Applied Mathematics and Computation 441 (2023) 127666

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

A fully quantization-based scheme for FBSDEs

Giorgia Callegaro [a,*], Alessandro Gnoatto [b], Martino Grasselli [a,c]

[a] University of Padova, Department of Mathematics, via Trieste 63, Padova 35121, Italy
[b] University of Verona, Department of Economics, via Cantarane 24, Verona 37129, Italy
[c] Devinci Research Center, Léonard de Vinci Pôle Universitaire, Paris La Défense 92916, France

ARTICLE INFO

Article history:
Received 17 April 2022
Revised 6 September 2022
Accepted 20 October 2022

JEL classification:
C02
C63

2010 MSC:
65C30
65C40
60H20

Keywords:
FBSDEs
Quantization
Numerical Scheme

ABSTRACT

We propose a quantization-based numerical scheme for a family of decoupled forward-backward stochastic differential equations. We simplify the scheme for the control in [1] so that our approach is fully based on recursive marginal quantization and does not involve any Monte Carlo simulation for the computation of conditional expectations. We analyse in detail the numerical error of our scheme and provide some examples of application to financial mathematics.

1. Introduction and motivation

In this paper we introduce an efficient scheme for the numerical approximation of the solution (Y,U,V) of a family of Forward-Backward Stochastic Differential Equations (FBSDEs hereafter)

$$\begin{cases} Y_t = y_0 + \int_0^t b(Y_s)ds + \int_0^t \sigma(Y_s)^\top dW_s, \quad y_0 \in \mathbb{R}^d \\ U_t = \xi + \int_t^T f(s, Y_s, U_s, V_s)ds - \int_t^T V_s^\top dW_s, \quad t \in [0,T], \end{cases} \tag{1.1}$$

where W is a Brownian motion, $T > 0$ is a deterministic terminal time and the functions b, σ, f, ξ satisfy some conditions specified in the sequel in order to grant that the solution of (1.1) exists and it is unique. FBSDEs of the form (1.1) are particularly popular in financial mathematics: in typical applications, the (forward) process Y describes the evolution of a financial asset, while the (backward) SDE for U is related to the value of the portfolio that hedges the terminal payoff ξ through the trading strategy V. BSDEs allow for the treatment of non-linear pricing problems and this originated their

* Corresponding author.
E-mail addresses: gcallega@math.unipd.it (G. Callegaro), alessandro.gnoatto@univr.it (A. Gnoatto), grassell@math.unipd.it (M. Grasselli).

https://doi.org/10.1016/j.amc.2022.127666

popularity in finance. More recently, in the aftermath of the 2007–2009 financial crisis, the valuation of financial products has been revisited in several aspects, often by means of advanced BSDEs treatments. The possibility of a default of both agents involved in the transaction and the presence of multiple sources of funding are represented at the level of valuation equations by introducing typically non-linear FBSDEs for value adjustments (xVA), see e.g. [2].

The history of BSDEs goes back to [3] and originates from the theory of stochastic optimal control. First existence and uniqueness results have been obtained in the seminal paper of [4] and have been further extended in several directions, including the presence of jumps, see [5] and reflection [6–10]. Applications in mathematical finance are abundant. We refer the reader to [11,12] and the surveys in [13,14] for numerous references and several applications in finance, both in complete and incomplete markets. In view of applications, an important issue concerns the approximation of the solution of a BSDE: the most relevant contribution is based on the dynamic programming approach, introduced by Briand et al. [15] in a Markovian setting. In this case, the rate of convergence for deterministic time discretization has been studied by Zhang [16], who transformed the problem to computing a sequence of conditional expectations. This opened the door to several approaches to attack the problem, as significant progress has been made in computing the conditional expectations: [17] adopted the Malliavin calculus approach, while [12] proposed the linear regression method based on the Least-Squares Monte Carlo approach in [18]. The approach of [19,20] was based on quantization, a technique that will be treated in the sequel as it represents the main source of inspiration for our work. Since then, the literature on BSDE flourished and attained high level of generality, including the non Markovian setting. In the case where the terminal condition is not necessarily Markovian, Briand and Labart [21] proposed a forward scheme based on Wiener chaos expansion, for which conditional expectations can be efficiently computed through the chaos decomposition formula.

In this paper, in view of financial applications to pricing and hedging, we consider the case of decoupled FBSDEs like (1.1) (that is, when the forward SDE for Y does not exhibit a dependence on U) in a Markovian setting. Even in the one dimensional case, the numerical procedures described above require several computations coupled with Monte Carlo simulations, which leads to algorithms that are too time consuming in view of concrete applications.

The aim of our study is to provide a new numerical scheme for the solution of FBSDEs that allows to improve the approximation of the solution of (1.1). We follow the spirit of [19,20], where Pagès and coauthors applied the optimal quantization technique to compute the conditional expectations. We extend their approach by considering an algorithm that is entirely based on fast quantization: in particular, our procedure does not rely on Monte Carlo simulation in any step of the algorithm.

We now give a brief overview on quantization. Quantization of random vectors provides the best possible discrete approximation to the original distribution, according to a distance that is commonly measured using the squared Euclidean norm. Many numerical procedures have been developed to obtain optimal quadratic quantizers of the Gaussian (and even non-Gaussian) distribution in high dimension, mostly based on stochastic optimization algorithms, see [22] and references therein. While theoretically sound and deeply investigated, optimal quantization typically suffers from the numerical burden that the algorithms involve. Indeed, the procedure to be performed to obtain the optimal grids is highly time-consuming, especially in the multi-dimensional case, where stochastic algorithms are necessary. Recursive marginal quantization, or fast quantization, introduced in [23] represents a very useful innovation in order to overcome the above-mentioned computational difficulties. Sub-optimal (stationary) quantizers of the stochastic process at fixed discretization dates (hence, of random variables) are obtained in a very fast recursive way, to the point that recursive marginal quantization has been successfully applied to many models for which a (time) discretization scheme is available, see e.g. the non exaustive list of papers: [1,24–26] for the multi dimensional case. We also mention [27] where recursive quantization has been applied outside the usual Euler scheme.

Here, we propose a scheme for the system (1.1) that is similar to the one in [1], based on recursive quantization, with a crucial difference: in a nutshell, we introduce a new discretization scheme for the control process V that we express in terms of U and Y instead of U and the Brownian motion W (details will be provided in the sequel). This apparently small difference leads to a simpler numerical procedure, as there will be no need to discretize the Brownian motion increments. This reduces the computational time required to solve the FBSDE. In fact, in the approximation of the conditional expectations required in our scheme, we only need the transition probabilities of the quantized process $\widehat{Y}$, while [1] need to additionally compute a conditional expectation involving the Brownian increments, that they estimate via Monte Carlo simulations. Such procedure implies an additional numerical effort which is not required in our case. In other words, once the process Y has been discretized in space via recursive marginal quantization to get $\widehat{Y}$, we apply our backward approximation scheme in order to get an explicit and fully quantization-based algorithm.

We provide two numerical experiments. The first involves a one dimensional BSDE with known analytical solution, so that we can test our approximated solution in a case where there exists a closed form for the control. Here our procedure reveals to be fast and accurate. The same accuracy is obtained when we consider a multi-dimensional example. We treat the case of a basket option, for which a semi-closed form solution for the price is known in terms of a multivariate Fourier integral, see [28].

The rest of the paper is organised as follows: in Section 2 we briefly introduce the FBSDE and we recall the main existence and uniqueness results in order for our working setting to be well-posed. In Section 3 we illustrate our new scheme for the control U. Section 4 provides the essentials on recursive marginal quantization that we apply in Section 5 to the computation of conditional expectations. In Section 6 we study the error, while in Section 7 we illustrate some numerical tests. Section 8 concludes.

2. Forward-Backward stochastic differential equations

We start by fixing some notations. Vectors will be column vectors and, for $x \in \mathbb{R}^d$, $|x|$ denotes the Euclidean norm and $\langle x, y\rangle$ denotes the inner product. Matrices are elements of $\mathbb{R}^{q\times d}$, with $|y| = \sqrt{Trace[yy^\top]}$ and $\langle x, y\rangle = Trace[xy^\top]$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space rich enough to support an $\mathbb{R}^q$-valued Brownian motion $W = (W_t)_{t\in[0,T]}$. Let $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ be the filtration generated by W, assumed to satisfy the standard assumptions. We consider the following spaces:

- L^2 is the space of all $\mathcal{F}_T$-measurable $\mathbb{R}^d$-valued random variables $X : \Omega \mapsto \mathbb{R}^d$ such that $\|X\|^2 = \mathbb{E}\left[|X|^2\right] < \infty$.
- $\mathbb{H}^{2,q\times d}$ is the space of all predictable $\mathbb{R}^{q\times d}$-valued processes $\phi : \Omega \times [0, T] \mapsto \mathbb{R}^{q\times d}$ such that $\mathbb{E}\left[\int_0^T |\phi_t|^2 dt\right] < \infty$.
- $\mathbb{S}^2$ the space of all adapted processes $\phi : \Omega \times [0, T] \mapsto \mathbb{R}^{q\times d}$ such that $\mathbb{E}\left[\sup_{0\le t\le T} |\phi_t|^2\right] < \infty$.

Let $Y = (Y_t)_{t\in[0,T]}$ be an $\mathbb{R}^d$-valued process solving the stochastic differential equation (henceforth SDE):

$$Y_t = y_0 + \int_0^t b(Y_s)ds + \int_0^t \sigma(Y_s)^\top dW_s, \quad y_0 \in \mathbb{R}^d \tag{2.1}$$

and let us consider the following standing assumption:

Assumption 2.1. The vector fields $b : \mathbb{R}^d \mapsto \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \mapsto \mathbb{R}^{q\times d}$ satisfy the following conditions

$$|b(y) - b(z)| \le \mathcal{L}_1|y - z|, \tag{2.2}$$

$$|\sigma(y) - \sigma(z)| \le \mathcal{L}_2|y - z|, \tag{2.3}$$

$$|\sigma(y)| \le \mathcal{L}_3(1 + |y|), \quad |b(y)| \le \mathcal{L}_3(1 + |y|), \tag{2.4}$$

for some positive constants $\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3$.

It is well known that under such regularity conditions there exists a unique adapted right continuous with left limits (henceforth RCLL) strong solution $Y^{y_0} = (Y_t^{y_0})_{t\in[0,T]}$ to (2.1) which is a homogeneous Markov process. It is also well known that the solution Y^{y_0} satisfies the following: for all couples $(t, y_0), (t, y_0') \in [0, T] \times \mathbb{R}^d$ and $p \ge 2$ we have

$$\mathbb{E}\left[\sup_{0\le t\le T} \left|Y_t^{y_0} - y_0\right|^p\right] \le \mathcal{L}_4(1 + |y_0|^p)T, \tag{2.5}$$

$$\mathbb{E}\left[\sup_{0\le t\le T} \left|Y_t^{y_0} - Y_t^{y_0'}\right|^p\right] \le \mathcal{L}_5\left(|y_0 - y_0'|^p\right), \tag{2.6}$$

where $\mathcal{L}_4, \mathcal{L}_5$ are positive constants. To alleviate notations we will simply write Y for the solution, omitting the dependence on the initial condition y_0. We investigate a backward SDE with a terminal condition and a generator that depends on the state process solving the forward SDE (2.1). More precisely, we consider the backward stochastic differential equation

$$U_t = \xi + \int_t^T f(s, Y_s, U_s, V_s)ds - \int_t^T V_s^\top dW_s, \quad t \in [0, T], \tag{2.7}$$

where $V = (V_t)_{t\in[0,T]}$ is a process in $\mathbb{H}^{2,q\times 1}$. We will also work under the following:

Assumption 2.2. (i) The function $f : [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^q \to \mathbb{R}$ is Lipschitz continuous, uniformly in $t \in [0, T]$:

$$\left|f(t, y, u, v) - f\left(t, y', u', v'\right)\right| \le \mathcal{L}_6\left(\left|y - y'\right| + \left|u - u'\right| + \left|v - v'\right|\right)$$

for a positive constant $\mathcal{L}_6$.
(ii) The terminal condition ξ is of the form $\xi = h(Y_T)$, for a given Borel function $h : \mathbb{R}^d \to \mathbb{R}$.

The system formed by the forward SDE (2.1) and the backward SDE (2.7) is a decoupled forward-backward SDE. Decoupled here means that the forward SDE for Y does not exhibit a dependence on U. The following result for FBSDE is standard, see e.g. [29] Theorem 3.1.1, Theorem 4.1.3.

Theorem 2.3. *Under assumptions 2.1 and 2.2 there exists a unique solution* $(Y, U, V) \in \mathbb{S}^2(\mathbb{R}^d) \times \mathbb{S}^2(\mathbb{R}) \times \mathbb{H}^{2,q\times 1}$ *to the FBSDE (2.1)–(2.7).*

In the study of the numerical error in Section 6 we will also need one last assumption, that we include here for completeness:

Assumption 2.4. We have $q=d$ and the matrix $\sigma\sigma^\top$ is uniformly elliptic, namely, for every $y\in\mathbb{R}^d$, denoting by $a_{ij}(y), i,j=1,\dots,d$ the elements of $[\sigma(y)\sigma(y)^\top]$, there exists $\lambda_0>0$ such that

$$\frac{1}{\lambda_0}||\xi||^2 \le \sum_{i,j=1}^{d} a_{ij}(y)\xi_i\xi_j \le \lambda_0||\xi||^2, \quad \xi\in\mathbb{R}^d.$$

Remark 2.5. This ensures that for every $y\in\mathbb{R}^d$, the matrix $\sigma(y)$ is positive definite, hence invertible, and bounded. The inverse matrix $\sigma(y)^{-1}$ is also bounded.

3. A generic scheme for FBSDEs

We introduce here the proposed numerical scheme to approximate the solution of the FBSDE (2.1)–(2.7). To do so, we fix a time discretization: let $n\in\mathbb{N}$, $\Delta=\Delta_n=\frac{T}{n}$ and set $t_k=\frac{Tk}{n}$. The scheme, whose derivation is detailed in Appendix Appendix A is:

$$\begin{cases}\widetilde{U}_{t_n} = h(\overline{Y}_{t_n}) \quad \text{and for} \quad k=0,\dots,n-1\\ \widetilde{U}_{t_k} = \mathbb{E}\left[\widetilde{U}_{t_{k+1}}\middle|\mathcal{F}_{t_k}\right] + \Delta\ f(t_k,\overline{Y}_{t_k},\mathbb{E}\left[\widetilde{U}_{t_{k+1}}\middle|\mathcal{F}_{t_k}\right],\widetilde{V}_{t_k})\\ \widetilde{V}_{t_k} = \frac{1}{\Delta}\left[\sigma\left(\overline{Y}_{t_k}\right)^\top\right]^{-1}\mathbb{E}\left[\widetilde{U}_{t_{k+1}}\left(\overline{Y}_{t_{k+1}}-\overline{Y}_{t_k}\right)\middle|\mathcal{F}_{t_k}\right]\\ \qquad -\left[\sigma\left(\overline{Y}_{t_k}\right)^\top\right]^{-1}\mathbb{E}\left[\widetilde{U}_{t_{k+1}}\middle|\mathcal{F}_{t_k}\right]b\left(\overline{Y}_{t_k}\right).\end{cases} \tag{3.1}$$

where $\widetilde{U}$ and $\widetilde{V}$ are approximations of U and V (for more details, see Section A.3) and where $\overline{Y}$ denotes a suitable (time) discretization of Y that, at this point, is left unspecified. The scheme is similar to the one proposed in [1], which we recall here for the reader's ease:

$$\begin{cases}\widetilde{U}_{t_n} = h(\overline{Y}_{t_n}) \quad \text{and for} \quad k=0,\dots,n-1\\ \widetilde{U}_{t_k} = \mathbb{E}\left[\widetilde{U}_{t_{k+1}}\middle|\mathcal{F}_{t_k}\right] + (t_{k+1}-t_k)f(t_k,\overline{Y}_{t_k},\mathbb{E}\left[\widetilde{U}_{t_{k+1}}\middle|\mathcal{F}_{t_k}\right],\widetilde{V}^{\text{PS}}_{t_k})\\ \widetilde{V}^{\text{PS}}_{t_k} = \frac{1}{\Delta}\mathbb{E}\left[\widetilde{U}_{t_{k+1}}\left(W_{t_{k+1}}-W_{t_k}\right)\middle|\mathcal{F}_{t_k}\right].\end{cases} \tag{3.2}$$

The novelty is a new discretization scheme for the control process: $\widetilde{V}_{t_k}, k=0,\dots,n-1$, in Eq. (3.1) is no longer a function of $(\widetilde{U}_{t_{k+1}}, W_{t_k}, W_{t_{k+1}})$, as it is the case for $\widetilde{V}^{\text{PS}}_{t_k}$, but depends only on $\widetilde{U}_{t_{k+1}}, \overline{Y}_{t_k}, \overline{Y}_{t_{k+1}}$. This leads to a simpler numerical procedure. Indeed, since Y is approximated indipendently from U and V, there will be no need to discretize the Brownian motion increments and this will result in a speed-up of the computational time required to solve the FBSDE. More details on this will be given in Remark 6.4. From a practical point of view, once the stochastic process $\overline{Y}$ has been discretized in space via recursive marginal quantization, thus obtaining $\widehat{Y}$, the scheme reads as in Eq. (5.1) and the backward recursion is explicit and fully quantization-based.

We stress that, for the moment, we introduced a general, yet original, discretization for the FBSDE (2.1)–(2.7). The role of recursive marginal quantization will become apparent as we approximate the conditional expectations appearing in the scheme above. For this we refer to Section 5.

4. A primer on recursive product marginal quantization

We provide some background on recursive marginal quantization. We consider a diffusion process Y as in (2.1) and its discretized version $\overline{Y}$ over a given time grid. Quantizing the diffusion process Y via recursive marginal quantization (henceforth RMQ) means the following: we consider the discretized analog $\overline{Y}$ of Y and, for each given point in time, we project every single random variable $\overline{Y}_{t_{k+1}}$ on a finite grid of points by exploiting the fact that the conditional law of $\overline{Y}_{t_{k+1}}$ given its value at time t_k is known. When the discretization $\overline{Y}$ is chosen to be the Euler scheme, the conditional law of $\overline{Y}_{t_{k+1}}$ given $\overline{Y}_{t_k}$ is Gaussian. This technique was first introduced in [23] and was further developed in [26] and applied in different settings, such as [24] among others.

Let us now provide a minimum insight on RMQ. The Euler scheme of Y, solution to Eq. (2.1), is defined via the recursion

$$\overline{Y}_{t_{k+1}} = \overline{Y}_{t_k} + \Delta\ b\left(\overline{Y}_{t_k}\right) + \sigma\left(\overline{Y}_{t_k}\right)^\top\left(W_{t_{k+1}}-W_{t_k}\right), \quad \bar{y}_0\in\mathbb{R}^d, \tag{4.1}$$

for $\Delta=\Delta_n=\frac{T}{n}$ and $t_k=\frac{kT}{n}$. For notational simplicity, in this section we set $\overline{Y}_k := \overline{Y}_{t_k}$.

Remark 4.1. Some extensions are possible:

a) The results presented here can be extended without any technical issue, yet with additional notational burden, to the case when the coefficients b and σ are no longer time homogeneous (this is the setting in [1]).
b) It is possible to consider higher order schemes such as e.g. the Milstein discretization as in [27]. This has an obvious implication on the shape of the conditional distribution of $\overline{Y}_{k+1}$ given $\overline{Y}_k$.

We define the Euler operator, which allows one to express the distribution of $\overline{Y}^{\ell}_{k+1}$ given $\{\overline{Y}^{\ell}_{k}=y\}$, with $\ell=1,\dots,d$

$$\mathcal{E}_k(y,z):=y+\Delta\ b(y)+\sqrt{\Delta}\ \sigma(y)^{\top}z,\quad y\in\mathbb{R}^d, z\in\mathbb{R}^q. \tag{4.2}$$

Our target is discretizing $\overline{Y}_{k+1}$ via a finite grid, Γ_{k+1}, under the constraint that the resulting approximating error has to be minimal. Namely, using the Euler operator, we consider the L^2-distortion at time t_{k+1}, $\overline{D}_{k+1}$, which is defined as the square of the L^2-distance between the random variable $\overline{Y}_{k+1}$ and the grid Γ_{k+1}

$$\overline{D}_{k+1}(\Gamma_{k+1})=\mathbb{E}\Big[\Big(\text{dist}\left(\overline{Y}_{k+1}\Gamma_{k+1}\right)^2\Big]=\mathbb{E}\Big[\text{dist}\left(\mathcal{E}_k\left(\overline{Y}_k,Z_{k+1}\right),\Gamma_{k+1}\right)^2\Big] \tag{4.3}$$

for $Z_{k+1}\sim\mathcal{N}\left(0,I_q\right)$, and we aim at finding a grid $\Gamma^{\star}_{k+1}$ that minimizes the distortion function. For a given size of the grid, it is known that an optimal quantizer exists (see e.g. [30]). Moreover, in the one-dimensional case, if the density of the random variable to be discretized is absolutely continuous and log-concave, then the optimal quantizer is unique.

Remark 4.2. The conditional distribution of the Euler process $\overline{Y}$ is Gaussian. Also, each component of a Gaussian vector is Gaussian.

To quantize the vector $\overline{Y}_k\in\mathbb{R}^d$, [26] consider each component of the vector separately: they quantize $\overline{Y}^{\ell}_k$ over a grid Γ^{ℓ}_k of size N^{ℓ}_k for $\ell=1,\dots,d$ and then they define its product quantization $\widehat{Y}_k$, i.e., the quantizer of the whole vector, on the product grid $\Gamma_k=\bigotimes_{\ell=1}^{d}\Gamma^{\ell}_k$ of size $N_k=N^1_k\times\dots\times N^d_k$ as $\widehat{Y}_k=\left(\widehat{Y}^1_k,\dots,\widehat{Y}^d_k\right)$.

More precisely, for any $k=0,\dots,n$ and any given $\ell\in\{1,\dots,d\}$, $\widehat{Y}^{\ell}_k$ denotes the quantization of $\overline{Y}^{\ell}_k$ on the grid $\Gamma^{\ell}_k=\left\{y^{\ell,i_\ell}_k, i_\ell=1,\dots,N^{\ell}_k\right\}$. The idea of [26] is as follows: assume we have access to Γ^{ℓ}_k, an N^{ℓ}_k-quantizer, for $\ell=1,\dots,d$, of the ℓ-th component $\overline{Y}^{\ell}_k$ of $\overline{Y}_k$. They define a componentwise recursive product quantizer $\Gamma_k=\bigotimes_{\ell=1}^{d}\Gamma^{\ell}_k$ of size $N_k=N^1_k\times\dots\times N^d_k$ of the vector $\overline{Y}_k=\left(\overline{Y}^{\ell}_k\right)_{\ell=1,\dots,d}$ via

$$\Gamma_k=\left\{\left(y^{1,i_1}_k,\dots,y^{d,i_d}_k\right),\quad y^{\ell,i_\ell}_k\in\Gamma^{\ell}_k\quad\text{for }\ell\in\{1,\dots,d\}\text{ and }i_\ell\in\left\{1,\dots,N^{\ell}_k\right\}\right\}. \tag{4.4}$$

To leverage the conditional normality feature, suppose now that $\overline{Y}_k$ has already been quantized and that we have the associated weights $\mathbb{P}\left(\widehat{Y}_k=\mathbf{y}^{\mathbf{i}}_{\mathbf{k}}\right)$, $\mathbf{i}\in I_k$, where $\mathbf{y}^{\mathbf{i}}_{\mathbf{k}}:=\left(y^{1,i_1}_k,\dots,y^{d,i_d}_k\right)$, $\mathbf{i}:=(i_1,\dots,i_d)\in I_k$ and

$$I_k=\left\{(i_1,\dots,i_d),i_\ell\in\left\{1,\dots,N^{\ell}_k\right\}\right\},\ k\in\{0,\dots,n\}. \tag{4.5}$$

By setting $\widetilde{Y}^{\ell}_k=\mathcal{E}^{\ell}_k\left(\widehat{Y}_k,Z_{k+1}\right)$, one can approximate each component of $\overline{D}_{k+1}\left(\Gamma_{k+1}\right)$ via $\widetilde{D}^{\ell}_{k+1}\left(\Gamma^{\ell}_{k+1}\right)$, $\ell=1,\dots,d$ where

$$\begin{aligned}\widetilde{D}^{\ell}_{k+1}\left(\Gamma^{\ell}_{k+1}\right)&:=\mathbb{E}\Big[\text{dist}\left(\widetilde{Y}^{\ell}_{k+1},\Gamma^{\ell}_{k+1}\right)^2\Big]=\mathbb{E}\Big[\text{dist}\left(\mathcal{E}^{\ell}_k\left(\widehat{Y}_k,Z_{k+1}\right),\Gamma^{\ell}_{k+1}\right)^2\Big]\\&=\sum_{\mathbf{i}\in I_k}\mathbb{E}\Big[\text{dist}\left(\mathcal{E}^{\ell}_k\left(\mathbf{y}^{\mathbf{i}}_{\mathbf{k}},Z_{k+1}\right),\Gamma^{\ell}_{k+1}\right)^2\Big]\mathbb{P}\left(\widehat{Y}_k=\mathbf{y}^{\mathbf{i}}_{\mathbf{k}}\right).\end{aligned} \tag{4.6}$$

Such approximation allows [26] to introduce the sequence of product recursive quantizations of $\widehat{Y}=\left(\widehat{Y}_k\right)_{k=0,\cdots,n}$, for $k=0,\dots,n-1$, as

$$\begin{cases}\widetilde{Y}_0=\widehat{Y}_0=y_0,\quad\widehat{Y}_k=\left(\widehat{Y}^1_k,\dots,\widehat{Y}^d_k\right)\\ \widehat{Y}^{\ell}_k=\text{Proj}_{\Gamma^{\ell}_k}\left(\widetilde{Y}^{\ell}_k\right)\quad\text{and}\quad\widetilde{Y}^{\ell}_{k+1}=\mathcal{E}^{\ell}_k\left(\widehat{Y}_k,Z_{k+1}\right),\ell=1,\dots,d\\ \mathcal{E}^{\ell}_k(y,z)=y^{\ell}+\Delta b^{\ell}(y)+\sqrt{\Delta}\left(\sigma^{\ell\bullet}(y)|z\right),z=\left(z^1,\dots,z^q\right)\in\mathbb{R}^q\\ y=\left(y^1,\dots,y^d\right),b=\left(b^1,\dots,b^d\right)\text{ and }\left(\sigma^{\ell\bullet}(y)|z\right)=\sum_{m=1}^{q}\sigma^{\ell m}(y)z^m\end{cases} \tag{4.7}$$

where for every matrix $A\in\mathcal{M}(d,q)$, $a^{\ell\bullet}=[a_{\ell j}]_{j=1,\dots,q}$.

We conclude this section by providing some information about the computation of transition probabilities. In [26] three methods are proposed: the first one, in their Proposition 3.1, concerns the computation of transition probabilities of the whole vector $\widehat{Y}$, the second covers the case where the diffusion matrix is diagonal and the third is a corollary to their Proposition 3.1 for the case where we are interested only in one component of the whole vector. We refer the reader to this paper for all the details on how to instantaneously compute these probabilities once the quantization grids have been obtained.

5. Computing the conditional expectations

Our numerical scheme (3.1) has been conceived in such a way that computing the conditional expectations, with respect to $(\mathcal{F}_{t_k})_{k=0,\dots,n}$, only requires the knowledge of the stochastic process $\overline{Y}$. We will see that this results, from the practical point of view, in a handy, easy to understand and ready-to-use numerical scheme.

Before proceeding, we now rigorously prove that for every $k = 0, \dots, n-1$

$$\widetilde{U}_{t_k} = u_k(\overline{Y}_{t_k}) \quad \text{and} \quad \widetilde{V}_{t_k} = v_k(\overline{Y}_{t_k})$$

for given Borel functions $u_k : \mathbb{R}^d \to \mathbb{R}$ and $v_k : \mathbb{R}^d \to \mathbb{R}^q$.

Proposition 5.1. *For every $k \in \{0, \dots, n-1\}$ the update rule for the control satisfies $\widetilde{V}_{t_k} = v_k(\overline{Y}_{t_k})$ and for every $l \in \{0, \dots, n\}$ we have $\widetilde{U}_{t_l} = u_l(\overline{Y}_{t_l})$, where $v_k : \mathbb{R}^d \to \mathbb{R}^q$ and $u_l : \mathbb{R}^d \to \mathbb{R}$ are*

$$\begin{cases} v_k(y) := \dfrac{1}{\Delta}(\sigma(y)^\top)^{-1}[g_{1,k}(y) + \Delta \cdot g_{2,k}(y)b(y)], \quad k = 0, \dots, n-1 \\ u_n(y) := h(y), \quad \text{and} \quad u_l(y) := g_{3,l}(y) + \Delta \cdot f\big(t_l, y, g_{3,l}(y), v_l(y)\big), \quad l = 0, \dots, n-1 \end{cases}$$

with $g_{1,k} : \mathbb{R}^d \to \mathbb{R}^d$, $g_{2,k} : \mathbb{R}^d \to \mathbb{R}$ and $g_{3,l} : \mathbb{R}^d \to \mathbb{R}$ as follows

$$\begin{cases} g_{1,k}(y) := \mathbb{E}[u_k(\mathcal{E}_{k-1}(y, Z_k))[\mathcal{E}_{k-1}(y, Z_k) - y]] \\ g_{2,k}(y) := \mathbb{E}[u_k(\mathcal{E}_{k-1}(y, Z_k))] \\ g_{3,l}(y) := \mathbb{E}[u_{l+1}(\mathcal{E}_l(y, Z_{l+1}))] \end{cases}$$

and for Z_j's i.i.d., $Z_j \sim \mathcal{N}(0, I_q)$.

Proof. First of all notice that, by definition of $\widetilde{U}_{t_n}$, we immediately have $\widetilde{U}_{t_n} = h(\overline{Y}_{t_n}) =: u_n(\overline{Y}_{t_n})$. We now work on the control at time t_{n-1}. By recalling (cf. Section A.1) that $(t_{k+1} - t_k) = \Delta$ we find:

$$\begin{aligned} \widetilde{V}_{t_{n-1}} &= \frac{1}{\Delta}\left[\sigma\left(\overline{Y}_{t_{n-1}}\right)^\top\right]^{-1}\mathbb{E}\left[\widetilde{U}_{t_n}\left(\overline{Y}_{t_n} - \overline{Y}_{t_{n-1}}\right)\middle|\mathcal{F}_{t_{n-1}}\right] - \left[\sigma\left(\overline{Y}_{t_{n-1}}\right)^\top\right]^{-1}\mathbb{E}\left[\widetilde{U}_{t_n}\middle|\mathcal{F}_{t_{n-1}}\right]b\left(\overline{Y}_{t_{n-1}}\right) \\ &= \frac{1}{\Delta}\left[\sigma\left(\overline{Y}_{t_{n-1}}\right)^\top\right]^{-1}\mathbb{E}\left[u_n(\overline{Y}_{t_n})\left(\overline{Y}_{t_n} - \overline{Y}_{t_{n-1}}\right)\middle|\mathcal{F}_{t_{n-1}}\right] \\ &\quad + \left[\sigma\left(\overline{Y}_{t_{n-1}}\right)^\top\right]^{-1}\mathbb{E}\left[u_n(\overline{Y}_{t_n})\middle|\mathcal{F}_{t_{n-1}}\right]b\left(\overline{Y}_{t_{n-1}}\right). \end{aligned}$$

Now, from Eq. (4.2), we have $\overline{Y}_{t_{k+1}} = \mathcal{E}_k(\overline{Y}_{t_k}, Z_{k+1})$, where $Z_{k+1} \sim \mathcal{N}(0, I_q)$ and the Z_k's, $k = 0, \dots, n-1$ are i.i.d., and so

$$\begin{aligned} \widetilde{V}_{t_{n-1}} &= \frac{1}{\Delta}\left[\sigma\left(\overline{Y}_{t_{n-1}}\right)^\top\right]^{-1}\mathbb{E}\left\{u_n\left(\mathcal{E}_{n-1}(\overline{Y}_{t_{n-1}}, Z_n)\right)\left[\mathcal{E}_{n-1}(\overline{Y}_{t_{n-1}}, Z_n) - \overline{Y}_{t_{n-1}}\right]|\mathcal{F}_{t_{n-1}}\right\} \\ &\quad + \left[\sigma\left(\overline{Y}_{t_{n-1}}\right)^\top\right]^{-1}\mathbb{E}\left[u_n(\mathcal{E}_{n-1}(\overline{Y}_{t_{n-1}}, Z_n))\middle|\mathcal{F}_{t_{n-1}}\right]b\left(\overline{Y}_{t_{n-1}}\right) \\ &= \frac{1}{\Delta}\left[\sigma\left(\overline{Y}_{t_{n-1}}\right)^\top\right]^{-1}\left[g_{1,n}(\overline{Y}_{t_{n-1}}) + \Delta\ g_{2,n}(\overline{Y}_{t_{n-1}})b\left(\overline{Y}_{t_{n-1}}\right)\right] =: v_{n-1}(\overline{Y}_{t_{n-1}}), \end{aligned}$$

where we have used the fact that $(\overline{Y}_{t_j})_{j=0,\dots,n}$ is a Markov process (see e.g. [1, Section 3.2.1]) measurable w.r.t $\mathcal{F}_{t_{n-1}}$ and Z_n is independent of $\mathcal{F}_{t_{n-1}}$ and, for $y \in \mathbb{R}^d$, $g_{1,n}(y)$ and $g_{2,n}(y)$ are defined as in the statement.

We now proceed by induction (notice that the values of k and l for which the claims hold are shifted): we assume that $\widetilde{U}_{t_{k+1}} = u_{k+1}(\overline{Y}_{t_{k+1}})$ and $\widetilde{V}_{t_k} = v_k(\overline{Y}_{t_k})$ and we prove that this implies that $\widetilde{U}_{t_k} = u_k(\overline{Y}_{t_k})$ and $\widetilde{V}_{t_{k-1}} = v_{k-1}(\overline{Y}_{t_{k-1}})$.

The proof on $\widetilde{V}_{t_{k-1}}$ is analogous to what we have just done for $\widetilde{V}_{t_{n-1}}$, so we omit it. It remains the $\widetilde{U}_{t_k}$-part, which we develop now:

$$\begin{aligned} \widetilde{U}_{t_k} &= \mathbb{E}\left[\widetilde{U}_{t_{k+1}}\middle|\mathcal{F}_{t_k}\right] + \Delta\ f\left(t_k, \overline{Y}_{t_k}, \mathbb{E}\left[\widetilde{U}_{t_{k+1}}\middle|\mathcal{F}_{t_k}\right], \widetilde{V}_{t_k}\right) \\ &= \mathbb{E}\left[u_{k+1}(\overline{Y}_{t_{k+1}})\middle|\mathcal{F}_{t_k}\right] + \Delta\ f\left(t_k, \overline{Y}_{t_k}, \mathbb{E}\left[u_{k+1}(\overline{Y}_{t_{k+1}})\middle|\mathcal{F}_{t_k}\right], v_k(\overline{Y}_{t_k})\right) \\ &= \mathbb{E}\left[u_{k+1}(\mathcal{E}_k(\overline{Y}_{t_k}, Z_{k+1}))\middle|\mathcal{F}_{t_k}\right] + \Delta\ f\left(t_k, \overline{Y}_{t_k}, \mathbb{E}\left[u_{k+1}(\mathcal{E}_k(\overline{Y}_{t_k}, Z_{k+1}))\middle|\mathcal{F}_{t_k}\right], v_k(\overline{Y}_{t_k})\right) \\ &= g_{3,k}(\overline{Y}_{t_k}) + \Delta\ f\left(t_k, \overline{Y}_{t_k}, g_{3,k}(\overline{Y}_{t_k}), v_k(\overline{Y}_{t_k})\right) =: u_k(\overline{Y}_{t_k}), \end{aligned}$$

where we have used the functions $g_{3,k}$ introduced in the statement. □

In summary, we can compute the conditional expectations in the discretization scheme (3.1) as follows, by exploiting Proposition 5.1 and the Markovianity of the discrete time stochastic process $(\overline{Y}_{t_k})_{k=0,\dots,n}$

$$\begin{cases} \mathbb{E}[\widetilde{U}_{t_{k+1}}|\mathcal{F}_{t_k}] = \mathbb{E}[u_{k+1}(\overline{Y}_{t_{k+1}})|\mathcal{F}_{t_k}] = \mathbb{E}[u_{k+1}(\overline{Y}_{t_{k+1}})|\overline{Y}_{t_k}] \\ \mathbb{E}[\widetilde{U}_{t_{k+1}}\left(\overline{Y}_{t_{k+1}} - \overline{Y}_{t_k}\right)|\mathcal{F}_{t_k}] = \mathbb{E}[u_{k+1}(\overline{Y}_{t_{k+1}})\left(\overline{Y}_{t_{k+1}} - \overline{Y}_{t_k}\right)|\mathcal{F}_{t_k}] \\ \qquad = \mathbb{E}[u_{k+1}(\overline{Y}_{t_{k+1}})\left(\overline{Y}_{t_{k+1}} - \overline{Y}_{t_k}\right)|\overline{Y}_{t_k}]. \end{cases}$$

5.1. Approximation via quantization

As a final step, now we approximate $\overline{Y}$ via $\widehat{Y}$, which is obtained as explained in Section 4 and we get the quantized final version of the recursive discretization scheme (3.1) (we recall that, for every $k = 0, \dots, n$, $\widehat{Y}_{t_k}$ is the quantization of $\overline{Y}_{t_k}$):

$$\begin{cases} \widehat{U}_{t_n} = h(\widehat{Y}_{t_n}) \quad \text{and for} \quad k = 0, \dots, n-1 \\ \widehat{U}_{t_k} = \mathbb{E}\big[\widehat{U}_{t_{k+1}}|\widehat{Y}_{t_k}\big] + \Delta\ f\big(t_k, \widehat{Y}_{t_k}, \mathbb{E}[\widehat{U}_{t_{k+1}}|\widehat{Y}_{t_k}], \widehat{V}_{t_k}\big) =: \widehat{u}_k(\widehat{Y}_{t_k}) \\ \widehat{V}_{t_k} = \frac{1}{\Delta}\Big[\sigma\big(\widehat{Y}_{t_k}\big)^\top\Big]^{-1}\mathbb{E}\big[\widehat{U}_{t_{k+1}}\big(\widehat{Y}_{t_{k+1}} - \widehat{Y}_{t_k}\big)|\widehat{Y}_{t_k}\big] \\ \qquad - \Big[\sigma\big(\widehat{Y}_{t_k}\big)^\top\Big]^{-1}\mathbb{E}\big[\widehat{U}_{t_{k+1}}|\widehat{Y}_{t_k}\big]b\big(\widehat{Y}_{t_k}\big) =: \widehat{v}_k(\widehat{Y}_{t_k}), \end{cases} \tag{5.1}$$

with $\widehat{u}_k : \Gamma_k \to \mathbb{R}$ and $\widehat{v}_k : \Gamma_k \to \mathbb{R}^q$ Borel functions, for $k = 0, \dots, n-1$. Namely, the conditional expectations in (3.1) are approximated as:

$$\begin{cases} \mathbb{E}[\widetilde{U}_{t_{k+1}}|\mathcal{F}_{t_k}] = \mathbb{E}[u_{k+1}(\overline{Y}_{t_{k+1}})|\overline{Y}_{t_k}] \approx \mathbb{E}\big[\widehat{U}_{t_{k+1}}|\widehat{Y}_{t_k}\big] = \mathbb{E}[\widehat{u}_{k+1}(\widehat{Y}_{t_{k+1}})|\widehat{Y}_{t_k}] \\ \mathbb{E}[\widetilde{U}_{t_{k+1}}\big(\overline{Y}_{t_{k+1}} - \overline{Y}_{t_k}\big)|\mathcal{F}_{t_k}] = \mathbb{E}[u_{k+1}(\overline{Y}_{t_{k+1}})\big(\overline{Y}_{t_{k+1}} - \overline{Y}_{t_k}\big)|\overline{Y}_{t_k}] \\ \qquad \approx \mathbb{E}[\widehat{U}_{t_{k+1}}\big(\widehat{Y}_{t_{k+1}} - \widehat{Y}_{t_k}\big)|\widehat{Y}_{t_k}] \\ \qquad = \mathbb{E}[\widehat{u}_{k+1}(\widehat{Y}_{t_{k+1}})\big(\widehat{Y}_{t_{k+1}} - \widehat{Y}_{t_k}\big)|\widehat{Y}_{t_k}]. \end{cases} \tag{5.2}$$

The approximating scheme is then fully explicit, since for every $\widetilde{\omega} \in \Omega$ such that $\widehat{Y}_{t_k}(\widetilde{\omega}) = \mathbf{y_k^i}$, for a given $\mathbf{i} \in I_k$, we have

$$\begin{cases} \mathbb{E}\big[\widehat{u}_{k+1}(\widehat{Y}_{t_{k+1}})|\widehat{Y}_{t_k}\big](\widetilde{\omega}) = \sum_{\mathbf{j} \in I_{k+1}} \widehat{u}_{k+1}(\mathbf{y^j_{k+1}})\mathbb{P}\big(\widehat{Y}_{t_{k+1}} = \mathbf{y^j_{k+1}}|\widehat{Y}_{t_k} = \mathbf{y^i_k}\big) \\ \mathbb{E}\big[\widehat{u}_{k+1}(\widehat{Y}_{t_{k+1}})\big(\widehat{Y}_{t_{k+1}} - \widehat{Y}_{t_k}\big)|\widehat{Y}_{t_k}\big](\widetilde{\omega}) = \sum_{\mathbf{j} \in I_{k+1}} \widehat{u}_{k+1}(\mathbf{y^j_{k+1}})\big(\mathbf{y^j_{k+1}} - \mathbf{y^i_k}\big)\mathbb{P}\big(\widehat{Y}_{t_{k+1}} = \mathbf{y^j_{k+1}}|\widehat{Y}_{t_k} = \mathbf{y^i_k}\big). \end{cases} \tag{5.3}$$

Remark 5.2. As expected and already announced, the proposed discretization scheme is fully driven by the process $\widehat{Y}$, which is the quantization of the Euler scheme process $\overline{Y}$.

We conclude the theoretical part of the paper by studying, In the next section, the error associated with our scheme (5.1).

6. The error

The analysis is divided in two parts:

- Section 6.1: error in time
- Section 6.2: error in space.

Here $\overline{Y}$ denotes the Euler scheme relative to the stochastic process Y. Moreover, recall that we work under Assumption 2.4.

Remark 6.1. a) Assumption 2.4 ensures that for every $y \in \mathbb{R}^d$, the matrix $\sigma(y)$ is positive definite, hence invertible, and bounded. The inverse matrix $\sigma(y)^{-1}$ is also bounded. More precisely, denoting by $||\cdot||_F$ the Frobenius norm[1], we have that for every $y \in \mathbb{R}^d$

$$||\sigma^{-1}(y)||_F^2 \le \lambda_0.$$

This will be crucial in Section 6.2. b) Assumption 2.4, together with a Lipschitz continuity condition on h with Lipschitz constant K, is required by [16, Lemma 2.5 (i)] to prove that the control process V admits a càdlàg version. This, in turn, is needed in [16, Theorem 3.1] to prove the following:

$$\sum_{i=1}^n \mathbb{E}\left\{\int_{t_i}^{t_{i+1}} \big[|V_t - V_{t_{i-1}}|^2 + |V_t - V_{t_i}|^2\big]dt\right\} \le C(1 + |y_0|^2)\Delta,$$

where C is a constant depending only on T and K and we recall that $\Delta = \Delta_n = \frac{T}{n}$. The results by Zhang are contained in [1, Theorem 3.1]. We adapt them below, in Theorem 6.3, to our setting.

[1] Recall that given a matrix $B := b_{ij}, i, j = 1, \dots, d$, $||B||_F := \sqrt{\mathrm{tr}(BB^\top)} = \sqrt{\sum_{i,j}(b_{i,j})^2}$.

6.1. Time discretization error

Studying the error of our scheme (3.1) with respect to time means computing a proper distance between (U, V) and $(\widetilde{U}, \widetilde{V})$. Inspired by [1], we will adapt to our setting their Theorem 3.1.

Before stating the time discretization error result, we need to introduce, as typically done in the literature, the continuous time extension of $(\widetilde{U}_{t_k})_{k=0,\dots,n}$, denoted by $(\widetilde{U}_t)_{t\in[0,T]}$. We introduce the random variable

$$M_T := \sum_{k=1}^{n} \widetilde{U}_{t_k} - \mathbb{E}\Big[\widetilde{U}_{t_k} | \mathcal{F}_{t_{k-1}}\Big]$$

and we notice that $M_T \in L^2$ and so, by the martingale representation theorem, M admits the following representation:

$$M_T = \int_0^T \widetilde{V}_s^\top dW_s,$$

for an $(\mathcal{F}_t)_{t\in[0,T]}$-progressively measurable stochastic process $\widetilde{V}$, with values in $\mathbb{R}^q$, such that $\mathbb{E}\Big[\int_0^T |\widetilde{V}_s|^2 ds\Big] < \infty$. As a consequence,

$$\widetilde{U}_{t_k} - \mathbb{E}\Big[\widetilde{U}_{t_k} | \mathcal{F}_{t_{k-1}}\Big] = \int_{t_{k-1}}^{t_k} \widetilde{V}_s^\top dW_s \tag{6.1}$$

and we introduce the continuous extension $(\widetilde{U}_t)_{t\in[0,T]}$ as follows: if $t \in [t_k, t_{k+1})$,

$$\widetilde{U}_t = \widetilde{U}_{t_k} - (t - t_k) f\Big(t_k, \overline{Y}_{t_k}, \mathbb{E}\Big[\widetilde{U}_{t_{k+1}} | \mathcal{F}_{t_k}\Big], \widetilde{V}_{t_k}\Big) + \int_{t_k}^{t} \widetilde{V}_s^\top dW_s. \tag{6.2}$$

Remark 6.2. Recalling Eq. (3.2) and exploiting Eq. (6.1), we get, for every $k = 0, \dots, n-1$,

$$\widetilde{V}_{t_k}^{\mathrm{PS}} = \frac{1}{\Delta}\mathbb{E}\Big[\widetilde{U}_{t_{k+1}}\big(W_{t_{k+1}} - W_{t_k}\big) \big| \mathcal{F}_{t_k}\Big] = \frac{1}{\Delta}\mathbb{E}\Bigg[\int_{t_k}^{t_{k+1}} \widetilde{V}_s ds | \mathcal{F}_{t_k}\Bigg].$$

This is extensively used in the proof of [1, Theorem 3.1], which we mimic here, to prove the error bounds with respect to time.

Theorem 6.3. *i) Under Assumption 2.2, let $f : [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^q \to \mathbb{R}$ be Lipschitz with respect to time and let $h : \mathbb{R}^d \to \mathbb{R}$ be Lipschitz. Then there exists a real constant $\overline{C} > 0$ only depending on b, σ, f, T such that, for every $n \geq 1$*

$$\max_{k=0,\dots,n} \mathbb{E}\Big[|U_{t_k} - \widetilde{U}_{t_k}|^2\Big] + \int_0^T \mathbb{E}\Big[|V_t - \widetilde{V}_t|^2\Big] dt \leq \overline{C}\Bigg(\Delta + \int_0^T \mathbb{E}\Big[|V_s - V_{\underline{s}}|^2\Big] ds\Bigg),$$

where $\underline{s} = t_k$ if $s \in [t_k, t_{k+1})$.

ii) Assume moreover that b, σ and f are continuously differentiable in their spatial variable with bounded partial derivatives and that f is $\frac{1}{2}$-Hölder continuous with respect to time. Then the process V admits a càdlàg modification and

$$\int_0^T \mathbb{E}\Big[|V_s - V_{\underline{s}}|^2\Big] ds \leq C' \Delta,$$

for a real positive constant C' (only depending on b, σ, f, T), so that we have

$$\max_{k=0,\dots,n} \mathbb{E}\Big[|U_{t_k} - \widetilde{U}_{t_k}|^2\Big] + \int_0^T \mathbb{E}\Big[|V_t - \widetilde{V}_t|^2\Big] dt \leq \widetilde{C}\Delta,$$

for a real positive constant $\widetilde{C}$ (only depending on b, σ, f, T).

Proof. Part *ii*) is a consequence of *i*) and it is obtained via [16, Lemma 2.5 (i) and Theorem 3.1].

The proof of *i*) is the same as the one in [1, Theorem 3.1 a)] relatively to Steps 2 and 3, while something has to be made precise relatively to Step 1. Indeed, the two schemes (3.2) and (3.1) differ in the control discretization. Nevertheless, since here $\overline{Y}$ is the Euler scheme of Y, namely $\overline{Y}_{t_{k+1}} - \overline{Y}_{t_k} = \Delta b(\overline{Y}_{t_k}) + \sigma(\overline{Y}_{t_k})^\top (W_{t_{k+1}} - W_{t_k})$, $k = 0, \dots, n-1$, also Step 1 in [1, Theorem 3.1] can be retraced straightforwardly. □

Remark 6.4. Despite the fact that the time error bounds for our proposed scheme are the same as in [1], our recursions only require the discretization of the process $\overline{Y}$ and this results in an increased numerical efficiency, namely in the speed-up of the computational time. What is more, in the approximation of the conditional expectations required in our scheme, we only need the transition probabilities of the quantized process $\widehat{Y}$, i.e., for $\mathbf{i} \in I_k, \mathbf{j} \in I_{k+1}$:

$$\mathbb{P}\Big(\widehat{Y}_{t_{k+1}} = \mathbf{y}_{\mathbf{k+1}}^{\mathbf{j}} | \widehat{Y}_{t_k} = \mathbf{y}_{\mathbf{k}}^{\mathbf{i}}\Big).$$

This is not the case in [1], where the authors, in their scheme (3.2) for $\widetilde{V}^{PS}$, need to additionally compute $\mathbb{E}[\widehat{Y}_{t_{k+1}}(W_{t_{k+1}} - W_{t_k})|\mathcal{F}_{t_k}]$, $k = 0, \dots, n-1$. So, they have to estimate, for $\mathbf{i} \in I_k, \mathbf{j} \in I_{k+1}$, the weights (we stick to their notation)

$$\pi_{ij}^{W,k} := \frac{1}{\mathbb{P}(\widehat{Y}_{t_k} = \mathbf{y}_\mathbf{k}^\mathbf{i})} \mathbb{E}\Big[(W_{t_{k+1}} - W_{t_k})\mathbf{1}_{\{\widehat{Y}_{t_{k+1}} = \mathbf{y}_{\mathbf{k+1}}^\mathbf{j}, \widehat{Y}_{t_k} = \mathbf{y}_\mathbf{k}^\mathbf{i}\}}\Big],$$

which is done via Monte Carlo simulation (see [1, Section 5]), hence requiring additional numerical effort, which is not needed in our case.

6.2. Space discretization (quantization) error

Here we study the quadratic quantization error induced by the approximation of $(\widetilde{U}_{t_k}, \widetilde{V}_{t_k})$ in (3.1) by $(\widehat{U}_{t_k}, \widehat{V}_{t_k})$ in (5.1), for every $k = 0, \dots, n$. We intuitively expect both the error components $||\widetilde{U}_{t_k} - \widehat{U}_{t_k}||_2^2$ and $||\widetilde{V}_{t_k} - \widehat{V}_{t_k}||_2^2$ to be written as functions of the quantization error $||\overline{Y}_{t_i} - \widehat{Y}_{t_i}||_2^2$ for $i = k, \dots, n$.

We start by discussing the error $||\widetilde{U}_{t_k} - \widehat{U}_{t_k}||_2^2$, for $k = 0, \dots, n$, for which a bound is given in the following Lemma, corresponding to [1, Theorem 3.2 a), Equation (31)]. We provide a sketch of its proof below.

Lemma 6.5. *Under Assumptions 2.1 and 2.2 and assuming that h is $[h]_{\mathrm{Lip}}$-continuous, we have that for every $k = 0, \dots, n$*

$$||\widetilde{U}_{t_k} - \widehat{U}_{t_k}||_2^2 \le \sum_{i=k}^{n} e^{(1+\mathcal{L}_6)(t_i - t_k)} K_i(b, \sigma, T, f) ||\overline{Y}_{t_i} - \widehat{Y}_{t_i}||_2^2 \tag{6.3}$$

where $K_n(b, \sigma, T, f) = [h]_{\mathrm{Lip}}^2$ and for every $k = 0, \dots, n-1$ the other K_k's are provided in [1, Theorem 3.2 a)].

Proof. The proof is essentially the same as the one of [1, Theorem 3.2 a), Equation (31)] and it has to be carried out in two steps: first, propagate the Lipschitz property through the functions $u_k, k = 0, \dots, n-1$ defined in Proposition 5.1, second, introduce the quantization scheme and find a backward recursive inequality satisfied by $||\widetilde{U}_{t_k} - \widehat{U}_{t_k}||_2^2$. Given the nature of the two numerical schemes in Eqs. (3.2) and (3.1), which differ on the control process side, the presence of $\widetilde{V}_{t_k}$ in the scheme for $\widetilde{U}_{t_k}$ requires additional care. More precisely, when retracing the proof of [1, Prop. 3.4 b)], to prove that u_k defined in Proposition 5.1 is $[u_k]_{\mathrm{Lip}}$-continuous, we find, for $y, y' \in \mathbb{R}^d$:

$$\begin{aligned} |u_k(y) - u_k(y')| = &\ |\mathbb{E}[u_{k+1}(\mathcal{E}_k(y, Z_{k+1}))] + \Delta \cdot f(t_k, y, \mathbb{E}[u_{k+1}(\mathcal{E}_k(y, Z_{k+1}))], v_k(y)) \\ &-\mathbb{E}\Big[u_{k+1}(\mathcal{E}_k(y', Z_{k+1}))\Big] - \Delta \cdot f\Big(t_k, y', \mathbb{E}\Big[u_{k+1}(\mathcal{E}_k(y', Z_{k+1}))\Big], v_k(y')\Big)| \end{aligned}$$

and exploting the Lipschitzianity Assumption 2.2 (i) on f it follows that

$$\begin{aligned} |u_k(y) - u_k(y')| \le &\ (1 + \Delta\mathcal{L}_6)\mathbb{E}\Big[|u_{k+1}(\mathcal{E}_k(y, Z_{k+1})) - u_{k+1}(\mathcal{E}_k(y', Z_{k+1}))|\Big](1 + \Delta\mathcal{L}_6) \\ &+\Delta\mathcal{L}_6|y - y'| + \Delta|v_k(y) - v_k(y')|. \end{aligned}$$

A priori this seems slightly different from the proof of [1, Prop. 3.4 b)] because of the presence of the term $\Delta|v_k(y) - v_k(y')|$. Nevertheless, it suffices to rewrite $\widetilde{V}_{t_k}$ as: $\widetilde{V}_{t_k} = \frac{1}{\Delta}\mathbb{E}\big[u_{k+1}\big(\mathcal{E}_k(\overline{Y}_{t_k}, Z_{k+1})\big)\sqrt{\Delta}Z_{k+1}\big|\mathcal{F}_{t_k}\big]$ to be able to proceed exactly as in [1]. □

We now focus on $||\widetilde{V}_{t_k} - \widehat{V}_{t_k}||_2^2$.

Theorem 6.6. *Under Assumptions 2.1 and 2.4 and if b is continuously differentiable with bounded derivative, there exist positive constants $\widehat{C}$ and $\overline{C}$, only depending on $(\lambda_0, \mathcal{L}_3, \mathcal{L}_4)$, such that for every $k = 0, \dots, n$:*

$$||\widetilde{V}_{t_k} - \widehat{V}_{t_k}||_2^2 \le \frac{1}{\Delta}[\Psi_k]_{\mathrm{Lip}}^2 ||\overline{Y}_{t_k} - \widehat{Y}_{t_k}||_2^2 + \left(\frac{\widehat{C}}{\Delta} + \overline{C}\right) ||\widetilde{U}_{t_{k+1}} - \widehat{U}_{t_{k+1}}||_2^2 \tag{6.4}$$

where $[\Psi_k]_{\mathrm{Lip}}$ is the Lipschitz constant of the function $\Psi_k(x) : \mathbb{R}^d \to \mathbb{R}$, $\Psi_k(x) := \mathbb{E}[u_{k+1}(\mathcal{E}_k(x, Z)) \cdot Z]$ for $Z \sim \mathcal{N}(0, I_q)$.

Proof. We start by noticing that

$$\begin{aligned} \Big|\Big|\widetilde{V}_{t_k} - \widehat{V}_{t_k}\Big|\Big|_2^2 &= \Big|\Big|\widetilde{V}_{t_k} - \mathbb{E}\Big[\widetilde{V}_{t_k}|\widehat{Y}_{t_k}\Big] + \mathbb{E}\Big[\widetilde{V}_{t_k}|\widehat{Y}_{t_k}\Big] - \widehat{V}_{t_k}\Big|\Big|_2^2 \\ &\le \underbrace{\Big|\Big|\widetilde{V}_{t_k} - \mathbb{E}\Big[\widetilde{V}_{t_k}|\widehat{Y}_{t_k}\Big]\ \Big|\Big|_2^2}_{(I)} + \underbrace{\Big|\Big|\mathbb{E}\Big[\widetilde{V}_{t_k}|\widehat{Y}_{t_k}\Big] - \widehat{V}_{t_k}\Big|\Big|_2^2}_{(II)}. \end{aligned}$$

Let us focus on (I). By exploting Eqs. (4.1), (4.2) and Proposition 5.1, we find

$$\begin{aligned}\widetilde{V}_{t_k} &= \frac{1}{\Delta}\mathbb{E}\Big[\widetilde{U}_{t_{k+1}}\big(W_{t_{k+1}} - W_{t_k}\big)\Big|\mathcal{F}_{t_k}\Big]\\ &= \frac{1}{\Delta}\mathbb{E}\Big[u_{k+1}\big(\mathcal{E}_k(\overline{Y}_{t_k}, Z_{k+1})\big)\sqrt{\Delta}Z_{k+1}\Big|\mathcal{F}_{t_k}\Big]\end{aligned}$$

where we used $\big(W_{t_{k+1}} - W_{t_k}\big) =: \sqrt{\Delta}Z_{k+1}$, with $Z_{k+1} \sim \mathcal{N}(0, I_q)$ independent of $\mathcal{F}_{t_k}$. So, we have, being $\sigma(\widehat{Y}_{t_k}) \subset \sigma(\overline{Y}_{t_k}) \subset \mathcal{F}_{t_k}$ and using the tower property of conditional expectation:

$$\begin{aligned}\Big|\Big|\widetilde{V}_{t_k} - \mathbb{E}\big[\widetilde{V}_{t_k}|\widehat{Y}_{t_k}\big]\ \Big|\Big|_2^2 &= \Big|\Big|\frac{1}{\Delta}\mathbb{E}\Big[u_{k+1}\big(\mathcal{E}_k(\overline{Y}_{t_k}, Z_{k+1})\big)\sqrt{\Delta}Z_{k+1}\Big|\mathcal{F}_{t_k}\Big] - \frac{1}{\Delta}\mathbb{E}\Big[u_{k+1}\big(\mathcal{E}_k(\overline{Y}_{t_k}, Z_{k+1})\big)\sqrt{\Delta}Z_{k+1}\Big|\widehat{Y}_{t_k}\Big]\Big|\Big|_2^2\\ &\le \frac{1}{\Delta}\Big|\Big|\Psi_k\big(\overline{Y}_{t_k}\big) - \Psi_k\big(\widehat{Y}_{t_k}\big)\Big|\Big|_2^2 \le \frac{1}{\Delta}[\Psi_k]^2_{\text{Lip}}\big|\big|\overline{Y}_{t_k} - \widehat{Y}_{t_k}\big|\big|_2^2\end{aligned}$$

where in the second passage we defined $\Psi_k(x) = \mathbb{E}[u_{k+1}\big(\mathcal{E}_k(x, Z_{k+1})\big)Z_{k+1}]$, we used the definition of conditional expectation as best L^2-approximation and exploited the Lipschitzianity of Ψ_k, for which we refer to [1, Prop. 3.4 (b)] (therein Ψ_k corresponds to z_k).

Now, consider (II): since $\sigma(\widehat{Y}_{t_k}) \subset \sigma(\overline{Y}_{t_k}) \subset \mathcal{F}_{t_k}$ and by recalling the definition of $\widetilde{V}_{t_k}$ and $\widehat{V}_{t_k}$ in Eqs. (3.1) and (5.1) we have

$$\begin{aligned}\Big|\Big|\mathbb{E}\big[\widetilde{V}_{t_k}|\widehat{Y}_{t_k}\big] - \widehat{V}_{t_k}\Big|\Big|_2^2 &= \Big|\Big|\mathbb{E}\big[\widetilde{V}_{t_k} - \widehat{V}_{t_k}|\widehat{Y}_{t_k}\big]\Big|\Big|_2^2\\ &= \Bigg|\Bigg|\mathbb{E}\Bigg[\frac{1}{\Delta}\Big[\sigma\big(\overline{Y}_{t_k}\big)^\top\Big]^{-1}\widetilde{U}_{t_{k+1}}\big(\overline{Y}_{t_{k+1}} - \overline{Y}_{t_k}\big) - \frac{1}{\Delta}\Big[\sigma\big(\widehat{Y}_{t_k}\big)^\top\Big]^{-1}\widehat{U}_{t_{k+1}}\big(\widehat{Y}_{t_{k+1}} - \widehat{Y}_{t_k}\big)\\ &\quad + \Big[\sigma\big(\overline{Y}_{t_k}\big)^\top\Big]^{-1}\widetilde{U}_{t_{k+1}}b\big(\overline{Y}_{t_k}\big) - \Big[\sigma\big(\widehat{Y}_{t_k}\big)^\top\Big]^{-1}\widehat{U}_{t_{k+1}}b\big(\widehat{Y}_{t_k}\big)\Big|\widehat{Y}_{t_k}\Bigg]\Bigg|\Bigg|_2^2\\ &\le \frac{2}{\Delta^2}\underbrace{\Bigg|\Bigg|\mathbb{E}\Bigg[\Big[\sigma\big(\overline{Y}_{t_k}\big)^\top\Big]^{-1}\widetilde{U}_{t_{k+1}}\big(\overline{Y}_{t_{k+1}} - \overline{Y}_{t_k}\big) - \Big[\sigma\big(\widehat{Y}_{t_k}\big)^\top\Big]^{-1}\widehat{U}_{t_{k+1}}\big(\widehat{Y}_{t_{k+1}} - \widehat{Y}_{t_k}\big)|\widehat{Y}_{t_k}\Bigg]\Bigg|\Bigg|_2^2}_{(IIa)}\\ &\quad + 2\underbrace{\Bigg|\Bigg|\mathbb{E}\Bigg[\Big[\sigma\big(\overline{Y}_{t_k}\big)^\top\Big]^{-1}\widetilde{U}_{t_{k+1}}b\big(\overline{Y}_{t_k}\big) - \Big[\sigma\big(\widehat{Y}_{t_k}\big)^\top\Big]^{-1}\widehat{U}_{t_{k+1}}b\big(\widehat{Y}_{t_k}\big)\Big|\widehat{Y}_{t_k}\Bigg]\Bigg|\Bigg|_2^2}_{(IIb)}.\end{aligned}$$

We are hence led to focus now on (IIa) and (IIb). We start by (IIb). By exploiting the definition of conditional expectation with respect to $\widehat{Y}_{t_k}$ as best L^2-approximation among square integrables $\sigma(\widehat{Y}_{t_k})$-measurables random vectors, we find:

$$\begin{aligned}&\Bigg|\Bigg|\mathbb{E}\Bigg[\Big[\sigma\big(\overline{Y}_{t_k}\big)^\top\Big]^{-1}\widetilde{U}_{t_{k+1}}b\big(\overline{Y}_{t_k}\big) - \Big[\sigma\big(\widehat{Y}_{t_k}\big)^\top\Big]^{-1}\widehat{U}_{t_{k+1}}b\big(\widehat{Y}_{t_k}\big)\Big|\widehat{Y}_{t_k}\Bigg]\Bigg|\Bigg|_2^2\\ &\le \Bigg|\Bigg|\Big[\sigma\big(\widehat{Y}_{t_k}\big)^\top\Big]^{-1}\mathbb{E}\Big[\widetilde{U}_{t_{k+1}}\Big|\widehat{Y}_{t_k}\Big]b\big(\widehat{Y}_{t_k}\big) - \Big[\sigma\big(\widehat{Y}_{t_k}\big)^\top\Big]^{-1}\mathbb{E}\Big[\widehat{U}_{t_{k+1}}\Big|\widehat{Y}_{t_k}\Big]b\big(\widehat{Y}_{t_k}\big)\Bigg|\Bigg|_2^2\\ &\le ||(\sigma(\cdot)^\top)^{-1}||_F^2\ \Big|\Big|\mathbb{E}\Big[\widetilde{U}_{t_{k+1}} - \widehat{U}_{t_{k+1}}\Big|\widehat{Y}_{t_k}\Big]b\big(\widehat{Y}_{t_k}\big)\Big|\Big|_2^2 \le \lambda_0\Big|\Big|\mathbb{E}\Big[\widetilde{U}_{t_{k+1}} - \widehat{U}_{t_{k+1}}\Big|\widehat{Y}_{t_k}\Big]b\big(\widehat{Y}_{t_k}\big)\Big|\Big|_2^2\end{aligned}$$

where we have used the fact that $||A||_2 \le ||A||_F$ for any matrix A and Remark 2.5 on the boundedness of the norm of $\sigma(\cdot)^{-1}$. Now notice that the boundedness of the derivative of b implies its uniform continuity and so, since the quantizer $\widehat{Y}_{t_k}$ takes values on a compact set, $b(\widehat{Y}_{t_k})$ is also bounded. Namely, there exists $\bar{c} > 0$, only depending on $\mathcal{L}_3$ and $\mathcal{L}_4$, such that $||b\big(\widehat{Y}_{t_k}\big)||_2^2 \le \bar{c}^2$ and we find

$$\begin{aligned}\lambda_0\ \Big|\Big|\mathbb{E}\Big[\widetilde{U}_{t_{k+1}} - \widehat{U}_{t_{k+1}}\Big|\widehat{Y}_{t_k}\Big]b\big(\widehat{Y}_{t_k}\big)\Big|\Big|_2^2 &\le \lambda_0\ \bar{c}^2\Big|\Big|\mathbb{E}\Big[\widetilde{U}_{t_{k+1}} - \widehat{U}_{t_{k+1}}\Big|\widehat{Y}_{t_k}\Big]\Big|\Big|_2^2\\ &= \lambda_0\ \bar{c}^2\Big|\Big|\widetilde{U}_{t_{k+1}} - \widehat{U}_{t_{k+1}}\Big|\Big|_2^2\end{aligned}$$

where the error $\big|\big|\widetilde{U}_{t_{k+1}} - \widehat{U}_{t_{k+1}}\big|\big|_2^2$ has already been studied in [1] and for this we refer to Eq. (6.3).

We now move to the last term, (IIa). Using again the definition of conditional expectation with respect to $\widehat{Y}_{t_k}$ as best L^2-approximation among square integrables $\sigma(\widehat{Y}_{t_k})$-measurables random vectors and the fact that the Frobenius norm of the matrix $\sigma(\cdot)^{-1}$ is bounded (see Remark 2.5), we have:

$$\begin{aligned}&\left|\left|\mathbb{E}\left[\left[\sigma\left(\overline{Y}_{t_k}\right)^\top\right]^{-1}\widetilde{U}_{t_{k+1}}\left(\overline{Y}_{t_{k+1}}-\overline{Y}_{t_k}\right)-\left[\sigma\left(\widehat{Y}_{t_k}\right)^\top\right]^{-1}\widehat{U}_{t_{k+1}}\left(\widehat{Y}_{t_{k+1}}-\widehat{Y}_{t_k}\right)|\widehat{Y}_{t_k}\right]\right|\right|_2^2\\
&\le\left|\left|\left[\sigma\left(\widehat{Y}_{t_k}\right)^\top\right]^{-1}\mathbb{E}\left[\widetilde{U}_{t_{k+1}}\left(\overline{Y}_{t_{k+1}}-\widehat{Y}_{t_k}\right)|\widehat{Y}_{t_k}\right]-\left[\sigma\left(\widehat{Y}_{t_k}\right)^\top\right]^{-1}\mathbb{E}\left[\widehat{U}_{t_{k+1}}\left(\widehat{Y}_{t_{k+1}}-\widehat{Y}_{t_k}\right)|\widehat{Y}_{t_k}\right]\right|\right|_2^2\\
&\le\lambda_0\left|\left|\mathbb{E}\left[\mathbb{E}\left[\widetilde{U}_{t_{k+1}}\left(\overline{Y}_{t_{k+1}}-\widehat{Y}_{t_k}\right)|\widehat{Y}_{t_{k+1}}\right]|\widehat{Y}_{t_k}\right]-\mathbb{E}\left[\widehat{U}_{t_{k+1}}\left(\widehat{Y}_{t_{k+1}}-\widehat{Y}_{t_k}\right)|\widehat{Y}_{t_k}\right]\right|\right|_2^2\end{aligned}$$

where in the last equality we used $\sigma(\widehat{Y}_{t_k})\subseteq\sigma(\widehat{Y}_{t_{k+1}})$. Now, by definition of conditional expectation with respect to $\widehat{Y}_{t_{k+1}}$ we find:

$$\begin{aligned}&\lambda_0\left|\left|\mathbb{E}\left[\mathbb{E}\left[\widetilde{U}_{t_{k+1}}\left(\overline{Y}_{t_{k+1}}-\widehat{Y}_{t_k}\right]|\widehat{Y}_{t_{k+1}}\right)|\widehat{Y}_{t_k}\right]-\mathbb{E}\left[\widehat{U}_{t_{k+1}}\left(\widehat{Y}_{t_{k+1}}-\widehat{Y}_{t_k}\right)|\widehat{Y}_{t_k}\right]\right|\right|_2^2\\
&\le\lambda_0\left|\left|\mathbb{E}\left[\left(\widetilde{U}_{t_{k+1}}-\widehat{U}_{t_{k+1}}\right)\left(\widehat{Y}_{t_{k+1}}-\widehat{Y}_{t_k}\right)|\widehat{Y}_{t_k}\right]\right|\right|_2^2\le\lambda_0\left|\left|\widetilde{U}_{t_{k+1}}-\widehat{U}_{t_{k+1}}\right|\right|_2^2\cdot\left|\left|\widehat{Y}_{t_{k+1}}-\widehat{Y}_{t_k}\right|\right|_2^2\end{aligned}$$

where in the last passage we have used conditional Cauchy-Schwartz inequality. Hence, it appears again the error term $||\widetilde{U}_{t_{k+1}}-\widehat{U}_{t_{k+1}}||_2^2$, for which we refer to Eq. (6.3). It remains, then, to deal with the error $||\widehat{Y}_{t_{k+1}}-\widehat{Y}_{t_k}||_2^2$.

We begin this last part of the proof by recalling that, for every $k=0,\dots,N$, $\widehat{Y}_{t_k}$ is the stationary quantizer relative to $\overline{Y}_{t_k}$ obtained via recursive marginal quantization as explained in Section 4. Namely, $\mathbb{E}(\overline{Y}_{t_k}|\widehat{Y}_{t_k})=\widehat{Y}_{t_k}$ and so, via conditionl Jensen's inequality and the tower property, in case when g is convex we have $\mathbb{E}[g(\widehat{Y}_{t_k})]\le\mathbb{E}[g(\overline{Y}_{t_k})]$. So when g is the square function, we find

$$\begin{aligned}\left|\left|\widehat{Y}_{t_{k+1}}-\widehat{Y}_{t_k}\right|\right|_2^2&=\mathbb{E}\left[(\widehat{Y}_{t_{k+1}}-\widehat{Y}_{t_k})^2\right]=\mathbb{E}\left[(\widehat{Y}_{t_{k+1}})^2+(\widehat{Y}_{t_k})^2-2\widehat{Y}_{t_k}\widehat{Y}_{t_{k+1}}\right]\\
&\le\mathbb{E}\left[(\overline{Y}_{t_{k+1}})^2+(\overline{Y}_{t_k})^2-2\mathbb{E}\left[\overline{Y}_{t_{k+1}}\mathbb{E}\left[\overline{Y}_{t_k}|\widehat{Y}_{t_k}\right]|\widehat{Y}_{t_{k+1}}\right]\right]\\
&=\left|\left|\overline{Y}_{t_{k+1}}-\overline{Y}_{t_k}\right|\right|_2^2\le\frac{\tilde{c}}{n},\end{aligned}$$

for a positive $\tilde{c}$ only depending on $\mathcal{L}_3$ and where we recalled the L^2-estimate associated to the increments in the Euler scheme. To conclude it suffices to collect all the terms. □

7. Numerical tests

In this section, we present two numerical experiments based on an implementation of our quantization-based BSDE solver. The first experiment involves a one-dimensional linear BSDE where the solution for the value process and the control is known in closed-form. This first test allows us to compare our newly proposed numerical approximation for the control with the closed-form solution. The second example focuses on a two-dimensional BSDE with known explicit solution, allowing us to validate the procedure also in the multivariate case. The implementation of the routines was performed by means of the Java programming language and it is available at https://github.com/AlessandroGnoatto. Numerical tests were performed on a laptop equipped with a 4 core 2.9 GHz Intel Core i7 processor with 16 GB of RAM.

7.1. A linear BSDE: hedging in the Black-Scholes model

We first consider the linear FSDE:

$$dY_t=rY_tdt+\sigma Y_tdW_t,\ Y_0=y_0>0,$$

where $r=0.04$, $\sigma=0.25$ and $y_0=100$. We associate to this forward process the BSDE

$$U_t=\xi+\int_t^T f(s,Y_s,U_s,V_s)ds-\int_t^T V_s^\top dW_s,\quad t\in[0,T],$$

with

$$\xi=(Y_T-K)^+,\quad f(t,y,u,v)=-rv,$$

for $K=100$ and $T=1$. This corresponds to the well-known Black-Scholes model for the evaluation of a European Call option on Y, maturity $T=1$ and strike price K. For this BSDE the solution is analytically known, namely the process Y is given by a direct application of the Black-Scholes formula, whereas the control satisfies

$$V_t=\frac{\partial U}{\partial y}(t,Y_t)\sigma Y_t=\mathcal{N}(d_1(t,Y_t))\sigma Y_t,$$

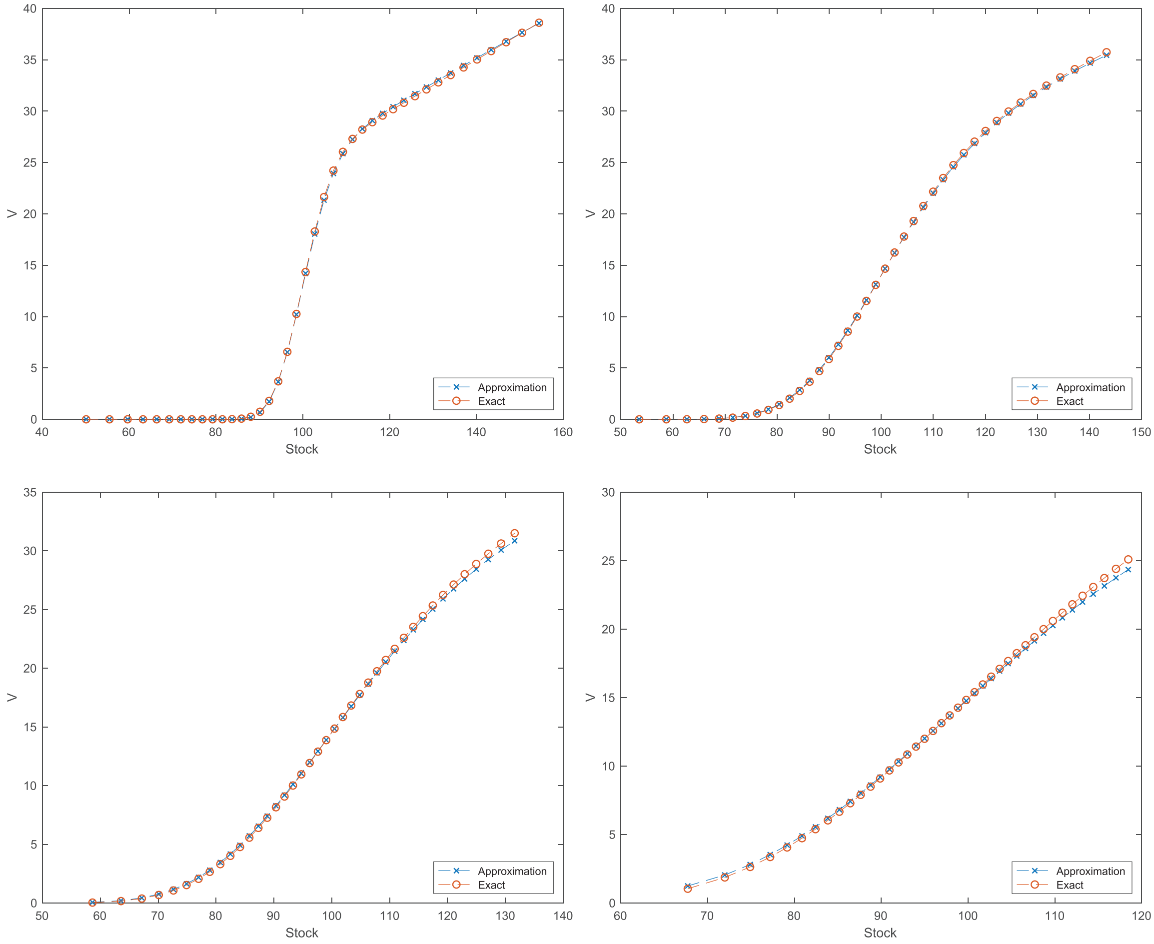

Fig. 1. Comparison between the exact and approximated hedging over the quantization grid at different time steps. Top left panel: 20th and terminal time step. Top right panel: 15*th* time step. Bottom left panel 10th time step. Bottom right panel 5*th* time step.

where $\mathcal{N}$ is the cumulative distribution function of the standard Gaussian and

$$d_1(t, y) := \frac{\log \frac{y}{K} + \left(r + \frac{\sigma^2}{2}\right)(T - t)}{\sigma\sqrt{T - t}}.$$

This example provides a validation of our proposed methodology in a simple case where a closed form solution to the BSDE is known. The exact solution for U, given the specified data, is $U_0 = 11.8370$. We apply our proposed algorithm by using a quantization grid consisting of 50 points, a time discretization with 20 points and a uniform time mesh. The approximate initial value for the price U is 11.7548. Since the novelty of our approach is given by the new scheme for the control, we show that the scheme produces a reliable approximation for the control by comparing our approximation with the exact solution. The reader is referred to Fig. 1, where we compare the exact and the quantization-based approximation for V over the quantization grid. We observe that the newly proposed scheme provides a very good approximation.

Concerning the execution time, we observe that the generation of the quantization grid for the present experiment requires 3447 milliseconds and the computation of the backward recursion is completed in 52 milliseconds. Such values represent an improvement with respect to [1] because we are able to approximate the control process by means of a single run of the algorithm. This is not the case in [1]: their quantized control is expressed in terms of the following weights $\pi_{ij}^{W,k}$ for $k = 0, \dots, n$ with (see Remark 6.4)

$$\pi_{ij}^{W,k} := \frac{1}{\mathbb{P}(\widehat{Y}_{t_k} = \mathbf{y}_\mathbf{k}^\mathbf{i})} \mathbb{E}\left[(W_{t_{k+1}} - W_{t_k})\mathbf{1}_{\{\widehat{Y}_{t_{k+1}} = \mathbf{y}_{\mathbf{k+1}}^\mathbf{j}, \widehat{Y}_{t_k} = \mathbf{y}_\mathbf{k}^\mathbf{i}\}}\right],$$

which need to be estimated by means of Monte Carlo simulations at each step of the backward recursion. In [1] they use 10^7 paths. To give a comparison, we performed a simulation of 10^7 Brownian paths in the context of our Java implementation and this required additional 3648 milliseconds. Such paths, once simulated, can be stored and reused on every step of the backward recursion, however one should add also the computational time required by the estimation of $\pi_{ij}^{W,k}$.

Table 1
Prices for the two-dimensional basket put option. $Quant_{n_1}^{n_2,n_3}$ means quantization with n_1 time steps and (n_2, n_3) quantization points for (Y^1, Y^2).

K	$Quant_{10}^{10,10}$	$Quant_{10}^{15,15}$	$Quant_{10}^{20,20}$ (rel. err.)	Fourier Transform
100.00	7.1230	7.3901	7.4910(1, 75%)	7.6246
105.00	9.6365	9.8937	9.9941(1, 55%)	10.151
110.00	12.522	12.773	12.872(1, 35%)	13.048
115.00	15.755	15.996	16.091(1, 16%)	16.280
120.00	19.298	19.522	19.609(1, 00%)	19.808

In summary, a conservative estimate of the computation time of the methodology of [1] is twice as large in this case.

7.2. A Multivariate BSDE: pricing a basket option

Here we consider a multi-dimensional example, arising in financial mathematics in the context of option pricing, where we consider two underlying assets. We first consider a linear SDE of the form:

$$d\begin{pmatrix} Y_t^1 \\ Y_t^2 \end{pmatrix} = r\begin{pmatrix} Y_t^1 \\ Y_t^2 \end{pmatrix}dt + \begin{pmatrix} \sigma^1 Y_t^1 & 0 \\ 0 & \sigma^2 Y_t^2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix} d\begin{pmatrix} W_t^1 \\ W_t^2 \end{pmatrix}$$

with initial data $(Y_0^1, Y_0^2)^\top = (y_0^1, y_0^2)^\top \in \mathbb{R}_+^2$ and where $r = 0.04$, $\sigma^1 = 0.3$, $\sigma^1 = 0.2$, $\rho = 0.7$ and $y_0^1 = y_0^2 = 100$ and the two Brownian motions W^1, W^2 are independent. We associate to this forward process the BSDE

$$U_t = \xi + \int_t^T f(s, Y_s, U_s, V_s)ds - \int_t^T V_s^\top dW_s, \quad t \in [0, T],$$

with

$$\xi = \left(K - 0.6Y_T^1 - 0.4Y_T^2\right)^+, \quad f(t, y, u, v) = -rv,$$

for $K = \{100, 105, 110, 115, 120\}$ and $T = 1$. This corresponds to pricing a basket put option, for which the price in semi closed-form is available in the literature and is given in terms of a two-dimensional Fourier integral. Indeed, by proceeding along the lines of [28], it is possible to prove the following lemma, which provides the Fourier transform of a basket put option in dimension d.

Lemma 7.1. *Let $x_j, j = 1, \dots, d$ be log-asset price processes. Let $\omega_j, j = 1, \dots, d$ be the corresponding quantities. Finally, denote by $K > 0$ the strike of a d-dimensional basket option having payoff*

$$\left(K - \sum_{j=1}^d \omega_j e^{x_j}\right)^+$$

Then the Fourier transform of the payoff function of the d-dimensional basket option above is given by

$$\Phi(z_1, \dots, z_d) = \frac{K^{1+\sum_{j=1}^d \mathrm{i}z_j}}{\prod_{j=1}^d \omega_j^{\mathrm{i}z_j}} \frac{\prod_{j=1}^d \Gamma(\mathrm{i}z_j)}{\Gamma(2+\sum_{j=1}^d \mathrm{i}z_j)}$$
$$\mathrm{Im}\left(z_j\right) < 0, j = 1, \dots, d$$

where $\Gamma(z), z \in \mathbb{C}$, is the complex gamma function.

We fix $d = 2$ and, by means of Lemma 7.1, we can recover the exact price at time t via

$$U_t = e^{-r(T-t)} \frac{1}{(2\pi)^2} \int_{-\infty+\mathrm{i}\,\mathrm{Im}\,(z_1)}^{\infty+\mathrm{i}\,\mathrm{Im}\,(z_1)} \int_{-\infty+\mathrm{i}\,\mathrm{Im}\,(z_2)}^{\infty+\mathrm{i}\,\mathrm{Im}\,(z_2)} \varphi(-z_1, -z_2)\Phi(z_1, z_2)dz_1dz_2,$$

where φ denotes the joint characteristic function of the logarithmic asset prices. In the case of the bivariate Black-Scholes model we have, for $(z_1, z_2) \in \mathbb{C}^2$

$$\varphi(z_1, z_2) = \exp\left\{ \mathrm{i}(z_1, z_2)\begin{pmatrix} \log Y_t^1 + \left(r - \frac{(\sigma^1)^2}{2}\right)(T-t) \\ \log Y_t^2 + \left(r - \frac{(\sigma^2)^2}{2}\right)(T-t) \end{pmatrix} - \frac{1}{2}(z_1, z_2)\begin{pmatrix} (\sigma^1)^2 & \rho\sigma^1\sigma^2 \\ \rho\sigma^1\sigma^2 & (\sigma^2)^2 \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \right\}.$$

The two-dimensional integral for the semi-closed form price has been truncated and approximated via a two-dimensional Fast Fourier Transform with 256 points per dimension. In Table 1 we report a comparison between the quantization price and the Fourier price, which is our benchmark. Each row provides the prices for a different choice of the strike price. We use 10 points to discretize the time dimension and we use different sizes of the quantization grid: 10, 15 and 20 points. Even though we are using a relatively low number of quantization points, we observe a satisfactory precision of the quantization-based approach: using just 20 points, the relative error is always below 2%, as shown in Table 1.

Let us comment on the calculation time also for the present experiment. The generation of the quantization grid is the most demanding part of the algorithm. For the two-dimensional quantization grids with 10,15,20 points per dimension the computation time is, respectively, 5913, 30284 and 91415 milliseconds. The subsequent computation of the backward recursion requires on average 450 milliseconds. We emphasize again that applying the methodology of [1] would require an additional Monte Carlo simulation: using 10^7 paths to generate the trajectories of a Brownian motion increases the execution time by 6780 milliseconds. We recall that, as an additional step, inside the backward recursion one needs to estimate the transition weights $\pi_{ij}^{W,k}$.

For the sake of completeness we report that the two-dimensional Fourier pricer requires on average 395 milliseconds, however let us stress that this refers only to the computation of the expectation at time zero without the price sensitivities. The quantization algorithm instead provides the value of the conditional expectation over the whole quantization tree at each point in time.

8. Conclusion

We provided a useful modification for the scheme of the control in [1] that allows to improve the algorithm for the approximation of the solution of a family of decoupled FBSDEs. Thanks to this simplification, we can apply a fully based recursive marginal quantization approach that does not involve any Monte Carlo simulation in any step of the procedure. We applied the scheme in some univariate and multidimensional FBSDE examples and we found very good results even with a parsimonious number of quantization and time discretization points. This opens the door to more ambitious applications, like the computation of xVA on single and multiple positions, along the lines of [31–33].

Data Availability

Data will be made available on request.

Acknowledgments

This research was partially supported by University of Padova research grants BIRD190200 and DOR1903290. The authors are grateful to Blanka Horvath for fruitful discussions.

Appendix A. Derivation of the numerical scheme

We provide here all the details relative to the scheme derivation.

A1. Scheme for the value process U

Following [34] we provide a step by step derivation of the numerical scheme for the process U. Let (Y, U, V) be the adapted solution to the FBSDE (2.1)–(2.7). Restricting ourselves to two consecutive points in time t_{k+1} and t_k, we write

$$U_{t_k} = U_{t_{k+1}} + \int_{t_k}^{t_{k+1}} f(s, Y_s, U_s, V_s)ds - \int_{t_k}^{t_{k+1}} V_s^\top dW_s \tag{A.1}$$

and taking $\mathcal{F}_{t_k}$-conditional expectations on both sides we get

$$U_{t_k} = \mathbb{E}\left[U_{t_{k+1}}\middle|\mathcal{F}_{t_k}\right] + \int_{t_k}^{t_{k+1}} \mathbb{E}\left[f(s, Y_s, U_s, V_s)|\mathcal{F}_{t_k}\right]ds.$$

Let us first concentrate on the integral term, using $\theta_1 \in [0, 1]$ we write

$$\begin{aligned}\int_{t_k}^{t_{k+1}} \mathbb{E}\left[f(s, Y_s, U_s, V_s)|\mathcal{F}_{t_k}\right]ds = {} & (t_{k+1} - t_k)\Big\{(1-\theta_1)\mathbb{E}\left[f(t_{k+1}, Y_{t_{k+1}}, U_{t_{k+1}}, V_{t_{k+1}})\middle|\mathcal{F}_{t_k}\right] \\ & + \theta_1 f(t_k, Y_{t_k}, U_{t_k}, V_{t_k})\Big\} + R^U\end{aligned}$$

where the error term R^U is defined as

$$\begin{aligned}R^U := {} & \int_{t_k}^{t_{k+1}} \Big(\mathbb{E}\left[f(s, Y_s, U_s, V_s)|\mathcal{F}_{t_k}\right] - (1-\theta_1)\mathbb{E}\left[f(t_{k+1}, Y_{t_{k+1}}, U_{t_{k+1}}, V_{t_{k+1}})\middle|\mathcal{F}_{t_k}\right] \\ & + \theta_1 f(t_k, Y_{t_k}, U_{t_k}, V_{t_k})\Big)ds.\end{aligned}$$

Hence we arrive at

$$\begin{aligned}U_{t_k} = {} & \mathbb{E}\left[U_{t_{k+1}}\middle|\mathcal{F}_{t_k}\right] + (t_{k+1} - t_k)\Big\{(1-\theta_1)\mathbb{E}\left[f(t_{k+1}, Y_{t_{k+1}}, U_{t_{k+1}}, V_{t_{k+1}})\middle|\mathcal{F}_{t_k}\right] \\ & + \theta_1 f(t_k, Y_{t_k}, U_{t_k}, V_{t_k})\Big\} + R^U.\end{aligned}$$

Remark A.1. In most situations, we do not have an exact simulation scheme for the solution of the forward SDE (2.1). This means in general that we are not able to simulate Y (i.e. the exact solution of (2.1)), and we need to introduce a suitable discretization $\overline{Y}$, such as the Euler-Maruyama scheme, the Milstein discretization or higher order scheme as presented e.g. in [35].

For the moment, let $\overline{Y}$ be a discretization scheme for Y, which is still left unspecified. We write

$$\begin{aligned} U_{t_k} &= \mathbb{E}\Big[U_{t_{k+1}}\Big|\mathcal{F}_{t_k}\Big] + (t_{k+1}-t_k)\Big\{(1-\theta_1)\Big(\mathbb{E}\Big[f(t_{k+1},\overline{Y}_{t_{k+1}},U_{t_{k+1}},V_{t_{k+1}})\Big|\mathcal{F}_{t_k}\Big]+R^{f1}\Big)\\ &\quad+\theta_1\Big(f(t_k,\overline{Y}_{t_k},U_{t_k},V_{t_k})+R^{f2}\Big)\Big\}+R^U, \end{aligned}$$

where

$$\begin{aligned} R^{f1} &:= \mathbb{E}\Big[f(t_{k+1},Y_{t_{k+1}},U_{t_{k+1}},V_{t_{k+1}})\Big|\mathcal{F}_{t_k}\Big] - \mathbb{E}\Big[f(t_{k+1},\overline{Y}_{t_{k+1}},U_{t_{k+1}},V_{t_{k+1}})\Big|\mathcal{F}_{t_k}\Big]\\ R^{f2} &:= f(t_k,Y_{t_k},U_{t_k},V_{t_k}) - f(t_k,\overline{Y}_{t_k},U_{t_k},V_{t_k}). \end{aligned}$$

Setting $R^f := (1-\theta_1)R^{f1}+\theta_1 R^{f2}$ we finally arrive at

$$\begin{aligned} U_{t_k} &= \mathbb{E}\Big[U_{t_{k+1}}\Big|\mathcal{F}_{t_k}\Big] + (t_{k+1}-t_k)\Big\{(1-\theta_1)\mathbb{E}\Big[f(t_{k+1},\overline{Y}_{t_{k+1}},U_{t_{k+1}},V_{t_{k+1}})\Big|\mathcal{F}_{t_k}\Big]\\ &\quad+\theta_1 f(t_k,\overline{Y}_{t_k},U_{t_k},V_{t_k})\Big\}+R^U+R^f. \end{aligned} \tag{A.2}$$

We observe that in (A.2) the discretization error is due to the time discretization and the choice of the numerical scheme for the forward process Y. Further sources of error will arise in the space dimension as we will approximate the conditional expectations appearing in (A.2).

A2. Scheme for the control

We derive the newly proposed scheme for the numerical approximation of the control process V. What is tipically done in the literature is obtaining a discretization scheme for V which involves the increments of the Brownian motion. This is done by multiplying Eq. (A.1) by $(W_{t_{k+1}}-W_{t_k})$ and then taking as usual conditional expectations and truncating the error terms. We will proceed here in a different way, which is new, up to our knowledge. Our objective, indeed, is to derive an update rule for the control that only involves Y (i.e. the process that we will quantize in the sequel) and not W. To this end we consider again the BSDE (A.1) and multiply both sides by $\int_{t_k}^{t_{k+1}}\sigma(Y_s)^\top dW_s$:

$$\begin{aligned} U_{t_k}\int_{t_k}^{t_{k+1}}\sigma(Y_s)^\top dW_s &= U_{t_{k+1}}\int_{t_k}^{t_{k+1}}\sigma(Y_s)^\top dW_s + \int_{t_k}^{t_{k+1}} f(r,Y_r,U_r,V_r)dr\int_{t_k}^{t_{k+1}}\sigma(Y_s)^\top dW_s\\ &\quad-\int_{t_k}^{t_{k+1}}V_s^\top dW_s\int_{t_k}^{t_{k+1}}\sigma(Y_s)^\top dW_s. \end{aligned} \tag{A.3}$$

We take then $\mathcal{F}_{t_k}$-conditional expectations on both sides, thus obtaining the following identity

$$\begin{aligned} \underbrace{U_{t_k}\mathbb{E}\left[\int_{t_k}^{t_{k+1}}\sigma(Y_s)^\top dW_s\middle|\mathcal{F}_{t_k}\right]}_{(A)} &= \underbrace{\mathbb{E}\left[U_{t_{k+1}}\int_{t_k}^{t_{k+1}}\sigma(Y_s)^\top dW_s\middle|\mathcal{F}_{t_k}\right]}_{(B)}\\ &\quad+\underbrace{\mathbb{E}\left[\int_{t_k}^{t_{k+1}} f(r,Y_r,U_r,V_r)dr\int_{t_k}^{t_{k+1}}\sigma(Y_s)^\top dW_s\middle|\mathcal{F}_{t_k}\right]}_{(C)}\\ &\quad-\underbrace{\mathbb{E}\left[\int_{t_k}^{t_{k+1}}V_s^\top dW_s\int_{t_k}^{t_{k+1}}\sigma(Y_s)^\top dW_s\middle|\mathcal{F}_{t_k}\right]}_{(D)}. \end{aligned} \tag{A.4}$$

We now analyze every conditional expectation in (A.4) starting from (D):

- (D) Via Itô isometry we find
$\mathbb{E}\left[\int_{t_k}^{t_{k+1}}V_s^\top dW_s\int_{t_k}^{t_{k+1}}\sigma(Y_s)^\top dW_s\middle|\mathcal{F}_{t_k}\right]=\mathbb{E}\left[\int_{t_k}^{t_{k+1}}\sigma(Y_s)^\top V_s ds\middle|\mathcal{F}_{t_k}\right]$
and using $\theta_2\in[0,1]$ we have
$$\mathbb{E}\left[\int_{t_k}^{t_{k+1}}\sigma(Y_s)^\top V_s ds\middle|\mathcal{F}_{t_k}\right] = (t_{k+1}-t_k)\left\{(1-\theta_2)\mathbb{E}\left[\sigma\left(Y_{t_{k+1}}\right)^\top V_{t_{k+1}}\middle|\mathcal{F}_{t_k}\right]+\theta_2\sigma\left(Y_{t_k}\right)^\top V_{t_k}\right\}+R^{V-\theta},$$
where
$R^{V-\theta} := \mathbb{E}\left[\int_{t_k}^{t_{k+1}}\left[\sigma(Y_s)^\top V_s-(1-\theta_2)\sigma\left(Y_{t_{k+1}}\right)^\top V_{t_{k+1}}-\theta_2\sigma\left(Y_{t_k}\right)^\top V_{t_k}\right]ds\middle|\mathcal{F}_{t_k}\right].$

We now take into account the impact of the numerical scheme to approximate Y, namely we insert $\overline{Y}$:

$$\mathbb{E}\left[\int_{t_k}^{t_{k+1}}\sigma(Y_s)^\top V_s ds\Big|\mathcal{F}_{t_k}\right] = (t_{k+1}-t_k)\left\{(1-\theta_2)\mathbb{E}\left[\sigma\left(\overline{Y}_{t_{k+1}}\right)^\top V_{t_{k+1}}\Big|\mathcal{F}_{t_k}\right]+\theta_2\sigma\left(\overline{Y}_{t_k}\right)^\top V_{t_k}\right\} + R^{V-\theta}+R^{V-Y},$$

with

$$R^{V-Y} := (t_{k+1}-t_k)\left\{(1-\theta_2)\left(\sigma\left(Y_{t_{k+1}}\right)^\top - \sigma\left(\overline{Y}_{t_{k+1}}\right)^\top\right)V_{t_{k+1}} + \theta_2\left(\sigma\left(Y_{t_{k+1}}\right)^\top - \sigma\left(\overline{Y}_{t_{k+1}}\right)^\top\right)V_{t_k}\right\}.$$

- (C) We have:

$$\mathbb{E}\left[\int_{t_k}^{t_{k+1}} f(r,Y_r,U_r,V_r)dr\int_{t_k}^{t_{k+1}}\sigma(Y_s)^\top dW_s\Big|\mathcal{F}_{t_k}\right]=0.$$

Indeed, we write:

$$\mathbb{E}\left[\int_{t_k}^{t_{k+1}} f(r,Y_r,U_r,V_r)dr\int_{t_k}^{t_{k+1}}\sigma(Y_s)^\top dW_s\Big|\mathcal{F}_{t_k}\right]=\mathbb{E}\left[\int_{t_k}^{t_{k+1}}\left(\int_{t_k}^{t_{k+1}} f(r,Y_r,U_r,V_r)dr\right)\sigma(Y_s)^\top dW_s\Big|\mathcal{F}_{t_k}\right]$$

and the claim follows since $\sigma(Y_s)$ is almost surely bounded thanks to Assumption 2.4 and given that $|\int_0^T f(r,Y_r,U_r,V_r)|$ is in L^2 (recall that by assumption f is a standard parameter).

- (A) Here, too:

$$U_{t_k}\mathbb{E}\left[\int_{t_k}^{t_{k+1}}\sigma(Y_s)^\top dW_s\Big|\mathcal{F}_{t_k}\right]=0.$$

- (B) A distinctive feature of our numerical scheme is based on the following simple observation: we can exploit the dynamics (2.1) to express the stochastic integral in (B) as follows

$$\mathbb{E}\left[U_{t_{k+1}}\int_{t_k}^{t_{k+1}}\sigma(Y_s)^\top dW_s\Big|\mathcal{F}_{t_k}\right]=\mathbb{E}\left[U_{t_{k+1}}\left(Y_{t_{k+1}}-Y_{t_k}-\int_{t_k}^{t_{k+1}}b(Y_s)ds\right)\Big|\mathcal{F}_{t_k}\right]. \tag{A.5}$$

Splitting the conditional expectation on the right hand side, we obtain two simple conditional expectations that can be suitably estimated, once we have an approximation for the transition probabilities of the forward process Y. We write

$$\mathbb{E}\left[U_{t_{k+1}}\left(Y_{t_{k+1}}-Y_{t_k}\right)\Big|\mathcal{F}_{t_k}\right]=\mathbb{E}\left[U_{t_{k+1}}\left(\overline{Y}_{t_{k+1}}-\overline{Y}_{t_k}\right)\Big|\mathcal{F}_{t_k}\right]+R^{U-Y},$$

where

$R^{U-Y} := \mathbb{E}\left[U_{t_{k+1}}\left(Y_{t_{k+1}}-\overline{Y}_{t_{k+1}}\right)\Big|\mathcal{F}_{t_k}\right]-\mathbb{E}\left[U_{t_{k+1}}\Big|\mathcal{F}_{t_k}\right]\left(Y_{t_k}-\overline{Y}_{t_k}\right)$,

while for the second expectation in (A.5) we have

$$\mathbb{E}\left[U_{t_{k+1}}\int_{t_k}^{t_{k+1}}b(Y_s)ds\Big|\mathcal{F}_{t_k}\right] = (t_{k+1}-t_k)(1-\theta_2)\mathbb{E}\left[U_{t_{k+1}}b\left(\overline{Y}_{t_{k+1}}\right)\Big|\mathcal{F}_{t_k}\right] + (t_{k+1}-t_k)\theta_2\mathbb{E}\left[U_{t_{k+1}}\Big|\mathcal{F}_{t_k}\right]b\left(\overline{Y}_{t_k}\right) + R^{b-\theta}+R^{b-Y},$$

where

$$\begin{aligned} R^{b-\theta} &:= \mathbb{E}\left[U_{t_{k+1}}\left(\int_{t_k}^{t_{k+1}}b(Y_s)ds-(t_{k+1}-t_k)\left\{(1-\theta_2)b\left(Y_{t_{k+1}}\right)+\theta_2 b\left(Y_{t_k}\right)\right\}\right)\Big|\mathcal{F}_{t_k}\right]\\ R^{b-Y} &:= (t_{k+1}-t_k)(1-\theta_2)\mathbb{E}\left[U_{t_{k+1}}\left(b\left(Y_{t_{k+1}}\right)-b\left(\overline{Y}_{t_{k+1}}\right)\right)\Big|\mathcal{F}_{t_k}\right]\\ &\quad + (t_{k+1}-t_k)\theta_2\mathbb{E}\left[U_{t_{k+1}}\Big|\mathcal{F}_{t_k}\right]\left(b\left(Y_{t_k}\right)-b\left(\overline{Y}_{t_k}\right)\right). \end{aligned}$$

By regrouping all terms (A), (B), (C) and (D) we obtain the following relation, providing an implicit update rule for the control process V (the explicit rule for the control V will be specified in the next subsection):

$$\begin{aligned} 0 &= \mathbb{E}\left[U_{t_{k+1}}\left(\overline{Y}_{t_{k+1}}-\overline{Y}_{t_k}\right)\Big|\mathcal{F}_{t_k}\right]+R^{U-Y}-(t_{k+1}-t_k)(1-\theta_2)\mathbb{E}\left[U_{t_{k+1}}b\left(\overline{Y}_{t_{k+1}}\right)\Big|\mathcal{F}_{t_k}\right]\\ &\quad -(t_{k+1}-t_k)\theta_2\mathbb{E}\left[U_{t_{k+1}}\Big|\mathcal{F}_{t_k}\right]b\left(\overline{Y}_{t_k}\right)-R^{b-\theta}-R^{b-Y}\\ &\quad -(t_{k+1}-t_k)\left\{(1-\theta_2)\mathbb{E}\left[\sigma\left(\overline{Y}_{t_{k+1}}\right)^\top V_{t_{k+1}}\Big|\mathcal{F}_{t_k}\right]+\theta_2\,\sigma\left(\overline{Y}_{t_k}\right)^\top V_{t_k}\right\}-R^{V-\theta}-R^{V-Y}. \end{aligned} \tag{A.6}$$

A3. The truncated scheme

Starting from Eqs. (A.2) and (A.6) and by truncating all error terms, we obtain the following system of two equations (for each k) for the couple $(\widetilde{U},\widetilde{V})$, where $(\widetilde{U},\widetilde{V})$ are approximations of (U,V) where we recall that $\overline{Y}$ is a suitable discretization of Y

$$\begin{aligned}\widetilde{U}_{t_k} &= \mathbb{E}\Big[\widetilde{U}_{t_{k+1}}\Big|\mathcal{F}_{t_k}\Big] + (t_{k+1}-t_k)\Big\{(1-\theta_1)\mathbb{E}\Big[f(t_{k+1},\overline{Y}_{t_{k+1}},\widetilde{U}_{t_{k+1}},\widetilde{V}_{t_{k+1}})\Big|\mathcal{F}_{t_k}\Big] \\ &\quad +\theta_1 f(t_k,\overline{Y}_{t_k},\widetilde{U}_{t_k},\widetilde{V}_{t_k})\Big\} \\ 0 &= \mathbb{E}\Big[\widetilde{U}_{t_{k+1}}\Big(\overline{Y}_{t_{k+1}}-\overline{Y}_{t_k}\Big)\Big|\mathcal{F}_{t_k}\Big] \\ &\quad -(t_{k+1}-t_k)(1-\theta_2)\mathbb{E}\Big[\widetilde{U}_{t_{k+1}}b\Big(\overline{Y}_{t_{k+1}}\Big)\Big|\mathcal{F}_{t_k}\Big] \\ &\quad -(t_{k+1}-t_k)\theta_2\mathbb{E}\Big[\widetilde{U}_{t_{k+1}}\Big|\mathcal{F}_{t_k}\Big]b\Big(\overline{Y}_{t_k}\Big) \\ &\quad -(t_{k+1}-t_k)\left\{(1-\theta_2)\mathbb{E}\left[\sigma\Big(\overline{Y}_{t_{k+1}}\Big)^\top\widetilde{V}_{t_{k+1}}\middle|\mathcal{F}_{t_k}\right]+\theta_2\sigma\Big(\overline{Y}_{t_k}\Big)^\top\widetilde{V}_{t_k}\right\}.\end{aligned} \tag{A.7}$$

Remark A.2. The second equation above (which is the truncation of Eq. (A.6)) provides an approximation scheme for $\widetilde{V}_{t_k}$ as a function of $\widetilde{V}_{t_{k+1}},\widetilde{U}_{t_{k+1}},\overline{Y}_{t_k},\overline{Y}_{t_{k+1}}$.

In particular, if we set $\theta_1=\theta_2=1$, we obtain the recursive scheme (which is not yet fully explicit):

$$\begin{cases}\widetilde{U}_{t_k} = \mathbb{E}\Big[\widetilde{U}_{t_{k+1}}\Big|\mathcal{F}_{t_k}\Big] + (t_{k+1}-t_k)f(t_k,\overline{Y}_{t_k},\widetilde{U}_{t_k},\widetilde{V}_{t_k}) \\ \widetilde{V}_{t_k} = \frac{1}{(t_{k+1}-t_k)}\left[\sigma\Big(\overline{Y}_{t_k}\Big)^\top\right]^{-1}\mathbb{E}\Big[\widetilde{U}_{t_{k+1}}\Big(\overline{Y}_{t_{k+1}}-\overline{Y}_{t_k}\Big)\Big|\mathcal{F}_{t_k}\Big] \\ \qquad\qquad -\left[\sigma\Big(\overline{Y}_{t_k}\Big)^\top\right]^{-1}\mathbb{E}\Big[\widetilde{U}_{t_{k+1}}\Big|\mathcal{F}_{t_k}\Big]b\Big(\overline{Y}_{t_k}\Big),\end{cases}$$

where $\left[\sigma\Big(\overline{Y}_{t_k}\Big)^\top\right]^{-1}$ denotes the $(q\times d)$ left-inverse of the matrix $\sigma\Big(\overline{Y}_{t_k}\Big)^\top$.

Remark A.3. In Section 6, focusing on the error analysis, for simplicity, we consider the case when $q=d$ and we work under Assumption 2.4, which guarantees the invertibility of σ.

In [1] the scheme is made fully explicit by performing a conditioning inside the driver, which results in the following

$$\begin{cases}\widetilde{U}_{t_n} = h(\overline{Y}_{t_n}) \quad \text{and for} \quad k=0,\dots,n-1 \\ \widetilde{U}_{t_k} = \mathbb{E}\Big[\widetilde{U}_{t_{k+1}}\Big|\mathcal{F}_{t_k}\Big] + (t_{k+1}-t_k)f(t_k,\overline{Y}_{t_k},\mathbb{E}\Big[\widetilde{U}_{t_{k+1}}\Big|\mathcal{F}_{t_k}\Big],\widetilde{V}^{\mathrm{PS}}_{t_k}) \\ \widetilde{V}^{\mathrm{PS}}_{t_k} = \frac{1}{(t_{k+1}-t_k)}\mathbb{E}\Big[\widetilde{U}_{t_{k+1}}\Big(W_{t_{k+1}}-W_{t_k}\Big)\Big|\mathcal{F}_{t_k}\Big].\end{cases} \tag{A.8}$$

So, borrowing this idea, we are now in a position to finally state our proposed scheme as (recall that $(t_{k+1}-t_k)=\Delta$ for every $k=0,\dots,n-1$):

$$\begin{cases}\widetilde{U}_{t_n} = h(\overline{Y}_{t_n}) \quad \text{and for} \quad k=0,\dots,n-1 \\ \widetilde{U}_{t_k} = \mathbb{E}\Big[\widetilde{U}_{t_{k+1}}\Big|\mathcal{F}_{t_k}\Big] + \Delta\ f(t_k,\overline{Y}_{t_k},\mathbb{E}\Big[\widetilde{U}_{t_{k+1}}\Big|\mathcal{F}_{t_k}\Big],\widetilde{V}_{t_k}) \\ \widetilde{V}_{t_k} = \frac{1}{\Delta}\left[\sigma\Big(\overline{Y}_{t_k}\Big)^\top\right]^{-1}\mathbb{E}\Big[\widetilde{U}_{t_{k+1}}\Big(\overline{Y}_{t_{k+1}}-\overline{Y}_{t_k}\Big)\Big|\mathcal{F}_{t_k}\Big] \\ \qquad\quad -\left[\sigma\Big(\overline{Y}_{t_k}\Big)^\top\right]^{-1}\mathbb{E}\Big[\widetilde{U}_{t_{k+1}}\Big|\mathcal{F}_{t_k}\Big]b\Big(\overline{Y}_{t_k}\Big).\end{cases} \tag{A.9}$$

References

[1] G. Pagès, A. Sagna, Improved error bounds for quantization based numerical schemes for BSDE and nonlinear filtering, Stoch. Processes Appl. 128 (3) (2018) 847–883, doi:10.1016/j.spa.2017.05.009. http://www.sciencedirect.com/science/article/pii/S0304414915300831.

[2] F. Biagini, A. Gnoatto, I. Oliva, A unified approach to xva with csa discounting and initial margin, SIAM J. Financ. Math. 12 (3) (2021) 1013–1053, doi:10.1137/20M1332153.

[3] J.M. Bismut, Conjugate convex functions in optimal stochastic control, J. Math. Anal. Appl. 44 (2) (1973) 384–404, doi:10.1016/0022-247X(73)90066-8. http://www.sciencedirect.com/science/article/pii/0022247X73900668.

[4] E. Pardoux, S. Peng, Adapted solution of a backward stochastic differential equation, Syst. Control Lett. 14 (1) (1990) 55–61, doi:10.1016/0167-6911(90)90082-6. http://www.sciencedirect.com/science/article/pii/0167691190900826.

[5] S. Tang, X. Li, Necessary conditions for optimal control of stochastic systems with random jumps, SIAM J. Control Optim. 32 (5) (1994) 1447–1475, doi:10.1137/s0363012992233858.

[6] N. El Karoui, C. Kapoudjian, E. Pardoux, S. Peng, M.C. Quenez, Reflected solutions of backward SDE's, and related obstacle problems for PDE's, Ann. Probab. 25 (2) (1997) 702–737, doi:10.1214/aop/1024404416.

[7] J. Cvitanić, I. Karatzas, Backward stochastic differential equations with reflection and Dynkin games, Ann. Probab. 24 (4) (1996) 2024–2056, doi:10.1214/aop/1041903216.

[8] B. Bouchard, J.-F. Chassagneux, Discrete-time approximation for continuously and discretely reflected BSDEs, Stoch. Processes Appl. 118 (2008) 2269–2293.
[9] J.-F. Chassagneux, A discrete-time approximation for doubly reflected BSDEs, Adv. Appl. Probab. 41 (2009) 101–130.
[10] J.-F. Chassagneux, R. Elie, I. Kharroubi, A note on the existence and uniqueness of solutions of reflected BSDEs associated to switching problems, Electron. Commun. Probab. 16 (2011) 120–128.
[11] N. El Karoui, S. Peng, M.C. Quenez, Backward stochastic differential equations in finance, Math. Financ. 7 (1) (1997) 1–71, doi:10.1111/1467-9965.00022.
[12] E. Gobet, J.-P. Lemor, A. Warin, A regression-based Monte Carlo method to solve backward stochastic differential equations, Ann. Appl. Probab. 15 (3) (2005) 2172–2202.
[13] S. Crépey, T.S. Bielecki, D. Brigo, Counterparty Risk and Funding: a Tale of Two Puzzles, Chapman and Hall/CRC Press Series in Financial Mathematics, 31, 1st, Chapman and Hall/CRC, 2014.
[14] S. Crépey, Financial Modeling: a Backward Stochastic Differential Equations Perspective, Springer Finance Textbooks, Springer-Nature, New York, 2013.
[15] P. Briand, B. Deylon, J. Mémin, On the robustness of backward stochastic differential equations, Stoch. Processes Appl. 97 (2) (2002) 229–253.
[16] J. Zhang, A numerical scheme for BSDEs, Ann. Appl. Probab. 14 (1) (2004) 459–488.
[17] B. Bouchard, N. Touzi, Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations, Stoch. Processes Appl. 111 (2) (2004) 175–206, doi:10.1016/j.spa.2004.01.001. http://www.sciencedirect.com/science/article/pii/S0304414904000031.
[18] F.A. Longstaff, E.S. Schwartz, Valuing American options by simulation: a simple least-squares approach, Rev. Financ. Stud. 14 (1) (2001) 113–147.
[19] V. Bally, G. Pagès, A quantization algorithm for solving multi-dimensional discrete-time optimal stopping problems, Bernoulli 9 (6) (2003) 1003–1049.
[20] V. Bally, G. Pagès, J. Printems, A quantization tree method for pricing and hedging multidimensional American options, Math. Financ. 15 (1) (2005) 119–168.
[21] P. Briand, C. Labart, Simulation of BSDEs by wiener chaos expansion, Ann. Appl. Probab. 24 (3) (2014) 1129–1171.
[22] G. Pagès, Introduction to vector quantization and its applications for numerics, ESAIM 48 (2015) 29–79.
[23] G. Pagès, A. Sagna, Recursive marginal quantization of the Euler scheme of a diffusion process, Appl. Math. Financ. 22 (2015) 463–498, doi:10.1080/1350486X.2015.1091741.
[24] G. Callegaro, L. Fiorin, M. Grasselli, Pricing via recursive quantization in stochastic volatility models, Quant. Financ. 17 (6) (2017) 855–872, doi:10.1080/14697688.2016.1255348.
[25] G. Callegaro, L. Fiorin, M. Grasselli, Quantization meets Fourier: a new technology for pricing options, Ann. Oper. Res. 282 (1–2) (2018) 59–86, doi:10.1007/s10479-018-3048-z.
[26] L. Fiorin, G. Pagès, A. Sagna, Product Markovian quantization of a diffusion process with applications to finance, Methodol. Comput. Appl. Probab. 21 (2019) 1087–1118, doi:10.1007/s11009-018-9652-1.
[27] T.A. McWalter, R. Rudd, J. Kienitz, E. Platen, Recursive marginal quantization of higher-order schemes, Quant. Financ. 18 (4) (2018) 693–706, doi:10.1080/14697688.2017.1402125.
[28] T.R. Hurd, Z. Zhou, A fourier transform method for spread option pricing, SIAM J. Financ. Math. 1 (1) (2010) 142–157, doi:10.1137/090750421.
[29] L. Delong, Backward Stochastic Differential Equations With Jumps and Their Actuarial and Financial Applications: BSDEs With Jumps, EAA Series, Springer, London, 2013.
[30] S. Graf, H. Luschgy, Foundations of Quantization for Probability Distributions, Springer-Verlag, Berlin, Heidelberg, 2000.
[31] A. Gnoatto, A. Picarelli, C. Reisinger, Deep xVA solver - a neural network based counterparty credit risk management framework, Preprint, available at https://arxiv.org/pdf/2005.02633.pdf (2020). https://arxiv.org/pdf/2005.02633.pdf.
[32] C. Albanese, S. Crépey, R. Hoskinson, B. Saadeddine, XVA analysis from the balance sheet, Quant. Financ. (2020). https://bit.ly/3fx5EpT.
[33] L. Abbas-Turki, S. Crépey, B. Saadeddine, Deep xVA analysis, presented at research in options 2020, available at https://bit.ly/3fx5EpT (2020). https://bit.ly/3fx5EpT.
[34] W. Zhao, L. Chen, S. Peng, A new kind of accurate numerical method for backward stochastic differential equations, SIAM J. Sci. Comput. 28 (4) (2006) 1563–1581, doi:10.1137/05063341x.
[35] P.E. Kloeden, E. Platen, Numerical Solution of Stochastic Differential Equations, Stochastic Modelling and Applied Probability, corrected, Springer, New York, 1992.