Stochastic Processes and their Applications 170 (2024) 104277

Contents lists available at ScienceDirect

Stochastic Processes and their Applications

journal homepage: www.elsevier.com/locate/spa

Time-delayed generalized BSDEs

Luca Di Persio [a], Matteo Garbelli [a,b,*], Lucian Maticiuc [c], Adrian Zălinescu [d,e]

[a] Department of Computer Science, University of Verona, Strada le Grazie, no. 15, Verona, 37134, Italy
[b] Department of Mathematics, University of Trento, Via Sommarive 14, Povo (Trento), 38123, Italy
[c] Faculty of Mathematics, "Alexandru Ioan Cuza" University, Carol I Blvd., no. 11, Iaşi, 700506, Romania
[d] Faculty of Computer Science, "Alexandru Ioan Cuza" University, Carol I Blvd., no. 11, Iaşi, 700506, Romania
[e] "Octav Mayer" Mathematics Institute of the Romanian Academy, Carol I Blvd., no. 8, Iaşi, 700506, Romania

ARTICLE INFO

MSC:
60H10
60H30

Keywords:
Generalized backward stochastic differential equations
Time–delayed generators
Stieltjes integral
Parameter dependence

ABSTRACT

We prove the existence and uniqueness of the solution of a BSDE with time-delayed generators in the small delay setting (or equivalently small Lipschitz constant), which employs the Stieltjes integral with respect to an increasing continuous stochastic process. Moreover, we obtain a result of continuity of the solution with regard to the increasing process, assuming only uniform convergence, but not in variation. We also prove the existence in the case of arbitrary delay by imposing monotonicity and linearity on generators. Lastly, we provide an application of the theoretical framework within an insurance based example.

1. Introduction

Backward stochastic differential equations (BSDEs for short) were introduced in the linear case by Bismut [1], as adjoint equations involved in the control of SDEs. The nonlinear case was considered by Pardoux and Peng first in [13] and then in [14,18,19], where they established a connection between BSDEs and semilinear parabolic partial differential equations (PDEs), by the so-called nonlinear Feynman–Kac formula. It was this kind of application which triggered an impressive amount of research on the subject. Concerning parabolic PDEs with Neumann boundary conditions, Pardoux and Zhang [17] discovered that their solutions can be linked to BSDEs involving the integral with respect to continuous increasing processes (Stieltjes integral). Moreover, the connection between BSDE and PDE with Neumann boundary condition has been further studied in [15,16] and also within a multivalued setting in [12,21].

This paper represents a first step in establishing a probabilistic representation formula of the solutions of delayed path-dependent parabolic PDEs with Neumann boundary conditions. It consists of studying the well posedness of the associated BSDEs, *i.e.* existence and uniqueness of solutions, as well as stability with respect to terminal data and coefficients. As already shown in [2] for the case of such PDEs considered on the whole space, the generator of the associated BSDE has to take into account the delayed-path of its solution. As a result, our present work is concerned with the following BSDE:

$$\begin{cases} dY(t) = -F(t, Y(t), Z(t), Y_t, Z_t)dt - G(t, Y(t), Y_t)dA(t) \\ \qquad\qquad +Z(t)dW(t), \quad t \in [0, T]\,; \\ Y(T) = \xi, \end{cases} \tag{1}$$

* Corresponding author at: Department of Mathematics, University of Trento, Via Sommarive 14, Povo (Trento), 38123, Italy.
E-mail addresses: luca.dipersio@univr.it (L. Di Persio), matteo.garbelli@unitn.it (M. Garbelli), lucian.maticiuc@uaic.ro (L. Maticiuc), adrian.zalinescu@info.uaic.ro (A. Zălinescu).

https://doi.org/10.1016/j.spa.2023.104277
Received 16 March 2023; Received in revised form 30 November 2023; Accepted 13 December 2023
Available online 15 December 2023

where the generators F and G depend also on the past of the solution (Y, Z). Here, if $\boldsymbol{x} : [-\delta, T] \to \mathbb{R}^n$ is a function and $t \in [0, T]$, $\boldsymbol{x}_t : [-\delta, 0] \to \mathbb{R}^n$ denotes the delayed-path of $\boldsymbol{x}$, defined as

$$\boldsymbol{x}_t(\theta) := \boldsymbol{x}(t+\theta),\ \theta \in [-\delta, 0],$$

where $\delta > 0$ is a fixed delay. The coefficient A is a continuous real valued increasing process.

We recall that time-delayed BSDEs were first introduced in [6,7] where the authors obtained the existence and uniqueness of the solution of the time–delayed BSDE

$$Y(t) = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z(s)dW(s), \quad 0 \le t \le T, \tag{2}$$

where

$$Y_s := (Y(r))_{r\in[0,s]} \quad \text{and} \quad Z_s := (Z(r))_{r\in[0,s]}.$$

In particular, the aforementioned existence and uniqueness result holds true if the time horizon T or the Lipschitz constant for the generator f are sufficiently small.

The motivation behind the introduction of a driving force dA and the corresponding integral goes beyond the link with PDE and can be traced in actuarial applications since [3,20]. In the context of insurance, a BSDE such as the one described in Eq. (1) can be used to model the evolution of a hedging strategy for an insurance portfolio over time. In this framework, the Riemann–Stieltjes integral is linked to the sum of claims with respect to an increasing continuous process that models the cumulative distribution of claims.

This paper is organized as follows. In the remaining part of this section, we introduce the notations and set the framework of our problem. In Section 2 we derive a result of existence and uniqueness for small delay (or small Lipschitz constant) for BSDE (1), based on Banach's fixed point theorem, expressed in Theorem 4. Moreover, we provide in Proposition 6, the well-posedness result for an arbitrary delay for a specific case assuming monotone (in the delayed term) and linear coefficients. Section 3 is devoted to the problem of stability of solutions with respect to terminal data ξ and coefficients F, G and A. Lastly, in Section 4, we present an insurance application dealing with a variable annuity investment that suits the theoretical setting. The main difficulty encountered in the article is to prove the convergence of the solutions of the approximating BSDEs when the increasing process A is approximated uniformly, but not in variation. In order to tackle this problem, we use a stochastic variant of Helly–Bray theorem, proved in the Appendix, as it may be an interesting result for use in other applications.

1.1. *Problem setting and notations*

On the Euclidean space $\mathbb{R}^n$ we consider the Euclidean norm and scalar product, denoted by $|\cdot|$ and $\langle\cdot,\cdot\rangle$, respectively. If $n, k \in \mathbb{N}^*$, $\mathbb{R}^{n\times k}$ denotes the space of real $n \times k$-matrices, equipped with the Frobenius norm (the Euclidean norm when this space is identified with $\mathbb{R}^{nk}$), denoted as well by $|\cdot|$.

For $s < t$, $C([s,t];\mathbb{R}^n)$ represents the set of continuous functions $\boldsymbol{x} : [s,t] \to \mathbb{R}^d$, endowed with the sup-norm: $\|\boldsymbol{x}\|_{C([s,t];\mathbb{R}^n)} := \sup_{r\in[s,t]} |\boldsymbol{x}(r)|$; $BV([s,t];\mathbb{R}^n)$ denotes the set of right-continuous functions with bounded variation $\boldsymbol{\eta} : [s,t] \to \mathbb{R}^n$, *i.e.* with a finite total variation. Recall that the total variation of $\boldsymbol{\eta}$ on $[s,t]$ is defined as

$$\mathrm{V}_s^t(\boldsymbol{\eta}) := \sup \textstyle\sum_{i=1}^n \left|\boldsymbol{\eta}(t_i) - \boldsymbol{\eta}(t_{i-1})\right|,$$

where the sup is taken on all the partitions $s = t_0 < t_1 < \cdots < t_n = t$. The standard norm on $BV([s,t];\mathbb{R}^n)$ is given by

$$\|\boldsymbol{\eta}\|_{BV([s,t];\mathbb{R}^n)} := |\boldsymbol{\eta}(s)| + \mathrm{V}_s^t(\boldsymbol{\eta}).$$

We will simply denote $C[s,t]$, $BV[s,t]$ instead of $C([s,t];\mathbb{R})$, $BV([s,t];\mathbb{R})$, respectively.

If $\boldsymbol{x} : [s,t] \to \mathbb{R}^n$ is a Borel-measurable function and $\boldsymbol{\eta} \in BV([s,t];\mathbb{R}^n)$, by $\int_s^t \langle \boldsymbol{x}(r)d\boldsymbol{\eta}(r)\rangle$ we denote the sum

$$\sum_{i=1}^n \int_s^t \langle \boldsymbol{x}_i(r)d\boldsymbol{\eta}_i(r)\rangle,$$

where $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ and $\boldsymbol{\eta}_1, \ldots, \boldsymbol{\eta}_n$ are the components of $\boldsymbol{x}$, respectively $\boldsymbol{\eta}$, in the case where the Lebesgue–Stieltjes integrals are well-defined and the sum makes sense.

We fix now the framework of our problem, to be utilized throughout the article.

Let $T > 0$ be a finite horizon of time, $d, m \in \mathbb{N}^*$ and $\delta \in (0,T]$ a fixed time-delay. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, W a d-dimensional Brownian motion and $\mathbb{F} = \{\mathcal{F}_t\}_{t\in[0,T]}$ the filtration generated by W, augmented by the null-probability subsets of Ω. The stochastic process $A : \Omega \times [0,T] \to \mathbb{R}$ is an increasing $\mathbb{F}$-adapted process with $A(0) = 0$, $\mathbb{P}$-a.s.

Definition 1. Let $p \ge 2$ and $\beta \ge 0$.

(i) $\mathcal{S}^{p,m}$ denotes the space of continuous $\mathbb{F}$–progressively measurable processes $Y : \Omega \times [0,T] \to \mathbb{R}^m$ such that

$$\mathbb{E}\left[\sup_{0\le s\le T} |Y(s)|^p\right] < +\infty.$$

(ii) $S_\beta^{p,m}$ denotes the space of continuous $\mathbb{F}$–progressively measurable processes $Y : \Omega \times [0,T] \to \mathbb{R}^m$ such that

$$\mathbb{E}\left[\sup_{0\le s\le T} e^{\beta A(s)}|Y(s)|^p\right] + \mathbb{E}\left[\int_0^T e^{\beta A(s)}|Y(s)|^2 dA(s)\right]^{p/2} < +\infty.$$

(iii) $\mathcal{H}_\beta^{p,m\times d}$ denotes the space of $\mathbb{F}$–progressively measurable processes $Z : \Omega \times [0,T] \to \mathbb{R}^{m\times d}$ such that

$$\mathbb{E}\left[\int_0^T e^{\beta A(s)}|Z(s)|^2 ds\right]^{p/2} < +\infty.$$

Instead of $\mathcal{H}_0^{p,m\times d}$ we will write $\mathcal{H}^{p,m\times d}$. The space $S_\beta^{p,m} \times \mathcal{H}_\beta^{p,m\times d}$ (in fact, its quotient with respect to $\mathbb{P}\times\mathbb{P}dt$-a.e. equality) is naturally equipped with the following norm

$$\begin{aligned}\|(Y,Z)\|_{p,\beta}^p = \mathbb{E}\left[\sup_{0\le s\le T} e^{\beta A(s)}|Y(s)|^p\right] + \mathbb{E}\left[\int_0^T e^{\beta A(s)}|Y(s)|^2 dA(s)\right]^{p/2}\\ +\mathbb{E}\left[\int_0^T e^{\beta A(s)}|Z(s)|^2 ds\right]^{p/2}.\end{aligned}$$

2. Existence and uniqueness

We consider the following BSDE

$$\begin{aligned}Y(t) = \xi + \int_t^T F(s,Y(s),Z(s),Y_s,Z_s)ds + \int_t^T G(s,Y(s),Y_s)dA(s)\\ - \int_t^T Z(s)dW(s), \quad t\in[0,T],\end{aligned} \tag{3}$$

with $\xi \in L^2(\Omega,\mathcal{F}_T,\mathbb{P};\mathbb{R}^m)$ and the generators $F : \Omega\times[0,T]\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\times L^2([-\delta,0];\mathbb{R}^m)\times L^2([-\delta,0];\mathbb{R}^{m\times d}) \to \mathbb{R}^m$, $G : \Omega\times[0,T]\times\mathbb{R}^m\times L^2([-\delta,0];\mathbb{R}^m)\to\mathbb{R}^m$ such that the functions $F(\cdot,y,z,\hat{y},\hat{z})$ and $G(\cdot,y,\hat{y})$ are $\mathbb{F}$–progressively measurable, for any $(y,z,\hat{y},\hat{z})\in\mathbb{R}^m\times\mathbb{R}^{m\times d}\times L^2([-\delta,0];\mathbb{R}^m)\times L^2([-\delta,0];\mathbb{R}^{m\times d})$, respectively for any $(y,\hat{y})\in\mathbb{R}^m\times L^2([-\delta,0];\mathbb{R}^m)$.

Recall that, for a function $\boldsymbol{x} : [-\delta,T]\to\mathbb{R}^n$ and some $t\in[0,T]$, $\boldsymbol{x}_t : [-\delta,0]\to\mathbb{R}^n$ denotes the delayed-path of $\boldsymbol{x}$, defined as

$$\boldsymbol{x}_t(\theta) := \boldsymbol{x}(t+\theta),\ \theta\in[-\delta,0].$$

In order to define Y_s and Z_s even for $s<\delta$, we prolong by convention, Y by $Y(0)$ and Z by 0 on the negative real axis.

In what follows we present the assumptions required in this section. We suppose that there exist constants β, L, $\tilde{L}>0$, bounded progressively measurable stochastic processes $K,\tilde{K} : \Omega\times[0,T]\to\mathbb{R}_+$ and ρ, $\tilde{\rho}$ probability measures on $([-\delta,0],\mathcal{B}([-\delta,0]))$ such that:

(A_0) $\mathbb{E}\left[e^{\beta A(T)}\left(1+|\xi|^2\right)\right] < +\infty$;

(A_1) $\mathbb{E}\left[\int_0^T e^{\beta A(t)}\left|F(t,0,0,0,0)\right|^2 dt + \int_0^T e^{\beta A(t)}\left|G(t,0,0)\right|^2 dA(t)\right] < +\infty$.

(A_2) for any $t\in[0,T]$, $(y,z),(y',z')\in\mathbb{R}^m\times\mathbb{R}^{m\times d}$, $\hat{y},\hat{y}'\in L^2([-\delta,0];\mathbb{R}^m)$ and $\hat{z},\hat{z}'\in L^2([-\delta,0];\mathbb{R}^{m\times d})$, we have

$$(i)\quad |F(t,y,z,\hat{y},\hat{z}) - F(t,y',z',\hat{y},\hat{z})| \le L(|y-y'|+|z-z'|),\ \mathbb{P}\text{-a.s.};$$

$$\begin{aligned}(ii)\quad &|F(t,y,z,\hat{y},\hat{z}) - F(t,y,z,\hat{y}',\hat{z}')|^2\\ &\le K(t)\int_{-\delta}^0\left(\left|\hat{y}(\theta)-\hat{y}'(\theta)\right|^2 + \left|\hat{z}(\theta)-\hat{z}'(\theta)\right|^2\right)\rho(d\theta),\ \mathbb{P}\text{-a.s.};\end{aligned}$$

(A_3) for any $t\in[0,T]$, $y,y'\in\mathbb{R}^m$ and $\hat{y},\hat{y}'\in L^2([-\delta,0];\mathbb{R}^m)$, we have

$$(i)\quad |G(t,y,\hat{y}) - G(t,y',\hat{y})| \le \tilde{L}|y-y'|,\ \mathbb{P}\text{-a.s.};$$

$$(ii)\quad |G(t,y,\hat{y}) - G(t,y,\hat{y}')|^2 \le \tilde{K}(t)\int_{-\delta}^0\left|\hat{y}(\theta)-\hat{y}'(\theta)\right|^2\tilde{\rho}(d\theta),\ \mathbb{P}\text{-a.s.};$$

Remark 2. Let us underline that the latter conditions differ from those used in [6], since we allow T to be arbitrary, but different from the delay $\delta\in[0,T]$. This allows to separate the Lipschitz constant L w.r.t. (y,z) from the Lipschitz constant K w.r.t. $(\hat{y},\hat{z})$; therefore the restriction on the coefficients can avoid the constant L.

Remark 3. Existence and uniqueness of a solution to the backward system (3) will be proved by exploiting a standard Banach's fixed point argument which requires K or δ to be small enough.

More precisely, by denoting $K_1 := \sup_{s\in[0,T]} K(s)$, $\tilde{K}_1 := \sup_{s\in[0,T]} \tilde{K}(s)$ and

$$\omega_\delta := \sup_{t\in[0,T-\delta]} (A(t+\delta) - A(t)),$$

we will assume the existence of a positive constant $c < c_{\beta,\tilde{L}} := \min\left\{\frac{\beta^2-8\tilde{L}^2}{4\beta^2}, \frac{1}{584}\right\}$ such that

(H$_1$) $K_1 \cdot \max\{1,T\} \cdot \dfrac{e^{(8L^2+\frac{1}{2})\delta+\beta\omega_\delta}}{4L^2} \le c, \quad \mathbb{P}\text{-a.s.};$

(H$_2$) $4\tilde{K}_1 \cdot A(T) \cdot \dfrac{e^{(8L^2+\frac{1}{2})\delta+\beta\omega_\delta}}{\beta} \le c, \quad \mathbb{P}\text{-a.s.}$

Our first result states existence and uniqueness of Eq. (3).

Theorem 4. *Let us assume that* (A$_0$)–(A$_3$) *hold true and* $\beta > 2\sqrt{2}\tilde{L}$. *If conditions* (H$_1$) *and* (H$_2$) *are satisfied then there exists a unique solution* $(Y,Z) \in \mathcal{S}^{2,m}_\beta \times \mathcal{H}^{2,m\times d}_\beta$ *for* (3).

Proof. The existence and uniqueness will be obtained by the Banach fixed point theorem.

Let us consider the map $\Gamma : \mathcal{S}^{2,m}_\beta \times \mathcal{H}^{2,m\times d}_\beta \to \mathcal{S}^{2,m}_\beta \times \mathcal{H}^{2,m\times d}_\beta$, defined in the following way: for $(U,V) \in \mathcal{S}^{2,m}_\beta \times \mathcal{H}^{2,m\times d}_\beta$, $\Gamma(U,V) = (Y,Z)$, where the couple of adapted processes (Y,Z) is the solution to the equation

$$\begin{aligned} Y(t) &= \xi + \int_t^T F(s,Y(s),Z(s),U_s,V_s)ds + \int_t^T G(s,U(s),U_s)dA(s) \\ &\quad - \int_t^T Z(s)dW(s), \quad t\in[0,T]. \end{aligned} \tag{4}$$

The existence of a unique solution $(Y,Z) \in \mathcal{S}^{2,m} \times \mathcal{H}^{2,m\times d}$ is guaranteed by [13]. Indeed, if we denote

$$\begin{aligned} B(t) &:= \int_0^t G(s,U(s),U_s)dA(s), \quad t\in[0,T]; \\ \hat{F}(t,y,z) &:= F(t,y-B(t),z,U_t,V_t), \quad t\in[0,T],\ (y,z)\in\mathbb{R}^m\times\mathbb{R}^{m\times d}, \end{aligned}$$

then (Y,Z) is a solution to Eq. (3) if and only if $(Y+B,Z)$ solves the equation

$$\hat{Y}(t) = \xi + B(T) + \int_t^T \hat{F}(s,\hat{Y}(s),Z(s),U_s,V_s)ds - \int_t^T Z(s)dW(s), \quad t\in[0,T].$$

Since $\hat{F}$ is Lipschitz with respect to (y,z), it remains to prove that $\mathbb{E}\int_0^T \left|\hat{F}(t,0,0)\right|^2 dt < +\infty$ and $\xi + B(T) \in L^2\left(\Omega,\mathcal{F}_T,\mathbb{P};\mathbb{R}^m\right)$. We have (remember that $K_1 := \sup_{s\in[0,T]} K(s)$ and $\tilde{K}_1 := \sup_{s\in[0,T]} \tilde{K}(s)$):

$$\begin{aligned} &\mathbb{E}\int_0^T \left|\hat{F}(t,0,0)\right|^2 dt = \mathbb{E}\int_0^T \left|F(t,-B(t),0,U_t,V_t)\right|^2 dt \le 3\mathbb{E}\int_0^T |F(t,0,0,0,0)|^2 dt \\ &\quad + 3L^2\mathbb{E}\int_0^T |B(t)|^2 dt + 3\mathbb{E}\int_0^T K(t)\int_{-\delta}^0 \left(|U(t+\theta)|^2 + |V(t+\theta)|^2\right)\rho(d\theta)dt \\ &\le 3\mathbb{E}\int_0^T |F(t,0,0,0,0)|^2 dt + 3L^2\mathbb{E}\int_0^T |B(t)|^2 dt \\ &\quad + 3T\mathbb{E}\left[K_1 \sup_{t\in[0,T]} |U(t)|^2\right] + 3\mathbb{E}K_1\int_0^T |V(t)|^2 dt. \end{aligned}$$

Since (A$_1$) holds and K_1 is bounded, we only have to show that $\mathbb{E}\int_0^T |B(t)|^2 dt < +\infty$ and $\mathbb{E}|B(T)|^2 < +\infty$. We have

$$\begin{aligned} &\mathbb{E}\int_0^T \left|\int_0^t G(s,U(s),U_s)dA(s)\right|^2 dt \\ &\le \mathbb{E}\int_0^T \left[\int_0^t e^{\beta A(s)} \left|G(s,U(s),U_s)\right|^2 dA(s) \cdot \int_0^t e^{-\beta A(s)} dA(s)\right] dt \\ &\le \frac{T}{\beta}\mathbb{E}\int_0^T e^{\beta A(t)} \left|G(t,U(t),U_t)\right|^2 dA(t) \le \frac{2T}{\beta}\mathbb{E}\int_0^T e^{\beta A(t)} |G(t,0,0)|^2 dA(t) \\ &\quad + \frac{2T}{\beta}\mathbb{E}\int_0^T e^{\beta A(t)} L^2 |U(t)|^2 dA(t) + \frac{2T}{\beta}\mathbb{E}\int_0^T e^{\beta A(t)}\tilde{K}(t)\int_{-\delta}^0 |U(t+\theta)|^2 \tilde{\rho}(d\theta)dA(t) \\ &\le \frac{2T}{\beta}\mathbb{E}\int_0^T e^{\beta A(t)} |G(t,0,0)|^2 dA(t) + \frac{2TL^2}{\beta}\mathbb{E}\int_0^T e^{\beta A(t)} |U(t)|^2 dA(t) \\ &+ \frac{2T}{\beta}\mathbb{E}\tilde{K}_1 A(T) e^{\beta\omega_\delta} \sup_{t\in[0,T]} e^{\beta A(t)} |U(t)|^2 < +\infty, \end{aligned}$$

by (A_1) and (H_2), which proves the claim (along the way we have also proven that $\mathbb{E}\,|B(T)|^2<+\infty$).

The proof that $(Y,Z)\in\mathcal{S}^{2,m}_{\beta}\times\mathcal{H}^{2,m\times d}_{\beta}$ is very similar to that of Proposition 1.1 from [17], so it is left to the reader.

Let us prove that Γ is a contraction with respect to the equivalent norm

$$\|(Y,Z)\|^2_{2,\alpha,\beta,a,b}:=\mathbb{E}\big(\sup_{t\in[0,T]}e^{\alpha t+\beta A(t)}|Y(t)|^2\big)+a\mathbb{E}\int_0^T e^{\alpha s+\beta A(s)}|Y(s)|^2dA(s)$$
$$+b\mathbb{E}\int_0^T e^{\alpha s+\beta A(s)}|Z(s)|^2ds.$$

where $\alpha:=8L^2+\frac{1}{2}$ and the constants $a,b>0$ are yet to be chosen.

Let us consider $\left(U^1,V^1\right),\left(U^2,V^2\right)\in\mathcal{S}^{2,m}_{\beta}\times\mathcal{H}^{2,m\times d}_{\beta}$ and $\left(Y^1,Z^1\right):=\Gamma\left(U^1,V^1\right)$, $\left(Y^2,Z^2\right):=\Gamma\left(U^2,V^2\right)$. For the sake of brevity, we will denote in what follows

$$\Delta F(s):=F(s,Y^1(s),Z^1(s),U^1_s,V^1_s)-F(s,Y^2(s),Z^2(s),U^2_s,V^2_s),$$
$$\Delta G(s):=G(s,U^1(s),U^1_s)-G(s,U^2(s),U^2_s),$$
$$\Delta U(s):=U^1(s)-U^2(s),\quad \Delta V(s):=V^1(s)-V^2(s),$$
$$\Delta Y(s):=Y^1(s)-Y^2(s),\quad \Delta Z(s):=Z^1(s)-Z^2(s).$$

Exploiting Itô's formula we have, for any $t\in[0,T]$

$$e^{\alpha t+\beta A(t)}|\Delta Y(t)|^2+\int_t^T e^{\alpha s+\beta A(s)}|\Delta Y(s)|^2(\alpha ds+\beta dA(s))+\int_t^T e^{\alpha s+\beta A(s)}|\Delta Z(s)|^2ds$$
$$=e^{\alpha T+\beta A(T)}|\Delta Y(T)|^2-2\int_t^T e^{\alpha s+\beta A(s)}\langle\Delta Y(s),\Delta Z(s)dW(s)\rangle$$
$$+2\int_t^T e^{\alpha s+\beta A(s)}\langle\Delta Y(s),\Delta F(s)\rangle ds+2\int_t^T e^{\alpha s+\beta A(s)}\langle\Delta Y(s),\Delta G(s)\rangle dA(s).$$

From assumptions (A_2)–(A_3) we obtain,

$$2\Big|\int_t^T e^{\alpha s+\beta A(s)}\langle\Delta Y(s),\Delta F(s)\rangle ds\Big|\le 2\int_t^T e^{\alpha s+\beta A(s)}\Big|\langle\Delta Y(s),\Delta F(s)\rangle\Big|ds$$
$$\le 8L^2\int_t^T e^{\alpha s+\beta A(s)}|\Delta Y(s)|^2ds+\frac{1}{8L^2}\int_t^T e^{\alpha s+\beta A(s)}|\Delta F(s)|^2ds$$
$$\le 8L^2\int_t^T e^{\alpha s+\beta A(s)}|\Delta Y(s)|^2ds+\frac{1}{2}\int_t^T e^{\alpha s+\beta A(s)}\left(|\Delta Y(s)|^2+|\Delta Z(s)|^2\right)ds$$
$$+\frac{K_1T}{4L^2}e^{\alpha\delta+\beta\omega_\delta}\cdot\sup_{s\in[0,T]}\left(e^{\alpha s+\beta A(s)}|\Delta U(s)|^2\right)$$
$$+\frac{K_1}{4L^2}e^{\alpha\delta+\beta\omega_\delta}\cdot\int_0^T e^{\alpha s+\beta A(s)}|\Delta V(s)|^2ds$$

and

$$2\Big|\int_t^T e^{\alpha s+\beta A(s)}\langle\Delta Y(s),\Delta G(s)\rangle dA(s)\Big|\le 2\int_t^T e^{\alpha s+\beta A(s)}\Big|\langle\Delta Y(s),\Delta G(s)\rangle\Big|dA(s)$$
$$\le\frac{\beta}{2}\int_t^T e^{\alpha s+\beta A(s)}|\Delta Y(s)|^2dA(s)+\frac{2}{\beta}\int_t^T e^{\alpha s+\beta A(s)}|\Delta G(s)|^2dA(s)$$
$$\le\frac{\beta}{2}\int_t^T e^{\alpha s+\beta A(s)}|\Delta Y(s)|^2dA(s)+\frac{4\tilde{L}^2}{\beta}\int_t^T e^{\alpha s+\beta A(s)}|\Delta U(s)|^2dA(s)$$
$$+\frac{4\tilde{K}_1\ A(T)}{\beta}e^{\alpha\delta+\beta\omega_\delta}\cdot\sup_{s\in[0,T]}\left(e^{\alpha s+\beta A(s)}|\Delta U(s)|^2\right).$$

By (H_1) and (H_2), we have

$$\left(\tfrac{K_1T}{4L^2}+\tfrac{4\tilde{K}_1\ A(T)}{\beta}\right)e^{\alpha\delta+\beta\omega_\delta}\le 2c,\quad \mathbb{P}\text{-a.s.};$$
$$\tfrac{K_1}{4L^2}e^{\alpha\delta+\beta\omega_\delta}\le c,\quad \mathbb{P}\text{-a.s},$$

(recall that $\alpha := 8L^2 + \frac{1}{2}$). Therefore,

$$\begin{aligned} & e^{\alpha t+\beta A(t)}|\Delta Y(t)|^2 + \frac{\beta}{2}\int_t^T e^{\alpha s+\beta A(s)}|\Delta Y(s)|^2 dA(s) \\ & +\frac{1}{2}\int_t^T e^{\alpha s+\beta A(s)}|\Delta Z(s)|^2 ds \\ & \le -2\int_t^T e^{\alpha s+\beta A(s)}\langle \Delta Y(s), \Delta Z(s)\, dW(s)\rangle + \frac{4\tilde{L}^2}{\beta}\int_t^T e^{\alpha s+\beta A(s)}|\Delta U(s)|^2 dA(s) \\ & +2c\sup_{s\in[0,T]}\left(e^{\alpha s+\beta A(s)}|\Delta U(s)|^2\right) + c\int_0^T e^{\alpha s+\beta A(s)}|\Delta V(s)|^2 ds. \end{aligned} \tag{5}$$

Since $e^{\alpha s+\beta A(s)}\Delta Y \in \mathcal{S}^{2,m}$ and $\Delta Z \in \mathcal{H}^{2,m\times d}$, one can show that

$$\mathbb{E}\left[\int_0^T e^{\alpha s+\beta A(s)}\langle \Delta Y(s), \Delta Z(s)\, dW(s)\rangle\right] = 0,$$

hence

$$\begin{aligned} & \frac{\beta}{2}\mathbb{E}\int_0^T e^{\alpha s+\beta A(s)}|\Delta Y(s)|^2 dA(s) + \frac{1}{2}\mathbb{E}\int_0^T e^{\alpha s+\beta A(s)}|\Delta Z(s)|^2 ds \\ & \le \frac{4\tilde{L}^2}{\beta}\mathbb{E}\int_0^T e^{\alpha s+\beta A(s)}|\Delta U(s)|^2 dA(s) + 2c\mathbb{E}\left[\sup_{s\in[0,T]}\left(e^{\alpha s+\beta A(s)}|\Delta U(s)|^2\right)\right] \\ & +c\mathbb{E}\left[\int_0^T e^{\alpha s+\beta A(s)}|\Delta V(s)|^2 ds\right]. \end{aligned} \tag{6}$$

On the other hand, by Burkholder–Davis–Gundy's inequality, we have

$$\begin{aligned} & 2\mathbb{E}\Big[\sup_{t\in[0,T]}\Big|\int_t^T e^{\alpha s+\beta A(s)}\langle \Delta Y(s), \Delta Z(s)\rangle dW(s)\Big|\Big] \\ & \le \frac{1}{2}\mathbb{E}\big(\sup_{t\in[0,T]} e^{\alpha t+\beta A(t)}|\Delta Y(t)|^2\big) + 72\,\mathbb{E}\int_0^T e^{\alpha s+\beta A(s)}|\Delta Z(s)|^2 ds. \end{aligned}$$

Hence, by (5),

$$\begin{aligned} & \frac{1}{2}\mathbb{E}\big(\sup_{t\in[0,T]} e^{\alpha t+\beta A(t)}|\Delta Y(t)|^2\big) \\ & \le 72\,\mathbb{E}\int_0^T e^{\alpha s+\beta A(s)}|\Delta Z(s)|^2 ds + \frac{4\tilde{L}^2}{\beta}\mathbb{E}\int_0^T e^{\alpha s+\beta A(s)}|\Delta U(s)|^2 dA(s) \\ & +2c\mathbb{E}\left[\sup_{s\in[0,T]}\left(e^{\alpha s+\beta A(s)}|\Delta U(s)|^2\right)\right] + c\mathbb{E}\left[\int_0^T e^{\alpha s+\beta A(s)}|\Delta V(s)|^2 ds\right]. \end{aligned}$$

Thus, with $a := \frac{\lambda\beta}{2}$, $b := \frac{\lambda}{2} - 144$ and some $\lambda > 288$, by taking into account (6), we obtain

$$\begin{aligned} & \mathbb{E}\big(\sup_{t\in[0,T]} e^{\alpha t+\beta A(t)}|\Delta Y(t)|^2\big) + a\int_0^T e^{\alpha s+\beta A(s)}|\Delta Y(s)|^2 dA(s) \\ & +b\mathbb{E}\int_0^T e^{\alpha s+\beta A(s)}|\Delta Z(s)|^2 ds \\ & \le 2c(2+\lambda)\,\mathbb{E}\left[\sup_{s\in[0,T]}\left(e^{\alpha s+\beta A(s)}|\Delta U(s)|^2\right)\right] + \frac{4\tilde{L}^2}{\beta}(2+\lambda)\mathbb{E}\int_0^T e^{\alpha s+\beta A(s)}|\Delta U(s)|^2 dA(s) \\ & +c(2+\lambda)\,\mathbb{E}\int_0^T e^{\alpha r+\beta A(r)}|\Delta V(r)|^2 dr, \end{aligned}$$

so

$$\|(\Delta Y, \Delta Z)\|_{2,\alpha,\beta,a,b}^2 \le \mu_\lambda\, \|(\Delta U, \Delta V)\|_{2,\alpha,\beta,a,b}^2,$$

where

$$\mu_\lambda := \max\left\{c(2+\lambda), \tfrac{8\tilde{L}^2(2+\lambda)}{\lambda\beta^2}, \tfrac{2c(2+\lambda)}{\lambda-288}\right\}.$$

Since $c < c_{\beta,\tilde{L}}$, we can take λ slightly bigger than $\frac{1}{2c_{\beta,\tilde{L}}} - 2$, such that $2c(2+\lambda) < 1$ and so $\mu_\lambda < 1$ (by the definition of $c_{\beta,\tilde{L}}$).

It follows that the application Γ is a contraction on the Banach space $\mathcal{S}_\beta^{2,m} \times \mathcal{H}_\beta^{2,m\times d}$. Therefore, by Banach fixed point theorem, there exists a unique fixed point $(Y, Z) = \Gamma(Y, Z)$ in the space $\mathcal{S}_\beta^{2,m} \times \mathcal{H}_\beta^{2,m\times d}$, which completes our proof. ■

Remark 5. Let us underline that the condition on A to be increasing can be relaxed assuming it to be a continuous bounded variation $\mathbb{F}$-adapted process with $A_0 = 0$, $\mathbb{P}$-a.s. Indeed, by considering the increasing process $\tilde{A}(t) := \|A\|_{BV([0,t])}$, $t \in [0,T]$ and the Radon–Nikodym derivative $\gamma(t) := \frac{dA(t)}{d\tilde{A}(t)}$, $t \in [0,T]$, we have that $|\gamma(t)| \le 1$, $\forall t \in [0,T]$, $\mathbb{P}$-a.s. and the BSDE (3) can be rewritten as

$$\begin{aligned} Y(t) = \xi + &\int_t^T F(s, Y(s), Z(s), Y_s, Z_s)ds + \int_t^T \tilde{G}(s, Y(s), Y_s)d\tilde{A}(s) \\ &- \int_t^T Z(s)dW(s), \quad t \in [0,T], \end{aligned}$$

where the new coefficient $\tilde{G} : \Omega\times[0,T]\times\mathbb{R}^m\times L^2([-\delta,0];\mathbb{R}^m) \to \mathbb{R}^m$ is defined as $\tilde{G}(t,y,\hat{y}) := \gamma(t)G(t,y,\hat{y})$, still satisfying the condition (A_3), by replacing A with $\tilde{A}$ and with the same $\tilde{L}$ and $\tilde{K}$.

As shown in [6], conditions as (H_1) and (H_2) restricting the magnitude of the delay are necessary. However, in the same paper, the authors provide some examples ($F \equiv KY(t-T)$ and $F \equiv K\int_0^t Y(s)ds$, with $K \le 0$) in which the delay can be considered of arbitrary length. The next result is a first attempt to get rid of the restrictive assumptions concerning the delay, by imposing monotonicity and linearity on generators F and G.

More precisely, we assume that $m = 1$, $\xi \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$ and we require F and G not depending on Z_s, namely $F : \Omega\times[0,T]\times\mathbb{R}\times\mathbb{R}^d\times L^2([-\delta,0]) \to \mathbb{R}$ and $G : \Omega\times[0,T]\times\mathbb{R}\times L^2([-\delta,0]) \to \mathbb{R}$.

Moreover, we require that:

(D_1) $\hat{y} \mapsto F(t,y,z,\hat{y})$ and $\hat{y} \mapsto G(t,y,\hat{y})$ are non-increasing with respect to the positive cone of $L^2([-\delta,0])$ for all $(t,y,z) \in [0,T]\times\mathbb{R}\times\mathbb{R}^d$, $\mathbb{P}$-a.s.;

(D_2) $F(t,y,z,\hat{y}) = F_0(t) + F_1(y,z,\hat{y})$, $G(t,y,\hat{y}) = G_0(t) + G_1(y,\hat{y})$, with F_1 and G_1 linear.

Thus, the BSDE (3) reduces to the following one:

$$\begin{aligned} Y(t) = \xi + &\int_t^T \left[F_0(s) + F_1(Y(s), Z(s), Y_s)ds\right] + \int_t^T \left[G_0(s) + G_1(Y(s), Y_s)\right] dA(s) \\ &- \int_t^T Z(s)dW(s), \quad t \in [0,T]. \end{aligned} \tag{7}$$

Proposition 6. *Assume conditions* (D_1), (D_2) *and* (A_0)–(A_3) *hold. If* $\beta > 2\sqrt{2}\tilde{L}$, *then there exists a solution* $(Y,Z) \in \mathcal{S}_\beta^{2,1} \times \mathcal{H}_\beta^{2,1\times d}$ *for* (7).

Proof. As in the proof of Theorem 4, we consider the map $\Gamma : \mathcal{S}_\beta^{2,1} \to \mathcal{S}_\beta^{2,1}$, defined in the following way: for $U \in \mathcal{S}_\beta^{2,1}$, $\Gamma(U) = Y$, where the couple of adapted processes (Y,Z) is the solution to the equation

$$\begin{aligned} Y(t) = \xi + &\int_t^T \left[F_0(s) + F_1(Y(s), Z(s), U_s)ds\right] + \int_t^T \left[G_0(s) + G_1(Y(s), U_s)\right] dA(s) \\ &- \int_t^T Z(s)dW(s), \quad t \in [0,T]. \end{aligned}$$

Using the same type of computations as in the above proof, it is easy to see that even without conditions (H_1) and (H_2), Γ is still a Lipschitz-continuous function. By a classical comparison theorem for BSDEs, if $U^1(t) \le U^2(t)$ $\mathbb{P}dt$-a.e., then $Y^1(t) \le Y^2(t)$, $\forall t \in [0,T]$, $\mathbb{P}$-a.s., with $Y^i(t) := \Gamma(U^i)$, $i = \overline{1,2}$. This shows that Γ is non-increasing with respect to the positive cone of $\mathcal{S}_\beta^{2,1}$.

One can use now an argument from [11, Theorem 2.2] to show that there exist $\underline{U}, \overline{U} \in \mathcal{S}_\beta^{2,1}$ such that $\Gamma([\underline{U}, \overline{U}]) \subseteq [\underline{U}, \overline{U}]$, where $[\underline{U}, \overline{U}] := \left\{U \in \mathcal{S}_\beta^{2,1} \mid \underline{U}(t) \le U(t) \le \overline{U}(t),\ \mathbb{P}dt\text{-a.e.}\right\}$. Obviously, $[\underline{U}, \overline{U}]$ is a closed, convex set of the Banach space $\mathcal{S}_\beta^{2,1}$.

Let $Y^0 := \underline{U}$ and, by recursion, $Y^{n+1} := \Gamma(Y^n)$. By the monotonicity property of Γ, it is easy to show that $\forall t \in [0,T]$, $\mathbb{P}$-a.s.,

$$\underline{U}(t) = Y^0(t) \le Y^2(t) \le \cdots \le Y^{2n}(t) \le \cdots \le Y^{2n+1}(t) \le \cdots \le Y^3(t) \le Y^1(t) \le \overline{U}(t).$$

Let $\underline{Y}(t) := \lim_{n\to\infty} Y^{2n}(t)$ and $\overline{Y}(t) := \lim_{n\to\infty} Y^{2n+1}(t)$. Since $\underline{U}, \overline{U} \in \mathcal{S}_\beta^{2,1}$, for any $H \in L^2(\Omega; BV[0,T])$ or $H \in L^2(\Omega\times[0,T], \mathbb{P}dA(\cdot))$ we have, by the dominated convergence theorem,

$$\begin{aligned} &\lim_{n\to\infty} \mathbb{E}\int_0^T e^{\beta A(t)} Y^{2n}(t)H(t)dt = \mathbb{E}\int_0^T e^{\beta A(t)}\underline{Y}(t)H(t)dt \text{ and} \\ &\lim_{n\to\infty} \mathbb{E}\int_0^T e^{\beta A(t)} Y^{2n+1}(t)H(t)dt = \mathbb{E}\int_0^T e^{\beta A(t)}\overline{Y}(t)H(t)dt. \end{aligned}$$

Hence $(e^{\beta A(\cdot)/2}Y^{2n})$ and $(e^{\beta A(\cdot)/2}Y^{2n+1})$ converge weakly to $e^{\beta A(\cdot)/2}\underline{Y}$, respectively $e^{\beta A(\cdot)/2}\overline{Y}$, in both $L^2(\Omega;C[0,T])$ and $L^2(\Omega\times[0,T],\mathbb{P}dA(\cdot))$. By Mazur's lemma (applied two times), for any $n\in\mathbb{N}$ there are convex combinations, let us call them $\underline{Y}^n$ and $\overline{Y}^n$, of the elements of $(Y^{2k})_{k\geq n}$, respectively $(Y^{2k+1})_{k\geq n}$, such that $(e^{\beta A(\cdot)/2}\underline{Y}^n)$ and $(e^{\beta A(\cdot)/2}\overline{Y}^n)$ converge strongly in both $L^2(\Omega;C[0,T])$ and $L^2(\Omega\times[0,T],\mathbb{P}dA(\cdot))$ to $e^{\beta A(\cdot)/2}\underline{Y}$, respectively $e^{\beta A(\cdot)/2}\overline{Y}$. Therefore, $(\underline{Y}^n)$ and $(\overline{Y}^n)$ converge strongly in $\mathcal{S}_\beta^{2,1}$ to $\underline{Y}$, respectively $e^{\beta A(\cdot)}\overline{Y}$; thus, $\lim_{n\to\infty}\Gamma(\underline{Y}^n)=\Gamma(\underline{Y})$ and $\lim_{n\to\infty}\Gamma(\overline{Y}^n)=\Gamma(\overline{Y})$.

On the other hand, by the linearity of F_1 and G_1, $\Gamma(\underline{Y}^n)$ and $\Gamma(\overline{Y}^n)$ are convex combinations of the elements of $(Y^{2k+1})_{k\geq n}$, respectively $(Y^{2k})_{k\geq n}$, so $e^{\beta A(\cdot)/2}\Gamma(\underline{Y}^n)$ and $e^{\beta A(\cdot)/2}\Gamma(\overline{Y}^n)$ converge pointwisely to $e^{\beta A(\cdot)/2}\overline{Y}$, respectively $e^{\beta A(\cdot)/2}\underline{Y}$. Consequently, $\Gamma(\underline{Y})=\overline{Y}$ and $\Gamma(\overline{Y})=\underline{Y}$. Then, setting $Y=\frac{1}{2}\underline{Y}+\frac{1}{2}\overline{Y}$, we have $\Gamma(Y)=Y$, which proves our claim. ■

3. Dependence on parameters

Let us consider, for all $n\in\mathbb{N}^*$, the following BSDEs which approximate (3):

$$
\begin{aligned}
Y^n(t) &= \xi^n + \int_t^T F^n(s,Y^n(s),Z^n(s),Y_s^n,Z_s^n)ds + \int_t^T G^n(s,Y^n(s),Y_s^n)dA^n(s)\\
&\quad - \int_t^T Z^n(s)dW(s), \quad t\in[0,T],
\end{aligned}
\tag{8}
$$

In order to unify the notations, we will sometimes denote ς^0 instead of ς, if ς is ξ, A, F, G, Y or Z. We suppose that the coefficients ξ^n, A^n, F^n, G^n, $n\geq 0$, satisfy conditions (A_2)–(A_3), (H_1), (H_2) with processes K^n, $\tilde{K}^n$, but the same constants β, c, L, $\tilde{L}$. Moreover, we have to impose that $\beta>2\sqrt{2}\tilde{L}$.

We suppose that there exists $p>1$ such that

(A_0') $\sup_{n\in\mathbb{N}}\mathbb{E}\left[e^{p\beta A^n(T)}\left|\xi^n\right|^{2p}\right]<+\infty$.

(A_0'') $\sup_{n\in\mathbb{N}}\mathbb{E}\left[e^{qA^n(T)}\right]<+\infty$, for any $q>0$.

(A_1') $\sup_{n\in\mathbb{N}}\mathbb{E}\left[\left(\int_0^T e^{\beta A^n(t)}\left|F^n(t,0,0,0,0)\right|^2dt\right)^p+\left(\int_0^T e^{\beta A^n(t)}\left|G^n(t,0,0)\right|^2dA^n(t)\right)^p\right]<+\infty$.

Under these assumptions, there exists a unique solution $(Y^n,Z^n)\in\mathcal{S}_\beta^{2,m}\times\mathcal{H}_\beta^{2,m\times d}$ to Eq. (8). In fact, one can now prove by standard computations that $(Y^n,Z^n)\in\mathcal{S}_\beta^{p,m}\times\mathcal{H}_\beta^{p,m\times d}$, $\forall n\in\mathbb{N}$ and

$$
\sup_{n\in\mathbb{N}}\|(Y^n,Z^n)\|_{p,\beta}<+\infty.
\tag{9}
$$

Our aim is to show that if the coefficients (ξ^n,A^n,F^n,G^n) of Eq. (8) converge to (ξ,A,F,G), then (Y^n,Z^n) converge to (Y,Z) in $\mathcal{S}^{2,m}\times\mathcal{H}^{2,m\times d}$. Let now specify in which sense the convergence of the coefficients takes place. We define

$$
\begin{aligned}
\Delta_nF &:= \sup_{t\in[0,T],\ (y,z)\in\mathbb{R}^m\times\mathbb{R}^{m\times d},\ (\hat{y},\hat{z})\in L^2([-\delta,0];\mathbb{R}^m\times\mathbb{R}^{m\times d})}\left|F^n(t,y,z,\hat{y},\hat{z})-F(t,y,z,\hat{y},\hat{z})\right|;\\
\Delta_nG &:= \sup_{t\in[0,T],\ y\in\mathbb{R}^m\times\mathbb{R}^{m\times d},\ \hat{y}\in L^2([-\delta,0];\mathbb{R}^m)}\left|G^n(t,y,\hat{y})-G(t,y,\hat{y})\right|
\end{aligned}
$$

and impose

(C_1) $\mathbb{E}\left[\left|\xi^n-\xi\right|^{2p}\right]\to 0$ as $n\to\infty$;

(C_2) $\mathbb{E}\sup_{t\in[0,T]}\left|A^n(t)-A(t)\right|\to 0$ as $n\to\infty$;

(C_3) $[\mathbb{E}\left(\Delta_nF\right)^{2p}+\left(\Delta_nG\right)^{2p}]\to 0$ as $n\to\infty$.

The uniform convergence from assumption (C_3) can be relaxed to a weaker type of convergence; however, we will work with this hypothesis for the sake of keeping computations as simple as possible.

Theorem 7. *Assume that the above assumptions are fulfilled. Then*

$$
\lim_{n\to\infty}\mathbb{E}\left[\sup_{t\in[0,T]}\left|Y^n(t)-Y(t)\right|^2+\int_0^T\left|Z^n(t)-Z(t)\right|^2dt\right]=0.
$$

Proof. Let us denote for short

$$
\begin{aligned}
&\Delta_n Y(t) := Y^n(t) - Y(t)\,, \quad \Delta_n Z(t) := Z^n(t) - Z(t)\,; \quad \Delta_n \xi := \xi^n(t) - \xi(t)\\
&\quad \omega^n_\delta := \sup_{t\in[0,T-\delta]} (A^n(t+\delta) - A^n(t))\,.
\end{aligned}
$$

Exactly as in the proof of Theorem 4, by (H_1) and (H_2), we have

$$
\begin{aligned}
\Big(\tfrac{K_1 T}{4L^2} + \tfrac{4\tilde K_1\, A(T)}{\beta}\Big) e^{\alpha\delta+\beta\omega_\delta} &\le 2c\,, \quad \mathbb{P}\text{-a.s.};\\
\tfrac{K_1}{4L^2} e^{\alpha\delta+\beta\omega_\delta} &\le c\,, \quad \mathbb{P}\text{-a.s,}
\end{aligned}
$$

for all $n \in \mathbb{N}$, where $\alpha = 8L^2 + \frac{1}{2}$. Let us apply Itô's formula to $e^{\alpha t+\beta A(t)}\,|Y^n(t) - Y(t)|^2$:

$$
\begin{aligned}
&e^{\alpha t+\beta A^n(t)}|\Delta_n Y(t)|^2 + \int_t^T e^{\alpha s+\beta A^n(s)}|\Delta_n Y(s)|^2\,(\alpha\, ds + \beta\, dA^n(s)) + \int_t^T e^{\alpha s+\beta A^n(s)}|\Delta_n Z(s)|^2 ds\\
&= e^{\alpha T+\beta A^n(T)}|\Delta_n \xi|^2 - 2\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y(s), \Delta_n Z(s) dW(s)\rangle\\
&\quad +2\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y(s), F^n(s, Y^n(s), Z^n(s), Y^n_s, Z^n_s) - F(s, Y(s), Z(s), Y_s, Z_s)\rangle ds\\
&\quad +2\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y\,(s)\,, G^n(s, Y^n(s), Y^n_s) dA^n(s) - G(s, Y(s), Y_s) dA(s)\rangle\,.
\end{aligned}
$$

From assumptions (A_2)–(A_3) and (A'_1), we have, with $K^n_1 := \sup_{t\in[0,T]} K^n$ and $\tilde K^n_1 := \sup_{t\in[0,T]} \tilde K^n$,

$$
\begin{aligned}
&2\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y(s), F^n(s, Y^n(s), Z^n(s), Y^n_s, Z^n_s) - F(s, Y(s), Z(s), Y_s, Z_s)\rangle ds\\
&\le 8L^2\int_t^T e^{\alpha s+\beta A^n(s)}|\Delta_n Y\,(s)\,|^2 ds + \frac{1}{2}\left|\Delta_n F\right|^2\int_t^T e^{\alpha s+\beta A^n(s)} ds\\
&\quad +\frac{1}{2}\int_t^T e^{\alpha s+\beta A^n(s)}\left(|\Delta_n Y\,(s)\,|^2 + |\Delta_n Z\,(s)\,|^2\right) dr\\
&\quad +\frac{K^n_1 T e^{\alpha\delta+\beta\omega^n_\delta}}{4L^2}\sup_{s\in[0,T]}\left(e^{\alpha s+\beta A^n(s)}|\Delta_n Y\,(s)\,|^2\right)\\
&\quad +\frac{K^n_1 e^{\alpha\delta+\beta\omega^n_\delta}}{4L^2}\int_0^T e^{\alpha r+\beta A^n(r)}|\Delta_n Z\,(r)\,|^2 dr
\end{aligned}
$$

and, for all $b > 0$,

$$
\begin{aligned}
&2\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y\,(s)\,, G^n(s, Y^n(s), Y^n_s) dA^n(s) - G(s, Y(s), Y_s) dA(s)\rangle\\
&= 2\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y\,(s)\,, G^n(s, Y^n(s), Y^n_s) - G^n(s, Y(s), Y_s)\rangle dA^n\,(s)\\
&\quad +2\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y\,(s)\,, G^n(s, Y^n(s), Y^n_s) - G(s, Y(s), Y_s)\rangle dA^n\,(s)\\
&\quad +2\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y\,(s)\,, G(s, Y(s), Y_s)\rangle\,\left(dA^n\,(s) - dA\,(s)\right)\\
&\le 2\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y\,(s)\,, G(s, Y(s), Y_s)\rangle\,\left(dA^n\,(s) - dA\,(s)\right)\\
&\quad +b\int_t^T e^{\alpha s+\beta A^n(s)}|\Delta_n Y\,(s)\,|^2 dA^n\,(s) + \frac{4\tilde L^2}{b}\left|\Delta_n G\right|^2\int_t^T e^{\alpha s+\beta A^n(s)} dA^n\,(s)\\
&\quad +\frac{\beta}{2}\int_t^T e^{\alpha s+\beta A^n(s)}|\Delta_n Y\,(s)\,|^2 dA^n\,(s) + \frac{4\tilde L^2}{\beta}\int_t^T e^{\alpha s+\beta A^n(s)}|\Delta_n Y\,(s)\,|^2 dA^n\,(s)\\
&\quad +\frac{4\tilde K^n_1\, A^n\,(T)\, e^{\alpha\delta+\beta\omega^n_\delta}}{\beta}\sup_{s\in[0,T]}\left(e^{\alpha s+\beta A^n(s)}|\Delta_n Y\,(s)\,|^2\right).
\end{aligned}
$$

Since $\alpha = 4L^2 + 1$ and $\beta > 2\sqrt{2}\tilde{L}$, one can choose $b := \frac{\beta}{2} - \frac{4\tilde{L}^2}{\beta}$ and so we obtain

$$
\begin{aligned}
& e^{\alpha t+\beta A^n(t)}|\Delta_n Y(t)|^2 + \frac{1}{2}\int_t^T e^{\alpha s+\beta A^n(s)}|\Delta_n Z(s)|^2 ds \\
& \leq e^{\alpha T+\beta A^n(T)}|\Delta_n \xi|^2 - 2\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y(s), \Delta_n Z(s) dW(s)\rangle \\
& \quad +2\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y(s), G(s, Y(s), Y_s)\rangle \, (dA^n(s) - dA(s)) \\
& \quad +\frac{1}{2}\left|\Delta_n F\right|^2 \int_t^T e^{\alpha s+\beta A^n(s)} ds + \frac{4\tilde{L}^2}{b}\left|\Delta_n G\right|^2 \int_t^T e^{\alpha s+\beta A^n(s)} dA^n(s) \\
& \quad +\frac{K_1 T e^{\alpha\delta+\beta\omega_\delta^n}}{4L^2} \sup_{s\in[0,T]} \left(e^{\alpha s+\beta A^n(s)}|\Delta_n Y(s)|^2\right) \\
& \quad +\frac{K_1 e^{\alpha\delta+\beta\omega_\delta^n}}{4L^2} \int_0^T e^{\alpha s+\beta A^n(s)}|\Delta_n Z(s)|^2 ds \\
& \quad +\frac{4\tilde{K}_1 A^n(T) e^{\alpha\delta+\beta\omega_\delta^n}}{\beta} \sup_{s\in[0,T]} \left(e^{\alpha s+\beta A^n(s)}|\Delta_n Y(s)|^2\right).
\end{aligned}
$$

Therefore, by conditions (H_1) and (H_2),

$$
\begin{aligned}
& \frac{1}{2}\int_t^T e^{\alpha s+\beta A^n(s)}|\Delta_n Z(s)|^2 ds \\
& \leq 2\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y(s), G(s, Y(s), Y_s)\rangle \, (dA^n(s) - dA(s)) \\
& \quad +\frac{1}{2}\left|\Delta_n F\right|^2 \int_t^T e^{\alpha s+\beta A^n(s)} ds + \frac{4\tilde{L}^2}{b}\left|\Delta_n G\right|^2 \int_t^T e^{\alpha s+\beta A^n(s)} dA^n(s) \\
& \quad +2c \sup_{s\in[0,T]} \left(e^{\alpha s+\beta A^n(s)}|\Delta_n Y(s)|^2\right) + c\int_0^T e^{\alpha s+\beta A^n(s)}|\Delta_n Z(s)|^2 ds.
\end{aligned}
$$

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

$$
\begin{aligned}
& 2\mathbb{E}\Big[\sup_{t\in[0,T]}\Big|\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y(s), \Delta_n Z(s) dW(s)\rangle\Big|\Big] \\
& \leq \frac{1}{4}\,\mathbb{E}\left(e^{\alpha s+\beta A^n(s)}|\Delta_n Y(s)|^2\right) + 144\,\mathbb{E}\int_0^T e^{\alpha s+\beta A^n(s)}|\Delta_n Z(s)|^2 ds\,.
\end{aligned}
$$

As in the proof of Theorem 4, we obtain

$$
\begin{aligned}
& \mathbb{E}\Big(\sup_{s\in[t,T]} e^{\alpha s+\beta A^n(s)}|\Delta_n Y(s)|^2\Big) + \mathbb{E}\int_0^T e^{\alpha s+\beta A^n(s)}|\Delta_n Z(s)|^2 ds \\
& \leq C\mathbb{E}\left[|\Delta_n \xi|^{2p} + |\Delta_n F|^{2p} + |\Delta_n G|^{2p}\right]\cdot \mathbb{E}e^{\beta q A^n(T)} \\
& \quad +C\mathbb{E}\sup_{t\in[0,T]}\left|\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y(s), G(s, Y(s), Y_s)\rangle \, (dA^n(s) - dA(s))\right|.
\end{aligned}
$$

where C is a positive constant and $q := \frac{p}{p-1}$.

By conditions (C_1) and (A_0''),

$$
\lim_{n\to\infty}\mathbb{E}\left[|\Delta_n \xi|^{2p} + |\Delta_n F|^{2p} + |\Delta_n G|^{2p}\right]\cdot \mathbb{E}e^{\beta q A^n(T)} = 0.
$$

It remains to prove that

$$
\lim_{n\to\infty}\mathbb{E}\sup_{t\in[0,T]}\left|\int_t^T X^n(s) dH^n(s)\right| = 0,
$$

where, for $s \in [0, T]$,

$$
\begin{aligned}
& X^n(s) := e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y(s), G(s, Y(s), Y_s)\rangle; \\
& H^n(s) := A^n(s) - A(s).
\end{aligned}
$$

One can prove that

$$
\mathbb{E}\sup_{t\in[0,T]}|X^n(t)|^p
$$

is uniformly bounded (with respect to n), by (9). Obviously, by (A_0''),

$$\sup_{n\in\mathbb{N}^*} \mathbb{E} \sup_{t\in[0,T]} |H^n(t)|^2 < +\infty.$$

Hence, the sequence $(X^n, H^n)_{n\in\mathbb{N}^*}$ is tight in $C[0,T]^2$. By Prokhorov's theorem, we can extract a sequence, say $(X^{n_k}, H^{n_k})_{k\in\mathbb{N}^*}$, convergent in distribution to some stochastic process (X, H) with continuous paths. Since, by (C_2), $\lim_{n\to\infty} \mathbb{E} \sup_{t\in[0,T]} |H^n(t)| = 0$, H must be $\mathbb{P}$-a.s. equal to 0. The condition (A_0'') also implies that $\sup_{n\in\mathbb{N}} \mathbb{E} \|H^n\|^a_{BV[0,T]} < +\infty$, for every $a > 1$, so $\|H^n\|_{BV[0,T]}$ is bounded in probability (*i.e.*, it satisfies condition (16)). We can now apply Proposition 8, proved as an auxiliary result in the Appendix, in order to derive the convergence in distribution to 0 of the process

$$\left(\int_0^t X^n(s) dH^n(s) \right)_{t\in[0,T]}.$$

Since, for some $\nu > 0$, the functional $\phi_\nu : C[0,T] \to \mathbb{R}$, defined by

$$\phi_\nu(\boldsymbol{x}) := \sup_{t\in[0,T]} |\boldsymbol{x}(T) - \boldsymbol{x}(t)| \wedge \nu,$$

is bounded and continuous, it follows that

$$\mathbb{E}\left[\sup_{t\in[0,T]} \left| \int_t^T X^n(s) dH^n(s) \right| \wedge \nu \right] = 0,$$

for every $\nu > 0$. Since, by Markov's inequality, for some $a \in (1, p)$

$$\mathbb{E} \sup_{t\in[0,T]} \left| \int_t^T X^n(s) dH^n(s) \right| \le \mathbb{E}\left[\sup_{t\in[0,T]} \left| \int_t^T X^n(s) dH^n(s) \right| \wedge \nu \right] + \frac{1}{\nu^a} \mathbb{E}\left[\sup_{t\in[0,T]} \left| \int_t^T X^n(s) dH^n(s) \right|^a \right]$$

and

$$\begin{aligned} \mathbb{E}\left[\sup_{t\in[0,T]} \left| \int_t^T X^n(s) dH^n(s) \right|^a \right] &\le \mathbb{E}\left[\left(\sup_{t\in[0,T]} |X^n(t)|^a \right) \|H^n\|^a_{BV[0,T]} \right] \\ &\le \left(\mathbb{E}\left[\sup_{t\in[0,T]} |X^n(t)|^p \right] \right)^{\frac{a}{p}} \left(\mathbb{E} \|H^n\|^{\frac{p}{a(p-a)}}_{BV[0,T]} \right)^{1-\frac{a}{p}}, \end{aligned}$$

it follows that

$$\lim_{n\to\infty} \mathbb{E} \sup_{t\in[0,T]} \left| \int_t^T X^n(s) dH^n(s) \right| = 0,$$

which concludes our proof. ■

4. Hedging a stream of payments with time-delayed GBSDE

In this last section, we present a risk management application for an insurance product, the so-called variable annuity instrument, whose composition can be controlled by the insurer selecting an appropriate strategy to reduce the overall risk of the policyholder's investment. This example is an extension of the work contained in [4–7], where the authors apply different classes of BSDEs with time–delayed generators to insurance and finance. Specifically, inspired by Section 7 in [4], we consider an insurance product where the policyholder withdraws some guaranteed amounts as a fraction of the maximum value of the investment and, additionally, is subjected to a continuous payment triggered by an increasing continuous process A modelling the cumulative function of claims (or, e.g. of fees for the management of the wealth). At maturity, the remaining value is converted into a life-time annuity with a guaranteed consumption rate C.

We consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with associated natural filtration $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le R}$ generated by a Brownian motion $W := (W(t),\, 0 \le t \le T)$ and a finite time horizon $T \le \infty$.

The goal of the investor is to replicate the insurance by investing in the assets and to quantify the risk of the investing activities. In the terminology of [8], we focus on an investment composed of a risk free asset S_0 and a risky asset D.

The price of the risk free asset $S_0 := (S_0(t),\, 0 \le t \le T)$ is given by the equation

$$\frac{dS_0(t)}{S_0(t)} = r(t)dt, \qquad S_0(0) = 1, \tag{10}$$

where r describes the risk free interest rate being a non-negative $\mathbb{F}$-progressively measurable stochastic process.

The price of the risky ass $D := (D(t),\ 0 \le t \le T)$ with maturity T is given by

$$\frac{dD(t)}{D(t)} = (r(t) + \sigma(t)\theta(t))\,dt + \sigma(t)dW(t), \qquad S(0) = x, \tag{11}$$

where the volatility $\sigma := (\sigma(t),\ 0 \le t \le T)$ and the risk premium $\theta := (\theta(t),\ 0 \le t \le T)$ are $\mathbb{F}$-progressively measurable processes.

On the other hand, the stream of liabilities $P(t) := (P(t),\ 0 \le t \le T)$ depends on the past value of the portfolio by the following:

$$P(t) = \gamma \sup_{s\in[0,t]} \{X(s)\}\,dt + \int_0^t X(s-\delta)dA(s). \tag{12}$$

The first term models a guaranteed withdrawal amount as a fraction $\gamma \in (0,1)$ of the running maximum value of the investment value. Instead, the second term models a Stieltjes integral representing the total amount of continuous claims that depend on a past value of the investment and that are triggered by the increasing continuous function A. We emphasize that if we consider no dependence on the value of the investment X, i.e. only $\int_0^t dA(s)$, we obtain the well-known case with A representing a cumulative consumption process. See, e.g., [8,9] for a detailed description or [10] for the problem of utility maximization under a drawdown constraint setting.

We consider a self financing investment portfolio $X := (X(t), 0 \le t \le T)$, while the admissible strategy $\pi := (\pi(t), 0 \le t \le T)$ denotes the amount invested in the risky bond D.

We denote $\mu(t) = r(t) + \theta(t)\sigma(t)$ and we write the dynamic of X by the following SDE

$$\begin{aligned} dX(t) =& \pi(t)\frac{dD(t)}{D(t)} + (X(t)-\pi(t))\,\frac{dS_0(t)}{S_0(t)} - dP(t) \\ =& \pi(t)\,(\mu(t)dt + \sigma(t)dW(t)) + (X(t)-\pi(t))\,r(t)dt \\ & - \gamma \sup_{s\in[0,t]} \{X(s)\}\,dt - \int_0^t X(s-\delta)dA(s) \\ X(T) =& Ca(T)\,, \end{aligned} \tag{13}$$

a being the annuity factor $a(T) = \mathbb{E}^{\mathbb{Q}}\left[\int_T^\infty e^{-\int_T^s r(u)du} ds|\mathcal{F}_T\right]$.

Eq. (13) models a variable annuity contract where the policyholder's contributions are invested into two assets (D and S_0). Positive returns are distributed to the policyholder account based on the maximum value of the investment and on a prescribed process A (hedging fee) while the remaining value at maturity is received as a life-time annuity.

From [4], we know that there exists a unique equivalent martingale measure $\mathbb{Q} \sim \mathbb{P}$ under which the discounted price process S is a $(\mathbb{Q}, \mathbb{F})$-martingale. Thus, we perform the following change of variables

$$Y(t) = X(t)e^{-\int_0^t r(s)ds}, \qquad Z(t) = \pi(t)\sigma(t)e^{-\int_0^t r(s)ds}, \qquad 0 \le t \le T \tag{14}$$

giving the following dynamic for the discounted portfolio process $Y := (Y(t))_{0\le t\le T}$ under the measure $\mathbb{Q}$

$$\begin{aligned} Y(t) = C\tilde{a}(T) + \int_t^T \gamma \sup_{u\in[0,s]} \left\{Y(u)e^{-\int_u^s r(v)dv}\right\} ds \\ + \int_t^T Y(s-\delta)e^{-\int_0^{s-\delta} r(v)dv} dA(s) - \int_t^T Z(s)dW^{\mathbb{Q}}(s)\,, \end{aligned} \tag{15}$$

$W^{\mathbb{Q}}$ being a $\mathbb{Q}$-Brownian motion.

Assuming that conditions (A_0)–(A_3) and (H_1)–(H_2) hold true and applying Theorem 4, we obtain existence and uniqueness of the solution of Eq. (15). Moreover, the stability of the investment under a perturbation (in uniform norm) of the distribution of the prescribed cumulative distribution is obtained by Theorem 7, letting to model robust hedging for the investment with respect to a modification of the prescribed cumulative distribution of future claims.

5. Conclusions and further developments

In this article, we develop a theoretical framework to study a BSDE with time-delayed generator whose dynamic depends also on Stieltjes integral term. Under regular assumptions of the coefficients and small delay, we prove the well-posedness of the problem in terms of existence, uniqueness and stability under a perturbation in uniform norm. We also provide an application of our results for a BSDE in insurance setting. Moreover, we obtain the global (in time) well posedness of the BSDE for an arbitrary delay that represent a novel result in the literature, representing a first attempt to handle (globally) time delayed BSDE. Providing a solid theoretical background for this setting could open up new directions for applications.

Concerning further direction of research, other extensions would consider the forward reflected SDE linked to the Stieltjes integral in (1) to investigate the corresponding FBSDE with delayed generator and possible connections with the nonlinear PDE with Neumann boundary conditions in the spirit of [17]. Another possibility concerns considering Stieltjes integration with respect to increasing functions that are not necessarily continuous, dealing with dynamics driven by Poisson random measure.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix

In this section, we state the result used in the proof of Theorem 7. It is a variant of the Helly–Bray theorem for the stochastic case and is also stronger than Proposition 3.4 from [21]:[1]

Proposition 8. *Let* $(X_n, H_n) : (\Omega_n, \mathcal{F}_n, \mathbb{P}_n) \to C([0,T];\mathbb{R}^d)^2$, $n \geq 1$, *be a sequence of random variables, converging in distribution to a random variable* $(X, H) : (\Omega, \mathcal{F}, \mathbb{P}) \to C([0,T];\mathbb{R}^d)^2$. *If for all* $n \geq 1$, H_n *is* $\mathbb{P}_n$*-a.s. with bounded variation and*

$$\lim_{\nu\to+\infty} \sup_{n\geq 1} \mathbb{P}_n \left(\|H_n\|_{BV([0,T];\mathbb{R}^d)} > \nu \right) = 0, \tag{16}$$

then H *is* $\mathbb{P}$*-a.s. with bounded variation and the sequence of* $C[0,T]$*-valued random variables* $\left(\int_0^\cdot \langle X_n(s), dH_n(s)\rangle\right)_{n\geq 1}$ *converges in distribution to* $\int_0^\cdot \langle X(s), dH_n(s)\rangle$.

As expected, the proof of this result uses a deterministic Helly–Bray type theorem aiming uniform convergence. For the reader's convenience, we will state and prove this result:

Lemma 9. *Let* $\left(\boldsymbol{x}_n\right)_{n\geq 1} \subseteq C([0,T];\mathbb{R}^d)$ *and* $\left(\boldsymbol{\eta}_n\right)_{n\geq 1} \subseteq BV([0,T];\mathbb{R}^d)$ *be two sequences of functions such that:*

(i) $\boldsymbol{x}_n$ *converges uniformly to a function* $\boldsymbol{x} \in C([0,T];\mathbb{R}^d)$*;*

(ii) $\boldsymbol{\eta}_n$ *converges uniformly to a function* $\boldsymbol{\eta}$*;*

(iii) $\sup_{n\geq 1} \|\boldsymbol{\eta}_n\|_{BV([0,T];\mathbb{R}^d)} < +\infty$.

Then $\boldsymbol{\eta} \in BV([0,T];\mathbb{R}^d)$, $\|\boldsymbol{\eta}\|_{BV([0,T];\mathbb{R}^d)} \leq \liminf_{n\to\infty} \|\boldsymbol{\eta}_n\|_{BV([0,T];\mathbb{R}^d)}$ *and the sequence of continuous functions* $\left(\int_0^\cdot \langle \boldsymbol{x}_n(s), d\boldsymbol{\eta}_n(s)\rangle\right)_{n\geq 1}$ *converges uniformly to* $\int_0^\cdot \langle \boldsymbol{x}(s), d\boldsymbol{\eta}(s)\rangle$.

Proof. The first two assertions are well-known, so we skip their proof.

Let us prove the last one. We say that a tuple $\pi = (t_0, \dots, t_k)$ is a *partition* of $[0,T]$ if $0 = t_0 < t_1 < \cdots < t_{k_N} = T$.

We consider $\pi^N = (t_0^N, \dots, t_{k_N}^N)$, $N \in \mathbb{N}^*$ partitions of the interval $[0,T]$ such that

$$\lim_{N\to\infty} \sup_{0\leq i<t_{k_N}^N} \left| t_{i+1}^N - t_i^N \right| = 0.$$

Let $\boldsymbol{x}^N : [0,T] \to \mathbb{R}^d$ be a step-function approximating $\boldsymbol{x}$, defined by

$$\boldsymbol{x}^N := \mathbf{1}_{\{0\}}\boldsymbol{x}(0) + \sum_{i=1}^{k_N} \mathbf{1}_{(t_{i-1},t_i]}\boldsymbol{x}(t_i).$$

Let $M := \sup_{n\geq 1} \|\boldsymbol{\eta}_n\|_{BV([0,T];\mathbb{R}^d)}$. Then

$$\begin{aligned}
&\left| \int_0^t \langle \boldsymbol{x}_n(s), d\boldsymbol{\eta}_n(s)\rangle - \int_0^t \langle \boldsymbol{x}(s), d\boldsymbol{\eta}(s)\rangle \right| \leq \left| \int_0^t \langle \boldsymbol{x}_n(s) - \boldsymbol{x}(s), d\boldsymbol{\eta}_n(s)\rangle \right| \\
&\quad + \left| \int_0^t \left\langle \boldsymbol{x}(s) - \boldsymbol{x}^N(s), d(\boldsymbol{\eta}_n - \boldsymbol{\eta})(s)\right\rangle \right| + \left| \int_0^t \left\langle \boldsymbol{x}^N(s), d(\boldsymbol{\eta}_n - \boldsymbol{\eta})(s)\right\rangle \right| \\
&\leq \|\boldsymbol{x}_n - \boldsymbol{x}\|_{C([0,T];\mathbb{R}^d)} \mathrm{V}_0^T(\boldsymbol{\eta}_n) + \left\|\boldsymbol{x}^N - \boldsymbol{x}\right\|_{C([0,T];\mathbb{R}^d)} \left(\mathrm{V}_0^T(\boldsymbol{\eta}_n) + \mathrm{V}_0^T(\boldsymbol{\eta})\right) \\
&\quad + \sum_{i=1}^{k_N} |\boldsymbol{x}(t_i)| \cdot |(\boldsymbol{\eta}_n - \boldsymbol{\eta})(t_i \wedge t) - (\boldsymbol{\eta}_n - \boldsymbol{\eta})(t_{i-1} \wedge t)|\,.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\sup_{t\in[0,T]} \left| \int_0^t \langle \boldsymbol{x}_n(s), d\boldsymbol{\eta}_n(s)\rangle - \int_0^t \langle \boldsymbol{x}(s), d\boldsymbol{\eta}(s)\rangle \right| \leq M \|\boldsymbol{x}_n - \boldsymbol{x}\|_{C([0,T];\mathbb{R}^d)} \\
&\quad + 2M \left\|\boldsymbol{x}^N - \boldsymbol{x}\right\|_{C([0,T];\mathbb{R}^d)} + 2\left(\sum_{i=1}^{k_N} |\boldsymbol{x}(t_i)|\right) \|\boldsymbol{\eta}^n - \boldsymbol{\eta}\|_{C([0,T];\mathbb{R}^d)}\,.
\end{aligned}$$

It follows that

$$\limsup_{n\to\infty} \sup_{t\in[0,T]} \left| \int_0^t \langle \boldsymbol{x}_n(s), d\boldsymbol{\eta}_n(s)\rangle - \int_0^t \langle \boldsymbol{x}(s), d\boldsymbol{\eta}(s)\rangle \right| \leq 2M \left\|\boldsymbol{x}^N - \boldsymbol{x}\right\|_{C([0,T];\mathbb{R}^d)}.$$

[1] In the same time, it corrects an error in the statement of that result: "Let X_n, K_n : $(\Omega n, \mathcal{F}n, Pn) \to \mathbf{W}, n \geq 1$, be two sequences of random variables, converging in distribution to X, respectively K ", should be replaced with "Let (X_n, K_n) : $(\Omega n, \mathcal{F}n, Pn) \to \mathbf{W}^2, n \geq 1$, be a sequence of random variables, converging in distribution to (X, K)". We emphasize that this does not affect in any way the validity of the other results in that paper, since the arguments involved use in fact this stronger assumption.

Since $\lim_{N\to\infty} \|x^N - x\|_{C([0,T];\mathbb{R}^d)} = 0$, we finally get

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \left| \int_0^t \langle x_n(s), d\eta_n(s)\rangle - \int_0^t \langle x(s), d\eta(s)\rangle \right| = 0. \qquad \blacksquare$$

Let us now proceed with the proof of the main result of this section, which follows the same steps as that of Proposition 3.4 from [21].

Proof of Proposition 8. Let $\mathbf{W} := C([0,T];\mathbb{R}^d)$, $\mathbf{V} := C([0,T];\mathbb{R}^d) \cap BV([0,T];\mathbb{R}^d)$ and, for $v > 0$,

$$\mathbf{V}_v := \left\{ \eta \in \mathbf{V} \mid \|\eta\|_{BV([0,T];\mathbb{R}^d)} \le v \right\}.$$

By the first part of Lemma 9, $\mathbf{V}_v$ is a closed subset of the Banach space $\mathbf{W}$.

Let us consider the function $\Lambda : \mathbf{W}\times\mathbf{W} \to \mathbf{W}$ defined by

$$\Lambda(x,\eta)(t) := \begin{cases} \int_0^t \langle x(s), d\eta(s)\rangle, & (x,\eta) \in \mathbf{W}\times\mathbf{V}; \\ 0, & (x,\eta) \in \mathbf{W}\times(\mathbf{W}\setminus\mathbf{V}). \end{cases}$$

By the last conclusion of Lemma 9, the restriction $\Lambda|_{\mathbf{W}\times\mathbf{V}_v}$ is continuous.

Let now $R_n := \mathbb{P}^n \circ (X^n, H^n)^{-1}$ and $R_0 := \mathbb{P}\circ(X,H)^{-1}$, the distribution probabilities of (X^n, H^n), respectively (X,H). By the assumptions of the theorem, $\left(R_n\right)_{n\ge 1}$ converges weakly to R_0, *i.e.*

$$\lim_{n\to\infty} \int_{\mathbf{W}\times\mathbf{W}} \Phi(x,\eta)\, R_n(dx, d\eta) = \int_{\mathbf{W}\times\mathbf{W}} \Phi(x,\eta)\, R_0(dx, d\eta), \tag{17}$$

for every bounded continuous functional $\Phi : \mathbf{W}\times\mathbf{W} \to \mathbb{R}$.

First of all, by the Portmanteau lemma,

$$\limsup_{n\to\infty} R_n\left(\mathbf{W}\times\mathbf{V}_v\right) \le R_0\left(\mathbf{W}\times\mathbf{V}_v\right), \ \forall v > 0.$$

Since, by condition (16),

$$\lim_{v\to+\infty} \inf_{n\ge 1} R_n\left(\mathbf{W}\times\mathbf{V}_v\right) = 1, \tag{18}$$

we get $\lim_{v\to+\infty} R_0\left(\mathbf{W}\times\mathbf{V}_v\right) = 1$, *i.e.* $R_0(\mathbf{W}\times\mathbf{V}) = 1$, meaning that H is $\mathbb{P}$-a.s. of bounded variation.

Let now $\phi : C[0,T] \to \mathbb{R}$ be an arbitrary bounded continuous functional. It remains to prove that $\lim_{n\to\infty} \mathbb{E}\phi(\Lambda(X^n, H^n)) = \mathbb{E}\phi(\Lambda(X,H))$, which can be written as

$$\lim_{n\to\infty} \int_{\mathbf{W}\times\mathbf{W}} (\phi\circ\Lambda)dR_n = \int_{\mathbf{W}\times\mathbf{W}} (\phi\circ\Lambda)dR_0.$$

Since $\phi\circ\Lambda|_{\mathbf{W}\times\mathbf{V}_v}$ is bounded and continuous, it can be extended to a continuous functional $\Phi_v : \mathbf{W}\times\mathbf{W} \to \mathbb{R}$, bounded by $M := \sup_{\mathbf{z}\in C[0,T]} \phi(\mathbf{z})$; hence, by (17),

$$\lim_{n\to\infty} \int_{\mathbf{W}\times\mathbf{W}} \Phi_v(x,\eta)\, R_n(dx, d\eta) = \int_{\mathbf{W}\times\mathbf{W}} \Phi_v(x,\eta)\, R_0(dx, d\eta).$$

Let us estimate the term

$$T_{n,v} := \left| \int_{\mathbf{W}\times\mathbf{W}} \left(\Phi_v(x,\eta) - \phi\circ\Lambda\right) R_n(dx, d\eta) \right|,$$

for $n \in \mathbb{N}$ (including then the case $n = 0$). We have

$$\begin{aligned} T_{n,v} &\le \int_{\mathbf{W}\times\mathbf{W}} \left|\Phi_v(x,\eta) - \phi\circ\Lambda\right| R_n(dx, d\eta) = \int_{\mathbf{W}\times(\mathbf{W}\setminus\mathbf{V}_v)} \left|\Phi_v(x,\eta) - \phi\circ\Lambda\right| R_n(dx, d\eta) \\ &\le 2M R_n\left(\mathbf{W}\times(\mathbf{W}\setminus\mathbf{V}_v)\right) = 2M\left(1 - R_n\left(\mathbf{W}\times\mathbf{V}_v\right)\right). \end{aligned}$$

Hence, by (18) and its consequence

$$\lim_{v\to+\infty} \sup_{n\ge 0} T_{n,v} = 0.$$

Finally, for all $n \ge 1$ and $v > 0$,

$$\begin{aligned} &\left| \int_{\mathbf{W}\times\mathbf{W}} (\phi\circ\Lambda)dR_n - \int_{\mathbf{W}\times\mathbf{W}} (\phi\circ\Lambda)dR_0 \right| \\ &\le \left| \int_{\mathbf{W}\times\mathbf{W}} \Phi_v(x,\eta)\, R_n(dx, d\eta) - \int_{\mathbf{W}\times\mathbf{W}} \Phi_v(x,\eta)\, R_0(dx, d\eta) \right| + T_{n,v} + T_{0,v} \end{aligned}$$

and therefore

$$\limsup_{n\to\infty} \left| \int_{\mathbf{W}\times\mathbf{W}} (\phi\circ\Lambda)dR_n - \int_{\mathbf{W}\times\mathbf{W}} (\phi\circ\Lambda)dR_0 \right| \le 2\sup_{n\ge 0} T_{n,v}, \ \forall v > 0$$

which, by passing to the limit as $v \to 0$, yields the desired conclusion. $\blacksquare$

References

[1] J.M. Bismut, An introductory approach to duality in optimal stochastic control, SIAM Rev. 20 (1978) 62–78.
[2] F. Cordoni, L. Di Persio, L. Maticiuc, A. Zălinescu, A stochastic approach to path-dependent nonlinear Kolmogorov equations via BSDEs with time-delayed generators and applications to finance, Stoch. Process. Appl. 130 (3) (2020) 1669–1712.
[3] B. de Finetti, M. Jacob, Sull' integrale di stieltjes-Riemann, G. Ist. Ital. Degli Attuari 6 (1935) 303–319.
[4] Ł. Delong, Applications of time–delayed backward stochastic differential equations to pricing, hedging and portfolio management in insurance and finance, Appl. Math. (Warsaw) 39 (2012a) 463–488.
[5] Ł. Delong, BSDEs with time-delayed generators of a moving average type with applications to non-monotone preferences, Stoch. Models 28 (2012b) 281–315.
[6] Ł. Delong, P. Imkeller, Backward stochastic differential equations with time delayed generators – results and counterexamples, Ann. Appl. Probab. 20 (2010a) 1512–1536.
[7] Ł. Delong, P. Imkeller, On Malliavin's differentiability of BSDE with time delayed generators driven by Brownian motions and Poisson random measures, Stochastic Process. Appl. 120 (2010b) 1748–1775.
[8] N. El Karoui, S. Peng, M.C. Quenez, Backward stochastic differential equations in finance, Math. Finance 7 (1997) 1–71.
[9] N. El Karoui, S. Peng, M.C. Quenez, A dynamic maximum principle for the optimization of recursive utilities under constraints, Ann. Appl. Probab. 11 (2001) 664–693.
[10] R. Elie, N. Touzi, Optimal lifetime consumption and investment under a drawdown constraint, Finance Stoch 12 (2008) 299–330.
[11] X. Li, Z. Wang, Fixed point theorems for decreasing operators in ordered Banach spaces with lattice structure and their applications, Fixed Point Theory Appl. 18 (2013).
[12] L. Maticiuc, A. Răşcanu, A stochastic approach to a multivalued Dirichlet–Neumann problem, Stochastic Process. Appl. 120 (2010) 777–800.
[13] E. Pardoux, S. Peng, Adapted solution of a backward stochastic differential equation, Systems Control Lett. 14 (1990) 55–61.
[14] E. Pardoux, S. Peng, Backward sde's and quasilinear parabolic PDE's, stochastic PDE and their applications, in: B.L. Rozovskii, R.B. Sowers (Eds.), LNCIS 176, Springer, 1992, pp. 200–217.
[15] E. Pardoux, A. Răşcanu, Stochastic Differential Equations, Backward SDEs, Partial Differential Equations, Springer Series: Stochastic Modelling and Applied Probability, Vol. 69, Springer, Berlin, 2014.
[16] E. Pardoux, A. Răşcanu, Continuity of the Feynman–Kac formula for a generalized parabolic equation, stochastics, Int. J. Probab. Stoch. Process. (2017) published online.
[17] E. Pardoux, S. Zhang, Generalized BSDE and nonlinear Neumann boundary value problems, Probab. Theory Related Fields 110 (1998) 535–558.
[18] S. Peng, Probabilistic interpretation for systems of quasilinear parabolic partial differential equation, Stochastics 37 (1991) 61–74.
[19] S. Peng, Backward stochastic differential equation, nonlinear expectation and their applications, in: Proceedings of the International Congress of Mathematicians Hyderabad, India, 2010.
[20] J. Steffensen, On stieltjes' integral and its applications to actuarial questions, J. Inst. Actuar. 63 (3) (1932) 443–483.
[21] A. Zălinescu, Weak solutions and optimal control for multivalued stochastic differential equations, NoDEA 15 (2008) 511–533.

Further reading

[1] S.-E.A. Mohammed, Stochastic Functional Differential Equations, in: Research Notes in Mathematics, vol. 99, Pitman, Boston, MA, 1984.
[2] S. Peng, F. Wang, BSDE, path–dependent PDE and nonlinear Feynman–Kac formula, Sci. China Math. 59 (2016) 19–36.