Check for
updates

Feynman–Kac formula for BSDEs with jumps and time delayed generators associated to path-dependent nonlinear Kolmogorov equations

Luca Di Persio, Matteo Garbelli and Adrian Zălinescu

Abstract. We consider a system of forward backward stochastic differential equations (FBSDEs) with a time-delayed generator driven by Lévy-type noise. We establish a non-linear Feynman–Kac representation formula associating the solution given by the FBSDEs system to the solution of a path dependent nonlinear Kolmogorov equation with both delay and jumps. Obtained results are then applied to study a generalization of the so-called large investor problem, where the stock price evolves according to a jump-diffusion dynamic.

Mathematics Subject Classification. Primary 34K50, 60H15, Secondary 60G51, 60H30.

Keywords. Feynman–Kac formula, Jump-diffusion, Delay, Non linear BSDE.

1. Introduction

Stochastic delay differential equations are derived as a natural generalisation of stochastic ordinary differential equations by allowing coefficients' evolution to depend not only on the present state but also on past values.

In this article, we analyze a stochastic process described by a system of forward–backward stochastic differential equations (FBSDEs). The forward path-dependent equation is driven by a Lévy process, while the backward one presents a path-dependent behaviour with dependence on a small delay δ.

We establish a non-linear Feynman–Kac representation formula associating the solution of the latter FBSDE system to the one of a path-dependent nonlinear Kolmogorov equation with delay and jumps. In particular, we prove

that the stochastic system allows to uniquely construct a solution to the parabolic partial differential equation (PDE), in the spirit of the Pardoux–Peng approach, see [31–33], i.e.:

$$\begin{cases} -\partial_t u(t,\phi) - \mathcal{L}u(t,\phi) - f(t,\phi, u(t,\phi), \partial_x u(t,\phi)\sigma(t,\phi), \mathcal{J}u(t,\phi), (u(\cdot,\phi))_t) = 0\,, \\ u(T,\phi) = h(\phi)\,, \qquad\qquad (t,\phi) \in [0,T] \times \Lambda\,, \end{cases} \quad (1.1)$$

$T < \infty$ being a fixed time horizon and Λ being $D([0,T]; \mathbb{R}^d)$, i.e. the space of càdlàg $\mathbb{R}^d$-valued functions defined on the interval $[0,T]$.

The integro-differential operator $\mathcal{J}$ is associated with the jump behaviour and is defined as

$$\mathcal{J}u(t,\phi) := \int_{\mathbb{R}\backslash\{0\}} \left[u(t, \phi^{t,\gamma(t,\phi,z)}) - u(t,\phi) \right] \lambda(z)\nu(dz)\,, \qquad (1.2)$$

where $\phi^{t,\gamma}$ models the vertical perturbation of the path ϕ defined by

$$\phi^{t,\gamma}(\theta) := \phi(\theta)\mathbb{1}_{\theta\in[0,t)} + [\phi(t) + \gamma]\,\mathbb{1}_{\theta\in[t,T]}\,, \qquad \theta, t \in [0,T]\,, \gamma \in \mathbb{R}^d\,, \; (1.3)$$

with $\gamma : [0,T] \times \Lambda \times \{\mathbb{R}\backslash\{0\}\} \to \mathbb{R}^d$ being a continuous, non-anticipative functional encoding the shifting at the (right) endpoint while λ models the intensity of the jumps and ν represents the associated Lévy measure.

The second order differential operator $\mathcal{L}$, associated to the diffusion, is defined by

$$\mathcal{L}u(t,\phi) := \frac{1}{2}\mathrm{Tr}\,[\sigma(t,\phi)\sigma^*(t,\phi)\partial^2_{xx}u(t,\phi)] + \langle b(t,\phi), \partial_x u(t,\phi)\rangle \qquad (1.4)$$

$$+ \int_{\mathbb{R}\backslash\{0\}} \left(u(t, \phi^{\gamma(t,\phi,z)}) - u(t,\phi) - \partial_x u(t,\phi)\gamma(t,\phi,z) \right) \nu(dz)\,,$$

with $b : [0,T] \times \Lambda \to \mathbb{R}^d$ and $\sigma : [0,T] \times \Lambda \to \mathbb{R}^{d\times l}$ being two non-anticipative functionals.

Also, for a fixed delay $\delta > 0$ we set

$$(u(\cdot,\phi))_t := (u((t+\theta)^+, \phi))_{\theta\in[-\delta,0]}\,. \qquad (1.5)$$

Let us underline that the study of a path-dependent Kolmogorov equation whose generator f depends on both a delayed term $(u(\cdot,\phi))_t$ and on a jump operator $\mathcal{J}u(t,\phi)$, represents the main novelty we provide in this paper.

Under appropriate assumptions on the coefficients, the deterministic non-anticipative functional $u : [0,T] \times \Lambda \to \mathbb{R}$ given by the representation formula

$$u(t,\phi) := Y^{t,\phi}(t)\,, \qquad (1.6)$$

is a mild solution of the Kolmorogov Equation (1.1).

Concerning the stochastic process $Y^{t,\phi}(t)$ in Eq. (1.6), provided conditions in Assumptions 2.2–3.1 are fulfilled, we prove that the quadruple

$$\left(X^{t,\phi}, Y^{t,\phi}, Z^{t,\phi}, U^{t,\phi}\right)_{s\in[t,T]}\,,$$

is the unique solution of the system of FBSDEs on $[t,T]$ given by

$$
\begin{cases}
X^{t,\phi}(s) = \phi(t) + \displaystyle\int_t^s b(r, X^{t,\phi})dr + \int_t^s \sigma(r, X^{t,\phi})dW(r) \\
\qquad + \displaystyle\int_t^s \int_{\mathbb{R}\setminus\{0\}} \gamma(r, X^{t,\phi}, z)\tilde{N}(dr, dz) \\
Y^{t,\phi}(s) = h(X^{T,\phi}) + \displaystyle\int_s^T f\big(r, X^{t,\phi}, Y^{t,\phi}(r), Z^{t,\phi}(r), \tilde{U}^{t,\phi}(r), Y_r^{t,\phi}\big)dr \\
\qquad - \displaystyle\int_s^T Z^{t,\phi}(r)dW(r) - \int_s^T \int_{\mathbb{R}\setminus\{0\}} U^{t,\phi}(r,z)\tilde{N}(dr,dz),
\end{cases}
\tag{1.7}
$$

where W stands for a l-dimensional standard Brownian motion. Assuming a delay $\delta \in \mathbb{R}^+$, the notation $Y_r^{t,\phi}$, appearing in the generator f of the backward dynamic in the system (1.7), stands for the delayed path of the process $Y^{t,\phi}$ restricted to $[r-\delta, r]$, namely

$$
Y_r^{t,\phi} := \big(Y^{t,\phi}((r+\theta)^+)\big)_{\theta \in [-\delta, 0]}.
\tag{1.8}
$$

In Eq. (1.7), $\tilde{N}$ models a compensated Poisson random measure, independent from W, with associated Lévy measure ν. This stochastic term appears also in the definition of the integral term $\tilde{U} : \Omega \times [0,T] \to \mathbb{R}$ related to the jump process by

$$
\tilde{U}^{t,\phi}(r) = \int_{\mathbb{R}\setminus\{0\}} U^{t,\phi}(r,z)\lambda(z)\nu(dz);
\tag{1.9}
$$

we need it in order to introduce to express the solution of (1.1) via the FBSDE system (1.7).

The connection between probability theory and PDEs is a widely analysed subject, the well-known Feynman–Kac formula [25] being one of its main turning points stating that solutions for a large class of second order PDEs of both elliptic and parabolic type, can be expressed in terms of expectation of a diffusion process. The latter result has been then generalised by Pardoux and Peng in [31,32] to show the connection between backward stochastic differential equations (BSDEs) and a system of semi-linear PDEs, then proving the nonlinear Feynman–Kac formula within the Markovian setting. Concerning the non-Markovian scenario, we know from [34,35] that a nonlinear Feynman–Kac formula can be still established, associating a path-dependent PDE to a non-Markovian BSDE. More recently, the introduction of horizontal and vertical derivatives of non-anticipative functionals on path spaces by Dupire [18] and by Cont and Fournié [8–10] facilitated the formulation of a new class of path dependent PDEs and the introduction of the so-called *viscosity solution* concept, see [20,21,35], for more details. For a complete overview about stochastic calculus with delay we refer to [29,30,39].

A connection between time-delayed BSDEs and path-dependent PDEs has been proved by an *infinite dimensional lifting* approach in [22,28]. In [3], the authors consider a BSDE driven by a Brownian motion and a Poisson random measure that provides a viscosity solution of a system of parabolic

integral-partial differential equations. In [11], the existence of a viscosity solution to a path-dependent nonlinear Kolmogorov equation (without jumps) and the corresponding nonlinear Feynman–Kac representation has been proved.

In this paper, we deal with the notion of *mild solution* which can be seen as an intermediate notion for solutions of a PDE lying in between the notions of classical and viscosity solutions. In [24], the authors provide the definition of mild solution for nonlinear Kolmogorov equations along with its link with a specific stochastic process. The latter has been also proved for semilinear parabolic equations in [23], where the definition of the *generalized directional gradient* is firstly introduced. The concept of mild solution together with the generalized directional gradient to handle path-dependent Kolmogorov equation with jumps and delay has been widely analyzed in the functional formulation, see, e.g., [12]. Moreover, a discrete-time approximation for solutions of a system of decoupled FBSDEs with jumps has been proved in [7] by means of Malliavin calculus tools.

Concerning the theory of BSDE with a dependence on a delay, in [16], the authors proved the existence of a solution for a BSDE with a time-delayed generator that depends on the past values of the solution. In particular, both existence and uniqueness are proved assuming a sufficiently small time horizon T or a sufficiently small Lipschitz constant for the generator. Let us underline that the latter has an equivalent within our setting, as we state in Remark 3.2. Moreover, in [15,17] the authors defined a path-dependent BSDE with time delayed generators driven by Brownian motions and Poisson random measures, with coefficients depending on the whole solution's path.

In [2], following a different approach, namely considering systems with memory and jumps, the authors provide a characterization of a strong solution for a delayed SDE with jumps, considering both L^p-type space and càdlàg processes to derive a non-linear Feynman–Kac representation theorem.

The present paper is structured as follows: we start stating notations and problem setting in Sect. 2, according to the theoretical framework developed by [15,16]; in Sect. 3 we study the well-posedness of the path-dependent BSDE mentioned appearing in the Markovian FBSDEs system (1.7) following the approach in [11] by additionally considering jumps; in Sect. 4 we provide a Feynman–Kac formula relating the BSDE to the Kolmogorov Equation defined in (1.1) to then generalise results in [12] by considering a dependence in the generator f of the backward dynamic on a delayed L^2 term, namely $Y_r^{t,\phi}$, for a small delay δ; in Sect. 5 we derive the existence of a mild solution for the Kolmogorov Equation within the setting developed in [23]; finally, in Sect. 6, we provide an application based on the analyzed theoretical setting, namely a version of the *Large Investor Problem* characterised by a jump-diffusion dynamic.

2. Notation and problem formulation

On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we consider a standard l-dimensional Brownian motion W and a homogeneous Poisson random measure N on $\mathbb{R}^+ \times (\mathbb{R} \setminus$

$\{0\}$), independent from W, with intensity ν. With the notation $\mathbb{R}_0 := \mathbb{R}\backslash\{0\}$, we also define the compensated Poisson random measure $\tilde{N}$ defined on $\mathbb{R}^+ \times \mathbb{R}_0$ by

$$\tilde{N}(dt, dz) := N(dt, dz) - \nu(dz)dt\,. \tag{2.1}$$

For the sake of completeness, let us recall that the term $\nu(dz)dt$ represents the *compensator* of the random measure N and we assume that

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

We refer to, e.g., [1,4] for further details about the stochastic integration in the presence of jumps. We remark that the assumption in (2.2) is a standard condition when dealing with financial applications.

2.1. The forward–backward delayed system

In this section we introduce the delayed forward-backward system, assuming path-dependent coefficients for the forward and the backward components, a dependence on a small delay into the generator f and the presence of jumps modelled via a compound Poisson measure. Furthermore, the equation is formulated on a general initial time t and initial values. Thus, we need to equip the backward equation with a suitable condition in $[0, t)$, as we introduced in Equation (2.6).

On previously defined probability space, we consider a filtration $\mathbb{F}^t = \{\mathcal{F}_s^t\}_{s\in[0,T]}$, which is nothing but the one jointly generated by $W(s \wedge t) - W(t)$ and $N(s \wedge t, \cdot) - N(t, \cdot)$, augmented by all $\mathbb{P}$-null sets. We emphasize that $\mathbb{F}^t$ depends explicitly on t, namely the arbitrary initial time in $[0, T]$ for the dynamic in Eq. (1.7).

Furthermore, the components of the solution of the backward dynamic are defined in the following Banach spaces:

- $\mathbb{S}_t^2(\mathbb{R})$ denotes the space of (equivalence class of) $\mathbb{F}^t$-adapted, product measurable càdlàg processes $Y : \Omega \times [0, T] \to \mathbb{R}$ satisfying

$$\mathbb{E}\left[\sup_{t\in[0,T]} |Y(t)|^2\right] < \infty\,;$$

- $\mathbb{H}_t^2(\mathbb{R}^l)$ denotes the space of (equivalence class of) $\mathbb{F}^t$-predictable processes $Z : \Omega \times [0, T] \to \mathbb{R}^l$ satisfying

$$\mathbb{E}\left[\int_0^T |Z(t)|^2 dt\right] < \infty\,;$$

- $\mathbb{H}_{t,\nu}^2(\mathbb{R})$ denotes the space of (equivalence class of) $\mathbb{F}^t$-predictable processes $U : \Omega \times [0, T] \times \mathbb{R}_0 \to \mathbb{R}$ satisfying

$$\mathbb{E}\left[\int_0^T \int_{\mathbb{R}_0} |U(t, z)|^2 \nu(dz)dt\right] < \infty\,.$$

The spaces $\mathbb{S}_t^2(\mathbb{R})$, $\mathbb{H}_t^2(\mathbb{R}^l)$ and $\mathbb{H}_{t,\nu}^2(\mathbb{R})$ are endowed with the following norms:

$$||Y||_{\mathbb{S}_t^2(\mathbb{R})}^2 = \mathbb{E}\left[\sup_{t \in [0,T]} e^{\beta t}|Y(t)|^2 \right],$$

$$||Z||_{\mathbb{H}_t^2(\mathbb{R}^l)}^2 = \mathbb{E}\left[\int_0^T e^{\beta t}|Z(t)|^2 dt \right],$$

and

$$||U||_{\mathbb{H}_{t,\nu}^2(\mathbb{R})}^2 = \mathbb{E}\left[\int_0^T \int_{\mathbb{R}_0} e^{\beta t}|U(t,z)|^2 \nu(dz)dt \right],$$

for some $\beta > 0$, to be precised later.

The main goal is to find a family of stochastic processes $\left(X^{t,\phi}, Y^{t,\phi}, Z^{t,\phi}, U^{t,\phi}\right)$ for $t,\phi \in [0,T] \times \Lambda$ adapted to $\mathbb{F}^t$ such that the following decoupled forward-backward system holds a.s.

$$
\begin{cases}
X^{t,\phi}(s) = \phi(t) + \displaystyle\int_t^s b(r, X^{t,\phi})dr + \int_t^s \sigma(r, X^{t,\phi})dW(r)+ \\
\qquad + \displaystyle\int_t^s \int_{\mathbb{R}_0} \gamma(r, X^{t,\phi}, z)\tilde{N}(dr, dz)\,, & s \in [t,T] \\
X^{t,\phi}(s) = \phi(s)\,, & s \in [0,t] \\
Y^{t,\phi}(s) = h(X^{t,\phi}) + \displaystyle\int_s^T f\left(r, X^{t,\phi}, Y^{t,\phi}(r), Z^{t,\phi}(r), \tilde{U}^{t,\phi}(r), Y_r^{t,\phi}\right)dr \\
\qquad - \displaystyle\int_s^T Z^{t,\phi}(r)dW(r) - \int_s^T \int_{\mathbb{R}_0} U^{t,\phi}(r,z)\tilde{N}(dr, dz)\,, & s \in [t,T] \\
Y^{t,\phi}(s) = Y^{s,\phi}(s)\,, \quad Z^{t,\phi}(s) = U^{t,\phi}(s,z) = 0\,, & s \in [0,t]\,,
\end{cases}
\tag{2.3}
$$

recalling that the term $\tilde{U}^{t,\phi}$ was introduced in (1.9).

It is worth mentioning that, differently from [15], we work in a non-Markovian setting, enforcing an initial condition over all the interval $[0,t]$. More precisely, for both forward and backward equations, the values of the solution $X^{t,\phi}$ need to be known in the time interval $[0,t]$. Analogously, regarding the backward component, the values of $Y^{t,\phi}$, $Z^{t,\phi}$ and $U^{t,\phi}$ need also to be prescribed for $s \in [0,t]$.

Remark 2.1. The δ-delayed feature concerns only Y, but we emphasize that it is possible to generalize this result to treat the case where both Z and U depend on their past values for a fixed delay δ.

For the sake of simplicity, we consider the case with Y_r, hence limiting ourselves to just one, the process Y, L^2 delayed term. As a consequence, the latter implies that we will have to consider a larger functional space to properly define the contraction which is an essential step to prove the fix point argument in Theorem 3.3.

2.1.1. The forward path-dependent SDE with jumps. We first study the forward component of X appearing in the system (2.3). It is defined according to the following equation:

$$
\begin{cases}
X^{t,\phi}(s) = \phi(t) + \displaystyle\int_t^s b(r, X^{t,\phi})dr + \int_t^s \sigma(r, X^{t,\phi})dW(r) + \\[2mm]
\qquad\qquad + \displaystyle\int_t^s \int_{\mathbb{R}_0} \gamma(r, X^{t,\phi}, z)\tilde{N}(dr, dz)\,, \qquad s \in [t, T] \qquad (2.4) \\[2mm]
X^{t,\phi}(s) = \phi(s)\,, \qquad\qquad\qquad\qquad\qquad s \in [0, t]\,.
\end{cases}
$$

More precisely, we say that $X^{t,\phi}$ is a solution to equation (2.4) if the process $s \mapsto X^{t,\phi}(s)$ is $\mathbb{F}^t$-adapted, $\mathbb{P}$-a.s. continuous and (2.4) is satisfied for any $s \in [0, T]$, $\mathbb{P}$-a.s.

Recalling that a Borel-measurable function $\varphi : [0, T] \times \Lambda \times S$ is *non-anticipative* if $\varphi(t, \phi, e) = \varphi(t, \phi(\cdot \wedge t), e)$, for all $(t, \phi, e) \in [0, T] \times \Lambda \times S$, where S is an arbitrary topological space, we assume the following assumptions to hold.

Assumption 2.2. Let us consider three non-anticipative functions $b : [0, T] \times \Lambda \to \mathbb{R}^d$, $\sigma : [0, T] \times \Lambda \to \mathbb{R}^{d \times l}$ and $\gamma : [0, T] \times \Lambda \times \mathbb{R} \to \mathbb{R}^d$ such that

(A_1) b, σ and γ are continuous;

(A_2) there exists $\ell > 0$ such that

$$
|b(t, \phi) - b(t, \phi')| + |\sigma(t, \phi) - \sigma(t, \phi')| + ||\gamma(t, \phi, \cdot) - \gamma(t, \phi', \cdot)||_{L^2}
$$
$$
\leq \ell ||\phi - \phi'||_{L^\infty}
$$

for any $t \in [0, T]$, $\phi, \phi' \in \Lambda$;

(A_3) the following bound

$$
\int_{\mathbb{R}_0} \sup_{\phi \in \Lambda} |\gamma(r, \phi, z)|^2 \nu(dz) < \infty\,,
$$

holds.

The existence and the pathwise uniqueness for a solution of forward SDE with jumps under Lipschitz coefficients is a known result, already classical for the case without path-dependent coefficients, see, e.g. [38]. For the sake of completeness, we report the following proposition:

Proposition 2.3. *If (A_1), (A_2), (A_3) hold, then there exists a solution to (2.4) and this solution is pathwise unique.*

The proof for an equivalent path-dependent setting is stated in Theorem 2.12 in [2]. Both existence and uniqueness are derived via a Picard iteration approach in the so-called Delfour–Mitter space that we introduce in Sect. 4.1. It turns out that our Eq. (2.4) can be reformulated with Delfour–Mitter coefficients such as Eq. (4.4) by transformations of the coefficients we define in Eq. (4.3). Hypotheses $(\mathbf{D_1})$ and $(\mathbf{D_2})$ required for the existence result can be deduced from our conditions (A_1)–(A_3). A similar approach, but for a different class of integrators is used in Theorem 5.2.15 in [6].

2.1.2. The backward delayed path-dependent SDE with jumps. We now focus on the BSDE appearing in the system (2.3), namely

$$
\begin{cases}
Y^{t,\phi}(s) = h(X^{t,\phi}) + \int_s^T f\big(r, X^{t,\phi}, Y^{t,\phi}(r), Z^{t,\phi}(r), \tilde{U}^{t,\phi}(r), Y_r^{t,\phi}\big)dr \\[2mm]
\qquad - \int_s^T Z^{t,\phi}(r)dW(r) - \int_s^T \int_{\mathbb{R}_0} U^{t,\phi}(r,z)\tilde{N}(dr,dz)\,, \qquad s \in [t,T] \quad (2.5) \\[2mm]
Y^{t,\phi}(s) = Y^{s,\phi}(s)\,, \qquad Z^{t,\phi}(s) = U^{t,\phi}(s,z) = 0\,, \qquad\qquad s \in [0,t]\,.
\end{cases}
$$

for a finite time horizon $T < \infty$ and $\phi \in \Lambda := D\big([0,T];\mathbb{R}^d\big)$. The path-dependent process $X^{t,\phi}$ represents the solution of the forward SDE with jumps of Eq. (2.4), while $\tilde{N}$ models the compensated Poisson random measure described in Eq. (2.1) and W is a l-dimensional Brownian motion.

We recall that, when we fix the delay term δ, the notation $Y_r^{t,\phi}$ stands for the path of the process restricted to $[r-\delta, r]$, according to Eq. (1.8). Notice that the terminal condition enforced by h depends on the solution of the forward SDE (2.4) as well as the solution (Y, Z, U) of the backward component considered in the time interval $[t,T]$.

Differently from the framework studied by Delong in [15], we consider a general initial time $s \in [0,t)$. As highlighted in [11], the Feynman–Kac formula would fail with standard prolongation.

Thus, an additional initial condition has to be satisfied over the interval $[0,t]$, given by

$$
Y^{t,\phi}(s) = Y^{s,\phi}(s)\,, \qquad s \in [0,t)\,. \tag{2.6}
$$

We remark that the supplementary initial condition stated in Eq. (2.6) represents one of the main differences between Theorem 3.3 and Theorem 14.1.1 in [15].

3. The well posedness of the BSDE

Concerning the delayed backward SDE (2.5), we will assume the following to hold.

Assumption 3.1. Let $f : [0,T] \times \Lambda \times \mathbb{R} \times \mathbb{R}^l \times \mathbb{R} \times L^2\left([-\delta,0];\mathbb{R}\right) \to \mathbb{R}$, $h : \Lambda \to \mathbb{R}$ and $\lambda : \mathbb{R}_0 \to \mathbb{R}_+$ (introduced in (1.9)) such that the following holds:

(A_4) There exist $L, K, M > 0$, $p \geq 1$ and a probability measure α on $\mathcal{B}\left([-\delta,0]\right)$ such that, for any $t \in [0,T]$, $\phi \in \Lambda$, $(y,z,u), (y',z',u') \in \mathbb{R} \times \mathbb{R}^l \times \mathbb{R}$ and $\hat{y}, \hat{y}' \in L^2\left([-\delta,0];\mathbb{R}\right)$, we have

(i) $\phi \mapsto f\left(t,\phi,y,z,u,\hat{y}\right)$ is continuous,

(ii) $|f(t,\phi,y,z,u,\hat{y}) - f(t,\phi,y',z',u',\hat{y})| \leq L\left(|y-y'| + |z-z'| + |u-u'|\right)$,

(iii) $|f(t,\phi,y,z,u,\hat{y}) - f(t,\phi,y,z,u,\hat{y}')|^2 \leq K \int_{-\delta}^0 \left(|\hat{y}(\theta) - \hat{y}'(\theta)|^2\right) \alpha(d\theta)$,

(iv) $|f\left(t,\phi,0,0,0,0\right)| < M(1 + \|\phi\|_T^p)$.

(A_5) The function f is non-anticipative.

(A_6) The function h is continuous and $|h(\phi)| \leq M(1 + \|\phi\|_T^p)$, for all $\phi \in \Lambda$.

(A_7) The function λ is measurable and $\lambda(z) \leq M(1 \wedge |z|)$, for all $z \in \mathbb{R}_0$.

The following remark generalizes a classical result, see, e.g., Theorem 14.1.1 in [15], Theorem 2.1 in [16] or Theorem 2.1 in [17].

Remark 3.2. In order to show both the existence and uniqueness of a solution to the backward part of the system (1.7) and to obtain the continuity of $Y^{t,\phi}$ with respect to ϕ, we need to impose K or δ to be small enough. More precisely, we will assume that there exists a constant $\chi \in (0,1)$, such that:

$$K \frac{\chi e^{\left(\chi + \frac{6L^2}{\chi}\right)\delta}}{(1-\chi)L^2} \max\{1, T\} < \frac{1}{578}. \tag{3.1}$$

The main difference between our result and Theorem 3.4 in [11] relies upon the presence of a jump component in the dynamics of the unknown process $Y^{t,\phi}$: this further term implies a stronger bound in the condition enforced in Eq. (3.1).

Hence, if K or δ are small enough to satisfy the condition stated in Eq. (3.1), then there exists a unique solution of (2.5) and the following theorem holds

Theorem 3.3. *Let assumptions (A_1)–(A_7) hold. If condition (3.1) is satisfied, then there exists a unique solution $(Y^{t,\phi}, Z^{t,\phi}, U^{t,\phi})_{(t,\phi)\in[0,T]\times\Lambda}$ of the BSDE (2.5) such that $(Y^{t,\phi}, Z^{t,\phi}, U^{t,\phi}) \in \mathbb{S}_t^2(\mathbb{R}) \times \mathbb{H}_t^2(\mathbb{R}^l) \times \mathbb{H}_{t,\nu}^2(\mathbb{R})$ for all $t \in [0,T]$ and the application $t \mapsto (Y^{t,\phi}, Z^{t,\phi}, U^{t,\phi})$ is continuous from $[0,T]$ into $\mathbb{S}_0^2(\mathbb{R}) \times \mathbb{H}_0^2(\mathbb{R}^l) \times \mathbb{H}_{0,\nu}^2(\mathbb{R})$.*

The proof of Theorem 3.3 is provided in "Appendix 7" and it is mainly based on the Banach fixed point theorem.

We emphasize that similar results hold also for multi-valued processes, namely $Y : \Omega \times [0,T] \to \mathbb{R}^m$, $Z : \Omega \times [0,T] \to \mathbb{R}^{m\times l}$ and $U : \Omega \times [0,T] \times (\mathbb{R}^n \backslash \{0\}) \to \mathbb{R}^m$. Further difficulties may arise, due to the presence of correlation between the different components of $Y \in \mathbb{S}_t^2(\mathbb{R}^m)$ or the necessity of introducing the *n-fold iterated stochastic integral*, see [5,7] or [12, Sec. 2.1] for further details.

4. The Feynman–Kac formula

In what follows we prove that the solution of Eq. (1.7), namely the path-dependent forward-backward system with delayed generator f and driven by a Lévy process, can be connected to the solution of path-dependent PIDE represented by the nonlinear Kolomogorov equation (1.1).

4.1. The Delfour–Mitter space

According to [2], we need the solution of the forward SDE (1.7) to be a Markov process to derive the Feynman–Kac formula. The Markov property of the solution is fully known for the SDE without jumps, i.e. when $\gamma = 0$, see [29] (Th. III. 1.1). Moreover, the Markov property also holds, by enlarging the state space, for the solution in a setting analogous to that of Eq. (2.4), see [12] (Prop. 2.6) where driving noises with independent increments are considered. Since $X^{t,\phi} : \Omega \times [0,T] \times D([0,T]; \mathbb{R}^d) \to \mathbb{R}^d$ in not Markovian, we enlarge the state space by considering the process X as a *process of the path*, by introducing a suitable Hilbert space, as described in [22,28], where they present a product-space reformulation of (2.4) splitting the present state $X(t)$ from the past trajectory X_t by a particular choice of the state space. Accordingly, we enlarge the state space of our interest, starting from paths defined on the Skorohod space $D\left([0,T]; \mathbb{R}^d\right)$ to then consider a *new* functional space, the so-called *Delfour–Mitter space* $M^2 := L^2([-T,0]; \mathbb{R}^d) \times \mathbb{R}^d$, by exploiting the continuous embedding of $D\left([-T,0]; \mathbb{R}^d\right)$ into $L^2([-T,0]; \mathbb{R}^d)$, as in-depth analyzed in, e.g., [2,12].

It is worth mentioning that M^2 has a Hilbert space structure, endowed with the following scalar product

$$\langle \phi, \psi \rangle_{M^2} = \langle \phi, \psi \rangle_{L^2} + \phi(0) \cdot \psi(0),$$

with associated norm

$$||\phi||^2_{M^2} = ||\phi||_{L^2} + |\phi(0)|^2,$$

where $\cdot$ and $|\cdot|$ stand for the scalar product in $\mathbb{R}^d$, resp. for the Euclidean norm in $\mathbb{R}^d$, while $\langle \cdot, \cdot \rangle_{L^2}$, resp. $||\cdot||_{L^2}$, indicates the scalar product, resp. the norm in $L^2 := L^2([-T,0]; \mathbb{R}^d)$.

For $t \in [0,T]$, $\phi \in \Lambda$ and $(\varphi, x) \in M^2$, let us set:

- compatible initial conditions $x^{t,\phi} := \phi(t)$ and $\eta^{t,\phi} \in D([-T,0]; \mathbb{R}^d)$ defined by

$$\eta^{t,\phi}(\theta) := \begin{cases} \phi(t+\theta), & \theta \in [-t,0]; \\ \phi(0), & \theta \in [-T,t); \end{cases} \tag{4.1}$$

- $(\varphi, x)^t \in \Lambda$ defined by

$$(\varphi, x)^t(\theta) := \begin{cases} \varphi(\theta - t), & \varphi \in D([-T,0]; \mathbb{R}^d), \ \theta \in [0,t); \\ x, & \varphi \in D([-T,0]; \mathbb{R}^d), \ \theta \in [t,T]; \\ 0, & \varphi \notin D([-T,0]; \mathbb{R}^d); \end{cases} \tag{4.2}$$

- $\tilde{b} : [0,T] \times M^2 \to \mathbb{R}^d$, $\tilde{\sigma} : [0,T] \times M^2 \to \mathbb{R}^{d \times l}$, $\tilde{\gamma} : [0,T] \times M^2 \times \mathbb{R} \to \mathbb{R}^d$ defined by

$$\begin{aligned} \tilde{b}(t, \varphi, x) &:= b(t, (\varphi, x)^t), \\ \tilde{\sigma}(t, \varphi, x) &:= \sigma(t, (\varphi, x)^t), \\ \tilde{\gamma}(t, \varphi, x, z) &:= \gamma(t, (\varphi, x)^t, z), \end{aligned} \tag{4.3}$$

where $b : [0,T] \times \Lambda \to \mathbb{R}^d$, $\sigma : [0,T] \times \Lambda \to \mathbb{R}^{d \times l}$ and $\gamma : [0,T] \times \Lambda \times \mathbb{R} \to \mathbb{R}^d$ are the given coefficients of Eq. (2.4).

We emphasize that $(\varphi, x)^t$ is well defined since it does not depend on the choice of the representative in the class of $\varphi \in L^2$, and the continuous embedding $D\left([-T,0];\mathbb{R}^d\right) \subset M^2$ is also injective.

We can now rewrite the forward equation (2.4) in the M^2-setting:

$$\begin{cases} X^{t,\eta,x}(s) = x + \int_t^s \tilde{b}(r, X_r^{t,\eta,x}, X^{t,\eta,x}(r))dr + \int_t^s \tilde{\sigma}(r, X_r^{t,\eta,x}, X^{t,\eta,x}(r))dW(r) \\ \qquad + \int_t^s \int_{\mathbb{R}_0} \tilde{\gamma}(r, X_r^{t,\eta,x}, X^{t,\eta,x}(r), z)\tilde{N}(dr,dz), \qquad s \in [t,T] \\ \left(X_t^{t,\eta,x}, X^{t,\eta,x}(t)\right) = (\eta, x). \end{cases} \qquad (4.4)$$

Here, for a process X, the term X_r means $(X((r+\theta)^+))_{\theta \in [-T,0]}$, which is slightly different from the delayed term introduced in (1.8). Since this notation occurs only in the context of the M^2-setting for the forward equation, the reader can clearly distinguish between the two uses.

The link between the solution of equation Eq. (2.4) and that of equation Eq. (4.4) is then provided by

$$X^{t,\phi} = X^{t,\eta^{t,\phi},x^{t,\phi}}, \forall (t,\phi) \in [0,T] \times \Lambda. \qquad (4.5)$$

In order to take advantage of the Feynman–Kac formula already derived in [12] in the case of non-delayed but still path-dependent BSDE, we have to additionally impose that b, σ, γ, f and h are locally Lipschitz-continuous with respect to $\phi \in \Lambda$ in the L^2-norm. Thus, in order to have the same regularity for the solution of the BSDE system with forward Eq. (4.4), we require that the coefficients b, σ and γ to be locally Lipschitz.

Assumption 4.1. There exists $K \geq 0$ and $m \geq 0$ such that:

(A_8) for all $t \in [0,T]$ and for all $\phi_1, \phi_2 \in \Lambda$,

$$|b(t,\phi_1) - b(t,\phi_2)|^2 + |\sigma(t,\phi_1) - \sigma(t,\phi_2)|^2 + \int_{\mathbb{R}_0} |\gamma(t,\phi_1,z) - \gamma(t,\phi_2,z)|^2\nu(dz)$$
$$\leq K||\phi_1 - \phi_2||_{L^2}^2 (1 + ||\phi_1||_{L^2}^2 + ||\phi_2||_{L^2}^2);$$

(A_9) for all $t \in [0,T]$, $y \in \mathbb{R}$, $z \in \mathbb{R}^l$, $u \in \mathbb{R}$, $\hat{y} \in L^2([-\delta,0];\mathbb{R})$ and for all $\phi_1, \phi_2 \in \Lambda$,

$$|f(t,\phi_1,y,z,u,\hat{y}) - f(t,\phi_2,y,z,u,\hat{y})| \leq$$
$$K(1 + ||\phi_1||_{L^2} + ||\phi_2||_{L^2} + |y|)^m \cdot (1 + |z| + |u|)||\phi_1 - \phi_2||_{L^2}.$$

(A_{10}) for all $\phi_1, \phi_2 \in \Lambda$,

$$|h(\phi_1) - h(\phi_2)| \leq K(1 + ||\phi_1||_{L^2} + ||\phi_2||_{L^2})^m||\phi_1 - \phi_2||_{L^2}.$$

We remark that $\tilde{b}$, $\tilde{\sigma}$ and $\tilde{\gamma}$, as defined by (4.3) are not necessarily locally Lipschitz, since $D\left([0,T];\mathbb{R}^d\right)$ is dense in $L^2\left([0,T];\mathbb{R}^d\right)$, but they are set to constants outside $D\left([0,T];\mathbb{R}^d\right)$. However, one can define these coefficients first on $[0,T] \times D\left([0,T];\mathbb{R}^d\right) \times \mathbb{R}^d$ and then extend to $[0,T] \times M^2$ by density, using

the uniform continuity provided by assumption (A_8). It is this version of the above functions that we will use in the sequel.

Within this setting, lifting the state space turns out to be particularly convenient in order to investigate differentiability properties of the solution and relate the solution of Eq. (4.4) (combined with the backward equation) to the solutions of the non-linear Kolmogorov equation defined by Eq. (1.1) on $[0, T] \times \Lambda$.

Remark 4.2. It is also possible to work directly on the Skorohod space D. However, since D is not a separable Banach space, one has to consider weaker topologies on D, following a semi-group approach like the one developed by Peszat and Zabczyk in [36].

4.2. Main theorem

In what follows we provide the main result, namely a nonlinear version of the Feynman–Kac formula in the case where the process $X^{t,\phi}$ has jumps and the generator of the backward dynamic f depends on the past values of Y.

Theorem 4.3. (Feynman–Kac formula) *Under hypotheses (A_1)–(A_{10}) with condition* (3.1) *being verified, let* $(X^{t,\phi}, Y^{t,\phi}, Z^{t,\phi}, U^{t,\phi})_{(t,\phi)\in[0,T]\times\Lambda}$ *be the solution of the forward-backward system* (2.3).

Let $u : [0, T] \times \Lambda \to \mathbb{R}$ *be the deterministic function defined by*

$$u(t, \phi) = Y^{t,\phi}(t), \quad (t, \phi) \in [0, T] \times \Lambda. \tag{4.6}$$

Then u is a non-anticipative function and there exist constants $C > 0$ and $m \geq 0$ such that, for all $t \in [0, T]$ and $\phi_1, \phi_2 \in \Lambda$,

$$|u(t, \phi_1) - u(t, \phi_2)| \leq C(1 + ||\phi_1||_{L^2} + ||\phi_2||_{L^2})^m ||\phi_1 - \phi_2||_{L^2}.$$

Moreover, the following formula holds:

$$Y^{t,\phi}(s) = u(s, X^{t,\phi}), \ \forall s \in [0, T], \tag{4.7}$$

for any $(t, \phi) \in [0, T] \times \Lambda$.

To prove the representation formula (4.6), we adapt the proof of Theorem 4.10 of [11] by adding the contribution of U and $\tilde{U}$, respectively modelling the process and the integral term connected to the jump component.

Proof. We follow the Picard iteration scheme, hence considering the iterative process of the BSDE with a delayed generator driven by Lévy process described by

$$Y^{n+1,t,\phi}(s) = h(X^{t,\phi})$$

$$+ \int_s^T f\big(r, X^{t,\phi}, Y^{n+1,t,\phi}(r), Z^{n+1,t,\phi}(r), \tilde{U}^{n+1,t,\phi}(r), Y_r^{n,t,\phi}\big) dr$$

$$- \int_s^T Z^{n+1,t,\phi}(r) dW(r) - \int_t^T \int_{\mathbb{R}} U^{n+1,t,\phi}(r, z) \tilde{N}(dr, dz),$$

with $Y^{0,t,\phi} \equiv 0$, $Z^{0,t,\phi} \equiv 0$ and $U^{0,t,\phi} \equiv 0$ and initial condition

$$Y^{n+1,t,\phi}(s) = Y^{n+1,s,\phi}(s), \quad Z^{n+1,t,\phi}(s) = U^{n+1,t,\phi}(s, z) = 0, \quad s \in [0, t].$$

Let us suppose that there exists a non-anticipative functional $u_n : [0,T] \times \Lambda \to \mathbb{R}$ such that u_n is locally Lipschitz and $Y^{n,t,\phi}(s) = u_n(s, X^{t,\phi})$ for every $t, s \in [0,T]$ and $\phi \in \Lambda$.

Since $Y^{n,t,\phi}(r + \theta) = u_n(r + \theta, X^{t,\phi})$ if $r + \theta \geq 0$ and $Y^{n,t,\phi}(r + \theta) = Y^{n,t,\phi}(0) = u_n(0, X^{t,\phi})$ if $r + \theta < 0$, by defining

$$\tilde{u}_n(t, \phi) := \left(u_n((t + \theta)^+, \phi) \right)_{\theta \in [-\delta, 0]},$$

the delayed term reads

$$Y^{n,t,\phi}_r = \tilde{u}_n(r, X^{t,\phi})$$

and the above equation becomes

$$\begin{cases} Y^{n+1,t,\phi}(s) = h(X^{t,\phi}) \\ \quad + \int_s^T f\big(r, X^{t,\phi}, Y^{n+1,t,\phi}(r), Z^{n+1,t,\phi}(r), \tilde{U}^{n+1,t,\phi}(r), \tilde{u}_n(r, X^{t,\phi})\big) dr \\ \quad - \int_s^T Z^{n+1,t,\phi}(r) dW(r) - \int_s^T \int_{\mathbb{R} \setminus \{0\}} U^{n+1,t,\phi}(r, z) \tilde{N}(dr, dz), \quad s \in [t, T] \\ Y^{n+1,t,\phi}(s) = Y^{n+1,s,\phi}(s), \quad Z^{n+1,t,\phi}(s) = U^{n+1,t,\phi}(s, z) = 0, \quad s \in [0, t]. \end{cases}$$

For fixed n, we define $\psi_n : [0,T] \times D([-T,0]; \mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^l \times \mathbb{R} \to \mathbb{R}$ as

$$\psi_n(t, \varphi, x, y, z, u) := f(t, (\varphi, x)^t, y, z, u, \tilde{u}_n(t, (\varphi, x)^t)), \tag{4.8}$$

and $\tilde{h} : D([-T,0]; \mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}$ as

$$\tilde{h}(\varphi, x) := h((\varphi, x)^T). \tag{4.9}$$

Since u_n is locally Lipschitz-continuous, one can show that ψ_n and $\tilde{h}$ are also locally Lipschitz in φ. Therefore, one can extend them by density to $[0,T] \times M^2 \times \mathbb{R} \times \mathbb{R}^l \times \mathbb{R}$, respectively to M^2, keeping the locally Lipschitz property.

After setting $\eta = \eta^{t,\phi}$, $x = x^{t,\phi}$ (see (4.1)), we enlarge the state space and write the forward–backward system in an equivalent way as (for simplicity, we omit the dependence of n)

$$\begin{cases} dX^{\tau,\eta,x}(t) = \tilde{b}(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t)) dt + \tilde{\sigma}(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t)) dW(t) \\ \quad + \int_{\mathbb{R}_0} \tilde{\gamma}(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t), z) \tilde{N}(dt, dz) \\ (X_\tau^{\tau,\eta,x}, X^{\tau,\eta,x}(\tau)) = (\eta, x) \\ dY^{\tau,\eta,x}(t) = \psi_n \left(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t), Y^{\tau,\eta,x}(t), Z^{\tau,\eta,x}(t), \tilde{U}^{\tau,\eta,x}(t) \right) dt \\ \quad + Z^{\tau,\eta,x}(t) dW(t) + \int_{\mathbb{R}_0} U^{\tau,\eta,x}(t, z) \tilde{N}(dt, dz) \\ Y^{\tau,\eta,x}(T) = \tilde{h}(X_T^{\tau,\eta,x}, X^{\tau,\eta,x}(T)). \end{cases}$$

It is easy to see that the two sets of solutions, $(X^{t,\eta,x}, Y^{t,\eta,x}, Z^{t,\eta,x}, U^{t,\eta,x})$ and $(X^{t,\phi}, Y^{n+1,t,\phi}, Z^{n+1,t,\phi}, U^{n+1,t,\phi})$ coincide.

We have already observed that the extended coefficients $\tilde{b}$, $\tilde{\sigma}$, $\tilde{\gamma}$ satisfy conditions (A1)–(A2) in [12, pp. 8–9]; in a similar manner, ψ_n and $\tilde{h}$ satisfy (B1) and (B2) in [12, p. 24]. Then we can exploit Theorem 4.5 in [12] and infer

that there exists a locally Lipschitz function $\bar{u}_n : [0, T] \times M^2 \to \mathbb{R}$ such that the following representation formula holds:

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

Let us define the following locally Lipschitz, non-anticipative functional $u_{n+1} : [0, T] \times \Lambda \to \mathbb{R}$ by

$$u_{n+1}(t, \phi) := \bar{u}_n(t, \eta^{t,\phi}, \phi(t)),$$

where the time shifting $\eta^{t,\phi}$ is defined according to (4.1). Then

$$Y^{n+1,t,\phi}(s) = u_{n+1}(s, X^{t,\phi}), \ \forall s \in [0, T].$$

Notice that $\left(Y^{n,t,\phi}, Z^{n,t,\phi}, U^{n,t,\phi}\right)$ is the Picard iterative sequence needed to construct the solution $\left(Y^{t,\phi}, Z^{t,\phi}, U^{t,\phi}\right)$:

$$\left(Y^{n+1,\cdot,\phi}, Z^{n+1,\cdot,\phi}, U^{n+1,\cdot,\phi}\right) = \Gamma(Y^{n,\cdot,\phi}, Z^{n,\cdot,\phi}, U^{n,\cdot,\phi}),$$

where Γ is the contraction defined in the proof of Theorem 3.3. By applying Theorem 3.3, we then have:

$$\lim_{n\to\infty} \mathbb{E}\Big(\sup_{s\in[0,T]} \big|Y^{n,t,\phi}(s) - Y^{t,\phi}(s)\big|^2 \Big) = 0.$$

Of course, $u_n(t, \phi)$ converges to $u(t, \phi) := \mathbb{E}\left[Y^{t,\phi}(t)\right]$, for every $t \in [0, T]$ and $\phi \in \Lambda$, hence implying that the nonlinear Feynman–Kac formula $Y^{t,\phi}(s) = u(s, X^{t,\phi})$ holds. From its definition, it is clear that u is non-anticipative. Regarding the locally Lipschitz property, it can be proven by applying Itô's formula to $|Y^{t,\phi_1} - Y^{t,\phi_2}|^2$ and resorting to standard calculus. □

5. Mild solution of the Kolmogorov equation

In this section, we prove the existence of a mild solution of the path-dependent partial integro-differential equation (PPIDE) Kolmorogov equation (1.1) showing a dependence both on a delayed term and on integral term modelling jumps.

Let us start by recalling from [12] the notion of the *Markov transition semigroup* corresponding to the operator $\mathcal{L}$ introduced by (1.4). From [12, Prop. 2.6], we know that the strong solution $X^{0,\eta,x} \in M^2$ of Eq. (4.4) is a *Markov process* in the sense that $\mathbb{P}$-a.s.,

$$\mathbb{P}((X_t^{0,\eta,x}, X^{0,\eta,x}(t)) \in B | \mathcal{F}_s) = \mathbb{P}((X_t^{0,\eta,x}, X^{0,\eta,x}(t)) \in B | (X_s^{0,\eta,x}, X^{0,\eta,x}(s))),$$

for all $(\eta, x) \in M^2$ and all Borel sets B of M^2, the proof of this fundamental result being given in several works, as, e.g. [2, Th. 3.9], [37, Prop. 3.3] or [36, Sec. 9.6]; for more details, see, e.g., [23].

Denoting by $B_p(S)$ the space of Borel functions with at most polynomial growth on a metric space S, the transition semigroup $\tilde{P}_{t,s}$, acting on $B_p(M^2)$ is then defined by

$$\tilde{P}_{t,s}[\varphi](\eta, x) := \mathbb{E}[\varphi(X_s^{t,\eta,x}, X^{t,\eta,x}(s))], \ \varphi \in B_p(M^2), \ (\eta, x) \in M^2.$$

Coming back to our setting, we define $P_{t,s} : B_p(\Lambda) \to B_p(\Lambda)$ by

$$P_{t,s}[\varphi](\phi) := \mathbb{E}[\varphi(X^{t,\phi}(\cdot \wedge s))], \ \varphi \in B_p(\Lambda), \ \phi \in \Lambda.$$

Obviously (see (4.1) for the notations), for $\varphi \in B_p(\Lambda)$ and $\phi \in \Lambda$,

$$P_{t,s}[\varphi](\phi) = \tilde{P}_{t,s}[\varphi](\eta^{t,\phi}, x^{t,\phi}). \tag{5.1}$$

In order to introduce the notion of *mild solution*, we will need to define the *generalized directional gradient* of a function $u : [0,T] \times \Lambda \to \mathbb{R}$, following the approach in [23], also described in [12]. Suppose that the function u satisfies the following locally Lipschitz-continuity condition:

$$|u(t, \phi_1) - u(t, \phi_2)| \le C\|\phi_1 - \phi_2\|_{L^2}(1 + \|\phi_1\|_{L^2} + \|\phi_2\|_{L^2})^m \,; \\ |u(t,0)| \le C \,; \tag{5.2}$$

Since the procedure of defining the generalized directional gradient takes place in Hilbert spaces, we will have to appeal again to the M^2-lifting, as we have already done for the coefficients of the forward equation. We define first

$$v(t, \varphi, x) := u(t, (\varphi, x)^t)$$

for $(t, \varphi, x) \in [0,T] \times D([-T,0]; \mathbb{R}^d) \times \mathbb{R}^d$ and then extend it to $[0,T] \times M^2$ by density. Again, this is possible due to condition (5.2). Then, according to [23, Th. 3.1], there exists a Borel function $\zeta : [0,T] \times M^2 \to \mathbb{R}^d$ such that

$$\langle v(\cdot, X^{t,\eta,x}_{\cdot}, X^{t,\eta,x}(\cdot)), W(\cdot)\rangle_{[t,\tau]} = \int_t^\tau \zeta(s, X^{t,\eta,x}_s, X^{t,\eta,x}(\cdot))ds, \tag{5.3}$$

for any $0 \le t \le \tau \le T$ and $(\eta, x) \in M^2$, where $\langle X, Y \rangle_{[t,\tau]}$ denotes the *joint quadratic variation* of a pair of real stochastic processes $(X(t), Y(t))_{t \in [0,T]}$ on the interval $[t, \tau]$,

$$\langle X, Y \rangle_{[t,\tau]} := \lim_{\epsilon \searrow 0} \frac{1}{\epsilon} \int_t^\tau (X(s+\epsilon) - X(s)) \cdot (Y(s+\epsilon) - Y(s))ds,$$

with the limit taken in probability and $X(s), Y(s)$ defined as $X(s)$, respectively $Y(s)$ for $s > T$.

The set of all functions ζ with the above property is called the *generalized directional gradient* of v and is denoted $\nabla^\sigma v$. Its name and notation come from the observation that if v (and, consequently u) and the coefficients of the forward equation are sufficiently regular, then $\nabla_{(\eta,x)} v \cdot \tilde{\sigma} \in \nabla^\sigma v$ (see [23, Remark 3.3]).

We come back to the Λ-setting and define $\nabla^\sigma u$, the generalized directional gradient of u as the set of all non-anticipative functions $\xi : [0,T] \times \Lambda \to \mathbb{R}$ such that there exists $\zeta \in \nabla^\sigma v$ satisfying

$$\zeta(t, \varphi, x) = \xi(t, (\varphi, x)^t)$$

for all $(t, \varphi, x) \in [0,T] \times D([-T,0]; \mathbb{R}^d) \times \mathbb{R}^d$. Such a function can be defined by setting $\xi(t, \phi) := \zeta(t, \eta^{t,\phi}, x^{t,\phi})$.

It is clear that relation (5.3) translates into

$$\langle u(\cdot, X^{t,\phi}), W(\cdot)\rangle_{[t,\tau]} = \int_t^\tau \xi(s, X^{t,\phi})ds,$$

for all $(t, \phi) \in [0, T] \times \Lambda$ and all $\tau \in [t, T]$, if $\xi \in \nabla^\sigma u$. From this relation, it is clear (see also Remark 3.2 in [23]) that if $\xi, \hat{\xi} \in \hat{\nabla}^\sigma u$, then

$$\xi(s, X^{t,\phi}) = \hat{\xi}(s, X^{t,\phi}) \ ds\text{-a.e.}, \ \mathbb{P}\text{-a.s.},$$

so there is no ambiguity if we write $\nabla^\sigma u(s, X^{t,\phi})$.

We have now all the ingredients for introducing the notion of a mild solution.

Definition 5.1. A non-anticipative function $u : [0, T] \times \Lambda \to \mathbb{R}$ is a mild solution to Eq. (1.1) if u satisfies (5.2) and the following equality holds true for all $(t, \phi) \in [0, T] \times \Lambda$ and $\xi \in \nabla^\sigma u$:

$$u(t, \phi) = P_{t,T}[h](\phi)$$
$$+ \int_t^T P_{t,s} \left[f\left(\cdot, u(s, \cdot), \xi(s, \cdot), \mathcal{J}u(s, \cdot)\right), (u(\cdot, \cdot))_s \right](\phi)ds, \quad (5.4)$$

where $(u(\cdot, \cdot))_s$ is the delayed term defined in (1.5).

Given the definition of $P_{t,T}$ and the above mentioned remark, we can reformulate relation (5.4) as

$$u(t, \phi) = \mathbb{E}[h(X_T^{t,\phi})]$$
$$+ \mathbb{E}\left[\int_t^T f\left(X^{t,\phi}, u(s, X^{t,\phi}), \nabla^\sigma u(s, X^{t,\phi}), \mathcal{J}u(s, X^{t,\phi}), (u(\cdot, X^{t,\phi}))_s \right) ds \right].$$
$$(5.5)$$

The next theorem represents the core result of this section.

Theorem 5.2. (Existence and uniqueness) *Let assumptions (A_1)–(A_{10}) and (3.1) hold true. Then the function u defined by (4.6) is the unique mild solution to the path-dependent partial integro-differential Eq. (1.1).*

Proof. Existence. Let us consider the backward component of the FBSDE described in Eq. (2.3) for $s \in [t, T]$

$$Y^{t,\phi}(s) = h(X^{t,\phi}) + \int_s^T f\big(r, X^{t,\phi}, Y^{t,\phi}(r), Z^{t,\phi}(r), \tilde{U}^{t,\phi}(r), Y_r^{t,\phi}\big)dr$$
$$- \int_s^T Z^{t,\phi}(r)dW(r) - \int_s^T \int_{\mathbb{R}\backslash\{0\}} U^{t,\phi}(r, z)\tilde{N}(dr, dz), \quad s \in [t, T].$$
$$(5.6)$$

By Theorem 4.3, the non-anticipative function u defined by (4.6) is satisfying (5.2) (the second part comes the continuity of $t \mapsto Y^{t,\phi}$, asserted in Theorem 3.3) and the representation formula (4.7) holds. Moreover, by means of Eq. (4.6), we can write the delayed term Y_r as a function of the path of solution of the forward dynamic $X^{t,\phi}$ and, thus, by defining

$$\tilde{u}(t, \phi) := \big(u((t + \theta)^+, \phi)\big)_{\theta \in [-\delta, 0]},$$

we can rewrite Eq. (5.6), leading to

$$
\begin{cases}
Y^{t,\phi}(s) = h(X^{t,\phi}) + \displaystyle\int_s^T f\big(r, X^{t,\phi}, Y^{t,\phi}(r), Z^{t,\phi}(r), \tilde{U}^{t,\phi}(r), \tilde{u}(r, X^{t,\phi})\big) dr \\
\qquad - \displaystyle\int_s^T Z^{t,\phi}(r) dW(r) - \int_s^T \int_{\mathbb{R}\setminus\{0\}} U^{t,\phi}(r,z) \tilde{N}(dr, dz), \quad s \in [t, T] \\
Y^{t,\phi}(s) = Y^{s,\phi}(s), \qquad Z^{t,\phi}(s) = U^{t,\phi}(s,z) = 0, \qquad s \in [0, t],
\end{cases}
\tag{5.7}
$$

At this point, we enlarge the state space going through M^2 coefficients analogously to the proof of Theorem 3.3, obtaining

$$
\begin{aligned}
Y^{t,\eta,x}(s) = \;& \tilde{h}(X_T^{s,\eta,x}, X^{s,\eta,x}(T)) \\
& + \int_s^T \psi\big(r, X_r^{t,\eta,x}, X^{t,\eta,x}(r), Y^{t,\eta,x}(r), Z^{t,\eta,x}(r), \tilde{U}^{t,\eta,x}(r)\big) dr \\
& - \int_s^T Z^{t,\eta,x}(r) dW(r) - \int_t^T \int_{\mathbb{R}_0} U^{t,\eta,x}(r,z) \tilde{N}(dr, dz),
\end{aligned}
\tag{5.8}
$$

where the map $\tilde{h}$ was defined in Eq. (4.9) and ψ is defined similarly to ψ_n introduced in Eq. (4.8):

$$
\psi(t, \varphi, x, y, z, u) := f(t, (\varphi, x)^t, y, z, u, \tilde{u}(t, (\varphi, x)^t))
$$

on $[0, T] \times D([-T, 0]; \mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^l \times \mathbb{R}$ and extending it by density. Since $\tilde{h}$ and ψ are locally Lipschitz, hence satisfying the conditions required by Theorem 4.8 in [12], we can apply this result in order to conclude that the function $v : [0, T] \times M^2 \to \mathbb{R}$ defined by

$$
v(t, \eta, x) := Y^{t,\eta,x}(t)
$$

is a mild solution (in the sense of [12], but similar to ours) of the following PPIDE

$$
\begin{cases}
-\partial_t v(t, \eta, x) - \tilde{\mathcal{L}} v(t, \eta, x) \\
\qquad - \psi\big(t, \eta, x, v(t, \eta, x), \partial_x v(t, \eta, x) \tilde{\sigma}(t, \eta, x), \tilde{\mathcal{J}} v(t, \eta, x)\big) = 0, \\
v(T, \eta, x) = \tilde{h}(\eta, x), \qquad (t, \eta, x) \in [0, T] \times M^2,
\end{cases}
$$

where $\tilde{\mathcal{L}}$ and $\tilde{\mathcal{J}}$ are the straightforward modifications of $\mathcal{L}$, respectively $\mathcal{J}$ in M^2. Since $v(t, \phi, x) = u(t, (\varphi, x)^t)$ for any $(t, \phi, x) \in [0, T] \times D([-T, 0]; \mathbb{R}^d)$, by playing on the relation (5.1) and the connection between the formulations of the generalized directional gradient in the càdlàg, respectively M^2 cases, it is straightforward to show that u is a mild solution of (1.1).

Uniqueness. Let us take two mild solutions u^1 and u^2 of the path-dependent PDE (1.1). We define

$$
f^i(t, \phi, y, z, w) := f(t, \phi, y, z, w, \big(u^i(\cdot, \phi)\big)_t), \quad i = \overline{1, 2}.
$$

Using these drivers we can consider the following BSDEs:

$$Y^{t,\phi}(s) = h(X^{t,\phi}) + \int_s^T f^i(r, X^{t,\phi}, Y^{t,\phi}(r), Z^{t,\phi}(r), \tilde{U}^{t,\phi}(r))dr$$

$$- \int_s^T Z^{t,\phi}(r) dW(r) - \int_s^T \int_{\mathbb{R}_0} U^{t,\phi}(r,z)\tilde{N}(dr,dz), \quad i = \overline{1,2}, \tag{5.9}$$

for which there exist unique solutions $\left(Y^{i,t,\phi}, Z^{i,t,\phi}, U^{i,t,\phi}\right) \in \mathbb{S}_t^2(\mathbb{R}) \times \mathbb{H}_t^2(\mathbb{R}^l) \times \mathbb{H}_{t,\nu}^2(\mathbb{R})$ for $i = \overline{1,2}$.

   By Theorem 4.3 we see that

$$Y^{i,t,\phi}(s) = v^i(s, X^{t,\phi}), \quad \text{for all } s \in [0,T], \quad \text{a.s.},$$

for any $(t, \phi) \in [0, T] \times \Lambda$, where $v^i : [0, T] \times \Lambda \to \mathbb{R}$, $i = \overline{1,2}$ are defined by

$$v^i(t, \phi) := Y^{i,t,\phi}(t), \ (t, \phi) \in [0, T] \times \Lambda.$$

Hence, by the existence part, we obtain that the functions v^i are solutions of the PDE of type (1.1), but without the delayed terms $\left(v^i(\cdot, \phi)\right)_t$:

$$\begin{cases} -\partial_t v^i(t, \phi) - \mathcal{L}v^i(t, \phi) - f^i(t, \phi, v^i(t, \phi), \partial_x v^i(t, \phi) \sigma(t, \phi), \mathcal{J}v^i(t, \phi)) = 0, \\ v^i(T, \phi) = h(\phi), \quad i = \overline{1,2}. \end{cases} \tag{5.10}$$

Since u^i is also solution to equation (5.10), by using the uniqueness part of Theorem 4.5 from [12] we get that (after embedding these equations in M^2, as we did in the previous part)

$$u^i(t, \phi) = v^i(t, \phi), \quad (t, \phi) \in [0, T] \times \Lambda, \quad i = \overline{1,2}.$$

Hence

$$Y^{i,t,\phi}(s) = v^i(s, X^{t,\phi}) = u^i(s, X^{t,\phi}), \ i = \overline{1,2},$$

so BSDEs (5.9) become a single equation,

$$Y^{i,t,\phi}(s) = h(X^{t,\phi}) + \int_s^T f(r, X^{t,\phi}, Y^{i,t,\phi}(r), Z^{i,t,\phi}(r), \tilde{U}^{i,t,\phi}(r), Y_r^{i,t,\phi})dr$$

$$- \int_s^T Z^{i,t,\phi}(r) dW(r) - \int_s^T \int_{\mathbb{R}_0} U^{t,\phi}(r,z)\tilde{N}(dr,dz), \tag{5.11}$$

with $i = \overline{1,2}$, for which we have uniqueness from Theorem 3.3.

   Therefore $Y^{1,t,\phi} = Y^{2,t,\phi}$ and, consequently

$$u^1(t, \phi) = Y^{1,t,\phi}(t) = Y^{2,t,\phi}(t) = u^2(t, \phi).$$

$\square$

Remark 5.3. From Theorem 4.5 in [12], besides relation (4.7), it can also be inferred that for every $(t, \phi) \in [0, T] \times \Lambda$ the following representation formulas hold, $\mathbb{P}$-a.s. and for a.e. $s \in [t, T]$:

$$Z^{t,\phi}(s) = \nabla^\sigma u(s, X^{t,\phi});$$

$$U^{t,\phi}(s, z) = u\left(s, (X^{t,\phi})^{t, \gamma(s, X^{t,\phi}, z)}\right) - u(s, X^{t,\phi}),$$

where we use the notation introduced in (1.3). These formulas could have been used to prove directly the existence part of the above result, by taking the expectation in equation (5.7) and using relation (5.5).

6. Financial application

In this section, we provide a financial application moving from the model studied in, e.g. [13], or [19]. We consider a generalization of the so-called Large Investor Problem, where a *large investor* wishes to invest in a given market, buying or selling a stock. The investor has the peculiarity that his actions on the market can affect the stock price. We refer to Example 14.1 in [15] for a detailed example of the time-delayed setting for the large investor problem.

6.1. A perfect replication problem for a large investor

Concerning the problem of the perfect replication strategy for a large investor, we generalize Example 14.1 in [15] by asking, in addition to path-dependent coefficients, the dynamic of the risky asset driven by a Poisson random measure. Similar results are also presented in [11,19].

We denote the investor's strategy by π and the investment portfolio by X^π and we assume that its past X_r^π may affect directly the stock coefficients μ, σ and γ and the bond rate. Consequently, we consider the following dynamic

$$\begin{cases} \dfrac{dS_0(t)}{S_0(t)} = r(t, X^\pi(t), X_t^\pi)dt, \\ S_0(0) = 1, \\ \dfrac{dS_i(t)}{S_i(t)} = \mu_i\left(t, X^\pi(t), X_t^\pi\right)dt + \sigma_i(t, X^\pi(t), X_t^\pi)dW(t) \\ \qquad\qquad + \displaystyle\int_{\mathbb{R}_0} \gamma_i(t, X^\pi(t), X_t^\pi, z)\tilde{N}(dt, dz), \\ S_i(0) = s_i > 0, \end{cases} \tag{6.1}$$

where r_i, μ_i, σ_i and γ_i, $i = \overline{1,2}$ are $\mathbb{F}^{W,\tilde{N}}$-predictable processes, $\mathbb{F}^{W,\tilde{N}}$ being the natural filtration associated to the Brownian motion W and to the Poisson random measure $\tilde{N}$, with compensator defined according to Eq. (2.1).

The total amount of the portfolio of the large investor is described by

$$dX^\pi(t) = \pi_1(t)\frac{dS_1(t)}{S_1(t)} + \pi_2(t)\frac{dS_2(t)}{S_2(t)} + (X^\pi(t) - \pi_1(t) - \pi_2(t))\frac{dS_0(t)}{S_0(t)}dt\,.$$

where, at any time $t \in [0, T]$, $\pi_i(t)$ represents the amount invested in the risky asset S_i, while $X^\pi(t) - \pi_1(t) - \pi_2(t)$ is the amount invested in the riskless one S_0.

Let us denote, for simplicity, $\phi := (\phi_1, \phi_2)$ for ϕ among the symbols S, π, μ, σ or γ and by $\cdot$ the scalar product in $\mathbb{R}^2$. The goal is to find an admissible replicating strategy $\pi \in \mathcal{A}$ for a claim $h(S(T))$.

We have that the portfolio X evolves according to

$$dX^\pi(t) = \frac{\pi_1(t)}{S_1(t)}dS_1(t) + \frac{\pi_2(t)}{S_2(t)}dS_2(t) + \frac{X^\pi(t) - \pi(t)}{S_0(t)}dS_0(t)$$

$$= \pi(t) \cdot \left[\mu(t, X^\pi, X_t^\pi) dt + \sigma(t, X^\pi, X_t^\pi)dW(t) + \int_{\mathbb{R}_0} \gamma(t, X^\pi, X_t^\pi, z)\tilde{N}(dt, dz)\right]$$

$$+ [X^\pi(t) - \pi_1(t) - \pi_2(t)] r(t, X^\pi(t), X_r^\pi)dt,$$

Hence, for $t \in [0, T]$, we have

$$X^\pi(t) = h(S(T)) + \int_t^T F(s, X^\pi(s), X_s^\pi, \pi(s)) \, ds - \int_t^T Z(s)dW(s)$$

$$- \int_t^T \int_{\mathbb{R}_0} U(s, z)\tilde{N}(ds, dz), \tag{6.2}$$

with final condition $X^\pi(T) = h(S)$ by denoting

$$F(s, X^\pi(s), X_s^\pi, \pi(s))$$
$$:= -r(s, X^\pi(s), X_s^\pi)(X^\pi(s) - \pi_1(s) - \pi_2(s)) - \pi(s) \cdot \mu(s, X^\pi(s), X_s^\pi);$$
$$Z^\pi(s) := \pi(s) \cdot \sigma(s, X^\pi(s), X_s^\pi); \tag{6.3}$$
$$U^\pi(s, z) := \pi(s) \cdot \gamma(s, X^\pi(s), X_s^\pi, z).$$

The generator F can be rewritten to accommodate the dependence of Z^π and U^π by introducing the following transformation

$$\bar{F}(s, X^\pi(s), X_s^\pi, Z^\pi(s), \tilde{U}^\pi(s)) = F(s, X^\pi(s), X_s^\pi, \pi(s)) \tag{6.4}$$

since from Eq. (6.3) we have

$$\begin{bmatrix} \pi_1(s) & \pi_2(s) \end{bmatrix} = \begin{bmatrix} Z^\pi(s) & U^\pi(s, z) \end{bmatrix} \begin{bmatrix} \sigma_1(s, X^\pi(s), X_s^\pi) & \gamma_1(s, X^\pi(s), X_s^\pi, z) \\ \sigma_2(s, X^\pi(s), X_s^\pi) & \gamma_2(s, X^\pi(s), X_s^\pi, z) \end{bmatrix}^{-1}$$

by imposing that the matrix $\begin{bmatrix} \sigma^\mathrm{T} & \gamma^\mathrm{T} \end{bmatrix}$ is invertible.

We then ask the coefficients r, μ, σ and γ to be such that the function $\bar{F} : [0, T] \times \mathbb{R} \times L^2([-\delta, 0]; \mathbb{R}) \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ satisfies assumptions (A_4), (A_5) and (A_9).

Furthermore, we need a couple of simplifying assumptions in order to fit the theoretical framework of the article.

First, we introduce a functional aiming at encoding the forward process in order to decouple the terminal condition in the BSDE (6.2) from the stock forward dynamic (6.1). The risky assets vector S can be explicitly written, by a variation of constants formula, as a functional of the random coefficients of the stock (W and $\tilde{N}$), on the wealth X and on the processes Z and U depending on the allocation strategy π. Hence $h(S(T))$ can be written as

$$h(S(T)) = \bar{h}(W, \tilde{N}, X^\pi, Z^\pi, U^\pi) \tag{6.5}$$

where $\bar{h}$ is asked to satisfy, besides conditions (A_6) and (A_{10}) in the first two coefficients, a Lipschitz condition in the last three with Lipschitz constant K_1.

Thus, we can rewrite (6.2) as

$$X(t) = \bar{h}\big(W, \tilde{N}, X, Z, U\big) + \int_t^T \bar{F}\big(s, X(s), X_s, Z(s), \tilde{U}(s)\big)ds - \int_t^T Z(s)\, dW(s)$$

$$- \int_t^T \int_{\mathbb{R}_0} U(s, z)\tilde{N}(ds, dz) \quad t \in [0, T] . \tag{6.6}$$

If we assume that $\delta = T$ and the Lipschitz constants K and K_1 are sufficiently small, then the time-delayed BSDE (6.6) has a unique solution. The setting is a little more general than ours, in the sense that the final condition also depends on the path of the solution of the backward equation (6.6), but one can use Theorem 14.1.1 in [15].

The further modelling concerns introducing the Markovian setting, i.e. allowing the initial time and value to vary. Hence, we introduce the forward processes $\bar{W}$ and $\bar{N}$ by

$$\begin{cases} (\bar{W}(s), \bar{N}(s)) = \left(\phi_1(t) + \int_t^s dW(r), \phi_2(t) + \int_t^s \int_{\mathbb{R}_0} \tilde{N}(dr, dz)\right), & s \in [t, T], \\ (\bar{W}(s), \bar{N}(s)) = (\phi_1(s), \phi_2(s)), & s \in [0, t), \end{cases} \tag{6.7}$$

with Brownian motion W, compensated Random measure $\tilde{N}$ and an initial càdlàg datum $\phi = (\phi_1, \phi_2) \in \Lambda$.

Finally, by recalling the coefficients $\bar{F}$, Z and U defined in Eq. (6.3) we may write the following decoupled forward–backward stochastic system:

$$\begin{cases} \bar{W}^{t,\phi}(s) = \phi_1(t) + \int_t^s dW(r), & s \in [t, T], \\ \bar{N}^{t,\phi}(s) = \phi_2(t) + \int_{\mathbb{R}_0}\int_t^s d\tilde{N}(r, z), & s \in [t, T], \\ \left(\bar{W}^{t,\phi}(s), \bar{N}^{t,\phi}\right) = \phi(s), & s \in [0, t), \\ X^{t,\phi}(s) = \bar{h}(\bar{W}^{t,\phi}, \bar{N}^{t,\phi}, X^{t,\phi}, Z^{t,\phi}, U^{t,\phi}) + \int_s^T \bar{F}(r, X^\pi(r), X_r^\pi, Z^{t,\phi}(r), \tilde{U}^{t,\phi}(r))dr \\ \qquad\qquad - \int_s^T Z^{t,\phi}(r)dW(r) - \int_s^T \int_{\mathbb{R}_0} U^{t,\phi}(r)\tilde{N}(dr, dz), & s \in [t, T], \\ X^{t,\phi}(s) = X^{s,\phi}(s), \quad Z^{t,\phi}(s) = U^{t,\phi}(s, z) = 0, & (s, z) \in [0, t) \times \mathbb{R}_0. \end{cases} \tag{6.8}$$

where the BSDE coefficients are defined according to (6.3).

Then, imposing the same assumptions as before (where $t = 0$) there exists a unique solution $\big(X^{t,\phi}, Z^{t,\phi}, U^{t,\phi}\big)_{(t,\phi)\in[0,T]\times\Lambda}$ for the BSDE in (6.8). Again, the final condition is a little more general that we have in our theoretical framework, but this can be easily modified by adapting the proof of Theorem 3.3, as it is done, for example in Theorem 14.1.1 in [15].
Moreover, by Theorem 4.3, the solution of the backward equation in the above system can be expressed as

$$X^{t,\phi}(s) = u(s, \bar{W}^{t,\phi}, \bar{N}^{t,\phi}), \quad \text{for all } s \in [t, T],$$

for every $(t, \phi) \in [0, T] \times \Lambda$, where $u(t, \phi) := X^{t,\phi}(t)$. Furthermore, u is the mild solution, according to Definition 5.1, of the following path-dependent

PDE:

$$\begin{cases} \partial_t u(t,\phi) + \frac{1}{2}\partial_{xx}^2 u(t,\phi) + \int_{\mathbb{R}_0} \Big(u(\cdot,\phi^{\cdot,(0,1)}) - u(t,\phi) - \partial_y u(t,\phi) \Big) \nu(dz) \\ \qquad\qquad + \bar{F}\left(t, u(t,\phi), (u(\cdot,\phi))_t, \partial_x u(t,\phi), \mathcal{J}u(t,\phi)\right) = 0, \\ u(T,\phi) = \bar{h}\big(\phi, u(\cdot,\phi), \partial_x u(\cdot,\phi), u(\cdot,\phi^{\cdot,(0,1)}) - u(\cdot,\phi)\big), \end{cases}$$

with $(t,\phi) \in [0,T] \times \Lambda$.

We also present a concrete example of a jump-diffusion model for option pricing that can help the reader to link the forward dynamic to a tractable application. This example has a clear limitation if applied to our setting, such as no path-dependence in the coefficients of the stock dynamic (6.9) but only in the terminal condition h of the BSDE. Moreover, if we need to assume suitable conditions on μ, σ and γ, e.g. asking μ bounded and σ, γ constant (w.r.t. X), so that the terminal condition also satisfies the required Lipschitz condition.

Example 6.1. (Forward SDE with a discrete number of jumps). The stock price may present a jump–diffusion dynamic with a discrete number of jumps triggered by a Poisson process, namely we consider the following equation

$$\frac{dS(t)}{S(t)} = \mu(t)\,dt + \sigma(t)dW(t) + d\sum_{i=1}^{N(t)} (V_i - 1), \tag{6.9}$$

in place of Eq. (6.1), where $N(t)$ is a standard Poisson process of fixed rate and jumps size $\{V_i\}$ modelled as a sequence of independent, identically, distributed non-negative random variables. We consider independence among all the sources of randomness, namely W and N. We refer to, e.g., [26] for a detailed treatment of this kind of jump-diffusion. Firstly, we notice that the forward Eq. (6.9) can be explicitly solved by the following

$$S(t) = s_0 \exp\left[\int_0^t \left(\mu(s) - \frac{1}{2}\sigma^2(s) \right) ds + \int_0^t \sigma(s)\,dW(s) \right] \sum_{i=1}^{N(t)} V_i. \tag{6.10}$$

If we assume no path-dependence in the coefficients, such as in Eq. (6.9), then we may encode in the terminal condition of the BSDE (6.2), a dependence only on W, N and $\{V_i\}$. Hence, by introducing the following functional

$$h(S) = \tilde{h}(W, N, \{V_i\}), \tag{6.11}$$

we can decouple the FBSDEs system to fit the setting of Theorem 3.3.

7. Conclusions and future development

The core result of this paper relies on deriving a stochastic representation for the solutions of a non-linear PDE and associating the PDE solution to a FBSDE with jumps and a time-delayed generator. The presence of jumps both in the forward and backward dynamic and, moreover, the dependence of the generator on a (small) time-delayed coefficient represents the main aspect of novelty arising in the analysis of this kind of FBSDE system. Furthermore, we present an application for a large investor problem admitting a jump–diffusion

dynamic. Throughout the article, we mention some possibilities to generalize the setting of our equations such as considering the dependence of f also on a delayed term for the processes Z and U, see Remark 2.1 for more details, or in-depth analyzing the choice of a weaker topology, see Remark 4.2. A different modelling choice deals with considering a further delay term affecting the forward process, see [27] for more details. Furthermore, it might deserve attention to investigate a discretization scheme for this equation, e.g. in line with [7], for the considered equations to obtain a numerical approach based on Neural Networks methods to efficiently compute an approximated solution for the considered FBSDE.

Declarations
Conflict of interest The authors declare they have no financial interests.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Proof of Theorem 3.3

Proof. The existence and the uniqueness are obtained by the Banach fixed point theorem. We consider ϕ fixed in Λ and we define the map Γ on $\mathcal{A}$ with $\mathcal{A} := \mathcal{C}\big([0,T]\,;\,\mathbb{S}_0^2(\mathbb{R})\big)$.

For $R \in \mathcal{A}$, we define $\Gamma(R) := Y$, where, for $t \in [0,T]$, the triple of adapted processes $\big(Y^t(s), Z^t(s), U^t(s,z)\big)_{s \in [t,T]}$ is the unique solution of the

following BSDE

$$
Y^t(s) = h(X^{t,\phi}) + \int_s^T F(r, X^{t,\phi}, Y^t(r), Z^t(r), \tilde{U}^t(r), R_r^t) dr
$$

$$
- \int_s^T Z^t(r) dW(r) - \int_s^T \int_{\mathbb{R}} U^t(r,z) \tilde{N}(dr, dz), \quad s \in [t, T] \ .
$$

$$(A.1)$$

For $s \in [0, t]$ we prolong the solution by taking $Y^t(s) := Y^s(s)$ and $Z^t(s) = U^{t,\phi}(s) := 0$.

Step 1. Let us first show that Γ takes values in the Banach spaces $\mathcal{A}$. We take $R \in \mathcal{A}$ and we will prove that $Y := \Gamma(R) \in \mathcal{A}$. Thus, for every $t \in [0, T]$ we have to show that

$$
Y^t \in \mathbb{S}_0^2(\mathbb{R}) \,,
$$

and that the application

$$
[0, T] \ni t \mapsto Y^t \in \mathbb{S}_0^2(\mathbb{R}) \,,
$$

is continuous.

Let $t \in [0, T]$ be fixed and $t' \in [0, T]$; with no loss of generality, we will suppose that $t < t'$ and $t' - t < \delta$.

Concerning the solution of the BSDE defined in (A.1), we obtain the following estimate

$$
\begin{aligned}
&\mathbb{E}\big(\sup_{s \in [0,T]} |Y^t(s) - Y^{t'}(s)|^2\big) \\
&\leq \mathbb{E}\big(\sup_{s \in [0,t']} |Y^t(s) - Y^{t'}(s)|^2\big) + \mathbb{E}\big(\sup_{s \in [t',T]} |Y^t(s) - Y^{t'}(s)|^2\big) \\
&\leq 2\mathbb{E}\big(\sup_{s \in [t,t']} |Y^t(s) - Y^t(t)|^2\big) + 2\mathbb{E}\big(\sup_{s \in [t,t']} |Y^t(t) - Y^s(s)|^2\big) \\
&\quad + \mathbb{E}\big(\sup_{s \in [t',T]} |Y^t(s) - Y^{t'}(s)|^2\big).
\end{aligned}
$$

We start by proving that

$$
\mathbb{E}\big[\sup_{s \in [t,t']} |Y^t(s) - Y^t(t)|^2 \big] \to 0,
$$

as $t' \to t$. By plugging the explicit solution and applying Doob's inequality, we get

$$
\mathbb{E}\Big(\sup_{s\in[t,t']} |Y^t(s) - Y^t(t)|^2 \Big) = \mathbb{E}\Big[\sup_{s\in[t,t']} \Big| \int_t^s F(r, X^{t,\phi}, Y^t(r), Z^t(r), \tilde{U}^t(r), R_r^t) dr +
$$
$$
- \int_t^s Z^t(r) dW(r) - \int_t^s \int_{\mathbb{R}} U^t(r,z) \tilde{N}(dr,dz) \Big|^2 \Big]
$$
$$
\leq 3\mathbb{E}\Big[\int_t^{t'} \Big| F(r, X^{t,\phi}, Y^t(r), Z^t(r), \tilde{U}^t(r), R_r^t) \Big|^2 dr \Big] +
$$
$$
+ 3\mathbb{E}\Big[\sup_{s\in[t,t']} \Big| \int_t^s Z^t(r) dW(r) \Big|^2 \Big]
$$
$$
+ 3\mathbb{E}\Big[\sup_{s\in[t,t']} \Big| \int_t^s \int_{\mathbb{R}} U^t(r,z) \tilde{N}(dr,dz) \Big|^2 \Big]
$$
$$
\leq 3\mathbb{E}\Big[\int_t^{t'} \Big| F(r, X^{t,\phi}, Y^t(r), Z^t(r), \tilde{U}^t(r), R_r^t) \Big|^2 dr \Big] +
$$
$$
+ 12\mathbb{E}\Big[\int_t^{t'} |Z^t(r)|^2 dr \Big] + 12\mathbb{E}\Big[\int_t^{t'} \int_{\mathbb{R}} |U^t(r,z)|^2 \nu(dz) dr \Big] ,
$$

From the absolute continuity of the Lebesgue integral, we deduce that

$$
\mathbb{E}\Big[\sup_{s\in[t,t']} |Y^t(s) - Y^t(t)|^2 \Big] \to 0 ,
$$

as $t' \to t$.

Concerning the term $\mathbb{E}\big(\sup_{s\in[t',T]} |Y^t(s) - Y^{t'}(s)|^2 \big)$ let us denote for short, only throughout this step,

$$
\Delta Y(r) := Y^t(r) - Y^{t'}(r), \qquad \Delta Z(r) := Z^t(r) - Z^{t'}(r),
$$
$$
\Delta U(r,z) := U^t(r,z) - U^{t'}(r,z), \qquad \Delta R_r(r) := R_r^t(r) - R_r^{t'}(r)
$$

and

$$
\Delta h := h(X^{t,\phi}) - h(X^{t',\phi}),
$$
$$
\Delta F(r) := F(r, X^{t,\phi}, Y^t(r), Z^t(r), \tilde{U}^t(r), R_r^t)
$$
$$
- F(r, X^{t',\phi}, Y^t(r), Z^t(r), \tilde{U}^t(r), R_r^t).
$$

We apply Itô's formula to $e^{\beta s} |\Delta Y(s)|^2$ and we derive, for any $\beta > 0$ and any $s \in [t', T]$,

$$e^{\beta s}|\Delta Y(s)|^2 + \int_s^T e^{\beta r}|\Delta Z(r)|^2 dr$$

$$+ \int_s^T \int_{\mathbb{R}} e^{\beta r}|\Delta U(r,z)|^2 \nu(dz)dr$$

$$= e^{\beta T}|\Delta Y(T)|^2 - 2\int_s^T e^{\beta r}\Delta Y(r)\Delta Z(r)\cdot dW(r)$$

$$-2\int_s^T \int_{\mathbb{R}} e^{\beta r}\Delta Y(r)\Delta U(r,z)\,\tilde{N}(dr,dz) + 2\int_s^T e^{\beta r}\Delta Y(r)$$

$$\left(F(r,X^{t,\phi}Y^t(r),Z^t(r),\tilde{U}^t(r),R_r^t) - F(r,X^{t',\phi},Y^{t'}(r),Z^{t'}(r),\tilde{U}^{t'}(r),R_r^{t'})\right)dr$$

$$-\int_s^T \int_{\mathbb{R}} e^{\beta r}|\Delta U(r,z)|^2\,\tilde{N}(dr,dz).$$

We note that the following estimate

$$\int_s^T e^{\beta r}\left(\int_{-\delta}^0 \left(|\Delta R(r+\theta)|^2\right)\alpha(d\theta)\right)dr = \int_{-\delta}^0 \left[\int_s^T e^{\beta r}\left(|\Delta R(r+\theta)|^2\right)dr\right]\alpha(d\theta)$$

$$\leq e^{\beta\delta}\cdot\int_{-\delta}^0 \alpha(d\theta)\cdot\int_0^T e^{\beta r}\left(|\Delta R(r)|^2\right)dr \leq Te^{\beta\delta}\sup_{r\in[0,T]}\left(e^{\beta r}|\Delta R(r)|^2\right),$$

holds. From assumptions (A_3)–(A_5), we have for any $a > 0$,

$$2\int_s^T e^{\beta r}\Delta Y(r)$$

$$\left(F(r,X^{t,\phi},Y^t(r),Z^t(r),\tilde{U}^t(r),R_r^t) - F(r,X^{t',\phi},Y^{t'}(r),Z^{t'}(r),\tilde{U}^{t'}(r),R_r^{t'})\right)dr$$

$$\leq a\int_s^T e^{\beta r}|\Delta Y(r)|^2 dr + \frac{3}{a}\int_s^T e^{\beta r}|\Delta F(r)|^2 dr$$

$$+\frac{6L^2}{a}\int_s^T e^{\beta r}\left(|\Delta Y(r)|^2 + |\Delta Z(r)|^2 + \int_{\mathbb{R}}|\Delta U(r,z)|^2\lambda(z)\nu(dz)\right)dr$$

$$+\frac{3TKe^{\beta\delta}}{a}\sup_{r\in[0,T]}\left(e^{\beta r}|\Delta R(r)|^2\right).$$

Therefore we have

$$e^{\beta s}|\Delta Y(s)|^2 + \left(\beta - a - \frac{6L^2}{a}\right)\int_s^T e^{\beta r}|\Delta Y(r)|^2 dr$$

$$+\left(1 - \frac{6L^2}{a}\right)\int_s^T e^{\beta r}|\Delta Z(r)|^2 dr + \left(1 - \frac{6L^2}{a}\right)$$

$$\int_s^T \int_{\mathbb{R}} e^{\beta r}|\Delta U(r,z)|^2\lambda(z)\nu(dz)dr$$

$$\leq e^{\beta T}|\Delta Y(T)|^2 + \frac{3}{a}\int_s^T e^{\beta r}|\Delta F(r)|^2 dr$$

$$-2\int_s^T e^{\beta r}\Delta Y(r)\Delta Z(r)\cdot dW(r) - 2\int_s^T \int_{\mathbb{R}} e^{\beta r}\Delta Y(r)\Delta U(r,z)\,\tilde{N}(dr,dz)$$

$$+\frac{3TKe^{\beta\delta}}{a}\sup_{r\in[0,T]}e^{\beta r}|\Delta R(r)|^2.$$

We now choose $\beta, a > 0$ such that

$$a + \frac{6L^2}{a} < \beta \quad \text{and} \quad \frac{6L^2}{a} < 1, \tag{A.2}$$

hence we obtain

$$\left(1 - \frac{6L^2}{a}\right) \mathbb{E}\left[\int_s^T e^{\beta r}|\Delta Z(r)|^2 dr + \int_s^T \int_{\mathbb{R}} e^{\beta r}|\Delta U(r,z)|^2 \lambda(z)\nu(dz)dr\right] \leq \tag{A.3}$$
$$\mathbb{E}\left(e^{\beta T}|\Delta h|^2\right) + \frac{3}{a}\mathbb{E}\int_s^T e^{\beta r}|\Delta F(r)|^2 dr + \frac{3TKe^{\beta\delta}}{a}\mathbb{E}\left(\sup_{r\in[0,T]} e^{\beta r}|\Delta R(r)|^2\right).$$

By Burkholder–Davis–Gundy's inequality, we have

$$2\mathbb{E}\left[\sup_{s\in[t',T]}\left|\int_s^T e^{\beta r}\Delta Y(r)\Delta Z(r)\cdot dW(r)\right|\right]$$
$$\leq \frac{1}{4}\mathbb{E}\left[\sup_{s\in[t',T]} e^{\beta s}|\Delta Y(s)|^2\right] + 144\mathbb{E}\int_{t'}^T e^{\beta r}|\Delta Z(r)|^2 dr.$$

and

$$2\mathbb{E}\left[\sup_{s\in[t',T]}\left|\int_s^T \int_{\mathbb{R}} e^{\beta r}\Delta Y(r)\Delta U(r)\tilde{N}(dr,dz)\right|\right]$$
$$\leq \frac{1}{4}\mathbb{E}\left[\sup_{s\in[t',T]} e^{\beta s}|\Delta Y(s)|^2\right] + 144\mathbb{E}\int_{t'}^T \int_{\mathbb{R}} e^{\beta r}|\Delta U(r,z)|^2 \lambda(z)\nu(dz)dr.$$

which immediately implies

$$\frac{1}{2}\mathbb{E}\left[\sup_{s\in[t',T]} e^{\beta s}|\Delta Y(s)|^2\right] \leq \mathbb{E}\left(e^{\beta T}|\Delta h|^2\right) + \frac{3}{a}\mathbb{E}\int_{t'}^T e^{\beta r}|\Delta F(r)|^2 dr$$
$$+ \frac{3TKe^{\beta\delta}}{a}\mathbb{E}\left(\sup_{r\in[0,T]} e^{\beta r}|\Delta R(r)|^2\right) + 144\,\mathbb{E}\int_{t'}^T |\Delta Z(r)|^2 dr +$$
$$+ 144\,\mathbb{E}\int_{t'}^T \int_{\mathbb{R}} e^{\beta r}|\Delta U(r,z)|^2 \lambda(z)\nu(dz)dr.$$

Hence, we have

$$\frac{1}{2}\mathbb{E}\left(\sup_{s\in[t',T]} e^{\beta s}|\Delta Y(s)|^2\right) \leq \mathbb{E}\left(e^{\beta T}|\Delta h|^2\right) + \frac{3}{a}C_1\mathbb{E}\int_{t'}^T e^{\beta r}|\Delta F(r)|^2 dr \tag{A.4}$$
$$+ \frac{3TKe^{\beta\delta}C_1}{a}\mathbb{E}\left(\sup_{r\in[0,T]} e^{\beta r}|\Delta R(r)|^2\right),$$

where

$$C_1 := 1 + \frac{144}{1 - 6L^2/a}.$$

Exploiting thus assumptions (A_3) and (A_5) together with the fact that $X^{\cdot,\phi}$ is continuous and bounded, we have

$$C_1\mathbb{E}\left(e^{\beta T}|\Delta h|^2\right) + \frac{3}{a}C_1\mathbb{E}\int_{t'}^T e^{\beta r}|\Delta F(r)|^2 dr \to 0 \quad \text{as } t' \to t.$$

Since $R \in \mathcal{A}$, and therefore we have

$$\mathbb{E}\left[\sup_{r \in [0,T]} e^{\beta r} |\Delta R (r)|^2 \right] \to 0,$$

as $t' \to t$, we have

$$\mathbb{E}\left[\sup_{s \in [t',T]} e^{\beta s} |\Delta Y (s)|^2 \right] \to 0, \tag{A.5}$$

as $t' \to t$.

We are left to show that the term $\mathbb{E}\big(\sup_{s \in [t,t']} |Y^t (t) - Y^s (s)|^2 \big)$ is also converging to 0 as $t' \to t$.

Since the map $t \mapsto Y^t (t)$ is deterministic, we have from equation (A.1),

$$
\begin{aligned}
Y^t (t) - Y^s (s) &= \mathbb{E}\big[Y^t (t) - Y^s (s) \big] \\
&= \mathbb{E}\big[h(X^{t,\phi}) - h(X^{s,\phi}) \big] + \mathbb{E}\int_t^T F(r, X^{t,\phi}, Y^t (r), Z^t (r), \tilde{U}^t (r), R_r^t) dr \\
&\quad - \mathbb{E}\int_s^T F(r, X^{s,\phi}, Y^s (r), Z^s (r), \tilde{U}^s (r), R_r^s) dr \\
&= \mathbb{E}\big[h(X^{t,\phi}) - h(X^{s,\phi}) \big] + \mathbb{E}\int_t^s F(r, X^{t,\phi}, Y^t (r), Z^t (r), \tilde{U}^t (r), R_r^t) dr \\
&\quad + \mathbb{E}\int_s^T \big[F(r, X^{t,\phi}, Y^t (r), Z^t (r), \tilde{U}^t (r), R_r^t) \\
&\quad - F(r, X^{s,\phi}, Y^s (r), Z^s (r), \tilde{U}^s (r), R_r^s) \big] dr.
\end{aligned}
$$

Using then the assumption (A_3) we have

$$
\begin{aligned}
|Y^t (t) - Y^s (s)| &\le \mathbb{E}|h(X^{t,\phi}) - h(X^{s,\phi})| + \mathbb{E}\int_t^s L \big(|Y^t (r)| + |Z^t (r)| \\
&\quad + \left| \int_{\mathbb{R}} U^t (r, z) \lambda(z) \nu(dz) \right| \big) dr \\
&\quad + \sqrt{K \int_t^s \mathbb{E}\left[\int_{-\delta}^0 (|R^t (r+\theta)|^2) \alpha(d\theta) \right] dr} \cdot \sqrt{s-t} + \mathbb{E}\int_t^s |F(r, X^{t,\phi}, 0,0,0,0,0,0)| dr \\
&\quad + \mathbb{E}\int_s^T |F(r, X^{t,\phi}, Y^t (r), Z^t (r), \tilde{U}^t (r), R_r^t) - F(r, X^{s,\phi}, Y^t (r), Z^t (r), \tilde{U}^t (r), R_r^t)| dr \\
&\quad + \mathbb{E}\int_s^T L \left(|Y^t (r) - Y^s (r)| + |Z^t (r) - Z^s (r)| + \left| \int_{\mathbb{R}} U^t (r, z) - U^s (r, z) \lambda(z) \nu(dz) \right| \right) dr \\
&\quad + \sqrt{K(T-s) \int_s^T \mathbb{E}\left[\int_{-\delta}^0 (|R^t (r+\theta) - R^s (r+\theta)|^2) \alpha(d\theta) \right] dr},
\end{aligned}
$$

and therefore we obtain

$$|Y^t(t) - Y^s(s)| \le \mathbb{E}|h(X^{t,\phi}) - h(X^{s,\phi})|$$

$$+ L\sqrt{s-t}\sqrt{T\mathbb{E}\sup_{r\in[0,T]}|Y^t(r)|^2 + \mathbb{E}\int_0^T |Z^t(r)|^2 dr + \mathbb{E}\int_0^T \int_{\mathbb{R}} |U^t(r,z)|^2 \lambda(z)\nu(dz)dr}$$

$$+ \sqrt{K}\sqrt{s-t}\sqrt{T\mathbb{E}\sup_{r\in[0,T]}|R^t(r)|^2 + (s-t)M(1+\mathbb{E}||X^{t,\phi}||_T^p)}$$

$$+ \mathbb{E}\int_s^T |F(r,X^{t,\phi},Y^t(r),Z^t(r),\tilde{U}^t(r),R_r^t) - F(r,X^{s,\phi},Z^t(r),\tilde{U}^t(r),R_r^t)|dr$$

$$+ L\sqrt{T-s}\left(T\mathbb{E}\sup_{r\in[s,T]}|Y^t(r) - Y^s(r)|^2 + \mathbb{E}\int_s^T |Z^t(r) - Z^s(r)|^2 dr \right.$$

$$+ \mathbb{E}\int_s^T \int_{\mathbb{R}} |U^t(r,z) - U^{s,z}(r,z)|^2 \lambda(z)\nu(dz)dr\Bigg)^{\frac{1}{2}}$$

$$+ \sqrt{K}\sqrt{T-s}\sqrt{T\mathbb{E}\sup_{r\in[0,T]}|R^t(r) - R^s(r)|^2}.$$

Taking again into account the fact that $R \in \mathcal{A}$, previous step and assumptions (A_3) and (A_5), we infer that

$$\mathbb{E}\left[\sup_{s\in[t,t']}|Y^t(t) - Y^s(s)|\right] \to 0, \quad \text{as } t' \to t. \tag{A.6}$$

Concerning the term $\mathbb{E}\int_0^T |Z^t(r) - Z^{t'}(r)|^2 dr$, we see that

$$\mathbb{E}\int_0^T |Z^t(r) - Z^{t'}(r)|^2 dr = \mathbb{E}\int_0^{t'} |Z^t(r) - Z^{t'}(r)|^2 dr + \mathbb{E}\int_{t'}^T |Z^t(r) - Z^{t'}(r)|^2 dr$$

$$= \mathbb{E}\int_t^{t'} |Z^t(r)|^2 dr + \mathbb{E}\int_{t'}^T |Z^t(r) - Z^{t'}(r)|^2 dr,$$

hence, by (A.3),

$$\mathbb{E}\int_0^T |Z^t(r) - Z^{t'}(r)|^2 dr \to 0, \quad \text{as } t' \to t. \tag{A.7}$$

Analogously, we can infer that

$$\mathbb{E}\int_0^T \int_{\mathbb{R}} |U^t(r,z) - U^{t'}(r,z)|^2 \lambda(z)\nu(dz)dr \to 0,$$

as $t' \to t$.

Step II. *Step 2.* We are going to prove that Γ is a contraction on $\mathcal{A}$ with respect to the norm

$$||Y||_{\mathcal{A}} := \left(\sup_{t\in[0,T]} \mathbb{E}\left[\sup_{r\in[0,T]} e^{\beta r}|Y^t(r)|^2\right]\right)^{1/2},$$

Let us recall that $\Gamma : \mathcal{A} \to \mathcal{A}$ is defined by $\Gamma(R) = Y$ being Y the process coming from the solution of the BSDE (A.1).

Let us consider $R^1, R^2 \in \mathcal{A}$ and $Y^1 := \Gamma(R^1)$, $Y^2 := \Gamma(R^2)$. For the sake of brevity, we will denote in what follows

$$
\begin{aligned}
\Delta F^t(r) &:= F(r, X^{t,\phi}, Y^{1,t}(r), Z^{1,t}(r), \tilde{U}^{1,t}(r), R_r^{1,t}) \\
&\quad - F(r, X^{t,\phi}, Y^{2,t}(r), Z^{2,t}(r), \tilde{U}^{2,t}(r), R_r^{2,t}), \\
\Delta R^t(r) &:= R^{1,t}(r) - R^{2,t}(r), \qquad \Delta Y^t(r) := Y^{1,t}(r) - Y^{2,t}(r) \\
\Delta Z^t(r) &:= Z^{1,t}(r) - Z^{2,t}(r), \qquad \Delta U^t(r) := U^{1,t}(r) - U^{2,t}(r).
\end{aligned}
$$

Proceeding as in **Step I**, we have from Itô's formula, for any $s \in [t, T]$ and $\beta > 0$,

$$
\begin{aligned}
& e^{\beta s} |\Delta Y^t(s)|^2 + \beta \int_s^T e^{\beta r} |\Delta Y^t(r)|^2 dr + \int_s^T e^{\beta r} |\Delta Z^t(r)|^2 dr \\
& + \int_s^T \int_{\mathbb{R}} e^{\beta r} |\Delta U^t(r,z)|^2 \lambda(z) \nu(dz) dr \\
& = 2 \int_s^T e^{\beta r} \Delta Y^t(r) \Delta F^t(r) dr - 2 \int_s^T e^{\beta r} \Delta Y^t(r) \Delta Z^t(r) \cdot dW(r) \\
& \quad -2 \int_s^T \int_{\mathbb{R}} e^{\beta r} \Delta Y^t(r) \Delta U^t(r,z) \tilde{N}(dr, dz).
\end{aligned} \tag{A.8}
$$

Noticing that it holds

$$
\begin{aligned}
& \frac{2K}{a} \int_s^T e^{\beta r} \left(\int_{-\delta}^0 \left(|\Delta R^t(r+\theta)|^2 \right) \alpha(d\theta) \right) dr \\
& \leq \frac{2K}{a} \int_{-\delta}^0 \left(\int_s^T e^{\beta r} \left(|\Delta R^t(r+\theta)|^2 \right) dr \right) \alpha(d\theta) \\
& \leq \frac{2K}{a} \int_{-\delta}^0 \left(\int_{s+r}^{T+r} e^{\beta(r'-\theta)} \left(|\Delta R^t(r')|^2 \right) dr' \right) \alpha(d\theta) \\
& \leq \frac{2K}{a} \int_{-\delta}^0 e^{-\beta\theta} \alpha(d\theta) \cdot \int_{s-\delta}^T e^{\beta r} \left(|\Delta R^t(r)|^2 \right) dr \\
& \leq \frac{2K e^{\beta\delta}}{a} \int_{s-\delta}^T e^{\beta r} \left(|\Delta R^t(r)|^2 \right) dr.
\end{aligned}
$$

we immediately have, from assumptions (A_4)–(A_6), that for any $a > 0$,

$$
\begin{aligned}
2 &\left| \int_s^T e^{\beta r} \Delta Y^t(r) \Delta F^t(r)\, dr \right| \le 2 \int_s^T e^{\beta r} |\Delta Y^t(r), \Delta F^t(r)|\, dr \\
&\le a \int_s^T e^{\beta r} |\Delta Y^t(r)|^2 + \frac{1}{a} \int_s^T e^{\beta r} |\Delta F^t(r)|^2 dr \\
&\le a \int_s^T e^{\beta r} |\Delta Y^t(r)|^2 + \frac{2}{a} \int_s^T e^{\beta r} L^2 \\
&\quad \left(|\Delta Y^t(r)| + |\Delta Z^t(r)| + \left| \int_{\mathbb{R}} \Delta U^t(r,z)\, \lambda(z)\nu(dz) \right| \right)^2 dr \\
&\quad + \frac{2}{a} \int_s^T e^{\beta r} \left(K \int_{-\delta}^0 \left(|\Delta R^t(r+\theta)|^2 \right) \alpha(d\theta) \right) dr \\
&\le a \int_s^T e^{\beta r} |\Delta Y^t(r)|^2 + \frac{4L^2}{a} \\
&\quad \int_s^T e^{\beta r} \left(|\Delta Y^t(r)|^2 + |\Delta Z^t(r)|^2 + \left| \int_{\mathbb{R}} \Delta U^t(r,z)\, \lambda(z)\nu(dz) \right|^2 \right) dr \\
&\quad + \frac{2K e^{\beta\delta}}{a} \int_{s-\delta}^T e^{\beta r} \left(|\Delta R^t(r)|^2 \right) dr.
\end{aligned}
\tag{A.9}
$$

Therefore equation (A.8) yields

$$
\begin{aligned}
e^{\beta s} &|\Delta Y^t(s)|^2 + \left(\beta - a - \frac{4L^2}{a} \right) \int_s^T e^{\beta r} |\Delta Y^t(r)|^2 dr + \left(1 - \frac{4L^2}{a} \right) \int_s^T e^{\beta r} |\Delta Z^t(r)|^2 dr \\
&+ \left(1 - \frac{4L^2}{a} \right) \int_s^T \int_{\mathbb{R}_0} e^{\beta r} |\Delta U^t(r,z)|^2 \lambda(z)\nu(dz) dr \\
&\le \frac{2K e^{\beta\delta}}{a} T \sup_{r\in[0,T]} e^{\beta r} |\Delta R^t(r)|^2 - 2 \int_s^T e^{\beta r} \Delta Y^t(r) \Delta Z^t(r) \cdot dW(r) \\
&\quad - 2 \int_s^T \int_{\mathbb{R}_0} e^{\beta r} \Delta Y^t(r), \Delta U^t(r,z)\, \tilde{N}(dr, dz).
\end{aligned}
\tag{A.10}
$$

Let now $\beta, a > 0$ satisfying

$$
\beta > a + \frac{4L^2}{a} \quad \text{and} \quad 1 > \frac{4L^2}{a},
\tag{A.11}
$$

we have

$$
\begin{aligned}
&\left(1 - \frac{4L^2}{a} \right) \mathbb{E} \int_s^T e^{\beta r} |\Delta Z^t(r)|^2 dr + \left(1 - \frac{4L^2}{a} \right) \mathbb{E} \\
&\int_s^T \int_{\mathbb{R}_0} e^{\beta r} |\Delta U^t(r,z)|^2 \lambda(z)\nu(dz) dr \\
&\quad \le \frac{2TK e^{\beta\delta}}{a} \mathbb{E} \left(\sup_{r\in[0,T]} e^{\beta r} |\Delta R^t(r)|^2 \right)
\end{aligned}
\tag{A.12}
$$

Exploiting now Burkholder–Davis–Gundy's inequality, we have

$$
2\mathbb{E}\left[\sup_{s\in[t,T]}\Big|\int_s^T e^{\beta r}\Delta Y^t(r)\,\Delta Z^t(r)\cdot dW(r)\Big|\right]
$$
$$
\leq \frac{1}{4}\mathbb{E}\left[\sup_{s\in[t,T]} e^{\beta s}|\Delta Y^t(s)|^2\right] + 144\mathbb{E}\int_t^T e^{\beta r}|\Delta Z^t(r)|^2 dr,
$$

and, analogously,

$$
2\mathbb{E}\left[\sup_{s\in[t',T]}\Big|\int_s^T\int_{\mathbb{R}} e^{\beta r}\Delta Y^t(r)\,\Delta U^t(r)\,\tilde{N}(dr,dz)\Big|\right]
$$
$$
\leq \frac{1}{4}\mathbb{E}\left[\sup_{s\in[t,T]} e^{\beta s}|\Delta Y^t(s)|^2\right] + 144\mathbb{E}\int_t^T\int_{\mathbb{R}\setminus\{0\}} e^{\beta r}|\Delta U^t(r,z)|^2\lambda(z)\nu(dz)dr.
$$

which implies

$$
\mathbb{E}\left[\sup_{s\in[t,T]} e^{\beta s}|\Delta Y^t(s)|^2\right] \leq \frac{2Ke^{\beta\delta}}{a}T\mathbb{E}\big(\sup_{s\in[0,T]} e^{\beta s}|\Delta R^t(s)|^2\big)
$$
$$
+2\mathbb{E}\left[\sup_{s\in[t,T]}\Big|\int_s^T e^{\beta r}\Delta Y^t(r)\,\Delta Z^t(r)\cdot dW(r)\Big|\right]
$$
$$
+2\mathbb{E}\left[\sup_{s\in[t',T]}\Big|\int_s^T\int_{\mathbb{R}} e^{\beta r}\Delta Y^t(r)\,\Delta U^t(r)\,\tilde{N}(dr,dz)\Big|\right]
$$
$$
\leq \frac{2Ke^{\beta\delta}}{a}T\mathbb{E}\left[\sup_{s\in[0,T]} e^{\beta s}|\Delta R^t(s)|^2\right]
$$
$$
+\frac{1}{4}\mathbb{E}\left[\sup_{s\in[t,T]} e^{\beta s}|\Delta Y^t(s)|^2\right] + 144\mathbb{E}\int_t^T e^{\beta r}|\Delta Z^t(r)|^2 dr
$$
$$
+\frac{1}{4}\mathbb{E}\left[\sup_{s\in[t',T]} e^{\beta s}|\Delta Y^t(s)|^2\right] + 144\mathbb{E}\int_t^T\int_{\mathbb{R}_0} e^{\beta r}|\Delta U^t(r,z)|^2\lambda(z)\nu(dz)dr.
$$

Hence, we have

$$
\mathbb{E}\left[\sup_{s\in[t,T]} e^{\beta s}|\Delta Y^t(s)|^2\right] \leq \frac{4TKe^{\beta\delta}}{a}C_1\mathbb{E}\left[\sup_{s\in[0,T]} e^{\beta s}|\Delta R^t(s]|^2\right] \quad \text{(A.13)}
$$

where we have denoted by $C_1 := 1 + \frac{144}{1-4L^2/a}$.

Let us now consider the term $\mathbb{E}\big(\sup_{s\in[0,t]} e^{\beta s}|\Delta Y(s)|^2\big)$. From equation (A.1), we see that,

$$
\mathbb{E}\big(\sup_{s\in[0,t]} e^{\beta s}|\Delta Y^t(s)|^2\big) = \mathbb{E}\big(\sup_{s\in[0,t]} e^{\beta s}|Y^{1,t}(s) - Y^{2,t}(s)|^2\big)
$$
$$
= \mathbb{E}\big(\sup_{s\in[0,t]} e^{\beta s}|Y^{1,s}(s) - Y^{2,s}(s)|^2\big) = \sup_{s\in[0,t]} e^{\beta s}|\Delta Y^s(s)|^2 = \quad \text{(A.14)}
$$
$$
= \sup_{s\in[0,t]} \mathbb{E}\big(e^{\beta s}|\Delta Y^s(s)|^2\big)
$$

so that, exploiting Itô's formula and proceeding as above, we obtain

$$\mathbb{E}\big(e^{\beta s}|\Delta Y^s\,(s)\,|^2\big) \le \frac{2TKe^{\beta\delta}}{a}\mathbb{E}\left[\sup_{r\in[0,T]} e^{\beta r}|\Delta R^s\,(r)\,|^2\right]. \qquad (A.15)$$

Thus from inequalities (A.12–A.15) we obtain

$$\mathbb{E}\left[\sup_{s\in[0,T]} e^{\beta r}|\Delta Y^t\,(s)\,|^2\right] + \mathbb{E}\int_0^T e^{\beta r}|\Delta Z^t\,(r)\,|^2 dr + \mathbb{E}\int_0^T\int_{\mathbb{R}} e^{\beta r}|\Delta U^t\,(r,z)\,|^2\lambda(z)\nu(dz)dr$$

$$\le \frac{4TKe^{\beta\delta}}{a}C_1\mathbb{E}\left[\sup_{s\in[0,T]} e^{\beta s}|\Delta R^t\,(s)\,|^2\right] + \frac{2TKe^{\beta\delta}}{a\,(1-4L^2/a)}\mathbb{E}\left[\sup_{r\in[0,T]} e^{\beta r}|\Delta R^t\,(r)\,|^2\right]$$

$$+ \frac{2TKe^{\beta\delta}}{a}\sup_{s\in[0,t]}\mathbb{E}\left[\sup_{r\in[0,T]} e^{\beta r}|\Delta R^s\,(r)\,|^2\right].$$

Then, passing to the supremum for $t\in[0,T]$ we get

$$||Y^1 - Y^2||_{\mathcal{A}}^2 \le \frac{2Ke^{\beta\delta}}{a}\left(3 + \frac{289}{1-4L^2/a}\right)\max\{1,T\}\left[||R^1 - R^2||_{\mathcal{A}}^2\right].$$

By choosing now $a := \frac{4L^2}{\chi}$ and β slightly bigger than $\chi + \frac{4L^2}{\chi}$, condition (A.11) is satisfied and, by restriction (C) we have

$$\frac{2Ke^{\beta\delta}}{a}\left(3 + \frac{289}{1-4L^2/a}\right)\max\{1,T\} < 1. \qquad (A.16)$$

Eventually, since R is chosen arbitrarily, it follows that the application Γ is a contraction on $\mathcal{A}$. Therefore, there exists a unique fixed point $\Gamma(R) = Y \in \mathcal{A}$ and this finishes the proof of the existence and uniqueness of a solution to BSDE with delay and driven by Lèvy process, described by Eq. (2.5). $\square$

References

[1] Applebaum, D.: Lèvy Processes and Stochastic Calculus. University Press, Cambridge (2000)

[2] Baños, D.R., Cordoni, F., Di Nunno, G., Di Persio, L., Røse, E.E.: Stochastic systems with memory and jumps. J. Differ. Equ. **266**(9), 5772–5820 (2019)

[3] Barles, G., Buckdahn, R., Pardoux, E.: Backward stochastic differential equations and integral-partial differential equations. Stoch. Stoch. Rep. **60**, 57–83 (1997)

[4] Bass, R.F.: Stochastic differential equations with jumps. Probab. Surv. **1**, 1–19 (2005)

[5] Bell, D.R., Mohammed, S.E.A.: The Malliavin calculus and stochastic delay equations. J. Funct. Anal. **99**(1), 75–99 (1991)

[6] Bichteler, K.: Stochastic Integration with Jumps Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge (2002). https://doi.org/10.1017/CBO9780511549878

[7] Bouchard, B., Elie, R.: Discrete-time approximation of decoupled forward–backward SDE with Jumps. Stoch. Process. Appl. **118**, 53–75 (2008)

[8] Cont, R.: Functional Ito calculus and functional Kolmogorov equations (2020)

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

[10] Cont, R., Fournié, D.A.: Functional Itô calculus and stochastic integral representation of martingales. Ann. Probab. **41**(1), 109–133 (2013)

[11] Cordoni, F., Di Persio, L., Maticiuc, L., Zălinescu, A.: A stochastic approach to path-dependent nonlinear Kolmogorov equations via BSDEs with time-delayed generators and applications to finance. Stoch. Process. Appl. **130**(3), 1669–1712 (2020)

[12] Cordoni, F., Di Persio, L., Oliva, I.: A nonlinear Kolmogorov equation for stochastic functional delay differential equations with jumps. Nonlinear Differ. Equ. Appl. **24**, 16 (2017)

[13] Cvitanic, J., Ma, J.: Hedging options for a large investor and forward-backward SDEs. Ann. Appl. Probab. **6**(2), 370–398 (1996)

[14] Delong, Ł.: Applications of time-delayed backward stochastic differential equations to pricing, hedging and portfolio management. arXiv e-prints (2010)

[15] Delong, Ł: Backward Stochastic Differential Equations with Jumps and Their Actuarial and Financial Applications: BSDEs with Jumps. EAA Series, Springer, London (2013)

[16] Delong, Ł, Imkeller, P.: Backward stochastic differential equations with time delayed generators—results and counterexamples. Ann. Appl. Probab. **20**(4), 1512–1536 (2010)

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

[18] Dupire, B.: Functional Itô Calculus. Portfolio Research Paper, Bloomberg (2010)

[19] El Karoui N., Peng S., Quenez M.C.: A dynamic maximum principle for the optimization of recursive utilities under constraints. Ann. Appl. Probab. 664–693 (2001)

[20] Ekren, I., Keller, C., Touzi, N., Zhang, J.: Viscosity solutions of fully nonlinear parabolic path dependent PDEs: part I. arXiv:1210.0006v3 (2014)

[21] Ekren, I., Keller, C., Touzi, N., Zhang, J.: Viscosity solutions of fully nonlinear parabolic path dependent PDEs: part II. arXiv: 1210.0007v3 (2014)

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

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

[24] Fuhrman, M., Tessitore, G.: Nonlinear Kolmogorov equations in infinite dimensional spaces: the backward stochastic differential equations approach and applications to optimal control. Ann. Probab. **30**(3), 1397–1465 (2002)

[25] Kac, M.: On some connections between probability theory and differential and integral equations (1951)

[26] Kou, S.G.: A jump-diffusion model for option pricing. Manag. Sci. **48**(8), 1086–1101 (2002)

[27] Ma, T., Xu, J., Zhang, H.: Explicit solution to forward and backward stochastic differential equations with state delay and its application to optimal control. Control Theory Technol. **20**, 303–315 (2022)

[28] Masiero, F., Orrieri, C., Tessitore, G., Zanco, G.: Semilinear Kolmogorov equations on the space of continuous functions via BSDEs. Stoch. Process. Appl. **136**, 1–56 (2021)

[29] Mohammed, S.E.A.: Stochastic Functional Differential Equations, Research Notes in Mathematics, vol. 99. Pitman (Advanced Publishing Program), Boston (1984)

[30] Mohammed, S.E.A.: Stochastic differential systems with memory: theory, examples and applications (1998)

[31] Pardoux, E., Peng, S.: Adapted solution of a backward stochastic differential equation. Syst. Control Lett. **14**, 55–61 (1990)

[32] Pardoux, E., Peng, S.: Backward SDEs and quasilinear PDEs. In: Rozovskii, B.L., Sowers, R. (eds.) Stochastic Partial Differential Equations and Their Applications. LNCIS, vol. 176. Springer, New York (1992)

[33] Pardoux, E., Pradeilles, F., Rao, Z.: Probabilistic interpretation of a system of semi-linear parabolic partial differential equations. Annales de l'I.H.P. Probabilités et statistiques, Tome **33**(4), 467–490 (1997)

[34] Peng, S.: Backward stochastic differential equation, nonlinear expectation and their applications. In: Proceedings of the International Congress of Mathematicians Hyderabad, India (2010)

[35] Peng, S., Wang, F.: BSDE, path-dependent PDE and nonlinear Feynman–Kac formula. Sci. China Math. **59**, 19–36 (2016)

[36] Peszat, S., Zabczyk, J.: Stochastic Partial Differential Equations with Lévy Noise: An Evolution Equation Approach, Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge (2007)

[37] Reib, M., Riedle, M., van Gaans, O.: Delay differential equations driven by Lévy processes: stationarity and Feller properties. Stochastic Processes and their Applications **116**(10), 1409–1432 (2006)

[38] Skorokhod, A.V.: Studies in the Theory of Random Processes. Inc., Addison-Wesley Publishing Co. Inc, Reading, Translated from the Russian by Scripta Technica (1965)

[39] Yan, F., Mohammed, S.E.A.: A stochastic calculus for systems with memory. Stoch. Anal. Appl. **23**(3), 613–657 (2005)

Luca Di Persio
Department of Computer Science
University of Verona
Strada le Grazie 15
37134 Verona
Italy
e-mail: luca.dipersio@univr.it

Matteo Garbelli
Department of Mathematics
University of Trento
Via Sommarive 14, Povo
38123 Trento
Italy
e-mail: matteo.garbelli@unitn.it

Adrian Zălinescu
Faculty of Computer Science
Alexandru Ioan Cuza University
General Berhtelot Street 16
700483 Iasi
Romania
e-mail: adrian.zalinescu@info.uaic.ro