A Monte Carlo method for backward stochastic differential equations with Hermite martingales

Research Article

Antoon Pelsser and Kossi Gnameho*

A Monte Carlo method for backward stochastic differential equations with Hermite martingales

https://doi.org/10.1515/mcma-2019-2028
Received June 27, 2018; revised December 20, 2018; accepted January 7, 2019

Abstract: Backward stochastic differential equations (BSDEs) appear in many problems in stochastic optimal control theory, mathematical finance, insurance and economics. This work deals with the numerical approximation of the class of Markovian BSDEs where the terminal condition is a functional of a Brownian motion. Using Hermite martingales, we show that the problem of solving a BSDE is identical to solving a countable infinite-dimensional system of ordinary differential equations (ODEs). The family of ODEs belongs to the class of stiff ODEs, where the associated functional is one-sided Lipschitz. On this basis, we derive a numerical scheme and provide numerical applications.

Keywords: Regression, BSDE, ODE, Hermite polynomials, martingale

MSC 2010: 65C05, 65C40, 60H10

1 Introduction

This paper deals with the numerical approximation of the class of Markovian backward stochastic differential equations (BSDEs) on the interval $[0, T]$, where the terminal condition is a functional of a Brownian motion. BSDEs were first introduced by Bismut [7] in the linear case and later developed by Pardoux and Peng [43]. In the past decade, BSDEs have attracted a lot of attention and have been intensively studied in mathematical finance, insurance and stochastic optimal control theory. For example, in a complete financial market, the price of a standard European option can be seen as the solution of a linear BSDE. Moreover, the price of an American option can be formulated as the solution of a reflected BSDE. These equations have also been widely applied for portfolio optimization, indifference pricing, modeling of convex risk measures and the modeling of ambiguity with respect to the stochastic drift and the volatility. See, for instance, [4, 5, 12, 14, 21, 32, 33]. In general, many of these equations do not have an explicit or closed form solution. Due to its importance, some efforts have been made to provide numerical solutions. For instance, a four-step scheme has been proposed by Ma, Protter and Yong in [38] to solve forward-backward SDEs. In [3], Bally has proposed a random time discretization scheme. Discrete time approximation schemes have been also proposed by Bouchard and Touzi in [8] and Chevance in [13], for instance. In Chevance's work [13], strong regularity assumptions of the coefficients of the BSDE are required for convergence results. In [15] a cubature method for BSDEs with application to nonlinear pricing was proposed. Gobet, Lemor and Warin [25] presented a discrete algorithm based

Antoon Pelsser, Deptartment of Quantitative Economics, Maastricht University and Netspar, Maastricht, Netherlands, e-mail: a.pelsser@maastrichtuniversity.nl
***Corresponding author: Kossi Gnameho,** Deptartment of Quantitative Economics, Maastricht University, Maastricht, Netherlands, e-mail: kossi.gnameho@gmail.com

on the Monte Carlo method to solve BSDEs. Fujii and Takahashi [22] proposed an analytical approximation for nonlinear FBSDEs with perturbation scheme.

Recently, Fourier methods for solving forward-backward stochastic differential equations (FBSDEs) were proposed in [29] and a convolution method in [31]. A new algorithm based on the regression-later approach is also proposed in [24]. In [26], a numerical scheme for solving BSDE with Malliavin weights was designed. Other recent references can be founded in [1, 6, 11, 23, 27, 35, 47], among others. In 2014, Briand and Labart [9] presented an algorithm to solve BSDEs based on Wiener chaos expansion and Picard's iterations. In their approach, the regression coefficients are computed with a weighted Monte Carlo method by Malliavin calculus. In our approach, the regression coefficients are determined exactly as solution of countable systems of ordinary differential equations (ODEs).

Most of the above numerical algorithms are not explicit. One of the major difficulties is to solve dynamic programming equations, which involves computing conditional expectations at each step across the time interval. This computation can be very costly in high-dimensional problems. As introduced above, the purpose of this paper is to develop a new probabilistic numerical scheme to solve Markovian forward-backward SDEs. We will discuss the particular case where the terminal condition is assumed to be a Gaussian functional. In our class, by developing the solution of a Markovian BSDE as a Fourier–Hermite expansion, we show that the problem of solving a BSDE is identical to solving an infinite countable system of ordinary differential equations (CODEs). The family of ODEs belongs to the class of stiff ODEs, where the associated functional is one-sided Lipschitz. On this basis, we derive a numerical algorithm for the BSDE via the standard Euler scheme, with respect to the solution of the countable system of ordinary differential equations.

This paper is structured as follows. We will introduce the basic theory of BSDEs, the *generalized* Hermite polynomials and develop the solution of Markovian BSDEs as a Fourier–Hermite expansion in a Hilbert space. We will show their connection to countable system of ODEs and derive a numerical algorithm to solve the corresponding BSDE. Finally, we will propose two numerical experiments to illustrate the performance of the scheme.

Notations and assumptions

We will use the notations of El Karoui, Hamadène and Matoussi [20]. We consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$, with $\mathcal{F} = \mathcal{F}_T$, $\mathbb{F} = (\mathcal{F}_t)_{0\le t\le T}$ a complete natural filtration of a d-dimensional Brownian motion W, and T a fixed finite horizon. For all $m \in \mathbb{N}^*$ and $x \in \mathbb{R}^m$, $|x|$ denotes the Euclidean norm of the vector x. For the matrix $A \in \mathbb{R}^{m\times d}$, we define its Frobenius norm by $|A| := \sqrt{\operatorname{Trace}(AA^*)}$. The matrix A can be considered as an element of the space $\mathbb{R}^{m\times d}$.

- We introduce the sets

$$\mathcal{L}^2_m(\mathcal{F}_t) := \{(X_t)_{t\in[0,T]} \in \mathbb{R}^m,\ \mathcal{F}_t\text{-measurable and } \|X\|_{\mathcal{L}^2} = \mathbb{E}[|X_t|^2]^{1/2} < \infty\},$$

$$\mathcal{S}^2(\mathbb{R}^m) := \Big\{(Y_t)_{t\in[0,T]} \in \mathbb{R}^m,\ \text{continuous and adapted such that } \|Y\|^2_{\mathcal{S}^2} = \mathbb{E}\Big[\sup_{t\in[0,T]} |Y_t|^2\Big] < \infty\Big\},$$

$$\mathcal{H}^2(\mathbb{R}^m) := \Big\{(Z_t)_{t\in[0,T]} \in \mathbb{R}^m,\ \text{continuous and adapted such that } \|Z\|^2_{\mathcal{H}^2} = \mathbb{E}\Big[\Big(\int_0^T |Z_s|^2\, ds\Big)\Big] < \infty\Big\}.$$

- All the equalities and inequalities between random variables are understood in the almost sure sense unless explicitly stated otherwise.
- The space $l^2(\mathbb{N}) := \{(x_i)_{i\in\mathbb{N}} : \sum_i |x_i|^2 < \infty\}$, equipped with its natural inner product, is a Hilbert space. We can identify each element of the space $l^2(\mathbb{N})$ as an infinite-dimensional vector. We will use this identification and clarify the cases unless explicitly stated otherwise.
- (x, y) denotes the usual inner product on $\mathbb{R}^m$ or on $l^2(\mathbb{N})$.
- For $x \in \mathbb{R}^m$, we define the gradient operator

$$\nabla_x := \Big(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_m}\Big).$$

2 Definitions and estimates

In this section, we introduce the general concept of backward stochastic differential equations (BSDEs) with respect to a standard Brownian motion. In the last part of this section, we recall classical estimates from the theory of BSDEs. In the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$, BSDEs are a special class of stochastic differential equations. The main difference is that these equations are specified with a prescribed terminal value as shown in the following equation:

$$\begin{cases} -dY_t = g(t, Y_t, Z_t)dt - Z_t\,dW_t, & 0 \le t < T, \\ Y_T = \xi. \end{cases} \tag{2.1}$$

The latter system can be written equivalently as the following stochastic integral

$$Y_t = \xi + \int_t^T g(s, Y_s, Z_s)\,ds - \int_t^T Z_s\,dW_s, \tag{2.2}$$

where

- ξ is the terminal condition of equation (2.1) and is assumed to be an $\mathcal{F}_T$-measurable and a square-integrable random variable,
- the mapping $(t, y, z) \mapsto g(t, y, z)$ is generally called the generator of the driver of (2.1).

A solution of the backward stochastic differential equation (2.1) is a couple of progressively measurable processes (Y, Z) such that:

(i) $\int_0^T |Z_s|^2\,ds < \infty$ and $\int_0^T |g(s, Y_s, Z_s)|\,ds < \infty$,

(ii) (Y_t, Z_t) satisfies equation (2.1).

In general, we do not have a unique solution to equation (2.1). The existence and uniqueness of a solution can be proved under the conditions given in [42], which involves the Lipschitz continuity of the generator function g. In this case,

$$(Y_t, Z_t)_{0\le t\le T} \in \mathcal{S}^2(\mathbb{R}^m) \times \mathcal{H}^2(\mathbb{R}^{m\times d}).$$

Remark 2.1. If the generator function g is identically equal to zero, the BSDE (2.2) is reduced to the following stochastic equation:

$$Y_t = \xi - \int_t^T Z_s\,dW_s.$$

This preceding simplification can be associated with the martingale representation theorem in the filtration generated by the Brownian motion. The process Y is a martingale and we have explicitly

$$Y_t = \mathbb{E}(\xi \mid \mathcal{F}_t).$$

Proposition 2.2 ([41]). *If the function g is continuous globally and Lipschitz in its second and third coordinate, then the couple (Y, Z) satisfies*

$$\mathbb{E}\Big(\sup_{0\le t\le T} |Y_t|^2 + \int_0^T |Z_s|^2\,ds\Big) \le C\mathbb{E}\Big(|\xi|^2 + \int_0^T |g(s, 0, 0)|^2\,ds\Big).$$

The preceding inequality shows how the solution of the BSDE is governed by the terminal condition ξ and the generator function g.

3 BSDEs and Hermite polynomials

The purpose of this paper is to develop a numerical scheme to solve the below Markovian forward-backward SDE (3.1), where the terminal condition is assumed to be a Gaussian functional. A forward backward stochastic differential equation (FBSDE) is a system which consists of two equations: the first is an Itô process and the second is a backward stochastic differential equation. The following BSDE is an example of a FBSDE where the forward component is a standard Brownian motion:

$$\begin{cases} -dY_t = g(t, W_t, Y_t, Z_t)dt - Z_t\, dW_t, & 0 \le t < T, \\ Y_T = \phi(W_T). \end{cases} \tag{3.1}$$

In this section, we will introduce Hermite polynomials, list some useful properties of these polynomials and highlight their connection to BSDEs via the conditional expectation operator. We assume the following general hypotheses:

(H1) There exists a positive constant $K > 0$ such that

$$|g(t_1, x_1, y_1, z_1) - g(t_2, x_2, y_2, z_2)| \le K(|x_1 - x_2| + |y_1 - y_2| + |z_1 - z_2|)$$

and

$$\sup_{t\in[0,T]} |g(t, 0, 0, 0)| \le K.$$

(H2) The function ϕ is Lipschitz.

From [43], one can represent the couple of processes (Y, Z) by the solution u of the following parabolic partial differential equation:

$$Y_t = u(t, W_t) \quad \text{and} \quad Z_t = (\nabla_x u)(t, W_t), \quad t \in [0, T], \tag{3.2}$$

where the function u solves the equation

$$\begin{cases} \dfrac{\partial u}{\partial t}(t, x) + \dfrac{1}{2}\Delta u(t, x) + g(t, x, u, \nabla_x u) = 0, \\ u(T, x) = \phi(x), \quad \text{with } (t, x) \in [0, T] \times \mathbb{R}^d. \end{cases} \tag{3.3}$$

We recall that $\nabla_x u$ denotes the gradient of u and the differential operator Δ denotes the Laplacian operator with respect to the space variable x. Explicitly,

$$\Delta u(t, x) =: \sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2} u(t, x).$$

Under some regularity assumptions, it is known that the previous PDE (3.3) has a bounded unique solution u with a bounded derivative in the space coordinate. This result establishes a direct connection between the solution of equation (3.1) and the solution of the PDE (3.3).

3.1 Hermite polynomials and martingales

Hermite polynomials belong to the family of orthogonal polynomials and appear in many areas such as physics, chemistry, mathematics, etc. These polynomials appear naturally in the study of the propagation of the heat equation and in the study of quantum harmonic oscillator, for instance. The most famous application of Hermite polynomials is in the Schrödinger theory of quantum physics. The system of the probabilists' Hermite polynomials $(H_n(x))_{n\in\mathbb{N}}$ can be easily defined by Rodrigues's formula. For $x \in \mathbb{R}$ and every positive integer n,

$$H_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2}, \quad \text{with } n \ge 1 \text{and } H_0(x) = 1.$$

The components of the sequence $(H_n(x))_{n\in\mathbb{N}}$ are orthogonal polynomials with respect to the Gaussian weight function

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad x \in \mathbb{R}.$$

Hence, for $(n, m) \in \mathbb{N}^2$, any pair $(H_n(x), H_m(x))$ satisfies the orthogonality relationship

$$\int_{-\infty}^{\infty} H_n(x) H_m(x) \phi(x)\, dx = n!\delta_{nm},$$

where δ_{nm} denotes the Kronecker symbol. The first few Hermite polynomials are:

$$\begin{aligned}H_0(x) &= 1,\\ H_1(x) &= x,\\ H_2(x) &= x^2 - 1,\\ H_3(x) &= x^3 - 3x,\\ H_4(x) &= x^4 - 6x^2 + 3.\end{aligned}$$

In general, Hermite polynomials satisfy the following recursion relation:

$$H_{n+1}(x) = xH_n(x) - nH_{n-1}(x). \tag{3.4}$$

This recursion relationship is very useful for generating values of $H_n(x)$ for a given x in a fast way. If we assume that H_n has the following representation:

$$H_n(x) = \sum_{k=0}^{n} a_{n,k}x^k,$$

where $a_{n,k} \in \mathbb{R}$, then one can deduce immediately from (3.4) the individual coefficients $a_{n,k}$ by identification. One can also integrate Hermite polynomials analytically against any Gaussian density. For this reason, we introduce the non-central moments of a Gaussian random variable $Z \sim \mathcal{N}(\mu, \sigma)$, which are given by

$$\int_{-\infty}^{\infty} z^n \frac{e^{-\frac{1}{2}(\frac{z-\mu}{\sigma})^2}}{\sqrt{2\pi\sigma^2}}\,dz = \mathbb{E}[Z^n] = (i\sigma)^n H_n\Big(\frac{\mu}{i\sigma}\Big) \quad \text{for all } n \geq 0, \tag{3.5}$$

and $i^2 = -1$. Although the last expression involves the imaginary unit i, the outcome is always a real number since the complex numbers cancel out for each integer value of n. We introduce a new notation for a "generalized" Hermite polynomial, that is,

$$H_n^{[\theta]}(x) := \theta^{\frac{n}{2}} H_n\Big(\frac{x}{\sqrt{\theta}}\Big), \tag{3.6}$$

which gives (as stated above) also a real value for any $\theta > 0$. Taking the limit as $\theta \to 0$, we obtain the result

$$H_n^{[0]}(x) = x^n. \tag{3.7}$$

The "generalized" Hermite polynomials satisfy the orthogonality relationship

$$\int_{-\infty}^{\infty} H_n^{[\theta]}(x)H_m^{[\theta]}(x)\frac{e^{-\frac{1}{2}\frac{x^2}{\theta}}}{\sqrt{2\pi\theta}}\,dx = \theta^n n!\,\delta_{nm},$$

with respect to the Gaussian density with mean 0 and variance θ. We can use formula (3.5) for the non-central moments to derive the result

$$\int_{-\infty}^{\infty} H_n(z)\frac{e^{-\frac{1}{2}(\frac{z-\mu}{\sigma})^2}}{\sqrt{2\pi\sigma^2}}\,dz = \mathbb{E}\Big[\sum_{k=0}^{n} a_{nk}Z^k\Big] = \sum_{k=0}^{n} a_{n,k}H_k^{[-\sigma^2]}(\mu).$$

In the last inequality, we have extended and used definition (3.6) for imaginary numbers. The last expression can be simplified by using the *umbral composition* formula for (generalized) Hermite polynomials

$$\sum_{k=0}^{n} a_{nk}H_k^{[-\sigma^2]}(\mu) = (H_n^{[1]} \circ H^{[-\sigma^2]})(\mu) = H_n^{[1-\sigma^2]}(\mu).$$

We point out that the last expression also yields the correct answer for $\sigma^2 > 1$. Using the same type of derivation, one can generalize this result to

$$\int_{-\infty}^{\infty} H_n^{[\theta]}(z)\frac{e^{-\frac{1}{2}(\frac{z-\mu}{\sigma})^2}}{\sqrt{2\pi\sigma^2}}\,dz = H_n^{[\theta-\sigma^2]}(\mu). \tag{3.8}$$

For the generalized Hermite polynomials, we have the following addition formula:

$$H_n^{[\theta]}(x+y) = \sum_{k=0}^{n} \binom{n}{k} y^{n-k} H_k^{[\theta]}(x). \tag{3.9}$$

The last formula can be proven directly via a Taylor expansion of $H_n^{[\theta]}(x+y)$. From the generalized Hermite polynomials expression, it is easy to see that $H_n^{[t]}(W_t)$ defines a sequence of martingales. Applying formula (3.8), we obtain immediately that

$$\mathbb{E}[H_n^{[T]}(W_T) \,|\, \mathcal{F}_t] = H_n^{[t]}(W_t), \quad 0 \le t \le T.$$

We will call these *Hermite martingales*. For each Hermite martingale, we obtain the explicit martingale representation formula

$$H_n^{[T]}(W_T) - H_n^{[t]}(W_t) = \int_t^T n H_{n-1}^{[s]}(W_s)\, dW_s,$$

where we have used the fact that $\partial H_n^{[t]}(W_t)/\partial x = nH_{n-1}^{[t]}(W_t)$. For a positive θ, we introduce below the normalized system $(\bar{H}^\theta)_{\theta>0}$. Each element $\bar{H}_n^{[\theta]}$ of the system $(\bar{H}^\theta)_{\theta>0}$ is defined by the following normalization:

$$\bar{H}_n^{[\theta]}(x) := \frac{1}{\sqrt{\theta^n n!}} H_n^{[\theta]}(x), \quad \theta > 0.$$

This system of generalized and normalized Hermite polynomials satisfies

$$\int_{-\infty}^{\infty} \bar{H}_n^{[\theta]}(x) \bar{H}_m^{\theta}(x) \frac{e^{-\frac{1}{2}\frac{x^2}{\theta}}}{\sqrt{2\pi\theta}}\, dx = \delta_{nm}, \quad m, n \in \mathbb{N},\ \theta > 0,$$

and we have the martingale equality

$$\mathbb{E}[\bar{H}_n^{[T]}(W_T) \,|\, \mathcal{F}_t] = \Big(\frac{t}{T}\Big)^{n/2} \bar{H}_n^{[t]}(W_t), \quad 0 \le t \le T. \tag{3.10}$$

It is not natural to obtain the preceding equality when t is equal to zero. In fact, we have

$$\Big(\frac{t}{T}\Big)^{n/2} \bar{H}_n^{[t]}(W_t) = \Big(\frac{1}{Tn!}\Big)^{n/2} H_n^{[t]}(W_t),$$

and the right side of the preceding equality is well defined when t goes to zero due to relation (3.7). The partial derivative with respect to the space variable of the normalized polynomials is given by

$$\partial_x \bar{H}_n^{[\theta]}(x) = \Big(\frac{n}{\theta}\Big)^{1/2} \bar{H}_{n-1}^{[\theta]}(x), \quad \theta > 0.$$

As we are in a Hilbert space, we can express the Gaussian functional random variable $Y(T, W_T)$ as

$$Y(T, W_T) = \sum_{k=0}^{\infty} \alpha_k(T) \bar{H}_k^{[T]}(W_T),$$

where

$$\alpha_k(T) := \mathbb{E}[Y(T, W_T) \bar{H}_k^{[T]}(W_T) H_k^{[T]}(W_T)]$$

and α belongs to the space $l^2(\mathbb{N})$. We also remark that the above series converges in the $\mathbb{L}_1^2(\mathcal{F}_T)$ sense. If we define $Y(t, W_t)$ as the conditional expectation of $Y(T, W_T)$, we obtain, due to the martingale equality (3.10),

$$Y(t, W_t) = \mathbb{E}\bigg[\sum_{k=0}^{\infty} \alpha_k(T) \bar{H}_k^{[T]}(W_T) \,\Big|\, F_t\bigg] = \sum_{k=0}^{\infty} \Big(\frac{t}{T}\Big)^{K/2} \alpha_k(T) \bar{H}_k^{[t]}(W_t)$$

and the relation

$$\alpha_k(t) = \Big(\frac{t}{T}\Big)^{k/2} \alpha_k(T) \quad \text{for every } t \in (0, T].$$

One can interpret this result as follows: for all $t \in (0, T]$ the coefficients $\alpha_k(t)$ of the conditional expectation process $Y(t, W_t)$ trace a *deterministic* path in our "Hermite space". Even though the random variable $Y(t, W_t)$ is stochastic, the coefficients $\alpha_k(t)$ are deterministic functions of time.

3.2 BSDEs in Hermite spaces

In this section and later on, we will consider the solution of the BSDE (3.1) in a one-dimensional framework. Each process of the couple (Y, Z) is Markovian in the state (t, W_t). In the normalized Hermite basis system $(\bar{H}^t)_{t\in[0,T]}$, we represent the solution of the BSDE (3.1) as the following series, for a fixed time instance t in the interval $[0, T]$:

$$Y_t = \sum_{k\geq 0} \alpha_k(t)\bar{H}_k^{[t]}(W_t), \quad Z_t = \sum_{k\geq 0} \beta_k(t)\bar{H}_k^{[t]}(W_t) = \sum_{k\geq 0}\Big(\frac{k+1}{t}\Big)^{1/2}\alpha_{k+1}(t)\bar{H}_k^{[t]}(W_t), \tag{3.11}$$

where

$$\alpha, \beta \in l^2(\mathbb{N}), \quad \alpha_k(t) = \mathbb{E}\big[Y_t\bar{H}_k^{[t]}(W_t)\big] \quad \text{and} \quad \beta_k(t) = \Big(\frac{k+1}{t}\Big)^{1/2}\alpha_{k+1}(t).$$

We will analyze the removable singularity problem (at $t = 0$) in Section 4.2.2. The expression of β_k comes from equality (3.2), where the process Z is given by the spatial derivative of the function u in (3.2). Furthermore, $g(t, W_t, Y_t, Z_t)$ can also be decomposed as

$$g(t, W_t, Y_t, Z_t) = \sum_{k=0}^{\infty} \gamma_k(t)\bar{H}_k^{[t]}(W_t),$$

where $\gamma_k(t) := \mathbb{E}[g(t, W_t, Y_t, Z_t)\bar{H}_k^{[t]}(W_t)]$. For each $t \in [0, T]$, the coefficients $\gamma_k(t)$ are deterministic functions (via Y_t and Z_t) of the $\alpha(t)$. We will denote these functions by $\gamma_k(t, \alpha(t))$ to highlight the dependence with respect to the coefficients $\alpha(t)$. By integrating equation (3.1) on the interval $[t, T]$ and taking the conditional expectation, we have

$$\mathbb{E}(Y_T \mid \mathcal{F}_t) - Y_t + \mathbb{E}\bigg(\int_t^T g(s, W_s, Y_s, Z_s)\, ds \,\Big|\, \mathcal{F}_t\bigg) = 0.$$

By the preceding decomposition of Y, Z and g in the Hermite space, we have

$$\sum_{k=0}^{\infty}\bigg(\frac{\alpha_k(T)}{\sqrt{T^k k!}} - \frac{\alpha_k(t)}{\sqrt{t^k k!}} + \int_t^T \frac{\gamma_k(s)}{\sqrt{s^k k!}}\, ds\bigg)H_k^{[t]}(W_t) = 0.$$

This equation can only be equal to zero for all $H_k^{[t]}$, $k \in \mathbb{N}$, if each of the coefficients in front of the Hermite basis-functions is equal to zero. Therefore, we obtain the result that the $\alpha_k(t)$ must be the solution to the following deterministic integral equation:

$$\Big(\frac{t}{T}\Big)^{k/2}\alpha_k(T) - \alpha_k(t) + \int_t^T \Big(\frac{t}{s}\Big)^{k/2}\gamma_k(s, \alpha(s))\, ds = 0, \quad k = 0, 1, 2, \ldots. \tag{3.12}$$

This system can also be expressed as a (countably infinite) system of ordinary differential equations by the partial differentiation of the latter equality. We then have

$$t\dot{\alpha}_k(t) - \frac{k}{2}\alpha_k(t) + t\gamma_k(t, \alpha(t)) = 0, \quad \text{with } k = 0, 1, 2, \ldots. \tag{3.13}$$

The function $\dot{\alpha}_k$ denotes the time-derivative of the function α_k. System (3.12) is defined on the whole interval $t \in [0, T]$ with the boundary condition $(\alpha_k(T))_{k\in\mathbb{N}}$ at the horizon time T when system (3.13) introduces a removable singularity time instance at zero. We will show in Lemma 4.5 that this singular point is removable for system (3.13). Furthermore, we use the notation $\gamma_k(t, \alpha(t))$ to emphasize that the γ_k are functions of the α_k. The couple (Y, g) can be represented by $(\alpha_k(t), \gamma_k(t))_{k\in\mathbb{N}}$ at the time instance $t \in [0, T]$. The solution (Y_t, Z_t) exists if and only if the countable differential system (3.12) has a solution. Countable systems of ordinary differential equations have been studied extensively. For early references, see [28, 37, 46].[1] For more recent texts, we refer to [18, 45]. The next section will illustrate and clarify our methodology with two examples.

1 It is interesting to note that [37, Part III] studies the existence and uniqueness of solutions of semilinear partial differential equations $u_t = u_{xx} + g(t, x, u, u_x)$, which we would now identify with BSDEs. Lewis considers the Fourier basis to express the solution in terms of a CODE.

3.3 Preliminaries examples

3.3.1 Quadratic case

In the spirit of [36], let us consider the example of the BSDE (3.1), where the driver is given by the quadratic function

$$g(t,x,y,z) = az^2, \quad a \in \mathbb{R}.$$

The preceding work is rich enough to consider this example. The couple (Y, Z) solves the BSDE

$$Y_t = \phi(W_T) + a\int_t^T Z_s^2\,ds - \int_t^T Z_s\,dW_s.$$

By the change of variable $\bar{Y}_t = \exp(2aY_t)$, there exists a process $\bar{Z}$ such that

$$\bar{Y}_t = \exp(2a\phi(W_T)) - \int_t^T \bar{Z}_s\,dW_s.$$

The solution $(\bar{Y}_t, \bar{Z}_t)_{t\in[0,T]}$ of the corresponding BSDE can be represented in the system the normalized Hermite basis $(\bar{H}^t)_{t\in(0,T]}$. Therefore, the countable systems of ordinary differential equations to solve are linear and given by

$$\dot{\alpha}_k(t) = \frac{k}{2t}\alpha_k(t), \quad k = 0,1,2,\ldots,$$

with the terminal boundary condition $(\alpha_k(T))_{k\in\mathbb{N}}$ at T. By solving the above linear CODEs, we can represent $\bar{Y}$ on the whole time interval $[0, T]$ as

$$\bar{Y}_t = \sum_{k=0}^{\infty} \alpha_k(T)\Big(\frac{t}{T}\Big)^{k/2}\bar{H}_j^{[t]}(W_t), \quad t \in [0,T].$$

By the martingale equality (3.10),

$$\bar{Y}_t = \mathbb{E}[\exp(2a\phi(W_T)) \mid \mathcal{F}_t].$$

By the change of variable $\bar{Y}_t = \exp(2aY_t)$, Y_t defines a super-martingale and we finally obtain

$$Y_t = \frac{1}{2a}\log(\mathbb{E}[\exp(2a\phi(W_T)) \mid \mathcal{F}_t]), \quad t \in [0,T].$$

3.3.2 Linear case

Let us consider the example of the BSDE (3.1) where the driver function is given by the linear function $g(t,x,y,z) = ay + bz$, where $a, b \in \mathbb{R}$. The solution $(Y_t, Z_t)_{t\in[0,T]}$ of the corresponding BSDE can be expressed in the system the normalized Hermite basis $(\bar{H}^t)_{t\in(0,T]}$. From (3.11), we obtain the following representation of $y_k(t)$:

$$y_k(t) = a\alpha_k(t) + b\Big(\frac{k+1}{t}\Big)^{1/2}\alpha_{k+1}(t), \quad t \in (0,T].$$

From (3.13), the countable system of ODEs to solve is given by

$$\dot{\alpha}_k(t) = \Big(a - \frac{k}{2t}\Big)\alpha_k(t) + b\Big(\frac{k+1}{t}\Big)^{1/2}\alpha_{k+1}(t), \quad t \in (0,T],\ k = 0,1,2,\ldots,$$

with the terminal boundary condition $(\alpha_k(T))_{k\in\mathbb{N}}$ at the horizon time T. This system of linear CODEs is row-infinite.

Let us consider the truncated system of ordinary differential equations with only N equations ($N > 1$). The $(N+1)$-dimensional vector of functions $\alpha^N(t)$ solves the following finite system of linear ordinary differential equations:

$$\dot{\alpha}_k^N(t) = \Big(a - \frac{k}{2t}\Big)\alpha_k^N(t) + b\Big(\frac{k+1}{t}\Big)^{1/2}\alpha_{k+1}^N(t), \quad k = 0,1,2,\ldots,N, \tag{3.14}$$

with $\alpha_{N+1}^N \equiv 0$. We solve this system backwardly from $k = N$ to $k = 0$ and obtain an explicit solution. We know that $\alpha_k(T) = \mathbb{E}[Y_T \bar{H}_k^{[T]}(W_T)]$.

- For $k = N$, we solve the corresponding ODE and obtain explicitly

$$\alpha_N^N(t) = \Big(\frac{t}{T}\Big)^{N/2} \exp\{a(T-t)\}\alpha_N(T), \quad 0 < t \le T.$$

- For $k = N-1$, the corresponding ordinary differential equation is given by

$$\dot{\alpha}_{N-1}^N(t) = \Big(a - \frac{N-1}{2t}\Big)\alpha_{N-1}^N(t) + b\Big(\frac{N}{t}\Big)^{1/2}\alpha_N^N(t), \quad 0 < t \le T.$$

By the method of variation of parameter, the explicit solution of the preceding ordinary differential equation is given by the following equality, for $t \in (0, T]$:

$$\alpha_{N-1}^N(t) = \Big(\frac{t}{T}\Big)^{(N-1)/2} \exp\{a(T-t)\}\alpha_{N-1}(T) + N\Big(\frac{t}{T}\Big)^{N/2} \exp\{a(T-t)\}\alpha_N(T).$$

We continue to solve system (3.14) backwardly and obtain the explicit solution

$$\alpha_j^N(t) = \sum_{k=j}^N \alpha_k(T)\binom{k}{j}\Big(\frac{t}{T}\Big)^{k/2}(b(T-t))^{k-j}\exp\{a(T-t)\}, \quad j = 0, 1, 2, \dots, N.$$

Based on this truncated system, we obtain the following representation for $Y_t^{(N)}$ (the approximated solution of the corresponding BSDE):

$$Y_t^{(N)} = \sum_{j=0}^N \Big(\sum_{k=j}^N \alpha_k(T)\binom{k}{j}\Big(\frac{t}{T}\Big)^{k/2}(b(T-t))^{k-j}\exp\{a(T-t)\}\Big)\bar{H}_j^{[t]}(W_t).$$

From the addition formula (3.9) and interchanging the order of summation,

$$Y_t^{(N)} = \exp\{a(T-t)\}\sum_{k=0}^N \alpha_k(T)\Big(\frac{t}{T}\Big)^{k/2}\bar{H}_k^{[t]}(W_t + b(T-t)).$$

When N goes to infinity, with the aid of Parseval's identity, it suffices to observe that

$$\sup_{t\in[0,T]}\Big(\frac{t}{T}\Big)^j \alpha_j(T)^2 \le \alpha_j(T)^2, \quad \text{where } \sum_{j=0}^\infty \alpha_j(T)^2 < \infty,$$

to conclude that the truncated solution $Y_t^{(N)}$ converges to the true solution Y_t via a suitable change of probability measure. The convergence is understood in the $\mathbb{L}_1^2(\mathcal{F}_t)$ sense. For these examples, we can remark that the singular time instance $t = 0$ is in fact a removable singular point of system (3.13). We will analyze this singularity problem in Section 4.2.2.

4 Uniqueness, existence and convergence

The existence and the uniqueness of the solution of the BSDE (3.1) is guaranteed by assumptions (H1)–(H2). The books [18] and [45] give sufficient conditions to study the existence of solutions of countable systems of ordinary differential equations (4.1) below or equivalently to system (3.12). In this part, we study the uniqueness of the solution of the countable systems of ordinary differential equations (4.1) below and its truncated solution in a finite-dimensional subspace which defines the Galerkin approximation of the solution of (4.1). In the system of the generalized orthonormal Hermite polynomials $(\bar{H}^t)_{t\in(0,T]}$, we formulate the following countable backward problem:

$$\begin{cases} \dot{\alpha}(t) = f(t, \alpha(t)), & 0 \le t < T, \\ \alpha(T) \text{ is the terminal condition}, \end{cases} \tag{4.1}$$

where $\alpha(T) = (\alpha_k(T))_{k\geq 0}$ and $f(t, \alpha(t))$ denotes an infinite-dimensional vector where each coordinate is defined by $f_k(t, \alpha(t)) = -\frac{k}{2t}\alpha_k(t) + \gamma_k(t, \alpha(t))$ for each $k \in \mathbb{N}$. The same definition is applied to $\alpha(t)$. The study of regular ODEs is well documented and developed in the case of a finite dimension. The classical Lipschitz condition and the Nagumo condition (see, e.g., [40, 44]) are the most known conditions of their studies. In an infinite-dimensional space, the problem of existence and uniqueness of ODEs are more delicate to obtain.

4.1 Uniqueness of solutions

The study of the uniqueness or the existence of the solution of (4.1) is equivalent to the study of the uniqueness or the existence of the solution of system (3.12) introduced previously.

Proposition 4.1. *If the solution of problem* (4.1) *exists, then the solution is unique on the time interval* $[0, T]$.

Proof. We remind that the uniqueness of the solution of system (4.1) is equivalent to the uniqueness of the solution of system (3.12). Let us consider two solutions $(\alpha^i(t))_{i=1,2}$ of the previous system, with $t \in [0, T]$ and $\alpha^1(T) = \alpha^2(T)$. We associate to each $\alpha^i(t)$, the solution Y_t^i of the BSDE (3.1) at the time instance t. Let us define

$$\Delta\alpha_k^{1,2}(t) = \alpha_k^1(t) - \alpha_k^2(t), \quad \Delta\beta_k^{1,2}(t) = \Big(\frac{k+1}{t}\Big)^{1/2}\Delta\alpha_{k+1}^{1,2}(t).$$

From system (3.12),

$$|\Delta\alpha_k^{1,2}(t)| = \bigg|\int_t^T \Big(\frac{t}{s}\Big)^{k/2}(\gamma_k^1(s) - \gamma_k^2(s))\,ds\bigg| = \bigg|\int_t^T \Big(\frac{t}{s}\Big)^{k/2}\mathbb{E}[(g(s, W_s, Y_s^1, Z_s^1) - g(s, W_s, Y_s^2, Z_s^2))\bar{H}_k^{[t]}(W_s)]\,ds\bigg|.$$

By the Lipschitz property of the function g,

$$|\Delta\alpha_k^{1,2}(t)| \leq K\int_t^T \Big(\frac{t}{s}\Big)^{k/2}|\Delta\alpha_k^{1,2}(s)|\,ds + K\int_t^T |\Delta\beta_k^{1,2}(s)|\,ds.$$

From Lemma A.3 and the Cauchy–Schwartz inequality,

$$|\Delta\alpha_k^{1,2}(t)|^2 \leq K^2 T \exp 2(T-t) \times \int_t^T |\Delta\beta_k^{1,2}(s)|^2\,ds. \tag{4.2}$$

By Itô's formula applied to $\Delta Y_t^{1,2} = |Y_t^1 - Y_t^2|$ and the Lipschitz property of g,

$$\mathbb{E}\bigg(|\Delta Y_t^{1,2}|^2 + \int_t^T |\Delta Z_s^{1,2}|^2\,ds\bigg) \leq 2K\int_t^T \mathbb{E}|\Delta Y_s^{1,2}|^2\,ds + 2K\int_t^T \mathbb{E}|\Delta Y_s^{1,2}||\Delta Z_s^{1,2}|\,ds,$$

where $\Delta Z_t^{1,2} = |Z_t^1 - Z_t^2|$. From the latter inequality and Young's inequality ($2ab \leq \frac{1}{\epsilon}a^2 + \epsilon b^2$, $a, b \in \mathbb{R}$), there exists a constant $C > 0$ such that

$$(1 - \epsilon K)\int_t^T \mathbb{E}|\Delta Z_s^{1,2}|^2\,ds \leq K\Big(2 + \frac{1}{\epsilon}\Big)\int_t^T \mathbb{E}|\Delta Y_s^{1,2}|^2\,ds \quad \text{for any } \epsilon > 0.$$

By choosing $\epsilon = 2/K$ in the preceding inequality, there exists a positive constant $C > 0$ such that

$$\sum_{k\geq 0}\int_t^T |\Delta\beta_k^{1,2}(s)|^2\,ds \leq C\int_t^T \sum_{k\geq 0}|\Delta\alpha_k^{1,2}(s)|^2\,ds.$$

By inserting the latter inequality into (4.2), with summation over the variable k, we obtain the estimate

$$\sum_{k\geq 0}|\Delta\alpha_k^{1,2}(t)|^2 \leq CK^2 T \exp 2(T-t) \times \int_t^T \sum_{k\geq 0}|\Delta\alpha_k^{1,2}(s)|^2\,ds.$$

Applying Lemma A.3 to the preceding inequality yields

$$\sum_{k\geq 0}|\Delta\alpha_k^{1,2}(t)|^2 = 0.$$

The latter equality implies that $\alpha_k^1(t) = \alpha_k^2(t)$, $k = 0, 1, 2, \ldots$. The preceding inequality concludes. □

4.2 Existence and convergence of the truncated solutions

As highlighted above, the study of ODEs is well documented in the case of a finite-dimensional system. In an infinite-dimensional system, the problems of existence and uniqueness of ODEs are more challenging. In the case of a Banach space, the study of CODEs is discussed in [28, 37, 46]. With the system $(\bar{H}^t)_{t\in[0,T]}$, let us consider the family of orthogonal projection operators $(\mathcal{P}_n)_{n\geq 1}$ in the span of the first n first normalized basis functions of $(\bar{H}^t)_{t\in[0,T]}$. We formulate the following n-dimensional problem:

$$\begin{cases} \dot{\alpha}^n(t) = \mathcal{P}_n f(t, \alpha^n(t)), & 0 \leq t < T, \\ \alpha^n(T) = \mathcal{P}_n \alpha(T), \end{cases} \tag{4.3}$$

where $\alpha(T) = (\alpha_k(T))_{k\geq 1}$. For every $t \in [0, T]$, $\alpha^n(t) \in \mathbb{R}^n$. The result of the following lemma is the cornerstone of the existence result of the truncation solution of the countable system of ordinary differential equations (4.1). The convergence result is based on [18, Theorem 7.1].

In a finite-dimensional space, let us introduce the following existence theorem.

Theorem 4.2 ([16]). *Let $I = [t_0, T + t_0]$ be an interval of $\mathbb{R}$ and f a continuous function,*

$$f : I \times \mathbb{R}^m \to \mathbb{R}^m, \quad (t, x) \mapsto f(t, x).$$

We assume also that there exists an integrable function ζ on I such that the function f satisfies

$$(f(t, x), x) \leq \zeta(t)(1 + \|x\|^2) \quad \text{for all } (t, x) \in I \times \mathbb{R}^m.$$

Then there exists a global solution of the following Cauchy problem:

$$\begin{cases} \dot{y}(t) = f(t, y(t)), & \text{where } (t, y(t)) \in I \times \mathbb{R}^m, \\ y(t_0) = y_0, & y_0 \in \mathbb{R}^m. \end{cases}$$

4.2.1 Existence results

The following lemma provides a quadratic inequality of the functional f. This inequality is a key ingredient to study the countable ordinary differential equation problem (4.3) or (4.1) .

Lemma 4.3. *On the set $[0, T] \times l^2(\mathbb{N})$ and for $\alpha, \beta \in l^2(\mathbb{N})$, the functional f satisfies the following inequality:*

$$(f(t, \alpha) - f(t, \beta), \alpha - \beta) \leq K\Big(1 + \frac{K}{2}\Big)|\alpha - \beta|^2 \quad \text{for all } t \in [0, T]. \tag{4.4}$$

Proof. For $\alpha, \beta \in l^2(\mathbb{N})$, we define the system

$$\begin{cases} \Delta Y_t^{\alpha,\beta} = Y_t^\alpha - Y_t^\beta, \\ \Delta Z_t^{\alpha,\beta} = Z_t^\alpha - Z_t^\beta, \\ \Delta g_t^{\alpha,\beta} = g(t, W_t, Y_t^\alpha, Z_t^\alpha) - g(t, W_t, Y_t^\beta, Z_t^\beta), \end{cases}$$

where

$$Y_t^\alpha = \sum_{k\geq 0} \alpha_k(t)\bar{H}_k^{[t]}(W_t) \quad \text{and} \quad Z_t^\alpha = \sum_{k\geq 0}\Big(\frac{k+1}{t}\Big)^{1/2}\alpha_{k+1}(t)\bar{H}_k^{[t]}(W_t).$$

The same definition is applied to the couple (Y_t^β, Z_t^β). By the orthogonality property of $(H_k^{[t]})_{k\in\mathbb{N}}$, we have

$$(f(t,\alpha) - f(t,\beta), \alpha - \beta) = -\frac{1}{2}\mathbb{E}(\Delta Z_t^{\alpha,\beta})^2 + \mathbb{E}(\Delta g_t^{\alpha,\beta}\Delta Y_t^{\alpha,\beta}). \tag{4.5}$$

Using the Lipschitz property of the function g, we obtain

$$\mathbb{E}\Delta g_t^{\alpha,\beta}\Delta Y_t^{\alpha,\beta} \le K\mathbb{E}|\Delta Y_t^{\alpha,\beta}|^2 + K\mathbb{E}(|\Delta Z_t^{\alpha,\beta}| \times |\Delta Y_t^{\alpha,\beta}|).$$

By Young's inequality, we have

$$\mathbb{E}(\Delta g_t^{\alpha,\beta}\Delta Y_t^{\alpha,\beta}) - \frac{1}{2}\mathbb{E}|\Delta Z_t^{\alpha,\beta}|^2 \le K\Big(1 + \frac{K}{2}\Big)\mathbb{E}|\Delta Y_t^{\alpha,\beta}|^2.$$

From the previous inequality and (4.5), we obtain

$$(f(t,\alpha) - f(t,\beta), \alpha - \beta) \le K\Big(1 + \frac{K}{2}\Big)|\alpha - \beta|^2.$$

The proof is complete. □

The inequality of the preceding lemma is similar to the monotonicity condition in [17]. In the theory of ODEs, from the quadratic inequality (4.4), we say that the functional f is *one-sided Lipschitz* with a positive coefficient $K(1 + \frac{K}{2})$. The Lipschitz property implies the one-sided Lipschitz property by Cauchy–Schwartz's inequality. The converse is not true in general.

Proposition 4.4. *The solution of problem* (4.3) *exists on the time interval* $[0, T]$.

Proof. The result is a direct application of Theorem 4.2. By adding and subtracting the term $(f(t,0),\alpha)$ to the inner product $(f(t,\alpha),\alpha)$, we have

$$(f(t,\alpha),\alpha) = (f(t,\alpha),\alpha) - (f(t,0),\alpha) + (f(t,0),\alpha).$$

As the functional f is one-sided Lipschitz,

$$(f(t,\alpha),\alpha) \le K\Big(1 + \frac{K}{2}\Big)|\alpha|^2 + |(f(t,0),\alpha)|, \quad \alpha \in l^2(\mathbb{N}).$$

By Cauchy–Schwartz's inequality and assumption (H1),

$$(f(t,\alpha),\alpha) \le K\Big(1 + \frac{K}{2}\Big)|\alpha|^2 + K|\alpha|.$$

Thus, we have

$$(f(t,\alpha),\alpha) \le \Big(2K + \frac{K^2}{2}\Big)|\alpha|^2 + K.$$

We conclude the result directly from Theorem 4.2. □

4.2.2 Convergence of the truncated solution

In order to analyze the convergence of the truncated solution of problem (4.1), we provide below a lemma and two key definitions (from Deimling [18]) to analyze the convergence result of the solution of problem (4.3). It is important to point out that the solution of the BSDE (3.1) is deterministic for $t = 0$. From Section 3, we can represent the solution of the BSDE (3.1) as $Y_t = u(t, W_t)$ for $t \in [0, T]$, where u solves the PDE (3.3). Recall that problem (4.1) or (4.3) introduces a singularity problem at the time instance $t = 0$. In order to solve this singularity problem, we introduce the following Lemma.

Lemma 4.5. *Under assumptions* (H1)–(H2), *the function* $\alpha_k(\cdot)$ *solves the following ordinary differential equation, for all* $(k,t) \in \mathbb{N} \times [0,T]$:

$$\dot{\alpha}_k(t) = f_k(t, \alpha(t)) = -\mathbb{E}\Big(\partial_t F_k(t, W_t) + \frac{1}{2}\partial_{xx}^2 F_k(t, W_t)\Big),$$

where $F_k(t,x) = u(t,x)\bar{H}_k^{[t]}(x)$ *and* u *is the unique solution of the PDE* (3.3).

Proof. The result of the lemma is an immediate consequence of Itô's formula. In line with the preceding work, by convention, we have

$$\dot{\alpha}_k(t) = \lim_{h\mapsto 0} \frac{1}{h}(\alpha_k(t-h) - \alpha_k(t)) = \lim_{h\mapsto 0} \frac{1}{h}\mathbb{E}[u(t-h, W_{t-h})\bar{H}_k^{[t-h]}(W_{t-h}) - u(t, W_t)\bar{H}_k^{[t]}(W_t)].$$

By Ito's formula,

$$\dot{\alpha}_k(t) = -\lim_{h\mapsto 0} \frac{1}{h}\mathbb{E}\bigg(\int_{t-h}^{t} \partial_s F_k(s, W_s) + \frac{1}{2}\partial_{xx}^2 F_k(s, W_s)\,ds + \int_{t-h}^{t} \partial_x F_k(s, W_s)\,dW_s\bigg).$$

As a martingale, the above stochastic integral vanishes with the expectation and we have

$$\dot{\alpha}_k(t) = f_k(t, \alpha(t)) = -\mathbb{E}\Big(\partial_t F_k(t, W_t) + \frac{1}{2}\partial_{xx}^2 F_k(t, W_t)\Big).$$

The proof is complete □

When t is different from zero, we know that the random variable $(W_t/\sqrt{t})$ and the standard Gaussian random variable $\mathcal{N}(0,1)$ has the same probability distribution. It is enough to repeat the arguments in the proof of Lemma 4.5 to obtain

$$\lim_{t\to 0} \dot{\alpha}_k(t) = -\frac{1}{\sqrt{k!}} \lim_{t\to 0} \mathbb{E}\Big[\Big(\partial_t u(t, W_t) + \frac{1}{2}\partial_{xx}^2 u(t, W_t)\Big) H_k(\mathcal{N}(0,1))\Big].$$

From Section 3, we know that u solves the PDE (3.3). Hence,

$$\lim_{t\to 0} \dot{\alpha}_k(t) = \lim_{t\to 0} \mathbb{E}\Big(\frac{H_k(\mathcal{N}(0,1))}{\sqrt{k!}} g(t, W_t, Y_t, Z_t)\Big).$$

By the continuity of γ_k, we have

$$\lim_{t\to 0} f_k(t, \alpha(t)) = \gamma_k(0, \alpha(0)).$$

In conclusion, the singular time instance 0 is in fact a regular singular point.

Definition 4.6 (Class $\mathcal{U}$). A function $\omega\colon (0,a] \times \mathbb{R}^+ \to \mathbb{R}$ is said to be of "*class* $\mathcal{U}$" (or $\omega \in \mathcal{U}$ in short) if for each ϵ, there exist $\delta > 0$, a sequence $t_i \to 0^+$ and a sequence of continuous functions $\rho_i\colon [t_i, a] \to \mathbb{R}^+$ with

$$\rho_i(t_i) \geq \delta t_i, \qquad D^- \rho_i(t) > \omega(t, \rho_i(t)), \quad 0 < \rho_i(t) \leq \epsilon \text{ in } (t_i, a],$$

where

$$D^- \rho_i(t) = \lim_{\tau\to 0^+} \sup_{h\in(0,\tau]} \frac{1}{h}(\rho_i(t) - \rho_i(t-h)).$$

The supremum in the preceding equality is taken on the variable h.

Definition 4.7 (Class $\mathcal{U}_1$). A function $\omega\colon (0,a] \times \mathbb{R}^+ \to \mathbb{R}$ is said to be of "*class* $\mathcal{U}_1$" (or $\omega \in \mathcal{U}_1$ in short) if $\omega \in \mathcal{U}$ and the function ρ_i from the previous definition satisfy the additional condition

$$D^- \rho_i(t) \geq \omega(t, \rho_i(t)) + \delta_i \quad \text{for some } \delta_i > 0.$$

From [18], the Lipschitz case where $\omega(t,x) = Lx$, $L > 0$, and the Nagumo condition $\omega(t,x) = \frac{x}{t}$ are of the class $\mathcal{U}$ and $\mathcal{U}_1$. The following theorem provides a convergence result of the solution of problem (4.3).

Theorem 4.8. *Let us consider the previous family of the orthogonal projection operators* $(\mathcal{P}_n)_{n\geq 1}$ *in the span of the first n basis functions of Section 3.1. The truncated solution* α^n *of system* (4.3) *converges punctually to the true solution on the interval* $[0,T]$ *when n goes to infinity.*

Proof. The proof of the result is based on [18, Theorem 7.1]. Let us make the change of variable $t' = T - t$, where $t \in [0,T]$. We know from Lemma 4.3 that f satisfies the inequality

$$(f(t', \alpha) - f(t', \beta), \alpha - \beta) \leq \omega(t', |\alpha - \beta|)|\alpha - \beta|, \quad t' \in [0,T] \text{ and } \alpha, \beta \in l^2(\mathbb{N}).$$

The linear function $(t', \rho) \mapsto \omega(t', \rho) = K(1 + \frac{K}{2})\rho$ is of the class $\mathcal{U}_1$. Let us define the function $\alpha_n^P(t') = \mathcal{P}_n\alpha(t')$, the error term $\epsilon_n^P(t') = \alpha^n(t') - \alpha_n^P(t')$ and the absolute error $\phi_n(t') = |\epsilon_n^P(t')|$. As the projection operator $\mathcal{P}_n$ is Lipschitz with the Lipschitz constant equal to 1, the preceding quadratic inequality is also satisfied for $\mathcal{P}_n f$. We then have

$$(\mathcal{P}_n f(t', \alpha) - \mathcal{P}_n f(t', \beta), \alpha - \beta) \leq \omega(t', |\alpha - \beta|)|\alpha - \beta|.$$

By the linearity of the inner product, for $h > 0$ and small enough, we have

$$(\epsilon_n^P(t') - \epsilon_n^P(t' - h), \epsilon_n^P(t')) = |\epsilon_n^P(t')|^2 - (\epsilon_n^P(t' - h), \epsilon_n^P(t')).$$

Using Cauchy–Schwartz's inequality, we obtain

$$|\epsilon_n^P(t')|^2 - (\epsilon_n^P(t' - h), \epsilon_n^P(t')) \geq |\phi_n(t')|^2 - \phi_n(t' - h)\phi_n(t').$$

Dividing by h the previous inequality and letting $h \to 0^+$, we have

$$\phi_n(t')D^-\phi_n(t') \leq (\dot{\overbrace{(\epsilon_n^P(t'))}}, \epsilon_n^P(t')),$$

where $\dot{\overbrace{(\epsilon_n^P(t'))}}$ denotes the derivative with respect to t' of the error term $\epsilon_n^P(t')$. From the latter inequality, we obtain

$$\phi_n(t')D^-\phi_n(t') \leq (\mathcal{P}_n f(t', \alpha^n(t')) - \mathcal{P}_n f(t', \alpha_n^P(t')), \epsilon_n^P(t')) + |(f(t', \alpha_n^P(t')) - f(t', \alpha(t'))|\phi_n(t')$$

and

$$\phi_n(t')D^-\phi_n(t') \leq (f(t', \alpha^n(t')) - f(t', \alpha_n^P(t')), \epsilon_n^P(t')) + |(f(t', \alpha_n^P(t')) - f(t', \alpha(t'))|\phi_n(t').$$

Clearly,

$$\phi_n(t')D^-\phi_n(t') \leq \omega(t', \phi_n(t'))\phi_n(t') + |(f(t', \alpha_n^P(t')) - f(t', \alpha(t'))|\phi_n(t'). \tag{4.6}$$

Moreover, the function f_k can be extended to the function $\hat{f}_k$ by continuity as follows:

$$\hat{f}_k(t) = f_k(t)\mathbb{1}_{\{t \neq 0\}} + \gamma_k(0, \alpha(0))\mathbb{1}_{\{t=0\}}, \quad t \in [0, T].$$

By using the function $\hat{f}_k$ and noticing that $\alpha_n^P(t') \to \alpha(t')$ as $n \to \infty$, we have $|(f(t', \alpha_n^P(t')) - f(t', \alpha(t'))| \to 0$ as n goes to ∞ punctually on $(0, T]$. The following convergence holds:

$$\frac{\phi_n(t')}{t'} \to \big|\mathcal{P}_n f(T, \mathcal{P}_n\alpha(T)) - \mathcal{P}_n f(T, \alpha(T))\big| \quad \text{as } t' \to 0^+. \tag{4.7}$$

Furthermore, $|\mathcal{P}_n f(T, \mathcal{P}_n\alpha(T)) - \mathcal{P}_n f(T, \alpha(T))| \to 0$ as n goes to ∞. From the evaluation of the latter result and (4.7), for any $\mu > 0$, there exist n_μ and $t_\mu > 0$ such that for all $n > n_\mu$,

$$\phi_n(t') \leq \Big(\frac{1}{n} + \mu\Big)t' \quad \text{for } t' \in [0, t_\mu].$$

We know that $\omega \in U_1$. For a given $\epsilon > 0$, we choose the constant $\delta > 0$, the sequence t_i, where $t_i < t_\mu$, the sequence ρ_i from Definition 4.7 such that $(\frac{1}{n} + \mu) \leq \frac{\delta}{2}$, for some n_0 such that $n > n_0$. This result implies that

$$\phi_n(t') \leq \frac{\delta}{2}t' \quad \text{for } t' \leq t_\mu.$$

For $t_i < t_\mu$, by considering the function ρ_i form Definition 4.7, we have

$$\phi_n(t_i) < \rho_i(t_i) \quad \text{for } t' \leq t_\mu.$$

Let us consider the time instance

$$t^* = \inf\{s > t_i,\ \phi_n(s) = \rho_i(s)\}.$$

The continuity of ϕ_n and the fact that $\phi_n(t_i) < \rho_i(t_i)$ imply that $\phi_n(t^*) > 0$. From the result (4.6), we have the inequality

$$D^-\phi_n(t^*) \le \omega(t^*, \phi_n(t^*)) + \delta.$$

As $\omega \in \mathcal{U}_1$, given the parameter $\delta_i > 0$ from Definition 4.7, we have, by choosing $\delta \le \delta_i$,

$$D^-\phi_n(t^*) \le \omega(t^*, \phi_n(t^*)) + \delta_i < D^-\phi_n(t^*).$$

The previous double inequality is impossible. Furthermore, before the time instance t^*, $\phi_n(s) < \rho_i(s)$ for $s \in [t_i, t^*)$. Therefore, $\phi_n(t') \le \epsilon$ for every $\epsilon > 0$. Hence, the sequence $\epsilon_n^P(t')$ converges to 0 as n goes to infinity. Since $\alpha_n^P(t')) \to \alpha(t')$ when $n \to \infty$, we conclude that α^n converges to α punctually on $[0, T]$. □

Solving BSDEs can be a challenging task. Theses equations appear in many mathematical problems. Unfortunately, as highlighted in the introduction, for a large class of BSDEs, we do not have an explicit solution. Due to their importance, we need robust approximation schemes to solve these equations. The following section describes in detail our numerical algorithm and provides two numerical experiments to illustrate its performance.

5 Euler scheme and numerical illustrations

Let us consider the unidimensional discrete-time approximation of equation (3.1). We build a partition π of the interval $[0, T]$, defined as follows:

$$\pi\colon 0 = t_0 < \cdots < t_N = T,$$

with $\Delta_i := t_{i+1} - t_i$ and $|\pi| := \max\{\Delta_i : 0 \le i \le N-1\}$. We will use decomposition (3.11) in order to solve the BSDE (3.1) and assume that we have at our disposal the trajectories of the Brownian motion W on the discretization grids of the partition π.

5.1 Euler scheme

We denote $(\bar{Y}, \bar{Z})$ the numerical approximation of the exact solution (Y, Z), deduced from the Euler approximation of the countable system of ODEs (3.13). In the Hermite basis, the couple (Y, Z) is represented by the couple (α, β) and its numerical approximation $(\bar{Y}, \bar{Z})$ by $(\bar{\alpha}, \bar{\beta})$. Following the work of Section 3.2, we have the following decomposition of the couple $(\bar{Y}_{t_i}, \bar{Z}_{t_i})_{t_i \in \pi}$:

$$\bar{Y}_{t_i} = \sum_{k \ge 0} \bar{\alpha}_k(t_i)\bar{H}_k^{[t_i]}(W_{t_i}), \quad \bar{Z}_{t_i} = \sum_{k \ge 0} \bar{\beta}_k(t_i)\bar{H}_k^{[t_i]}(W_{t_i}), \quad \text{with } \bar{\beta}_k(t_i) = \Big(\frac{k+1}{t_i}\Big)^{1/2} \bar{\alpha}_{k+1}(t_i).$$

We also define

$$\bar{\gamma}_{t_i} = \sum_{k \ge 0} \bar{\gamma}_k(t_i)\bar{H}_k^{[t_i]}(W_{t_i}), \quad \text{where } \bar{\gamma}_k(t_i) = \mathbb{E}[g(t_i, W_{t_i}, \bar{Y}_{t_i}, \bar{Z}_{t_i})\bar{H}_k^{[t]}(W_t)].$$

The decomposition of the couple (Y_{t_i}, Z_{t_i}) follows the same structure as in Section 3.2. Let us remark that by decomposition (3.11), the computation of $\bar{\beta}$ is completely determined by $\bar{\alpha}$. Integrating equality (3.12) from t_i to t_{i+1}, we have

$$\alpha_k(t_i) = \Big(\frac{t_i}{t_{i+1}}\Big)^{k/2} \alpha_k(t_{i+1}) + \int_{t_i}^{t_{i+1}} \Big(\frac{t_i}{s}\Big)^{k/2} \gamma_k(s, \alpha(s))\, ds = 0, \quad k = 0, 1, 2, \dots,$$

where $[t_i, t_{i+1}] \subset [0, T]$. It will be enough to compute the couple $(\bar{\alpha}, \bar{\gamma})$ on the discretization grid to have the Euler approximation of the couple (Y, Z). We remark that the computation of Z is completely determined by the computation of Y and its terminal condition.

Description of the algorithm.

- *Initialization*: Approximate $\bar{Y}_T = \phi(W_T)$ and for $k = 0, 1, 2, \dots$, compute the coefficients
$$\bar{\alpha}_k(T) = \alpha_k(T), \quad \bar{\beta}_k(T) = \beta_k(T) = \alpha_{k+1}(T)\Big(\frac{k+1}{T}\Big)^{1/2}.$$
- For $i = (N-1)$ to 0, on each sub-interval $[t_i, t_{i+1}] \subset [0, T]$, with $t_i, t_{i+1} \in \pi$,
  - compute $\bar{y}^{\star}_{t_{i+1}}$ by the following optimization problem: find
  $$\bar{y}^{\star}_{t_{i+1}} = (\bar{y}_1(t_{i+1}), \bar{y}_2(t_{i+1}), \bar{y}_2(t_{i+1}), \bar{y}_3(t_{i+1}), \dots)$$
  such that
  $$J(\bar{y}^{\star}_{t_{i+1}}) = \inf_{\xi} \mathbb{E}\big|\xi \bar{H}_i(W_{t_{i+1}}) - g(t_{i+1}, W_{t_{i+1}}, \bar{Y}_{t_{i+1}}, \bar{Z}_{t_{i+1}})\big|^2,$$
  where $\bar{H}_i := (\bar{H}_0^{[t_{i+1}]}, \bar{H}_1^{[t_{i+1}]}, \bar{H}_2^{[t_{i+1}]}, \dots)$,
  - for $k = 0, 1, 2, \dots$, compute
  $$\bar{\alpha}_k(t_i) = \Big(\frac{t_i}{t_{i+1}}\Big)^{k/2} \bar{\alpha}_k(t_{i+1}) + \Delta_i\Big(\frac{t_i}{t_{i+1}}\Big)^{k/2} \bar{y}_k(t_{i+1}) = 0, \quad \bar{\beta}_k(t_i) = \bar{\alpha}_{k+1}(t_i)\Big(\frac{k+1}{t_i}\Big)^{1/2},$$
  - compute
  $$\bar{Y}_{t_i} = \sum_{k \geq 0} \bar{\alpha}_k(t_i)\bar{H}_k^{[t_i]}(W_{t_i}) \quad \text{and} \quad \bar{Z}_{t_i} = \sum_{k \geq 0} \bar{\beta}_k(t_i)\bar{H}_k^{[t_i]}(W_{t_i}).$$

The couple of coefficients $(\bar{\alpha}, \bar{y})$ denotes the Euler approximation of the couple (α, y).

5.2 Convergence result

From Lemma 4.3, we recall that the functional f is not Lipschitz and satisfies only the quadratic inequality (4.4). We will analyze the convergence results via the corresponding CODEs by combining standard probabilistic techniques.

Theorem 5.1. *Under assumptions* (H1)–(H2) *and considering the previous subdivision π of the interval* $[0, T]$, *there exists a positive constant C, independent of the partition π, such that*
$$\max_{0 \leq i < N} \mathbb{E}|Y_{t_i} - \bar{Y}_{t_i}|^2 + \mathbb{E} \sum_{i=0}^{N-1} \int_{t_i}^{t_i+1} |Z_t - \bar{Z}_{t_i}|^2 \, ds \leq C|\pi|.$$

Proof. The proof of the theorem is deduced from the couples (α, β) and $(\bar{\alpha}, \bar{\beta})$, which represent, respectively, the couples (Y, Z) and $(\bar{Y}, \bar{Z})$. During the proof, the strictly positive constant C may take different values from line to line, but independent of the partition π. With $0 \leq i \leq N-1$, we define for every $t_i \in \pi$ and $s \in [t_i, t_{i+1}]$, the quantities
$$\begin{aligned} \Delta\alpha_k(t_i) &= \alpha_k(t_i) - \bar{\alpha}_k(t_i), \\ \Delta\beta_k(s, t_i) &= \beta_k(s) - \bar{\beta}_k(t_i), \\ \Delta y(s, t_i) &= y_k(s) - \bar{y}_k(t_i), \end{aligned}$$
where $\bar{\beta}_k(t_i) = (\frac{k+1}{t_i})^{1/2}\bar{\alpha}_{k+1}(t_i)$ and $\beta_k(t) = (\frac{k+1}{t})^{1/2}\alpha_{k+1}(t)$. We also define the positive quantity
$$U_i = \sum_{k \geq 0} |\Delta\alpha_k(t_i)|^2 + \int_{t_i}^{t_{i+1}} |\Delta\beta_k(s, t_i)|^2 \, ds.$$

From the following system:
$$\begin{cases} \alpha_k(t_i) = \Big(\dfrac{t_i}{t_{i+1}}\Big)^{k/2} \alpha_k(t_{i+1}) + \displaystyle\int_{t_i}^{t_{i+1}} \Big(\frac{t_i}{s}\Big)^{k/2} y_k(s) \, ds, \\ \bar{\alpha}_k(t_i) = \Big(\dfrac{t_i}{t_{i+1}}\Big)^{k/2} \bar{\alpha}_k(t_{i+1}) + \Delta\Big(\dfrac{t_i}{t_{i+1}}\Big)^{k/2} \bar{y}_k(t_{i+1}), \end{cases}$$

we deduced that

$$\alpha_k(t_i) - \bar{\alpha}_k(t_i) = (\alpha_k(t_{i+1}) - \bar{\alpha}_k(t_{i+1}))\Big(\frac{t_i}{t_{i+1}}\Big)^{k/2} + \int_{t_i}^{t_{i+1}} \Big(\Big(\frac{t_i}{s}\Big)^{k/2}\gamma_k(s) - \Big(\frac{t_i}{t_{i+1}}\Big)^{k/2}\bar{\gamma}_k(t_{i+1})\Big)\,ds.$$

By adding and subtracting in the integral term of the above equality the term $(\frac{t_i}{s})^{k/2}\bar{\gamma}_k(t_{i+1})$ and using Jensen's inequality, we get

$$|\Delta\alpha_k(t_i)| \le |\Delta\alpha_k(t_{i+1})| + \int_{t_i}^{t_{i+1}} \Big(\frac{t_i}{s}\Big)^{k/2} |(\gamma_k(s) - \bar{\gamma}_k(t_{i+1}))|\,ds + \int_{t_i}^{t_{i+1}} \Big|\Big(\frac{t_i}{s}\Big)^{k/2} - \Big(\frac{t_i}{t_{i+1}}\Big)^{k/2}\Big| |\bar{\gamma}_k(t_{i+1})|\,ds.$$

By the inequality $ab \le \frac{1}{2\epsilon}a^2 + \frac{1}{2}\epsilon b^2$, $\epsilon > 0$, we have

$$\begin{aligned}|\Delta\alpha_k(t_i)|^2 + \int_{t_i}^{t_{i+1}} |\Delta\beta_k(s,t_i)|^2\,ds &\le \Big(1 + \frac{\Delta_i}{\epsilon}\Big)|\Delta\alpha_k(t_{i+1})|^2 + 2\Delta_i|\Delta\beta_k(t_i,t_i)|^2\\ &\quad + \Big(1 + \frac{\epsilon}{\Delta_i}\Big)\Big(\int_{t_i}^{t_{i+1}} \Big|\Big(\frac{t_i}{s}\Big)^{k/2}(\gamma_k(s) - \bar{\gamma}_k(t_{i+1}))\Big|\,ds\\ &\quad + \int_{t_i}^{t_{i+1}} \Big|\Big(\frac{t_i}{s}\Big)^{k/2} - \Big(\frac{t_i}{t_{i+1}}\Big)^{k/2}\Big| |\bar{\gamma}_k(t_{i+1})|\,ds\Big)^2 + 2\int_{t_i}^{t_{i+1}} |\beta_k(s) - \beta_k(t_i)|^2\,ds.\end{aligned}$$

By noticing that $\gamma_k(s) - \bar{\gamma}_k(t_{i+1}) = \gamma_k(s) - \gamma_k(s+\Delta_i) + \Delta\gamma_k(s+\Delta_i, t_{i+1})$ and using the Hölder inequality, there exists a positive generic constant C such that for every $\epsilon > 0$,

$$\begin{aligned}|\Delta\alpha_k(t_i)|^2 + \int_{t_i}^{t_{i+1}} |\Delta\beta_k(s,t_i)|^2\,ds &\le \Big(1 + \frac{\Delta_i}{\epsilon}\Big)|\Delta\alpha_k(t_{i+1})|^2 + 2\Delta_i|\Delta\beta_k(t_i,t_i)|^2\\ &\quad + 2\int_{t_i}^{t_{i+1}} |\beta_k(s) - \beta_k(t_i)|^2\,ds + (C\Delta_i + C\epsilon)\Big(\int_{t_i}^{t_{i+1}} |\gamma_k(s) - \gamma_k(s+\Delta_i)|^2\,ds\Big)\\ &\quad + (C\Delta_i + C\epsilon)\Big(\int_{t_i}^{t_{i+1}} |\Delta\gamma_k(s+\Delta_i, t_{i+1})|^2\,ds + \Delta_i^2|\bar{\gamma}_k(t_{i+1})|^2\Big).\end{aligned} \tag{5.1}$$

Remark 5.2. The following hold:

$$\int_{t_i}^{t_{i+1}} |\Delta\gamma_k(s+\Delta_i, t_{i+1})|^2\,ds = \int_{t_{i+1}}^{t_{i+2}} |\Delta\gamma_k(s, t_{i+1})|^2\,ds, \tag{5.2}$$

$$|\gamma_k(s) - \bar{\gamma}_k(t_{i+1})| \le K(|\alpha_k(s) - \alpha_k(t_{i+1})| + |\Delta\alpha_k(t_{i+1})| + |\Delta\beta_k(s, t_{i+1})|). \tag{5.3}$$

Equation (5.2) is obtained from the Lipschitz property of the driver function g. Inserting (5.2) and (5.3) into inequality (5.1), there exists a positive constant C such that for every $\epsilon > 0$ and Δ_i small enough, we have

$$\begin{aligned}&|\Delta\alpha_k(t_i)|^2 + \int_{t_i}^{t_{i+1}} |\Delta\beta_k(s,t_i)|^2\,ds\\ &\quad\le c_1^{i,\epsilon}|\Delta\alpha_k(t_{i+1})|^2 + 2\Delta_i|\Delta\beta_k(t_i,t_i)|^2 + 2\int_{t_i}^{t_{i+1}} |\beta_k(s) - \beta_k(t_i)|^2\,ds + c_2^{i,\epsilon}\Big(\int_{t_i}^{t_{i+1}} |\gamma_k(s) - \gamma_k(s+\Delta_i)|^2\,ds\Big)\\ &\qquad + c_2^{i,\epsilon}\Big(\int_{t_{i+1}}^{t_{i+2}} |\alpha_k(s) - \alpha_k(t_{i+1})|^2\,ds + \int_{t_{i+1}}^{t_{i+2}} |\Delta\beta_k(s, t_{i+1})|^2\,ds + \Delta_i^2|\bar{\gamma}_k(t_{i+1})|^2\Big),\end{aligned} \tag{5.4}$$

where $c_1^{i,\epsilon} = (1 + \Delta_i/\epsilon + C\Delta_i(\Delta_i + \epsilon))$ and $c_2^{i,\epsilon} = (C\Delta_i + C\epsilon)$.

From inequality (5.4), taking $\epsilon = \frac{1}{C}$ and π small enough, there exists a positive constant C, independent of π, such that

$$U_i \le (1+\Delta_i C)U_{i+1} + (C|\pi|+1)\sum_{k\ge 0}\int_{t_i}^{t_{i+1}} |\gamma_k(s)-\gamma_k(s+\Delta_i)|^2\,ds + 2\sum_{k\ge 0}\Delta_i|\Delta\beta_k(t_i,t_i)|^2$$
$$+ 2\sum_{k\ge 0}\int_{t_i}^{t_{i+1}} |\beta_k(s)-\beta_k(t_i)|^2\,ds + (C|\pi|+1)\sum_{k\ge 0}\bigg(\int_{t_{i+1}}^{t_{i+2}} |\alpha_k(s)-\alpha_k(t_{i+1})|^2\,ds + \Delta_i^2|\bar{\gamma}_k(t_{i+1})|^2\bigg).$$

By the Lipschitz property of g, we deduced the following decomposition:

$$|\gamma_k(s)-\gamma_k(s+\Delta_i)| \le K\big(|\alpha_k(s)-\alpha_k(s+\Delta_i)| + |\beta_k(s)-\beta_k(s+\Delta_i)|\big).$$

We then have

$$U_i \le (1+\Delta_i C)U_{i+1} + C(|\pi|+1)\sum_{k\ge 0}\int_{t_i}^{t_{i+1}} |\beta_k(s)-\beta_k(s+\Delta_i)|^2\,ds + |\alpha_k(s)-\alpha_k(s+\Delta_i)|^2\,ds$$
$$+ 2\sum_{k\ge 0}\Delta_i|\Delta\beta_k(t_i,t_i)|^2 + 2\sum_{k\ge 0}\int_{t_i}^{t_{i+1}} |\beta_k(s)-\beta_k(t_i)|^2\,ds$$
$$+ (C|\pi|+1)\sum_{k\ge 0}\bigg(\int_{t_{i+1}}^{t_{i+2}} |\alpha_k(s)-\alpha_k(t_{i+1})|^2\,ds + \Delta_i^2|\bar{\gamma}_k(t_{i+1})|^2\bigg).$$

By Lemma A.2 applied to the preceding inequality, there exists a constant $C > 0$ such that for $|\pi|$ small enough,

$$\max_{0\le i\le N} U_i \le \sum_{k\ge 0}|\alpha_k(T)-\bar{\alpha}_k(T)|^2 + \sum_{i=0}^{N-1} A_{t_i}(\alpha) + B_{t_i}(\beta) + 2|\pi|\sum_{i=0}^{N-1}\sum_{k\ge 0}|\Delta\beta_k(t_i,t_i)|^2, \tag{5.5}$$

where

$$B_{t_i}(\beta) = C(|\pi|+1)\sum_{k\ge 0}\int_{t_i}^{t_{i+1}} |\beta_k(s)-\beta_k(s+\Delta_i)|^2\,ds + 2\sum_{k\ge 0}\int_{t_i}^{t_{i+1}} |\beta_k(s)-\beta_k(t_i)|^2\,ds$$

and

$$A_{t_i}(\alpha) = C(|\pi|+1)\sum_{k\ge 0}\int_{t_i}^{t_{i+1}} |\alpha_k(s)-\alpha_k(s+\Delta_i)|^2\,ds + (C|\pi|+1)\sum_{k\ge 0}\bigg(\int_{t_i}^{t_{i+1}} |\alpha_k(s)-\alpha_k(t_{i+1})|^2\,ds + \Delta_i^2|\bar{\gamma}_k(t_{i+1})|^2\bigg).$$

Remark 5.3.
- By Bessel's inequality and with the aid of [48, Lemma 3.2], there exists a constant $C > 0$ such that for $|\pi|$ small enough,
$$\sum_{i=0}^{N-1} A_{t_i}(\alpha) \le C|\pi|.$$
- Using the L_2-regularity result (see [48, Lemma 3.1]), there exists $C > 0$ such that for $|\pi|$ small enough,
$$\sum_{i=0}^{N-1} B_{t_i}(\beta) \le C|\pi|.$$

From the previous remark and Lemma A.2, we derive from (5.5)

$$\max_{0\le i\le N}\sum_{k\ge 0}|\alpha_k(t_i)-\bar{\alpha}_k(t_i)|^2 \le C\sum_{k\ge 0}|\alpha_k(T)-\bar{\alpha}_k(T)|^2 + C|\pi| + 2|\pi|\sum_{i=0}^{N-1}\sum_{k\ge 0}|\Delta\beta_k(t_i,t_i)|^2.$$

From inequality (5.1) and choosing $\epsilon = \frac{1}{2C}$, there exists a positive constant C such that

$$
\begin{aligned}
U_{i-1} + \frac{1}{2}\sum_{k\geq 0}\int_{t_i}^{t_{i+1}} |\beta_k(s) - \bar{\beta}_k(t_i)|^2\,ds \leq (1 + C\Delta_i)U_i &+ 2\sum_{k\geq 0}\int_{t_{i-1}}^{t_i} |\beta_k(s) - \beta_k(t_{i-1})|^2\,ds \\
&+ c_i\bigg(\int_{t_{i-1}}^{t_i} |\gamma_k(s) - \gamma_k(s+\Delta_i)|^2\,ds\bigg) + \sum_{k\geq 0}\Delta_i|\Delta\beta_k(t_{i-1}, t_{i-1})|^2 \\
&+ c_i\bigg(\sum_{k\geq 0}\int_{t_i}^{t_{i+1}} |\alpha_k(s) - \alpha_k(t_i)|^2\,ds + \Delta_i^2|\bar{\gamma}_k(t_i)|^2\bigg),
\end{aligned}
$$

where $c_i = (C\Delta_i + 1)$.

Summing both sides of the preceding inequality over the variable i from 1 to $N-1$ and from Remark 5.2, there exists a positive constant $C > 0$ such that

$$
\sum_{i=1}^{N-1} U_{i-1} + \frac{1}{2}\sum_{i=1}^{N-1}\sum_{k\geq 0}\int_{t_i}^{t_{i+1}} |\beta_k(s) - \bar{\beta}_k(t_i)|^2\,ds \leq \sum_{i=1}^{N-1}(1 + C\Delta_i)U_i + 2|\pi|\sum_{i=1}^{N-1}\sum_{k\geq 0}|\Delta\beta_k(t_{i-1}, t_{i-1})|^2 + C|\pi|.
$$

Applying inequality (5.5) to the latter inequality, there exists a constant $C > 0$, independent of π, such that

$$
\sum_{i=1}^{N-1}\sum_{k\geq 0}\int_{t_i}^{t_{i+1}} |\beta_k(s) - \bar{\beta}_k(t_i)|^2\,ds \leq C|\pi| + C\sum_{k\geq 0}|\alpha_k(T) - \bar{\alpha}_k(T)|^2 + 2|\pi|\sum_{i=1}^{N-1}\sum_{k\geq 0}|\Delta\beta_k(t_{i-1}, t_{i-1})|^2. \tag{5.6}
$$

From (5.6) and (5.5), the theorem follows. □

5.3 Applications and numerical illustrations

In this section, we illustrate our scheme with two examples. As highlighted above, several problems in finance or in insurance can be formulated as a solution of BSDEs. For realistic applications in finance and insurance, we refer to [19, 21] and the references therein for further details.

5.3.1 Application 1

The solution of problem (4.1) in Section 4, dwells in the infinite-dimensional space. For numerical purposes, it is desirable to consider the solution into a finite-dimensional space. In our numerical implementation, we consider the orthogonal projection operator $(\mathcal{P}_k)_{k\geq 1}$ in the span of the k first orthonormal basis functions as introduced in Section 3.2.

Let us consider the unidimensional discrete-time approximation of equation (3.1). We build a partition π of the interval $[0, T]$ defines as follows:

$$
0 = t_0 < t_1 < \ldots < t_N = T, \quad \Delta_i := t_{i+1} - t_i \quad \text{and} \quad |\pi| = \max\{\Delta_i : 0 \leq i \leq N-1\}.
$$

For our numerical simulation, we define the following parameters:
- k is the number of the basis functions,
- M is the number of simulated paths of the Brownian motion,
- N is the number of the discretization instances on π.

Due the propagation of the error during the backward approximation of the solution of the BSDE, we will be interested in the initial value of the solution. We will assume that we have at our disposal all the trajectories of the Brownian motion at the time instances of the partition π.

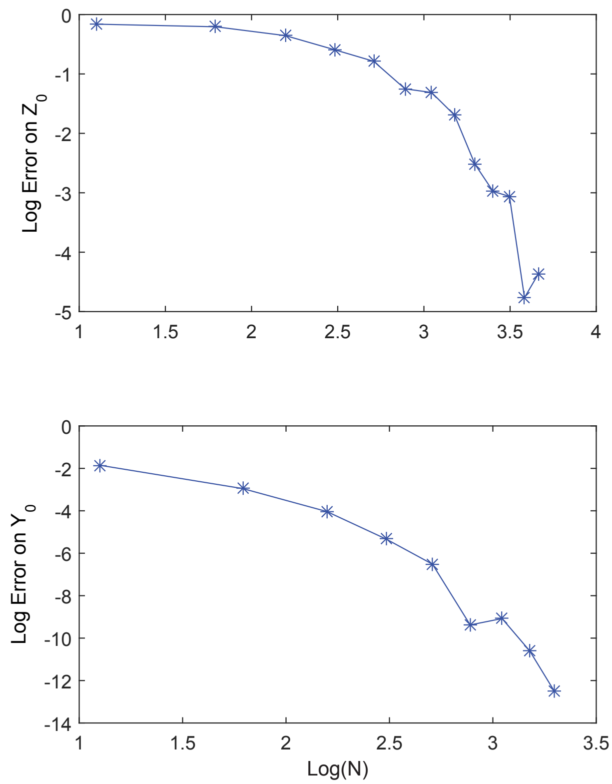

Figure 1: Error curve on (Y_0, Z_0), with $k = 7$, $M = 10000$.

The first example is defined by the following BSDE, inspired from the work of Ruijter and Oosterlee [29]. The underlying process is the Brownian motion $(W_t)_{0\le t\le T}$. We consider the system

$$\begin{cases} -dY_t = g(t, W_t, Y_t, Z_t)\,dt - Z_t\,dW_t, & 0 \le t < 1, \\ Y_1 = \phi(W_1), \end{cases}$$

where the functions g and ϕ are defined as follows;

$$\phi(x) = \cos(x + 1), \qquad x \in \mathbb{R},$$
$$g(t, X_t, Y_t, Z_t) = Z_t(Y_t + 1) - \frac{1}{2}(Y_t - \sin(2(t + W_t)) + \cos(t + W_t)).$$

The exact unique solution of the above BSDE is defined almost surely by the couple

$$(Y_t, Z_t) = (\cos(W_t + t), -\sin(W_t + t)).$$

The exact value of the couple (Y, Z) at zero is $(Y_0, Z_0) = (1, 0)$. By the comparison theorem of BSDEs, the couple of processes (Y_t, Z_t) is included in the bounded domain $[-1, 1] \times [-1, 1]$. Figure 1 shows the log-representation of the relative error curve induced by the numerical approximation of (Y_0, Z_0).

5.3.2 Application 2

The second example is defined by the following BSDE. As in the first example, the underlying process is a Wiener process W and the terminal condition is a functional of W. We consider

$$\begin{cases} -dY_t = g(t, W_t, Y_t, Z_t)dt - Z_t\,dW_t, & 0 \le t < 1, \\ Y_1 = \phi(W_1), \end{cases}$$

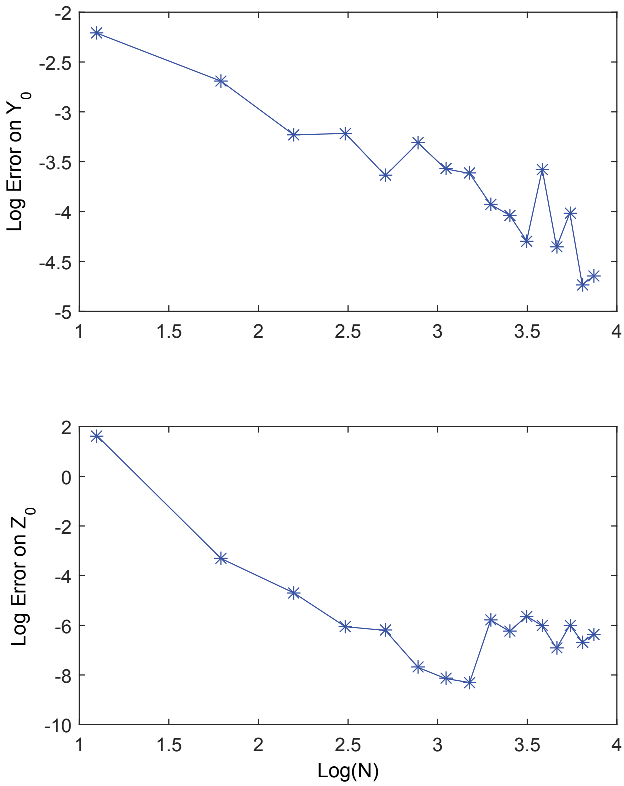

Figure 2: Error curve on (Y_0, Z_0) with $k = 6$, $M = 10000$.

where the terminal condition and the driver functions are defined as follows:

$$\phi(x) = x\arctan(x) - \ln(\sqrt{1+x^2}),$$
$$g(t, W_t, Y_t, Z_t) = -\frac{1}{2(1+\tan^2(Z_t))}.$$

It is easy to check that the solution of the above system is

$$(Y_t, Z_t) = \Big(-\frac{1}{2}\ln(1+W_t^2) + W_t\arctan(W_t), \arctan(W_t)\Big).$$

By noticing that the function $x \mapsto \ln(x)$ satisfies the linear growth condition and the function $x \mapsto \arctan(x)$ is bounded, one has

$$(Y_t, Z_t)_{0\le t\le T} \in \mathcal{S}^2(\mathbb{R}) \times \mathcal{H}^2(\mathbb{R}).$$

Figure 2 shows the log-representation of the relative error curve induced by the numerical approximation of the couple (Y_0, Z_0). Modulo the choice of $|\pi|$ and the number of basis functions k, the numerical illustrations show a stable convergence order regarding the approximation of the couple (Y_0, Z_0). Nonetheless, the convergence regarding the estimation of the initial value Y_0 is more stable and quicker than the approximation of Z_0 in the first example. The convergence of the method could be accelerated by two-step schemes or Runge–Kutta methods (cf., e.g., [2, 10, 30, 34]).

6 Conclusion

This paper covers the numerical approximation of the class of Markovian backward stochastic differential equations (BSDEs), wherein the terminal condition is a functional of Brownian motion. BSDEs appear in many problems in finance, insurance, and their numerical solutions can be challenging to compute especially in high dimensions when several risk factors are involved. The main difficulty is to solve a dynamic programming problem, which involves computing conditional expectations at each step across the time interval. This computation can be very costly especially in high-dimensional problems. In our class, by developing the solution of a Markovian BSDE as a Fourier–Hermite expansion, we show that the problem of solving the BSDE is identical to solving a countable infinite-dimensional system of ordinary differential equations (CODEs). The family of ODEs belongs to the class of stiff ODEs, where the associated functional is one-sided Lipschitz. The use of Hermite polynomials is very useful for calculating the conditional expectations in an exact way during each time-step, thereby eliminating a potential source of error in our algorithm. On this basis, we derive a numerical algorithm for the BSDE via the standard Euler scheme with respect to the solution of the countable system of ordinary differential equations. It is interesting to note the simplicity of our algorithm. The two examples show a stable convergence regarding the computation of the solution of the BSDE. This research could be developed further by investigating non-Markovian cases with applications to pricing problems and risk management issues.

A Appendix

Lemma A.1. *For any $\alpha > 0$ and for any $a, b \in \mathbb{R}$,*

$$(a+b)^2 \leq (1+\alpha)a^2 + \Big(1+\frac{1}{\alpha}\Big)b^2.$$

Proof. The result is a direct consequence of Young's inequality. □

Let us recall the classical discrete Gronwall lemma (see, e.g., [48] or [39]) .

Lemma A.2 (Gronwall inequality (1)). *Let us consider the partition*

$$\pi\colon 0 = t_0 < \cdots < t_N = T$$

of the interval $[0,T]$, and let Δ_i be its mesh. We consider the families $(a_k)_{0\leq k\leq N}$ and $(b_k)_{0\leq k\leq N}$, assumed to be non-negative, such that for some positive constant $\gamma > 0$, we have

$$a_{k-1} \leq (1+\gamma\Delta_i)a_k + b_k, \quad k = 1,\dots,N.$$

Then

$$\max_{0\leq i\leq N} a_i \leq e^{\gamma T}\Big(a_N + \sum_{i=1}^{N} b_i\Big).$$

Lemma A.3 (Gronwall inequality (2)). *Let $y, b, a\colon \mathbb{R}^+ \mapsto \mathbb{R}$ be three continuous functions such that b is non-negative and*

$$y(t) \leq a(t) + \int_0^t b(s)y(s)\,ds.$$

Then

$$y(t) \leq a(t) + \int_0^t a(s)b(s)\exp\Big(\int_s^t b(u)\,du\Big)\,ds.$$

Moreover, if the function a is non-decreasing or monotone, we have

$$y(t) \leq a(t)\exp\Big(\int_0^t b(s)\,ds\Big).$$

Funding: The research leading to these results has received funding from the European Union Seventh Framework Programme ([FP7/2007-2013]), [FP7/2007-2011]) under grant agreement no. 289032.

References

[1] S. Alanko and M. Avellaneda, Reducing variance in the numerical solution of BSDEs, *C. R. Math. Acad. Sci. Paris* **351** (2013), no. 3–4, 135–138.

[2] U. M. Ascher, S. J. Ruuth and R. J. Spiteri, Implicit-explicit Runge–Kutta methods for time-dependent partial differential equations, *Appl. Numer. Math.* **25** (1997), no. 2–3, 151–167-

[3] V. Bally, Approximation scheme for solutions of BSDE, in: *Backward Stochastic Differential Equations* (Paris 1995–1996), Pitman Res. Notes Math. Ser. 364, Longman, Harlow (1997), 177–191.

[4] G. Barles and E. Lesigne, SDE, BSDE and PDE, in: *Backward Stochastic Differential Equations* (Paris 1995–1996), Pitman Res. Notes Math. Ser. 364, Longman, Harlow (1997), 47–80.

[5] C. Bender and M. Kohlmann, Optimal superhedging under non-convex constraints—a BSDE approach, *Int. J. Theor. Appl. Finance* **11** (2008), no. 4, 363–380.

[6] C. Bender and J. Steiner, A posteriori estimates for backward SDEs, *SIAM/ASA J. Uncertain. Quantif.* **1** (2013), no. 1, 139–163.

[7] J.-M. Bismut, Conjugate convex functions in optimal stochastic control, *J. Math. Anal. Appl.* **44** (1973), 384–404.

[8] B. Bouchard and N. Touzi, Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations, *Stochastic Process. Appl.* **111** (2004), no. 2, 175–206.

[9] P. Briand and C. Labart, Simulation of BSDEs by Wiener chaos expansion, *Ann. Appl. Probab.* **24** (2014), no. 3, 1129–1171.

[10] J. Butcher, Runge–Kutta methods for ordinary differential equations, in: *Numerical Analysis and Optimization* (Oman 2014), Springer Proc. Math. Stat. 134, Springer, Cham (2015), 37–58.

[11] J.-F. Chassagneux and A. Richou, Numerical simulation of quadratic BSDEs, *Ann. Appl. Probab.* **26** (2016), no. 1, 262–304.

[12] P. Cheridito, H. M. Soner, N. Touzi and N. Victoir, Second-order backward stochastic differential equations and fully nonlinear parabolic PDEs, *Comm. Pure Appl. Math.* **60** (2007), no. 7, 1081–1110.

[13] D. Chevance, *Résolution numérique des équations différentielles stochastiques rétrogrades*, PhD thesis, Universtité de Provence, 1997.

[14] S. Crépey, R. Gerboud, Z. Grbac and N. Ngor, Counterparty risk and funding: The four wings of the TVA, *Int. J. Theor. Appl. Finance* **16** (2013), no. 2, Article ID 1350006.

[15] D. Crisan and K. Manolarakis, Solving backward stochastic differential equations using the cubature method: Application to nonlinear pricing, *SIAM J. Financial Math.* **3** (2012), no. 1, 534–571.

[16] M. Crouzeix and A. L. Mignot, *Analyse numérique des équations différentielles*, Coll. Math. Appl. Maîtrise, Masson, Paris, 1984.

[17] R. W. R. Darling and E. Pardoux, Backwards SDE with random terminal time and applications to semilinear elliptic PDE, *Ann. Probab.* **25** (1997), no. 3, 1135–1159.

[18] K. Deimling, *Ordinary Differential Equations in Banach Spaces*, Lecture Notes in Math. 596, Springer, Berlin, 1977.

[19] L. U. Delong, *Backward Stochastic Differential Equations with Jumps and Their Actuarial and Financial Applications*, EAA Ser., Springer, London, 2013.

[20] N. El Karoui, S. Hamadène and A. Matoussi, Backward stochastic differential equations and applications, in: *Indifference Pricing: Theory and Applications*, Princeton University, Princeton (2008), 267–320.

[21] N. El Karoui, S. Peng and M. C. Quenez, Backward stochastic differential equations in finance, *Math. Finance* **7** (1997), no. 1, 1–71.

[22] M. Fujii and A. Takahashi, Analytical approximation for non-linear FBSDEs with perturbation scheme, *Int. J. Theor. Appl. Finance* **15** (2012), no. 5, Article ID 1250034.

[23] K. Gnameho, *Numerical solution of backward stochastic differential equations and volatility risk premium*, PhD thesis, Maastricht University, 2016.

[24] K. Gnameho, M. Stadje and A. Pelsser, A regress-later algorithm for backward stochastic differential equations, preprint (2017), https://arxiv.org/abs/1706.07986.

[25] E. Gobet, J.-P. Lemor and X. Warin, A regression-based Monte Carlo method to solve backward stochastic differential equations, *Ann. Appl. Probab.* **15** (2005), no. 3, 2172–2202.

[26] E. Gobet and P. Turkedjiev, Approximation of backward stochastic differential equations using Malliavin weights and least-squares regression, *Bernoulli* **22** (2016), no. 1, 530–562.

[27] B. Gong and H. Rui, One order numerical scheme for forward-backward stochastic differential equations, *Appl. Math. Comput.* **271** (2015), 220–231.

[28] W. L. Hart, Linear differential equations in infinitely many variables, *Amer. J. Math.* **39** (1917), no. 4, 407–424.

[29] T. P. Huijskens, M. J. Ruijter and C. W. Oosterlee, Efficient numerical Fourier methods for coupled forward-backward SDEs, *J. Comput. Appl. Math.* **296** (2016), 593–612.

[30] T. E. Hull, W. H. Enright, B. M. Fellen and A. E. Sedgwick, Comparing numerical methods for ordinary differential equations, *SIAM J. Numer. Anal.* **9** (1972), no. 4, 603–637.

[31] C. B. Hyndman and P. O. Ngou, Global convergence and stability of a convolution method for numerical solution of BSDEs, preprint (2014), https://arxiv.org/abs/1410.8595.

[32] P. Imkeller, A. Réveillac and J. Zhang, Solvability and numerical simulation of BSDEs related to BSPDEs with applications to utility maximization, *Int. J. Theor. Appl. Finance* **14** (2011), no. 5, 635–667.
[33] M. Jeanblanc, T. Mastrolia, D. Possamaï and A. Réveillac, Utility maximization with random horizon: A BSDE approach, *Int. J. Theor. Appl. Finance* **18** (2015), no. 7, Article ID 1550045.
[34] P. Kaps and P. Rentrop, Generalized Runge-Kutta methods of order four with stepsize control for stiff ordinary differential equations, *Numer. Math.* **33** (1979), no. 1, 55–68.
[35] A. Khedher and M. Vanmaele, Discretisation of FBSDEs driven by càdlàg martingales, *J. Math. Anal. Appl.* **435** (2016), no. 1, 508–531.
[36] M. Kobylanski, Backward stochastic differential equations and partial differential equations with quadratic growth, *Ann. Probab.* **28** (2000), no. 2, 558–602.
[37] D. C. Lewis, Jr., Infinite systems of ordinary differential equations with applications to certain second-order partial differential equations, *Trans. Amer. Math. Soc.* **35** (1933), no. 4, 792–823.
[38] J. Ma, P. Protter and J. M. Yong, Solving forward-backward stochastic differential equations explicitly—a four step scheme, *Probab. Theory Related Fields* **98** (1994), no. 3, 339–359.
[39] J. Mémin, S.-g. Peng and M.-y. Xu, Convergence of solutions of discrete reflected backward SDE's and simulations, *Acta Math. Appl. Sin. Engl. Ser.* **24** (2008), no. 1, 1–18.
[40] O. G. Mustafa and D. O'Regan, On the Nagumo uniqueness theorem, *Nonlinear Anal.* **74** (2011), no. 17, 6383–6386.
[41] E. Pardoux, BSDEs, weak convergence and homogenization of semilinear PDEs, in: *Nonlinear Analysis, Differential Equations and Control* (Montreal 1998), NATO Sci. Ser. C Math. Phys. Sci. 528, Kluwer Academic, Dordrecht (1999), 503–549.
[42] E. Pardoux and S. Peng, Adapted solution of a backward stochastic differential equation, *Systems Control Lett.* **14** (1990), no. 1, 55–61.
[43] E. Pardoux and S. Peng, Backward stochastic differential equations and quasilinear parabolic partial differential equations, in: *Stochastic Partial Differential Equations and Their Applications* (Charlotte 1991), Lecture Notes in Control and Inform. Sci. 176, Springer, Berlin (1992), 200–217.
[44] T. Rogers, On Nagumo's condition, *Canad. Math. Bull.* **15** (1972), 609–611.
[45] A. M. Samoilenko and Y. V. Teplinskii, *Countable Systems of Differential Equations*, VSP, Utrecht, 2003.
[46] A. Tychonoff, Über unendliche Systeme von Differentialgleichungen, *Math. Sbornik* **41** (1934), no. 4, 551–560.
[47] W. A. Ventura and A. Korzeniowski, On discretely reflected backward stochastic differential equations, *Stoch. Anal. Appl.* **34** (2016), no. 1, 1–23.
[48] J. Zhang, A numerical scheme for BSDEs, *Ann. Appl. Probab.* **14** (2004), no. 1, 459–488.