M^2-Markov process, and denoting by $\mathbb{E}^t$ the conditional expectation w.r.t. the filtration $\mathcal{F}_t$, we can write

$$
\begin{aligned}
u(t, X_t^{t,\eta,x}, X^{t,\eta,x}(t)) &= \mathbb{E}^t\left[\phi(X_T^{t,\eta,x}, X^{t,\eta,x}(T))\right] \\
&+\mathbb{E}^t\left[\int_\tau^T \psi\left(X_s^{t,\eta,x}, X^{t,\eta,x}(s), Y^{t,\eta,x}(s), Z^{t,\eta,x}(s), \tilde{U}^{t,\eta,x}(s)\right) ds\right] \\
&-\mathbb{E}^t\left[\int_\tau^t \psi\left(X_s^{\tau,\eta,x}, X^{\tau,\eta,x}(s), Y^{\tau,\eta,x}(s), Z^{\tau,\eta,x}(s), \tilde{U}^{\tau,\eta,x}(s)\right) \mathrm{d}s\right]\,.
\end{aligned}
$$

We set, for short,

$$
\begin{aligned}
&\xi := \phi(X_T^{t,\eta,x}, X^{t,\eta,x}(T)) \\
&+ \int_\tau^T \psi\left(X_s^{t,\eta,x}, X^{t,\eta,x}(s), Y^{t,\eta,x}(s), Z^{t,\eta,x}(s), \tilde{U}^{t,\eta,x}(s)\right) \mathrm{d}s\,.
\end{aligned}
$$

Thanks to the martingales representation theorem, see, e.g., [3, Thm. 5.3.5], there exist two predictable processes $\bar{Z} \in L^2(\Omega\times[0,T])$ and $\bar{U} \in L^2(\Omega\times[0,T]\times\mathbb{R}_0)$ such that

$$
\begin{aligned}
u(t,&X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t)) = u(\tau,\eta,x) \\
&+ \int_\tau^t \bar{Z}^{\tau,\eta,x}(s)\mathrm{d}W(s) + \int_\tau^t \int_{\mathbb{R}_0} \bar{U}^{\tau,\eta,x}(s,z)\tilde{N}(\mathrm{d}s,\mathrm{d}z) \\
&- \int_\tau^t \psi\left(X_s^{\tau,\eta,x}, X^{\tau,\eta,x}(s), Y^{\tau,\eta,x}(s), Z^{\tau,\eta,x}(s), \tilde{U}^{\tau,\eta,x}(s)\right) \mathrm{d}s\,.
\end{aligned}
$$

Applying theorem 3.2, we have

$$
\begin{aligned}
&u(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t)) = \phi(X_T^{\tau,\eta,x}, X^{\tau,\eta,x}(T)) + \\
&- \int_t^T \partial_x^\sigma u(s, X_s^{\tau,\eta,x}, X^{\tau,\eta,x}(s)) \mathrm{d}W(s) \\
&- \int_t^T \int_{\mathbb{R}_0} [u(s, X_s^{\tau,\eta,x}, X^{\tau,\eta,x}(s) + \gamma(s, X_s^{\tau,\eta,x}, X^{\tau,\eta,x}(s), z)) \\
&- u(s, X_s^{\tau,\eta,x}, X^{\tau,\eta,x}(s))] \, \tilde{N}(\mathrm{d}s, \mathrm{d}z) \\
&+ \int_t^T \psi \left(X_s^{\tau,\eta,x}, X^{\tau,\eta,x}(s), Y^{\tau,\eta,x}(s), Z^{\tau,\eta,x}(s), \tilde{U}^{\tau,\eta,x}(s) \right) \mathrm{d}s \, .
\end{aligned}
$$

By comparing last equation with the backward component of equation (4.1), we note that $(Y^{\tau,\eta,x}(t), Z^{\tau,\eta,x}(t), U^{\tau,\eta,x}(t,z))$ and the following three functions

$$
\begin{aligned}
&(u(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t)), \partial_x^\sigma u(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t)) \, , \\
&u\left(t, X_s^{\tau,\eta,x}, X^{\tau,\eta,x}(s) + \gamma(s, X_s^{\tau,\eta,x}, X^{\tau,\eta,x}(s), z)\right) - u(t, X_s^{\tau,\eta,x}, X^{\tau,\eta,x}(s))) \, ,
\end{aligned}
$$

solve the same equation. Therefore, due to the uniqueness of the solution, we have that

$$
Y^{\tau,\eta,x}(t) = u(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t)) \, .
$$

Setting $\tau = t$, we obtain $Y^{\tau,\eta,x}(t) = u(t, \eta, x)$ and the proof is complete. □

5. Application to optimal control

We are to apply previously derived results to a general class on non-linear control problem. The present section closely follows in [35, Section. 7], in particular we will consider weak control problems, we refer to [32] for a general treatment of the present notion of control, or [18, 19, 50, 51].

Let us therefore consider the following $\mathbb{R}-$valued controlled delay equation,

$$
\begin{cases}
dX(t) & = (\mu(t, X_t, X(t)) + F(t, X(t), \alpha(t))) \, dt + \\
& \quad + \sigma(t, X(t)) dW(t) + \int_{\mathbb{R}_0} \gamma(t, X(t), z) \tilde{N}(dt, dz) \, , \\
(X_{t_0}, X(t_0)) & = (x, \eta) \, ,
\end{cases} \tag{5.1}
$$

where we have denoted by $\alpha : \Omega \times [0,T] \to \mathcal{A}$ a $(\mathcal{F}_t)_{t \geq 0}$ −predictable process representing the control, being $\mathcal{A} \subset \mathbb{R}^N$ a convex set, $N \in \mathbb{N}$.

In what follows we assume μ, σ and γ to satisfy assumptions 2.2, we also require that it exists a constant $C_\sigma > 0$ such that, for any $t \in [0,T]$ and $x \in \mathbb{R}$,

$$
|\sigma^{-1}(t,x)| \leq C_\sigma \, .
$$

We remark that a possibly choice for the coefficient μ in equation (5.1) is of the form

$$
\mu(t, X_t, X(t)) = \int_{-r}^0 X(t+\theta) \varpi(d\theta) \, ,
$$

for ϖ a Borel measure of bounded variation on the interval $[-r, 0]$.

Following [35, Section 7], we will say therefore that an *admissible control system* (acs) is given by $\mathbb{U} = (\Omega, \mathcal{F}, \mathbb{P}, W, \nu, \alpha, X)$, where $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space, with an associated filtration satisfying usual conditions, W is a Wiener process whereas ν is a Lévy measure also satisfying usual assumptions introduced in previous sections, α is the control defined above and X is the unique solution to equation (5.1). Then we wish to minimize, over all control $\alpha \in \mathcal{A}$, the following functional

$$J\left(t_0, (x, \eta), \mathbb{U}\right) = \int_0^T h(s, X(s), \alpha) ds + g(X(T)) \,. \tag{5.2}$$

We thus assume the following to hold.

Assumption 5.1. **(i).** let $F : [0, T] \times \mathbb{R} \times \mathcal{A} \to \mathbb{R}$ be measurable and such that there exist $C_F > 0$ and $m \geq 0$ such that, for any $t \in [0, T]$, x, x_1, $x_2 \in \mathbb{R}$ and $\alpha \in \mathcal{A}$,

$$|F(t, x, \alpha)| \leq C \,,$$
$$|F(t, x_1, \alpha) - F(t, x_2, \alpha)| \leq C_F (1 + |x_1| + |x_2|)^m |x_1 - x_2| \,.$$

(ii). let $h : [0, T] \times \mathbb{R} \times \mathcal{A} \to \mathbb{R} \cup \{+\infty\}$ be measurable and such that there exist $C_h > 0$ and $m \geq 0$ such that, for any $t \in [0, T]$, x, x_1, $x_2 \in \mathbb{R}$ and $\alpha \in \mathcal{A}$,

$$h(t, 0, \alpha) \geq -C_h \,, \quad \inf_{\alpha \in \mathcal{A}} h(t, 0, \alpha) \leq C_h \,,$$
$$|h(t, x_1, \alpha) - h(t, x_2, \alpha)| \leq C_h (1 + |x_1| + |x_2|)^m |x_1 - x_2| + h(t, x_2, \alpha) \,.$$

(iii). let $g : \mathbb{R} \to \mathbb{R}$ be measurable and such that there exist $C_g > 0$ and $m > 0$ such that, for any x, x_1, $x_2 \in \mathbb{R}$ it holds

$$|g(x_1) - g(x_2)| \leq C_g (1 + |x_1| + |x_2|)^m |x_1 - x_2| \,.$$

The particular form for equation (5.1) leads to consider an associated *Hamilton-Jacobi-Bellman* (HJB) equation which is a semilinear partial integro-differential equation of the form of equation (4.7) studied in previous sections. Noticed that the particular form for equation (5.1), in particular the presence of the control in the drift, is imposed by the techniques we will use.

Then the controlled equation (5.1), together with the functional J introduced in equation (5.2), lead to define in a classical way the *Hamiltonian* associated to the above problem as

$$\psi(t, x, z) = -\inf_{\alpha \in \mathcal{A}} \left\{ h(t, x, \alpha) + z\sigma^{-1}(s, x) F(s, x, \alpha) \right\} \,,$$
$$\Gamma(t, x, z) = \left\{ \alpha \in \mathcal{A} \,:\, \psi(t, x, z) + h(t, x, \alpha) + z\sigma^{-1}(s, x) F(s, x, \alpha) = 0 \right\} \,.$$

Let us stress that under above assumptions we have that ψ satisfies assumptions 4.1. Eventually we can formulate the HJB equation associated to the above stated non-linear control problem to be

$$\begin{cases} \frac{\partial}{\partial t} u(t, \eta, x) + \mathcal{L}_t u(t, \eta, x) & = \psi(t, x, \partial_x^{\sigma^{-1} F} u(t, \eta, x)) \,, \\ u(T, \eta, x) = g(x) \,, \end{cases} \tag{5.3}$$

where the notation is as above introduced. From Theorem 4.8, it follows that equation (5.3) admits a unique solution in the sense of generalized direction gradient.

Eventually, from [34, Theorem 7.2] or [35, Theorem 7.2] which follow in a straightforward manner in the present case, we have that an acs system is optimal if and only if

$$\alpha(t) \in \Gamma\left(t, X(t), \zeta(t, X_t, X(t))\right) ,$$

being $\zeta : [0,T] \times M^2 \to \mathbb{R}$ an element of the directional generalized gradient.

Theorem 5.2. *Let u be a mild solution to the HJB equation* (5.3)*, and choose ζ to be an element of the generalized directional gradient $\partial_x^{\sigma^{-1}F} u$. Then, for all acs, we have that $J(t_0, x, \eta, \mathbb{U}) \geq u(t_0, x, \eta)$, and the equality holds if and only if*

$$\alpha(t) \in \Gamma\left(t, X(t), \zeta(t, X_t, X(t))\right) , \quad \mathbb{P} - a.s. \text{ for a.a. } t \in [t_0, T] .$$

Moreover, if there exists a measurable function $\varsigma : [0,T] \times \mathbb{R} \to \mathcal{A}$ with

$$\varsigma(t, x, z) \in \Gamma(t, x, z) ,$$

then there also exists at least one acs such that

$$\bar{\alpha}(t) = \varsigma(t, X^\alpha(t), \zeta(t, X_t^\alpha, X^\alpha(t))) , \quad \mathbb{P} - a.s. \text{ for a.a. } t \in [t_0, T] ,$$

where $(X_t^\alpha, X^\alpha(t))$ is the solution to equation

$$\begin{cases} dX^\alpha(t) & = \mu(t, X_t^\alpha, X^\alpha(t))dt + \\ & + F(t, X^\alpha(t), \varsigma(t, X^\alpha(t), \zeta(t, X^\alpha(t))))dt + \\ & + \sigma(t, X^\alpha(t))dW(t) + \int_{\mathbb{R}_0} \gamma(t, X^\alpha(t), z)\tilde{N}(dt, dz) , \\ \left(X_{t_0}^\alpha, X^\alpha(t_0)\right) & = (x, \eta) , \end{cases}$$

Proof. See [35, Th. 7.2] or also [50, Th. 4.7, Cor. 4.8]. □

References

[1] S. Albeverio, L. Di Persio, E. Mastrogiacomo and B. Smii *A Class of Lévy Driven SDEs and their Explicit Invariant Measures*, Potential Analysis Volu. 45, Issue 2, (2016): 229-259

[2] W.G. Aiello, H.I. Freedman, *A time-delay model of single-species growth with stage structure,* Math. Biosci. **101(2)** (1990), 139–153.

[3] D. Applebaum *Lévy processes and stochastic calculus.* Cambridge University Press, 2009.

[4] D. Applebaum *Universal Malliavin Calculus in Fock and Lévy-Itô Spaces.* arXiv preprint arXiv:0808.2593 (2008).

[5] M. Arriojas, Y. Hu, S.-E.A. Mohammed, G. Pap, *A delayed Black and Scholes formula,* Stochastic Analysis and Applications **25.2** (2007): 471-492.

[6] D. Banos, F. Cordoni, G. Di Nunno, L. Di Persio, E.E. Rose, *Stochastic systems with memory and jumps,*, arXiv:1603.00272 (submitted) 2016.

[7] G. Barles, R. Buckdahn, E. Pardoux. *Backward stochastic differential equations and integral-partial differential equations*, Stochastics: An International Journal of Probability and Stochastic Processes **60.1-2** (1997): 57-83.

[8] D.R. Bell. *The Malliavin Calculus.* Courier Corporation, 2012.

[9] I. Boutle, R.H.S. Taylor, R.A. Römer, *El Niño and the delayed action oscillator,* Am. J. Phys., **75(1)** (2007), http://dx.doi.org/10.1119/1.2358155.

[10] Y.K. Chang and J. J. Nieto. *Existence of solutions for impulsive neutral integro-differential inclusions with nonlocal initial conditions via fractional operators.* Numerical Functional Analysis and Optimization 30.3-4 (2009): 227-244.

[11] M.H. Chang, R. Youree, *The European option with hereditary price structures: basic theory.* Appl. Math. Comput., **102** (1999), 279–296.

[12] A. Chojnowska–Michalik. *Representation theorem for general stochastic delay equations.* Bulletin de l'academie Polonaise des sciences– serie des sciences mathematiques astronomiques et physiques 26.7 (1978): 635-642.

[13] R. Cont, D. Fourniè, *Change of variable formulas for non-anticipative functionals on path space,* J. Funct. Anal., **259(4)** (2010), 1043-1072.

[14] R. Cont, D. Fourniè, *Functional Itô calculus and stochastic integral representation of martingales,* Ann. Prob., **41(1)** , 109–133, (2013).

[15] F. Cordoni, L. Di Persio. *Backward Stochastic Differential Equations Approach to Hedging, Option Pricing, and Insurance Problems,* International Journal of Stochastic Analysis 2014 (2014).

[16] F. Cordoni, L. Di Persio, *Delay stochastic differential equations approach to the pricing of path dependent options in stochastic volatility models with jumps,* (2015) (submitted).

[17] F. Cordoni, L. Di Persio, *A BSDE with delayed generator approach to pricing under counterparty risk and collateralization,* International Journal of Stochastic Analysis (2016).

[18] F. Cordoni and L. Di Persio, *Gaussian estimates on networks with dynamic stochastic boundary conditions*, Infinite Dimensional Analysis and Quantum Probability, in press, to be published in the first 2017 issue of IDAQP (2017).

[19] F. Cordoni and L. Di Persio, *Stochastic reaction-diffusion equations on networks with dynamic time-delayed boundary conditions*, Journal of Mathematical Analysis and Applications (JMAA); in Press, available online: 10 February 2017 (2017)

[20] F. Cordoni, L. Di Persio, *Invariant measure for the Vasicek interest rate model in the Heath-Jarrow-Morton-Musiela framework,* Infinite Dimensional Analysis, Quantum Probability and Related Topics, 18 (3) (2015).

[21] F. Cordoni, L. Di Persio, L. Maticiuc and A. Zălinescu, A. *A stochastic approach to path–dependent nonlinear Kolmogorov equations via BSDEs with time-delayed generators and applications to finance.* arXiv:1602.05793 (submitted) (2016).

[22] A. Cosso and F. Russo. *Functional Itô versus Banach space stochastic calculus and strict solutions of semilinear path-dependent equations.* Infinite Dimensional Analysis, Quantum Probability and Related Topics, 19 (4) (2016).

[23] K.R. Dahl, S.-E. A. Mohammed, B. Øksendal, E.E. Røse, *Optimal control of systems with noisy memory and BSDEs with Malliavin derivatives,* Journal of Functional Analysis 271.2 (2016): 289-329.

[24] L. Delong, *Backward stochastic differential equations with jumps and their actuarial and financial applications*, EAA Series, Springer-Verlag, 2013.

[25] L. Delong, P. Imkeller, *On Malliavin's differentiability of BSDEs with time delayed generators driven by Brownian motions and Poisson random measures,* Stoch. Proc. Appl., **120(9)** (2010), 1748–1775.

[26] G. Di Nunno, B. Øksendal, F. Proske, *Malliavin Calculus for Lévy Processes with Applications to Finance.* Springer, Berlin, 2009.

[27] I. Ekren, C. Keller,N. Touzi and J. Zhang. *On viscosity solutions of path dependent PDEs.* The Annals of Probability, 42(1), 204-236 (2014).

[28] I. Ekren, N. Touzi and J. Zhang. Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part I. The Annals of Probability, 44(2), 1212-1253 (2016).

[29] I. Ekren, N. Touzi and J. Zhang. Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part II. The Annals of Probability, 44(4), 2507-2553 (2016).

[30] F. Flandoli and G. Zanco, *An infinite-dimensional approach to path-dependent Kolmogorov equations,* Ann. Probab. 44, 4, 2643–2693, (2016).

[31] F. Flandoli, F. Russo and G. Zanco. *Infinite-dimensional calculus under weak spatial regularity of the processes,* Journal of Theoretical Probability (2015): 1-38.

[32] W. H. Fleming and H. M. Soner, Controlled Markov processes and viscosity solutions. Vol. 25. Springer Science & Business Media, (2006).

[33] M. Fuhrman, F. Masiero, G. Tessitore, *Stochastic Equations with Delay: Optimal Control via BSDEs and Regular Solutions of Hamilton-Jacobi-Bellman Equations,* Siam J. Control Optim., **48(7)** (2010), 4624–4651.

[34] M. Fuhrman, G. Tessitore, *Nonlinear Kolmogorov Equations in Infinite Dimensional Spaces: the Backward Stochastic Differential Equations Approach and Applications to Optimal Control,* Ann. Prob., **30(3)** (2002), 1397–1465.

[35] M. Fuhrman, G. Tessitore, *Generalized directional gradients, backward stochastic differential equations and mild solutions of semilinear parabolic equations,* Appl. Math. Optim., **51(3)** (2005) 279–332.

[36] L. Hu, Yong Ren, and R. Sakthivel. *Existence and uniqueness of mild solutions for semilinear integro-differential equations of fractional order with nonlocal initial conditions and delays.* Semigroup Forum. Vol. 79. No. 3. Springer-Verlag, 2009.

[37] J. A. Leòn, J. L. Solé, F. Utzet, J. Vives, *On Lévy processes, Malliavin calculus and market models with jumps,* Financ. Stoch., 6(2) (2002), 197–225.

[38] U. Kuchler, E. Platen, *Time delay and noise explaining cyclical fluctuations in prices of commodities,* University of Technology, Sydney, 2007.

[39] J.A. León, J.L. Solè, F. Utzet, J. Vives, *On Lévy processes, Malliavin calculus and market models with jumps,* Financ. Stoch., **6(2)** (2002), 197–225.

[40] C. Lizama and G. M. N'Guérékata. *Bounded mild solutions for semilinear integro differential equations in Banach spaces.* Integral Equations and Operator Theory 68.2 (2010): 207-227.

[41] S.-E.A. Mohammed, *Stochastic Functional Differential Equations.* Research Notes in Mathematics No. 99. Pitman, Boston, 1984.

[42] S.-E.A. Mohammed, *Stochastic differential equation with memory: theory, examples, and appli- cations*, in: *Decreusefond, L., Gjerde, J., Oksendal, B., Ustunel, A.S. (eds.) Stochastic Analysis. Progress in Probability,* **42**, 1–77. Birkhäuser, Basel, 1998.

[43] S.-E.A. Mohammed, F. Yan. *A stochastic calculus for systems with memory.* Stochastic analysis and applications 23.3 (2005): 613-657.

[44] S. Peszat, J. Zabczyk *Stochastic partial differential equations with Lévy noise. An evolution equation approach,* Encyclopedia Math. Appl., **113**, Cambridge University Press, Cambridge, 2007.

[45] E. Petrou, *Malliavin calculus in Lévy spaces and applications in finance.* Electron. J. Probab. Volume 13 (2008), paper no. 27, 852-879.

[46] M. Reiß, M. Riedle, O. van Gaans, *Delay differential equations driven by Lévy processes: Stationarity and Feller properties,* Stoch. Proc. Appl., **116(10)** (2006), 1409–1432.

[47] F. Russo, P. Vallois, *Forward, backward and symmetric stochastic integration,* Probab. Theory Related Fields, **97(3)** (1993), 403–421.

[48] F. Russo, P. Vallois, *The generalized covariation process and Itô formula,* Stoch. Proc. Appl., **59** (1995), 81–104.

[49] F. Russo, P. Vallois, *Itô formula for C^1-functions of semimartingales,* Probab. Theory Related Fields, **104(1)** (1996), 27–41.

[50] K-L. Su, *Semilinear parabolic equations driven by cylindrical stable processes and its optimality applications*, Journal of Business, 21 (2013): 125-154.

[51] K-L. Su, *Stochastic Optimal Control Problems for Elliptic Equations Driven by α–stable Noise in Hilbert Spaces* (preprint).

[52] K-L. Su, *A Note of Girsanov's Theorem in Banach Spaces*, J. Business, 19, 133-148, NOU Press, Taipei, Taiwan (2011).

[53] Swishchuk, Anatoliy. *Modelling and pricing of variance swaps for multi-factor stochastic volatilities with delay.* Canadian Applied Mathematics Quarterly 14.4 (2006): 439-467.

F. Cordoni
Dept. of Mathematics,
University of Trento,
V. Sommarive 14,
38123 Povo, Italy
e-mail: francesco.cordoni@unitn.it

L. Di Persio†
Dept. of Computer Sciences,
University of Verona,
Strada Le Grazie 15,
37134 Verona, Italy
e-mail: luca.dipersio@univr.it

I. Oliva
Dept. of Economics,
University of Verona,
Vicolo Campofiore 2,
37129 Verona, Italy
e-mail: immacolata.oliva@univr.it