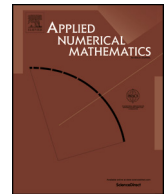




Contents lists available at ScienceDirect

Applied Numerical Mathematics

www.elsevier.com/locate/apnum


Numerical Fourier method and second-order Taylor scheme for backward SDEs in finance

M.J. Ruijter^{a,*}, C.W. Oosterlee^{a,b}^a *Centrum Wiskunde & Informatica, Amsterdam, The Netherlands*^b *Delft University of Technology, Delft, The Netherlands*

ARTICLE INFO

Article history:

Received 9 January 2015

Received in revised form 31 August 2015

Accepted 7 December 2015

Available online 15 January 2016

Keywords:

Fourier cosine expansion method

European and Bermudan options

CEV process

CIR process

Local volatility

Characteristic function

Backward stochastic differential equations

Milstein scheme

Order 2.0 weak Taylor scheme

ABSTRACT

We develop a Fourier method to solve quite general backward stochastic differential equations (BSDEs) with second-order accuracy. The underlying forward stochastic differential equation (FSDE) is approximated by different Taylor schemes, such as the Euler, Milstein, and Order 2.0 weak Taylor schemes, or by exact simulation. A θ -time-discretization of the time-integrands leads to an induction scheme with conditional expectations. The computation of the conditional expectations appearing relies on the availability of the characteristic function for these schemes. We will use the characteristic function of the discrete forward process. The expected values are approximated by Fourier cosine series expansions. Numerical experiments show rapid convergence of our efficient probabilistic numerical method. Second-order accuracy is observed and also proved. We apply the method to, among others, option pricing problems under the Constant Elasticity of Variance and Cox–Ingersoll–Ross processes.

© 2016 IMACS. Published by Elsevier B.V. All rights reserved.

1. Introduction

Backward stochastic differential equations (BSDEs) form an interesting recent concept in financial mathematics. Their range of applicability has increased, for example, by counterparty credit exposure and also insurance applications [6]. The asset dynamics have also been generalized, for example to jump diffusion processes [2]. BSDEs are directly connected to semilinear partial differential equations (PDEs) as the solution to these PDEs can be found by solving the corresponding decoupled forward–backward stochastic differential equation (FBSDE) problem. Recently, several advanced probabilistic numerical methods have been developed for FBSDEs, like least-squares Monte Carlo regression, like the Longstaff–Schwartz method, to approximate the conditional expectations [18,12,1], integration methods [28] and also Fourier methods [14,26]. A rich literature exists on other methods, for example based on chaos decomposition formulas [3]. Our method distinguishes from this in the way the expected values are approximated. Of course, a natural aim for these solution methods is to make them highly efficient so that they can compete with semilinear PDE discretization and solution methods. With the BCOS method (Backward Stochastic Differential Equations COS method), in [26], we developed an efficient FB-SDE solution method for asset dynamics for which a characteristic function of the continuous process can easily be derived.

* Corresponding author.

E-mail addresses: marjonruijter@gmail.com (M.J. Ruijter), c.w.oosterlee@cwi.nl (C.W. Oosterlee).

In the present paper we extend the BCOS method. We apply the well-known Euler and Milstein schemes and an Order 2.0 weak Taylor scheme to the forward stochastic differential equation. Then we explicitly derive the characteristic function for the *discrete form* of a forward stochastic differential equation (FSDE), which will give us opportunities and generalizations. Traditionally, we know that a characteristic function can be derived, as the Fourier transform of the probability density function for models from the class of regular affine processes of [7], which also includes the exponentially affine jump-diffusion class of and some stochastic volatility models, and for exponentially Lévy models. When FSDE dynamics are not affine, like in the case of local volatility dynamics of Constant Elasticity of Variance (CEV) process, on which the Stochastic Alpha Beta Rho (SABR) model is based, we cannot traditionally employ Fourier techniques. The characteristic function corresponding to the discrete process can be however used, as we will show, to price options by a Fourier technique. This feature forms a natural generalization of our BCOS method. In [26] we focused on normal, lognormal and jump processes for which we could relatively easily achieve second-order convergence, in the number of timesteps, in an FBSDE context by means of a θ -time-discretization scheme [15] and the COS Fourier cosine expansion technique. In the present paper the second-order accuracy is achieved, in combination with a fast scheme, also for the CEV model, for time- and spatially-dependent processes, as well as for the Cox–Ingersoll–Ross (CIR) process. In the traditional context, it has been shown in [9] that Fourier cosine expansions will lead to exponentially convergent computational methods for smooth density functions. Here, in the discrete case, we will not observe exponential convergence in the number of Fourier coefficients, but second-order convergence, due to the use of discrete Fourier cosine transforms.

We start in Section 2 with definitions and notation of forward and backward SDEs and present the link between coupled FBSDEs and semilinear partial differential equations. Section 2.2 discusses discrete schemes for the FSDE, such as the Euler, Milstein, and Order 2.0 weak Taylor schemes and exact simulation schemes. Our numerical algorithm is described in Section 3, where we start with the characteristic function of the discretized FSDE and an example of Bermudan options (under the risk-neutral measure). A general time-discretization of a BSDE results in expressions with conditional expectations (Section 3.2). These conditional expectations are computed by the COS formulas (Section 3.3) and the problem is then solved backwards in time. Section 3.4 presents the overall algorithm. An error analysis is performed in Section 4. We demonstrate the numerical method by extensive experiments in Section 5. Section 6 concludes.

2. Backward and forward stochastic differential equations

We start with some notation and definitions for BSDEs, for which we follow the survey paper [8]. Let $\omega = (\omega_t)_{0 \leq t \leq T}$ be a standard one-dimensional Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, with $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ the natural filtration of the Brownian motion ω , and T a fixed finite time horizon. We denote by $\mathbb{H}_T^2(\mathbb{R})$ the set of predictable processes $\eta : \Omega \times [0, T] \rightarrow \mathbb{R}$ such that $\mathbb{E} \left[\int_0^T |\eta_t|^2 dt \right] < \infty$ and by $\mathbb{L}_T^2(\mathbb{R})$ the set of \mathcal{F}_T -measurable random variables $X : \Omega \rightarrow \mathbb{R}$ that are square integrable. We consider the *backward stochastic differential equation*

$$dY_s = -f(s, Y_s, Z_s)ds + Z_s d\omega_s, \quad Y_T = \xi, \quad (2.1)$$

where function $f : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ is $\mathcal{P} \otimes \mathcal{B} \otimes \mathcal{B}$ -measurable. \mathcal{P} is the set of \mathcal{F}_t -progressively measurable scalar processes on $\Omega \times [0, T]$. $f(\cdot)$ is the *generator* or *driver* of the process and the *terminal condition* $\xi : \Omega \rightarrow \mathbb{R}$ is an \mathcal{F}_T -measurable random variable. Here we consider one-dimensional processes, but the BSDE theory can be extended to higher dimensions, with d -dimensional processes ω_t and Y_t and an $n \times d$ -dimensional Z_t process, as described in [8].

The solution to BSDE (2.1) is given by a pair of processes (Y, Z) , with Y a continuous real-valued adapted process and Z a real-valued predictable process satisfying $\int_0^T |Z_t|^2 dt < \infty$, \mathbb{P} a.s., satisfying

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s d\omega_s, \quad 0 \leq t \leq T. \quad (2.2)$$

Unlike a forward stochastic differential equation, the solution of a BSDE is thus defined as a *pair* of adapted processes (Y, Z) . Note that BSDEs cannot be considered as time-reversed FSDEs, because at time t the pair (Y_t, Z_t) is \mathcal{F}_t -measurable and the process does not yet ‘know’ the terminal condition.

Function f and terminal condition ξ are called *standard parameters* for the BSDE, if $\xi \in \mathbb{L}_T^2(\mathbb{R})$, $f(\cdot, 0, 0) \in \mathbb{H}_T^2(\mathbb{R})$, and f is uniformly Lipschitz in y and z , with Lipschitz constant L_f . A result from [8,25,22] is that, given a pair of standard parameters (f, ξ) , there exists a unique solution $(Y, Z) \in \mathbb{H}_T^2(\mathbb{R}) \times \mathbb{H}_T^2(\mathbb{R})$ to BSDE (2.1).

2.1. Decoupled FBSDEs

It is well-known that a *linear* parabolic PDE has a probabilistic representation by means of the Feynman–Kac theorem. Here, we consider a *semilinear* parabolic PDE of the form

$$\frac{\partial v}{\partial t}(t, x) + \mu(x)D_x v(t, x) + \frac{1}{2}\sigma^2(x)D_x^2 v(t, x) + f(t, x, v(t, x), \sigma(x)D_x v(t, x)) = 0, \tag{2.3a}$$

$$(t, x) \in [0, T] \times \mathbb{R},$$

$$v(T, x) = g(x), \quad x \in \mathbb{R}. \tag{2.3b}$$

D_x and D_x^2 denote the first and second derivative of a function with respect to the x -variable, respectively. This PDE also has a probabilistic representation, by means of the following FSDE,

$$X_t = x, \quad dX_s = \mu(X_s)ds + \sigma(X_s)d\omega_s, \quad t \leq s \leq T. \tag{2.4}$$

and BSDE

$$dY_s = -f(s, X_s, Y_s, Z_s)ds + Z_s d\omega_s, \quad Y_T = g(X_T), \tag{2.5}$$

whose terminal condition is determined by the terminal value of FSDE (2.4). Notice that also the driver function now depends on the state of the FSDE. Let $X_s^{t,x}$ denote the solution to (2.4), starting from x at time t , and $(Y_s^{t,x}, Z_s^{t,x})$ be the corresponding solution to the BSDE.

The coefficients $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ and $\mu : \mathbb{R} \rightarrow \mathbb{R}$ in (2.4) are assumed to be Lipschitz in x and satisfy a linear growth condition in x . Functions $f : [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ and $g : \mathbb{R} \rightarrow \mathbb{R}$ are assumed to be uniformly continuous with respect to x . Moreover, f satisfies a Lipschitz condition in (y, z) and there exists a constant C such that $|f(t, x, y, z)| + |g(x)| \leq C(1 + |x|^p + |y| + |z|)$, $p \geq 1/2$. The conditions on f and ξ guarantee the existence of a unique solution (Y, Z) to the BSDE (2.5). The following results hold

Result 2.1. (See [23,25].) Let $v \in C^{1,2}$ be a classical solution to (2.3a) and suppose there exists a constant $C \geq 0$ such that, for all (t, x) , $|v(t, x)| + |\sigma(x)D_x v(t, x)| \leq C(1 + |x|)$. Then the pair (Y, Z) , defined by

$$Y_s^{t,x} = Y_s(X_s^{t,x}) = v(s, X_s^{t,x}), \quad Z_s^{t,x} = Z_s(X_s^{t,x}) = \sigma(X_s^{t,x})D_x v(s, X_s^{t,x}), \quad t \leq s \leq T, \tag{2.6}$$

is the solution to BSDE (2.5) (a so-called verification result).

The converse result states: Suppose (Y, Z) is the solution to the BSDE, then the function defined by $v(t, x) = Y_t^{t,x}$ is a viscosity solution to the PDE.

A PDE can be solved by applying numerical discretization techniques and for FBSDEs probabilistic numerical methods, like Monte Carlo simulation techniques, are available. For our numerical method, we wish to discretize both the forward and backward stochastic processes by schemes higher than order one and aim to find a flexible and efficient solution method, competitive in performance. In the following subsection we recap the Itô–Taylor expansion and discretization schemes for the FSDEs.

2.2. Itô–Taylor expansion and discretization schemes

For an extensive survey on stochastic Taylor expansions we refer to [17]. We consider the integral form of the FSDE,

$$X_t = X_{t_0} + \int_{t_0}^t \mu(X_s)ds + \int_{t_0}^t \sigma(X_s)d\omega_s, \quad t_0 \leq t \leq T. \tag{2.7}$$

Drift function $\mu(\cdot)$ and volatility function $\sigma(\cdot)$ are assumed to be sufficiently smooth. Itô’s formula for a general sufficiently smooth function $h(t, x)$ gives

$$h(t, X_t) = h(t_0, X_{t_0}) + \int_{t_0}^t L^0 h(s, X_s)ds + \int_{t_0}^t L^1 h(s, X_s)d\omega_s, \quad t_0 \leq t \leq T, \tag{2.8}$$

with diffusion operators

$$L^0 := \frac{\partial}{\partial t} + \mu \frac{\partial}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial x^2}, \quad L^1 := \sigma \frac{\partial}{\partial x}. \tag{2.9}$$

By applying Itô’s formula to the functions $\mu(X_s)$ and $\sigma(X_s)$ in (2.7) we find

$$\begin{aligned} X_t = X_{t_0} + \mu(X_{t_0})I_{(0),t_0,t} + \sigma(X_{t_0})I_{(1),t_0,t} + \int_{t_0}^t \int_{t_0}^s L^0 \mu(X_u)duds + \int_{t_0}^t \int_{t_0}^s L^1 \mu(X_u)d\omega_u ds \\ + \int_{t_0}^t \int_{t_0}^s L^0 \sigma(X_u)dud\omega_s + \int_{t_0}^t \int_{t_0}^s L^1 \sigma(X_u)d\omega_u d\omega_s, \end{aligned} \tag{2.10}$$

with Itô integrals

$$I_{(0),\rho,\tau} := \int_{\rho}^{\tau} ds = \tau - \rho, \quad I_{(1),\rho,\tau} := \int_{\rho}^{\tau} d\omega_s = \omega_{\tau} - \omega_{\rho}. \quad (2.11)$$

The general multiple Itô integral is defined recursively by (see [17], Chapter 5.2)

$$I_{\alpha}[h(\cdot, X)]_{\rho,\tau} := \begin{cases} h(\tau, X_{\tau}), & l(\alpha) = 0, \\ \int_{\rho}^{\tau} I_{\alpha-}[h(\cdot, X)]_{\rho,s} ds, & l(\alpha) \geq 1, \alpha_l = 0, \\ \int_{\rho}^{\tau} I_{\alpha-}[h(\cdot, X)]_{\rho,s} d\omega_s, & l(\alpha) \geq 1, \alpha_l = 1, \end{cases} \quad (2.12)$$

with $l(\alpha)$ the size of multi-index vector $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_{l(\alpha)})$, $\alpha^- := (\alpha_1, \alpha_2, \dots, \alpha_{l(\alpha)-1})$, and $-\alpha := (\alpha_2, \dots, \alpha_l)$. α_{\emptyset} denotes the multi-index of length zero. For notational convenience we write $I_{\alpha,\rho,\tau} = I_{\alpha}[1]_{\rho,\tau}$. The Itô coefficient functions are defined by $h_{\alpha} = L^{\alpha_1} L^{\alpha_2} \dots L^{\alpha_l} h$. Let \mathcal{M} denote the set of all multi-indices. Subset $\mathcal{A} \subset \mathcal{M}$ is called a hierarchical set if $\mathcal{A} \neq \emptyset$, $\sup_{\alpha \in \mathcal{A}} l(\alpha) < \infty$, and $-\alpha \in \mathcal{A}$ for all $\alpha \in \mathcal{A} \setminus \alpha_{\emptyset}$. For any hierarchical set \mathcal{A} the corresponding remainder set $\mathcal{B}(\mathcal{A})$ is defined by $\mathcal{B}(\mathcal{A}) = \{\alpha \in \mathcal{M} \setminus \mathcal{A} : -\alpha \in \mathcal{A}\}$. By iterating Itô's formula we obtain the Itô-Taylor expansion, as follows

$$h(t, X_t) = \sum_{\alpha \in \mathcal{A}} h_{\alpha}(t_0, X_{t_0}) I_{\alpha,t_0,t} + \sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_{\alpha}[h_{\alpha}(\cdot, X)]_{t_0,t}. \quad (2.13)$$

Notice that for the Itô-Taylor expansion of (2.7) we use $h(t, x) = x$.

We now briefly discuss some discretization schemes for FSDE (2.7). We define a time-grid $t_0, t_1, \dots, t_m, \dots, t_M = T$, with fixed timesteps $\Delta t := t_{m+1} - t_m$. For notational convenience we write $X_m = X_{t_m}$, $\omega_m = \omega_{t_m}$, $\Delta\omega_{m+1} := \omega_{m+1} - \omega_m$, and $I_{\alpha,m+1} := I_{\alpha,t_m,t_{m+1}}$. The approximated process is denoted by $X_m^{\Delta} = X_{t_m}^{\Delta}$. We start with $X_0^{\Delta} = X_0$ and one of the following forward schemes is used to determine the values X_{m+1}^{Δ} , for $m = 0, \dots, M-1$. The strong convergence rate γ_s and weak convergence rate γ_w satisfy, for sufficiently small Δt , ([17], Chapter 9.6 and 9.7)

$$\mathbb{E}_0[|X_T - X_T^{\Delta}|] \leq C(\Delta t)^{\gamma_s}, \quad |\mathbb{E}_0[P(X_T) - P(X_T^{\Delta})]| \leq C(\Delta t)^{\gamma_w}, \quad (2.14)$$

with $C > 0$ a constant, which does not depend on Δt , and $P(\cdot)$ any $2(\gamma_w + 1)$ times continuously differentiable function of polynomial growth.

Well-known schemes include the **Euler scheme** ($\gamma_s = \frac{1}{2}$, $\gamma_w = 1$), i.e.,

$$\begin{aligned} X_{m+1}^{\Delta} &= X_m^{\Delta} + \mu(X_m^{\Delta}) I_{(0),m+1} + \sigma(X_m^{\Delta}) I_{(1),m+1} \\ &= X_m^{\Delta} + \mu(X_m^{\Delta}) \Delta t + \sigma(X_m^{\Delta}) \Delta\omega_{m+1} \end{aligned} \quad (2.15)$$

and the **Milstein scheme** ($\gamma_s = 1$, $\gamma_w = 1$), i.e.,

$$\begin{aligned} X_{m+1}^{\Delta} &= X_m^{\Delta} + \mu(X_m^{\Delta}) I_{(0),m+1} + \sigma(X_m^{\Delta}) I_{(1),m+1} + L^1 \sigma(X_m^{\Delta}) I_{(1,1),m+1} \\ &= X_m^{\Delta} + \mu(X_m^{\Delta}) \Delta t + \sigma(X_m^{\Delta}) \Delta\omega_{m+1} + \sigma(X_m^{\Delta}) \sigma_x(X_m^{\Delta}) \frac{1}{2} [(\Delta\omega_{m+1})^2 - \Delta t]. \end{aligned} \quad (2.16)$$

Here we use the short-hand notation $\sigma_x = D_x \sigma$.

We also consider the **Order 2.0 weak Taylor scheme** ($\gamma_s = 1$, $\gamma_w = 2$), defined as,

$$\begin{aligned} X_{m+1}^{\Delta} &= X_m^{\Delta} + \mu(X_m^{\Delta}) I_{(0),m+1} + \sigma(X_m^{\Delta}) I_{(1),m+1} + L^1 \sigma(X_m^{\Delta}) I_{(1,1),m+1} \\ &\quad + L^1 \mu(X_m^{\Delta}) I_{(1,0),m+1} + L^0 \sigma(X_m^{\Delta}) I_{(0,1),m+1} + L^0 \mu(X_m^{\Delta}) I_{(0,0),m+1} \\ &= X_m^{\Delta} + \mu(X_m^{\Delta}) \Delta t + \sigma(X_m^{\Delta}) \Delta\omega_{m+1} + \sigma(X_m^{\Delta}) \sigma_x(X_m^{\Delta}) \frac{1}{2} [(\Delta\omega_{m+1})^2 - \Delta t] \\ &\quad + \mu_x(X_m^{\Delta}) \sigma(X_m^{\Delta}) \Delta\varpi_{m+1} + \left(\mu(X_m^{\Delta}) \sigma_x(X_m^{\Delta}) + \frac{1}{2} \sigma_{xx}(X_m^{\Delta}) \sigma^2(X_m^{\Delta}) \right) [\Delta\omega_{m+1} \Delta t - \Delta\varpi_{m+1}] \\ &\quad + \left(\mu(X_m^{\Delta}) \mu_x(X_m^{\Delta}) + \frac{1}{2} \mu_{xx}(X_m^{\Delta}) \sigma^2(X_m^{\Delta}) \right) \frac{1}{2} (\Delta t)^2, \end{aligned} \quad (2.17)$$

with

$$\Delta\varpi_{m+1} := I_{(1,0),m+1} \sim \mathcal{N}(0, \frac{1}{3}(\Delta t)^3), \quad (2.18)$$

$$\mathbb{E}[\Delta\varpi_{m+1}] = 0, \quad \text{Var}(\Delta\varpi_{m+1}) = \frac{1}{3}(\Delta t)^3, \quad \text{Cov}(\Delta\omega_{m+1}, \Delta\varpi_{m+1}) = \frac{1}{2}(\Delta t)^2. \quad (2.19)$$

If we replace the Wiener process ω_t by a trinomial tree with increments $\Delta\hat{\omega}_{m+1} \in \{-\sqrt{3}\sqrt{\Delta t}, 0, +\sqrt{3}\sqrt{\Delta t}\}$ and replace $\Delta\varpi_{m+1}$ by $\frac{1}{2}\Delta\hat{\omega}_{m+1}\Delta t$, this leads to the so-called Order 2.0 *simplified* weak Taylor scheme [17]. However, we do not replace $\Delta\omega_{m+1}$ by a discrete tree, but keep the continuous random variable. We replace $\Delta\varpi_{m+1}$ by $\Delta\tilde{\varpi}_{m+1} := \frac{1}{2}\Delta\omega_{m+1}\Delta t$. The moments do not change up to order two:

$$\mathbb{E}[\Delta \tilde{\omega}_{m+1}] = 0, \quad \text{Var}(\Delta \tilde{\omega}_{m+1}) = \frac{1}{2}(\Delta t)^3, \quad \text{Cov}(\Delta \omega_{m+1}, \Delta \tilde{\omega}_{m+1}) = \frac{1}{2}(\Delta t)^2. \quad (2.20)$$

This gives us the **Order 2.0 ‘continuous’ simplified weak Taylor scheme**

$$\begin{aligned} X_{m+1}^\Delta &= X_m^\Delta + \mu(X_m^\Delta)\Delta t + \sigma(X_m^\Delta)\Delta \omega_{m+1} \\ &\quad + \sigma(X_m^\Delta)\sigma_x(X_m^\Delta)\frac{1}{2}[(\Delta \omega_{m+1})^2 - \Delta t] \\ &\quad + \left(\mu_x(X_m^\Delta)\sigma(X_m^\Delta) + \mu(X_m^\Delta)\sigma_x(X_m^\Delta) + \frac{1}{2}\sigma_{xx}(X_m^\Delta)\sigma^2(X_m^\Delta)\right)\frac{1}{2}\Delta \omega_{m+1}\Delta t \\ &\quad + \left(\mu(X_m^\Delta)\mu_x(X_m^\Delta) + \frac{1}{2}\mu_{xx}(X_m^\Delta)\sigma^2(X_m^\Delta)\right)\frac{1}{2}(\Delta t)^2. \end{aligned} \quad (2.21)$$

We abbreviate the Order 2.0 ‘continuous’ simplified weak Taylor scheme by **2.0-weak-Taylor scheme**. With the theory in [17] it can be proved that $\gamma_w = 2$.

2.2.1. Exact simulation schemes

For, for example, the Cox–Ingersoll–Ross (CIR) process and Heston model **exact simulation schemes** exist to simulate

$$X_{m+1} = X_m + \int_{t_m}^{t_{m+1}} \mu(X_s)ds + \int_{t_m}^{t_{m+1}} \sigma(X_s)d\omega_s, \quad (2.22)$$

by sampling directly the explicit transitional density function. For the CIR process we have

$$\mu(x) = x(\bar{x} - x), \quad \sigma(x) = \eta\sqrt{x}. \quad (2.23)$$

There holds $X_t \geq 0$. If the Feller condition, $2x\bar{x} \geq \eta^2$, is satisfied, then the process is strictly positive. If the Feller condition is not satisfied then the process can become zero. We define $q_F := 2x\bar{x}/\eta^2 - 1$ and $\zeta := 2x / ((1 - e^{-x\Delta t})\eta^2)$. Then $2\zeta X_{m+1}$ is noncentral chi-squared distributed with degrees of freedom $2(q_F + 1)$ and noncentrality parameter $2\zeta X_m e^{-x\Delta t}$ (see [11,16]). The characteristic function of X_{m+1} is known, i.e.,

$$\varphi_{X_{m+1}}(u | X_m = x) = \exp\left(\frac{iux e^{-x\Delta t}}{1 - \frac{i u(1 - e^{-x\Delta t})\eta^2}{2x}}\right) \left(1 - \frac{i u(1 - e^{-x\Delta t})\eta^2}{2x}\right)^{-2x\bar{x}/\eta^2}. \quad (2.24)$$

3. BCOS method

In Section 3.1 we derive the characteristic function of the underlying discretized FSDE. In Section 3.2 we discuss the θ - and Δ -time-discretization schemes for the coupled FBSDE and in Section 3.3 we derive COS formulas to approximate the occurring expectations by using the characteristic function. Section 3.4 presents the overall BCOS algorithm.

3.1. Characteristic function discretization schemes FSDE

We can write the Euler, Milstein, and 2.0-weak-Taylor discretization schemes from the previous section in the following general form

$$X_{m+1}^\Delta = x + m(x)\Delta t + s(x)\Delta \omega_{m+1} + \kappa(x)(\Delta \omega_{m+1})^2, \quad X_m^\Delta = x. \quad (3.1)$$

For the Euler scheme we have

$$m(x) = \mu(x), \quad s(x) = \sigma(x), \quad \kappa(x) = 0, \quad (3.2)$$

for the Milstein scheme

$$m(x) = \mu(x) - \frac{1}{2}\sigma(x)\sigma_x(x), \quad s(x) = \sigma(x), \quad \kappa(x) = \frac{1}{2}\sigma(x)\sigma_x(x), \quad (3.3)$$

and for the 2.0-weak-Taylor scheme

$$m(x) = \mu(x) - \frac{1}{2}\sigma(x)\sigma_x(x) + \frac{1}{2}\left(\mu(x)\mu_x(x) + \frac{1}{2}\mu_{xx}(x)\sigma^2(x)\right)\Delta t, \quad (3.4a)$$

$$s(x) = \sigma(x) + \frac{1}{2}\left(\mu_x(x)\sigma(x) + \mu(x)\sigma_x(x) + \frac{1}{2}\sigma_{xx}(x)\sigma^2(x)\right)\Delta t, \quad (3.4b)$$

$$\kappa(x) = \frac{1}{2}\sigma(x)\sigma_x(x). \quad (3.4c)$$

Lemma 3.1. The characteristic function of X_{m+1}^Δ , given $X_m^\Delta = x$, in equation (3.1) is given by

$$\begin{aligned}\varphi_{X_{m+1}^\Delta}(u|X_m^\Delta = x) &= \mathbb{E} \left[\exp(iuX_{m+1}^\Delta) \middle| X_m^\Delta = x \right] \\ &= \exp \left(iux + ium(x)\Delta t - \frac{\frac{1}{2}u^2s^2(x)\Delta t}{1 - 2iu\kappa(x)\Delta t} \right) (1 - 2iu\kappa(x)\Delta t)^{-1/2}.\end{aligned}\quad (3.5)$$

For $\kappa(x) = 0$ it follows that¹

$$\varphi_{X_{m+1}^\Delta}(u|X_m^\Delta = x) = e^{iux + ium(x)\Delta t - \frac{1}{2}u^2s^2(x)\Delta t}.\quad (3.6)$$

Proof. With polynomial factorization we can rewrite equation (3.1) as (for $\kappa(x) \neq 0$)

$$\begin{aligned}X_{m+1}^\Delta &= x + m(x)\Delta t + \kappa(x) \left(\Delta\omega_{m+1} + \frac{1}{2} \frac{s(x)}{\kappa(x)} \right)^2 - \frac{1}{4} \frac{s^2(x)}{\kappa(x)} \\ &\stackrel{d}{=} x + m(x)\Delta t - \frac{1}{4} \frac{s^2(x)}{\kappa(x)} + \kappa(x)\Delta t \left(\varepsilon_{m+1}^\mathcal{N} + \sqrt{\lambda(x)} \right)^2,\end{aligned}\quad (3.7)$$

with $\lambda(x) := \frac{1}{4} \frac{s^2(x)}{\kappa^2(x)\Delta t}$ and $\Delta\omega_{m+1} \stackrel{d}{=} \sqrt{\Delta t} \varepsilon_{m+1}^\mathcal{N}$, $\varepsilon_{m+1}^\mathcal{N} \sim \mathcal{N}(0, 1)$. Random variable $(\varepsilon_{m+1}^\mathcal{N} + \sqrt{\lambda(x)})^2 \sim \chi_1^2(\lambda(x))$ is governed by a noncentral chi-squared distribution with degrees of freedom $\nu = 1$ and $\lambda(x)$ the noncentrality parameter. The characteristic function of a noncentral chi-squared distributed random variable reads

$$\varphi_{\chi_1^2(\lambda)}(u) = \exp \left(\frac{i\lambda u}{1 - 2iu} \right) (1 - 2iu)^{-\nu/2}.\quad (3.8)$$

The characteristic function of X_{m+1}^Δ , given $X_m^\Delta = x$, is then given by

$$\begin{aligned}\varphi_{X_{m+1}^\Delta}(u|X_m^\Delta = x) &= \exp \left(iux + ium(x)\Delta t - iu \frac{1}{4} \frac{s^2(x)}{\kappa(x)} \right) \varphi_{\chi_1^2(\lambda(x))}(u\kappa(x)\Delta t) \\ &= \exp \left(iux + ium(x)\Delta t - \frac{\frac{1}{2}u^2s^2(x)\Delta t}{1 - 2iu\kappa(x)\Delta t} \right) (1 - 2iu\kappa(x)\Delta t)^{-1/2}.\end{aligned}\quad (3.9)$$

3.1.1. Intermezzo: Bermudan put option – CEV – \mathbb{Q} -measure

To test the discretization schemes of the FSDE and the discrete characteristic function, we perform an option pricing experiment with the COS method for a Bermudan put option. We take ten early-exercise dates τ_j , $j = 1, \dots, 10$, with fixed time intervals $T/10$. The underlying asset price under the risk-neutral \mathbb{Q} -measure follows a Constant Elasticity of Variance (CEV) process,

$$dX_s = rX_s ds + \bar{\sigma} X_s^\gamma d\omega_s.\quad (3.10)$$

The option price $v(t, X_t)$ is given by the risk-neutral valuation formula

$$v(t, x) = \sup_{\tau_j, j=1, \dots, 10} e^{-r(\tau_j - t)} \mathbb{E}^\mathbb{Q} [g(X_{\tau_j}) | X_t = x],\quad (3.11)$$

with payoff function $g(x) = \max(K - x, 0)$. This problem can also be represented by a linear parabolic PDE variational inequality by means of the Feynman–Kac theorem [25]. We take the number of timesteps of the time-discretization schemes, that is M , equal to a multiple of the number of early-exercise dates. The dynamic programming principle gives

$$v(t_m, x) = \begin{cases} g(x), & \text{for } m = M, \\ e^{-r\Delta t} \mathbb{E} [v(t_{m+1}, X_{m+1}) | X_m = x], & \text{for } m < M, t_m \neq \tau_j, \\ \max \left[e^{-r\Delta t} \mathbb{E} [v(t_{m+1}, X_{m+1}) | X_m = x], g(x) \right], & \text{for } m < M, t_m = \tau_j. \end{cases}\quad (3.12)$$

¹ In Section 5.1 we discuss an example with $\kappa(0) = 0$.

Table 1
Error and CPU time for 2.0-weak-Taylor scheme ($\gamma = 0.2, M = 1000$).

N	2^4	2^5	2^6	2^7	2^8	2^9
2.0-weak-Taylor	3.2235e-02	2.5416e-03	2.9661e-04	1.1736e-04	7.9437e-05	1.6482e-05
CPU time (s)	0.0127	0.0128	0.0132	0.0143	0.0169	0.0352

The COS formula (see [9]) gives us

$$\begin{aligned}
 \mathbb{E} \left[v(t_{m+1}, X_{m+1}) \mid X_m = x \right] &\approx \mathbb{E} \left[v(t_{m+1}, X_{m+1}^\Delta) \mid X_m^\Delta = x \right] \\
 &\approx \sum_{k=0}^{N-1} \mathcal{V}_k(t_{m+1}) \mathbb{E} \left[\cos \left(k\pi \frac{X_{m+1}^\Delta - a}{b-a} \right) \mid X_m^\Delta = x \right] \\
 &= \sum_{k=0}^{N-1} \mathcal{V}_k(t_{m+1}) \Re \left\{ \mathbb{E} \left[\exp \left(ik\pi \frac{X_{m+1}^\Delta}{b-a} \right) \mid X_m^\Delta = x \right] \exp \left(ik\pi \frac{-a}{b-a} \right) \right\} \\
 &= \sum_{k=0}^{N-1} \mathcal{V}_k(t_{m+1}) \Re \left\{ \varphi_{X_{m+1}^\Delta} \left(\frac{k\pi}{b-a} \mid X_m^\Delta = x \right) \exp \left(ik\pi \frac{-a}{b-a} \right) \right\}, \tag{3.13}
 \end{aligned}$$

where $\mathcal{V}_k(t_{m+1})$ denote the Fourier cosine coefficients of $v(t_{m+1}, x)$, i.e.,

$$\mathcal{V}_k(t_{m+1}) = \frac{2}{b-a} \int_a^b v(t_{m+1}, x) \cos \left(k\pi \frac{x-a}{b-a} \right) dx. \tag{3.14}$$

\sum' in (3.13) means that the first term of the summation has half weight. The coefficients are recovered recursively backwards in time (similar as in Section 3.4). The characteristic function of the discretized FSDE, X_{m+1}^Δ , is known, whereas the characteristic function of the FSDE, X_{m+1} , is only available in closed form for $\gamma \in \{0, 0.5, 1, 1.5\}$ [19]. For the tests, we use the following parameter values

$$X_0 = 100, K = 100, r = 0.1, T = 0.1. \tag{3.15}$$

We take the elasticity of variance equal to $\gamma = 0.2$ and $\gamma = 0.8$ and choose $\bar{\sigma}$ so that $\sigma(X_0) = 25$. The exact solutions for corresponding European options [13] are given by $v(t_0, X_0) = 2.6650$ and $v(t_0, X_0) = 2.6655$, respectively. The Bermudan values are $v(t_0, X_0) = 2.7353$ and $v(t_0, X_0) = 2.7373$, where we computed our reference values by taking high values of M and N .

The results for the Euler, Milstein, and 2.0-weak-Taylor schemes are shown in Fig. 3.1 for $N = 2^9$. The 2.0-weak-Taylor scheme results in second-order convergence in M . It is interesting to see that the 2.0-weak-Taylor scheme does not only have a better convergence rate, but also the absolute value of the error is lower even for small M . Because of that, we only need a small number of timesteps to achieve a small error. For comparison, with the 2.0-weak-Taylor scheme we need only 20 timesteps to get errors smaller than 10^{-5} , whereas the Euler scheme requires approximately 900 timesteps.

Table 1 presents the errors for the 2.0-weak-Taylor scheme for different values of N , with $M = 1000$ timesteps. The convergence in N , the number of Fourier coefficients, is of second order, due to the use of discrete Fourier cosine transforms (see Section 3.4).

The CPU time is shown in the last row. The computation time for the Euler scheme is only slightly shorter. The computation of the expected values, with the COS formula (equation (3.13)), on an x -grid with length N is $\mathcal{O}(N^2)$, due to matrix-vector multiplications. The usage of discrete Fourier cosine transforms (DCTs) for the recovery of the coefficients is of order $\mathcal{O}(N \log N)$, see equation (3.37) in Section 3.4. As the use of the DCTs is the most time-consuming part we do not observe quadratic complexity for these values of N .

We would like to mention that this method can also be applied to time-dependent drift and diffusion terms and to other local volatility models.

Remark 3.1. We can write equation (3.13) as

$$\mathbb{E} \left[v(t_{m+1}, X_{m+1}^\Delta) \mid X_m^\Delta = x \right] = \int_{\mathbb{R}} v(t_{m+1}, y) p(y \mid X_m^\Delta = x) dy, \tag{3.16}$$

with $p(\cdot \mid X_m^\Delta = x)$ the conditional density function of X_{m+1}^Δ . The density function corresponding to the 2.0-weak-Taylor scheme is known, as it involves a noncentral chi-squared distributed random variable. However, the density function of a noncentral chi-squared distributed random variable involves some mathematical special functions. This makes it impractical and time-consuming to calculate. Besides, numerical integration algorithms have in general algebraic convergence. The

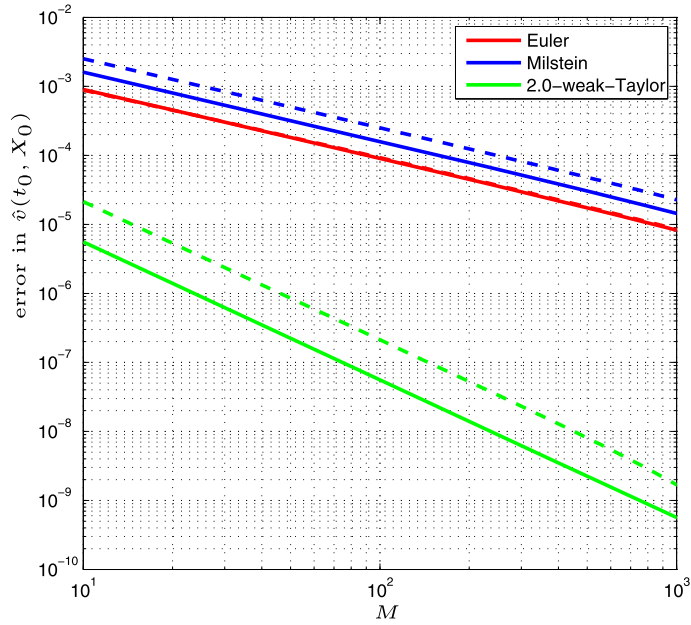


Fig. 3.1. Results Bermudan put, $\gamma = 0.2$ (lines) and $\gamma = 0.8$ (dashes) ($N = 2^9$).

strength of the COS formula lies in the availability of the corresponding characteristic function, which can be derived for models from the class of regular affine processes and for exponentially Lévy models. This results in exponential convergence in the number of Fourier cosine terms for smooth densities (in absence of the use of discrete Fourier cosine transforms).

3.2. Δ -time-discretization scheme

In this section, we focus on the discretization scheme of the decoupled FBSDE system. First, we consider the integral form of the BSDE,

$$Y_t = g(X_T) + \int_t^T f(s, \mathbf{\Lambda}_s) ds - \int_t^T Z_s d\omega_s, \quad (3.17)$$

with $\mathbf{\Lambda}_s := (X_s, Y_s, Z_s)$. For the discretization of the BSDE at time-point t_m , we start with

$$Y_m = Y_{m+1} + \int_{t_m}^{t_{m+1}} f(s, \mathbf{\Lambda}_s) ds - \int_{t_m}^{t_{m+1}} Z_s d\omega_s. \quad (3.18)$$

Taking conditional expectations at both sides of equation (3.18) and applying the θ -method [15] results in

$$\begin{aligned} Y_m &= \mathbb{E}_m[Y_{m+1}] + \int_{t_m}^{t_{m+1}} \mathbb{E}_m[f(s, \mathbf{\Lambda}_s)] ds \\ &\approx \mathbb{E}_m[Y_{m+1}] + \Delta t \theta f(t_m, \mathbf{\Lambda}_m) + \Delta t(1 - \theta) \mathbb{E}_m[f(t_{m+1}, \mathbf{\Lambda}_{m+1})], \end{aligned} \quad (3.19)$$

where $\mathbb{E}_m[\cdot]$ represents the conditional expectation $\mathbb{E}[\cdot | \mathcal{F}_{t_m}]$. Multiplying both sides of equation (3.18) by $\Delta\omega_{m+1}$, taking the conditional expectations, and applying the θ -method gives us

$$\begin{aligned} 0 &= \mathbb{E}_m[Y_{m+1} \Delta\omega_{m+1}] + \int_{t_m}^{t_{m+1}} \mathbb{E}_m[f(s, \mathbf{\Lambda}_s)(\omega_s - \omega_{t_m})] ds - \int_{t_m}^{t_{m+1}} \mathbb{E}_m[Z_s] ds \\ &\approx \mathbb{E}_m[Y_{m+1} \Delta\omega_{m+1}] + \Delta t(1 - \theta) \mathbb{E}_m[f(t_{m+1}, \mathbf{\Lambda}_{m+1}) \Delta\omega_{m+1}] \\ &\quad - \Delta t \theta Z_m - \Delta t(1 - \theta) \mathbb{E}_m[Z_{m+1}]. \end{aligned} \quad (3.20)$$

For the approximation of the FSDE we will use the Euler, Milstein, or 2.0-weak-Taylor schemes, as described in Section 2.2. Then the following Δ -time-discretization scheme algorithm is used to approximate the BSDE

$$Y_M^\Delta = g(X_M^\Delta), \quad Z_M^\Delta = \sigma(X_M^\Delta) D_x g(X_M^\Delta), \quad (3.21a)$$

for $m = M - 1, \dots, 0$:

$$Z_m^\Delta = \frac{1}{\Delta t} \theta^{-1} \mathbb{E}_m [Y_{m+1}^\Delta \Delta \omega_{m+1}] + \theta^{-1} (1 - \theta) \mathbb{E}_m [f(t_{m+1}, \Lambda_{m+1}^\Delta) \Delta \omega_{m+1}] - \theta^{-1} (1 - \theta) \mathbb{E}_m [Z_{m+1}^\Delta], \quad (3.21b)$$

$$Y_m^\Delta = \mathbb{E}_m [Y_{m+1}^\Delta] + \Delta t \theta f(t_m, \Lambda_m^\Delta) + \Delta t (1 - \theta) \mathbb{E}_m [f(t_{m+1}, \Lambda_{m+1}^\Delta)]. \quad (3.21c)$$

The values Y_m^Δ and Z_m^Δ depend on the value of the forward process. Then it is easily seen, using an induction argument, that deterministic functions $y_m^\Delta(x)$ and $z_m^\Delta(x)$ exists, so that

$$Y_m^\Delta = y_m^\Delta(X_m^\Delta), \quad Z_m^\Delta = z_m^\Delta(X_m^\Delta). \quad (3.22)$$

X_m denotes the exact solution of the FSDE (2.4) at time t_m and X_m^Δ is its discrete approximation. Let $X_k^{m,x}$ denote the value of X_k , given $X_m = x$. Similarly, $X_k^{\Delta,m,x}$ denotes the value of X_k^Δ , given $X_m^\Delta = x$. We use the following notation

- $(y_m(x), z_m(x))$ is the exact solution of BSDE (3.17) at time t_m , given $X_m = x$.
- $(y_m^\theta(x), z_m^\theta(x))$ is the discrete approximation of the BSDE with the θ -time-discretization scheme, given exact solution $X_m = x$, and, for $m = M - 1, \dots, 0$,

$$z_m^\theta(x) = \frac{1}{\Delta t} \theta^{-1} \mathbb{E} [y_{m+1}^\theta(X_{m+1}^{m,x}) \Delta \omega_{m+1}] + \theta^{-1} (1 - \theta) \mathbb{E} [f(t_{m+1}, \Lambda_{m+1}^\theta(X_{m+1}^{m,x})) \Delta \omega_{m+1}] - \theta^{-1} (1 - \theta) \mathbb{E} [z_{m+1}^\theta(X_{m+1}^{m,x})], \quad (3.23a)$$

$$y_m^\theta(x) = \mathbb{E} [y_{m+1}^\theta(X_{m+1}^{m,x})] + \Delta t \theta f(t_m, \Lambda_m^\theta(x)) + \Delta t (1 - \theta) \mathbb{E} [f(t_{m+1}, \Lambda_{m+1}^\theta(X_{m+1}^{m,x}))]. \quad (3.23b)$$

- $(y_m^\Delta(x), z_m^\Delta(x))$ is the discrete approximation with the Δ -time-discretization scheme, given discrete approximation $X_m^\Delta = x$, and, for $m = M - 1, \dots, 0$,

$$z_m^\Delta(x) = \frac{1}{\Delta t} \theta^{-1} \mathbb{E} [y_{m+1}^\Delta(X_{m+1}^{\Delta,m,x}) \Delta \omega_{m+1}] + \theta^{-1} (1 - \theta) \mathbb{E} [f(t_{m+1}, \Lambda_{m+1}^\Delta(X_{m+1}^{\Delta,m,x})) \Delta \omega_{m+1}] - \theta^{-1} (1 - \theta) \mathbb{E} [z_{m+1}^\Delta(X_{m+1}^{\Delta,m,x})], \quad (3.24a)$$

$$y_m^\Delta(x) = \mathbb{E} [y_{m+1}^\Delta(X_{m+1}^{\Delta,m,x})] + \Delta t \theta f(t_m, \Lambda_m^\Delta(x)) + \Delta t (1 - \theta) \mathbb{E} [f(t_{m+1}, \Lambda_{m+1}^\Delta(X_{m+1}^{\Delta,m,x}))]. \quad (3.24b)$$

The values $y_m^\theta(x)$ and $y_m^\Delta(x)$ are implicit for $\theta > 0$ and are determined by performing *Picard iterations*, starting with initial guesses $\mathbb{E} [y_{m+1}^\theta(X_{m+1}^{m,x})]$ and $\mathbb{E} [y_{m+1}^\Delta(X_{m+1}^{\Delta,m,x})]$, respectively. Picard iterations are terminated when the difference between subsequent estimates is less than error tolerance 10^{-12} . For the numerical experiments in Section 5, mostly 5 iterations suffice. The driver function f is assumed to be Lipschitz in y and z , with Lipschitz constant L_f . For Δt small enough, i.e., $\Delta t \theta L_f < 1$, a unique fixed-point exists, and the Picard iterations converge towards that point for any initial guess. The fixed-point technique converges to the true solution at the geometric rate $\Delta t \theta L_f$, which depends on the Lipschitz condition of the driver function.

In the BCOS method we here use the characteristic function of the underlying discretized FSDE, as discussed in Section 3.1, to approximate the appearing conditional expectations $\mathbb{E} [y_{m+1}^\Delta(X_{m+1}^{\Delta,m,x})]$, $\mathbb{E} [z_{m+1}^\Delta(X_{m+1}^{\Delta,m,x})]$, and $\mathbb{E} [f(t_{m+1}, \Lambda_{m+1}^\Delta(X_{m+1}^{\Delta,m,x}))]$. Besides, we need to approximate the expected values $\mathbb{E} [y_{m+1}^\Delta(X_{m+1}^{\Delta,m,x}) \Delta \omega_{m+1}]$ and $\mathbb{E} [f(t_{m+1}, \Lambda_{m+1}^\Delta(X_{m+1}^{\Delta,m,x})) \Delta \omega_{m+1}]$. Details are discussed in the following subsection.

3.3. Expected values Δ -time-discretization scheme FBSDE

In this section we derive an equation for the conditional expectations $\mathbb{E} [h(t_{m+1}, X_{m+1}^{\Delta,m,x})]$ and $\mathbb{E} [h(t_{m+1}, X_{m+1}^{\Delta,m,x}) \Delta \omega_{m+1}]$ under the general discrete dynamics (3.1). Here $h(t, x)$ is a general function.

Let $\mathcal{H}_k(t_{m+1})$ denote the Fourier cosine coefficients of $h(t_{m+1}, x)$, i.e.,

$$\mathcal{H}_k(t_{m+1}) = \frac{2}{b-a} \int_a^b h(t_{m+1}, x) \cos\left(k\pi \frac{x-a}{b-a}\right) dx, \quad (3.25)$$

With the COS formula we get

$$\begin{aligned} \mathbb{E} \left[h(t_{m+1}, X_{m+1}^{\Delta, m, x}) \right] &\approx \sum_{k=0}^{N-1} \mathcal{H}_k(t_{m+1}) \mathbb{E} \left[\cos \left(k\pi \frac{X_{m+1}^{\Delta, m, x} - a}{b - a} \right) \right] \\ &= \sum_{k=0}^{N-1} \mathcal{H}_k(t_{m+1}) \Re \left\{ \varphi_{X_{m+1}^{\Delta}} \left(\frac{k\pi}{b - a} \middle| X_m^{\Delta} = x \right) \exp \left(ik\pi \frac{-a}{b - a} \right) \right\} \end{aligned} \tag{3.26}$$

and Fourier cosine series give

$$\begin{aligned} \mathbb{E} \left[h(t_{m+1}, X_{m+1}^{\Delta, m, x}) \Delta\omega_{m+1} \right] &\approx \sum_{k=0}^{N-1} \mathcal{H}_k(t_{m+1}) \mathbb{E} \left[\cos \left(k\pi \frac{X_{m+1}^{\Delta, m, x} - a}{b - a} \right) \Delta\omega_{m+1} \right] \\ &= \sum_{k=0}^{N-1} \mathcal{H}_k(t_{m+1}) \Re \left\{ \mathbb{E} \left[\exp \left(ik\pi \frac{X_{m+1}^{\Delta, m, x}}{b - a} \right) \Delta\omega_{m+1} \right] \exp \left(ik\pi \frac{-a}{b - a} \right) \right\}. \end{aligned} \tag{3.27}$$

Integration by parts gives

$$\begin{aligned} &\mathbb{E} \left[\exp(iu X_{m+1}^{\Delta, m, x}) \Delta\omega_{m+1} \right] \\ &= \mathbb{E} \left[\exp \left(iux + ium(x)\Delta t + ius(x)\Delta\omega_{m+1} + iu\kappa(x)(\Delta\omega_{m+1})^2 \right) \Delta\omega_{m+1} \right] \\ &= \frac{1}{\sqrt{2\pi}\sqrt{\Delta t}} \int_{\mathbb{R}} \exp \left(iux + ium(x)\Delta t + ius(x)\zeta + iu\kappa(x)\zeta^2 \right) \zeta e^{-\frac{1}{2} \left(\frac{\zeta}{\sqrt{\Delta t}} \right)^2} d\zeta \\ &= \frac{1}{\sqrt{2\pi}\sqrt{\Delta t}} \int_{\mathbb{R}} \{s(x) + 2\kappa(x)\zeta\} \Delta t D_x \exp \left(iux + ium(x)\Delta t + ius(x)\zeta + iu\kappa(x)\zeta^2 \right) e^{-\frac{1}{2} \left(\frac{\zeta}{\sqrt{\Delta t}} \right)^2} d\zeta \\ &= s(x)\Delta t \mathbb{E} \left[D_x \exp(iu X_{m+1}^{\Delta, m, x}) \right] + 2\kappa(x)\Delta t \mathbb{E} \left[D_x \exp(iu X_{m+1}^{\Delta, m, x}) \Delta\omega_{m+1} \right]. \end{aligned} \tag{3.28}$$

Using the same procedure for the last term in (3.28) and iterating recursively gives us

$$\begin{aligned} \mathbb{E} \left[\exp(iu X_{m+1}^{\Delta, m, x}) \Delta\omega_{m+1} \right] &= s(x)\Delta t \mathbb{E} \left[D_x \exp(iu X_{m+1}^{\Delta, m, x}) \right] \\ &\quad + s(x)\Delta t (2\kappa(x)\Delta t) \mathbb{E} \left[D_x^2 \exp(iu X_{m+1}^{\Delta, m, x}) \right] \\ &\quad + s(x)\Delta t (2\kappa(x)\Delta t)^2 \mathbb{E} \left[D_x^3 \exp(iu X_{m+1}^{\Delta, m, x}) \right] \\ &\quad + s(x)\Delta t (2\kappa(x)\Delta t)^3 \mathbb{E} \left[D_x^4 \exp(iu X_{m+1}^{\Delta, m, x}) \right] + \dots \end{aligned} \tag{3.29}$$

It holds that

$$\begin{aligned} \mathbb{E} \left[\frac{\partial^{(\ell)}}{\partial (X_{m+1}^{\Delta, m, x})^{(\ell)}} \exp \left(iu X_{m+1}^{\Delta, m, x} \right) \right] &= (iu)^\ell \mathbb{E} \left[\exp \left(iu X_{m+1}^{\Delta, m, x} \right) \right] \\ &= (iu)^\ell \varphi_{X_{m+1}^{\Delta}} (u | X_m^{\Delta} = x). \end{aligned} \tag{3.30}$$

For the numerical experiments, it appears sufficient to take only the first two terms in (3.29), as the other terms are at least of order $\mathcal{O}((\Delta t)^3)$, and we end up with

$$\begin{aligned} \mathbb{E} \left[h(t_{m+1}, X_{m+1}^{\Delta, m, x}) \Delta\omega_{m+1} \right] &\approx \sum_{k=0}^{N-1} \mathcal{H}_k(t_{m+1}) \Re \left\{ \left(s(x)\Delta t \left(i \frac{k\pi}{b - a} \right) \varphi_{X_{m+1}^{\Delta}} \left(\frac{k\pi}{b - a} \middle| X_m^{\Delta} = x \right) \right. \right. \\ &\quad \left. \left. + s(x)\Delta t 2\kappa(x)\Delta t \left(i \frac{k\pi}{b - a} \right)^2 \varphi_{X_{m+1}^{\Delta}} \left(\frac{k\pi}{b - a} \middle| X_m^{\Delta} = x \right) \right) \exp \left(ik\pi \frac{-a}{b - a} \right) \right\}, \end{aligned} \tag{3.31}$$

which enables us to approximate the conditional expectations in (3.21c).

3.3.1. Exact simulation scheme

For, for example, the CIR process and Heston stochastic volatility models an **exact simulation scheme** exists based on directly sampling from the available transitional density function. In this section, we explain how to use exact simulation schemes to solve the discrete problem (3.23b). The characteristic function of $X_{m+1}^{m,x}$ can be found as the Fourier transform of the density function and is denoted by $\varphi_{X_{m+1}}(u|X_m = x)$. With the COS formula we get

$$\mathbb{E}[h(t_{m+1}, X_{m+1}^{m,x})] \approx \sum_{k=0}^{N-1} \mathcal{H}_k(t_{m+1}) \Re \left\{ \varphi_{X_{m+1}} \left(\frac{k\pi}{b-a} \middle| X_m = x \right) \exp \left(ik\pi \frac{-a}{b-a} \right) \right\}. \tag{3.32}$$

The question is how to approximate the expected value $\mathbb{E}[h(t_{m+1}, X_{m+1}^{m,x})\Delta\omega_{m+1}]$. Notice that $X_{m+1}^{m,x}$ and $\Delta\omega_{m+1}$ are correlated. We first use Fourier cosine series to get

$$\mathbb{E}[h(t_{m+1}, X_{m+1}^{m,x})\Delta\omega_{m+1}] \approx \sum_{k=0}^{N-1} \mathcal{H}_k(t_{m+1}) \Re \left\{ \mathbb{E} \left[\exp \left(ik\pi \frac{X_{m+1}^{m,x}}{b-a} \right) \Delta\omega_{m+1} \right] \exp \left(ik\pi \frac{-a}{b-a} \right) \right\}, \tag{3.33}$$

and then with the theory in Section 2.2 and [17] we find

$$\begin{aligned} & \mathbb{E}[\exp(iuX_{m+1}^{m,x})\Delta\omega_{m+1}] \\ &= \mathbb{E} \left[\sum_{\alpha \in \mathcal{A}} L^\alpha \exp(iux) I_{\alpha, m+1} \Delta\omega_{m+1} \right] + \mathbb{E} \left[\sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_\alpha [L^\alpha \exp(iuX_{m+1}^{m,x})]_{m+1} \Delta\omega_{m+1} \right] \\ &= L^1 \exp(iux) \Delta t + L^1 L^0 \exp(iux) \frac{1}{2} (\Delta t)^2 + L^0 L^1 \exp(iux) \frac{1}{2} (\Delta t)^2 + \mathcal{O}((\Delta t)^3). \end{aligned} \tag{3.34}$$

Notice that $\exp(iux)$ does not depend on time and

$$L^1 \exp(iux) = \sigma(x) iu \exp(iux), \tag{3.35a}$$

$$L^1 L^0 \exp(iux) = \left[\sigma(x) \mu_x(x) iu + (\sigma(x) \mu(x) + \sigma^2(x) \sigma_x(x)) (iu)^2 + \frac{1}{2} \sigma(x)^3 (iu)^3 \right] \exp(iux), \tag{3.35b}$$

$$L^0 L^1 \exp(iux) = \left[(\mu(x) \sigma_x(x) + \frac{1}{2} \sigma^2(x) \sigma_{xx}(x)) iu + (\mu(x) \sigma(x) + \sigma^2(x) \sigma_x(x)) (iu)^2 + \frac{1}{2} \sigma(x)^3 (iu)^3 \right] \exp(iux). \tag{3.35c}$$

This enables us to approximate the conditional expectations in (3.23b).

3.4. Recovery of coefficients and algorithm

Suppose we would like to approximate the Fourier cosine coefficients $\mathcal{H}_k(t_{m+1})$ of function $h(t_{m+1}, x)$, see equation (3.25). For this, we take N grid-points and define an equidistant x -grid

$$x_n := a + (n + \frac{1}{2}) \frac{b-a}{N} \quad \text{and} \quad \Delta x := \frac{b-a}{N}. \tag{3.36}$$

We determine the value of function $h(t_{m+1}, x)$ on the N grid-points. The midpoint-rule integration gives us

$$\begin{aligned} \mathcal{H}_k(t_{m+1}) &= \frac{2}{b-a} \int_a^b h(t_{m+1}, x) \cos \left(k\pi \frac{x-a}{b-a} \right) dx \\ &\approx \sum_{n=0}^{N-1} \frac{2}{b-a} h(t_{m+1}, x_n) \cos \left(k\pi \frac{x_n-a}{b-a} \right) \Delta x = \sum_{n=0}^{N-1} h(t_{m+1}, x_n) \cos \left(k\pi \frac{2n+1}{2N} \right) \frac{2}{N}. \end{aligned} \tag{3.37}$$

The appearing discrete Fourier cosine transform (DCT) (Type II) can be calculated efficiently by, for example, MATLAB's function `dct`. The error of the numerical integration is second order in N .

Now we return to FBSDE problem (3.24b), where we obtained deterministic functions $y_m^\Delta(x)$ and $z_m^\Delta(x)$. Let $\mathcal{Y}_k^\Delta(t_m)$ be the Fourier cosine coefficients of $y_m^\Delta(x)$ in (3.24b), i.e.,

$$\mathcal{Y}_k^\Delta(t_m) = \frac{2}{b-a} \int_a^b y_m^\Delta(x) \cos \left(k\pi \frac{x-a}{b-a} \right) dx, \tag{3.38}$$

$\mathcal{Z}_k^\Delta(t_m)$ the Fourier cosine coefficients of function $z_m^\Delta(x)$ in (3.24a), i.e.,

$$\mathcal{Z}_k^\Delta(t_m) = \frac{2}{b-a} \int_a^b z_m^\Delta(x) \cos\left(k\pi \frac{x-a}{b-a}\right) dx, \quad (3.39)$$

and $\mathcal{F}_k^\Delta(t_m)$ the Fourier cosine coefficients of driver function $f(t_m, \Lambda_m^\Delta(x))$, i.e.,

$$\mathcal{F}_k^\Delta(t_m) = \frac{2}{b-a} \int_a^b f(t_m, \Lambda_m^\Delta(x)) \cos\left(k\pi \frac{x-a}{b-a}\right) dx. \quad (3.40)$$

The computation of functions $z_m^\Delta(x)$ and $y_m^\Delta(x)$ at time-point t_m requires the Fourier cosine coefficients $\mathcal{Z}_k^\Delta(t_{m+1})$, $\mathcal{Y}_k^\Delta(t_{m+1})$, and $\mathcal{F}_k^\Delta(t_{m+1})$ at time-point t_{m+1} , with COS formulas (3.26) and (3.31). The coefficients can be computed recursively backwards in time, as we explain in this section.

We start with the coefficients at the terminal time

$$\mathcal{Y}_k^\Delta(t_M) = \frac{2}{b-a} \int_a^b g(x) \cos\left(k\pi \frac{x-a}{b-a}\right) dx, \quad (3.41a)$$

$$\mathcal{Z}_k^\Delta(t_M) = \frac{2}{b-a} \int_a^b \sigma(x) D_x g(x) \cos\left(k\pi \frac{x-a}{b-a}\right) dx, \quad (3.41b)$$

$$\mathcal{F}_k^\Delta(t_M) = \frac{2}{b-a} \int_a^b f(t_M, x, g(x), \sigma(x) D_x g(x)) \cos\left(k\pi \frac{x-a}{b-a}\right) dx. \quad (3.41c)$$

For some problems the above integrals can be computed analytically. Otherwise we may approximate them, for example by computing the function on the x -grid and using the discrete Fourier cosine transform or another numerical integration method.

We then compute functions $z_{M-1}^\Delta(x)$, $f(t_{M-1}, \Lambda_{M-1}^\Delta(x))$, and $y_{M-1}^\Delta(x)$, see equations (3.24b), on the equidistant x -grid with N grid points. For this we use the Fourier cosine coefficients at time t_M and COS formulas (3.26) and (3.31). Subsequently the Fourier cosine coefficients $\mathcal{Z}_k^\Delta(t_{M-1})$, $\mathcal{F}_k^\Delta(t_{M-1})$, and $\mathcal{Y}_k^\Delta(t_{M-1})$ are approximated by a discrete Fourier cosine transform.

We repeat this procedure for all times t_m . So, we estimate the Fourier cosine coefficients $\mathcal{Z}_k^\Delta(t_m)$, $\mathcal{F}_k^\Delta(t_m)$, and $\mathcal{Y}_k^\Delta(t_m)$ by using the Fourier cosine coefficients at time t_{m+1} and the COS formulas. The approximation of the Fourier cosine coefficients introduces an additional error, which has been discussed in detail in [10]. The final approximations of the functions $y_m^\Delta(x)$ and $z_m^\Delta(x)$ by the BCOS method are denoted by $\hat{y}_m^\Delta(x)$ and $\hat{z}_m^\Delta(x)$, respectively. The overall algorithm to solve the FBSDE backwards in time by Δ -time-discretization scheme (3.21c) can be summarized as:

BCOS method:

Initial step: Compute, or approximate, the terminal coefficients $\mathcal{Y}_k^\Delta(t_M)$, $\mathcal{F}_k^\Delta(t_M)$, and $\mathcal{Z}_k^\Delta(t_M)$, equations (3.41a), (3.41b), and (3.41c).

Loop: For $m = M - 1$ to $m = 1$: Compute the functions $\hat{z}_m^\Delta(x)$, $f(t_m, \hat{\Lambda}_m^\Delta(x))$, and $\hat{y}_m^\Delta(x)$, see equations (3.24b), with COS formulas (3.26) and (3.31). Determine the corresponding Fourier cosine coefficients $\mathcal{Z}_k^\Delta(t_m)$, $\mathcal{F}_k^\Delta(t_m)$, and $\mathcal{Y}_k^\Delta(t_m)$ by using a DCT (3.37).

Terminal step: Compute $\hat{z}_0^\Delta(X_0)$ and $\hat{y}_0^\Delta(X_0)$.

The algorithm for the exact simulation scheme, with final values $\hat{z}_0^\theta(X_0)$ and $\hat{y}_0^\theta(X_0)$, is similar, except that we use COS formulas (3.32) and (3.33)–(3.34).

Remark 3.2. For constant drift μ and volatility σ terms we can apply the efficient Fast Fourier transform (FFT) algorithm to recover the Fourier cosine coefficients, as explained in [26].

4. Error analysis

In this section we perform an error analysis of the discretization with timesteps Δt for the 2.0-weak-Taylor scheme and parameter $\theta = 1/2$. The convergence in N , the number of Fourier coefficients, is of second order, due to the use of discrete Fourier cosine transforms. The error of the Fourier cosine formulas has been analyzed in [9,10,26] and we refer to these

articles for more details. We start with Itô–Taylor expansions and related expected values in Section 4.1. In Section 4.2 we look at the local errors related to the discretization scheme for the FSDE. In Section 4.3 we discuss the local errors of the θ -time-discretization. A global error result is presented in Section 4.4.

The authors of [29] apply also an Order 2.0 weak Taylor scheme and a slightly different θ -scheme and they obtain second-order convergence in their numerical experiments, for which they use a Gauss–Hermite quadrature rule to approximate the conditional expectation.

4.1. Itô–Taylor expansion

The Itô–Taylor expansion of a general sufficiently smooth function $h(t, x)$ reads [17]

$$h(t_{m+1}, X_{m+1}^{m,x}) = \sum_{\alpha \in \mathcal{A}} h_{\alpha}(t_m, x) I_{\alpha, m+1} + \sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_{\alpha} [h_{\alpha}(\cdot, X_{\cdot}^{m,x})]_{m+1}, \tag{4.1}$$

$$\begin{aligned} h(t_{m+2}, X_{m+2}^{m,x}) &= \sum_{\beta \in \mathcal{A}} \sum_{\alpha \in \mathcal{A}} h_{\alpha\beta}(t_m, x) I_{\alpha, m+1} I_{\beta, m+2} + \sum_{\beta \in \mathcal{A}} \sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_{\alpha} [h_{\alpha\beta}(X_{\cdot}^{m,x})]_{m+1} I_{\beta, m+2} \\ &+ \sum_{\beta \in \mathcal{B}(\mathcal{A})} I_{\beta} [h_{\beta}(\cdot, X_{\cdot}^{m,x})]_{m+2}, \end{aligned} \tag{4.2}$$

with \mathcal{A} a hierarchical set.

Lemma 4.1. *For a sufficiently smooth function $h(t, x)$ we have the following conditional expectations of the Itô–Taylor expansion*

$$\mathbb{E} [h(t_{m+1}, X_{m+1}^{m,x})] = h(t_m, x) + h_{(0)}(t_m, x) \Delta t + h_{(0,0)}(t_m, x) \frac{1}{2} (\Delta t)^2 + \mathcal{O}((\Delta t)^3), \tag{4.3a}$$

$$\mathbb{E} [h(t_{m+1}, X_{m+1}^{m,x}) \Delta \omega_{m+1}] = h_{(1)}(t_m, x) \Delta t + [h_{(1,0)}(t_m, x) + h_{(0,1)}(t_m, x)] \frac{1}{2} (\Delta t)^2 + \mathcal{O}((\Delta t)^3), \tag{4.3b}$$

$$\mathbb{E} [h(t_{m+2}, X_{m+2}^{m,x})] = h(t_m, x) + h_{(0)}(t_m, x) 2\Delta t + h_{(0,0)}(t_m, x) 2(\Delta t)^2 + \mathcal{O}((\Delta t)^3), \tag{4.3c}$$

$$\mathbb{E} [h(t_{m+2}, X_{m+2}^{m,x}) \Delta \omega_{m+1}] = h_{(1)}(t_m, x) \Delta t + h_{(0,1)}(t_m, x) \frac{1}{2} (\Delta t)^2 + h_{(1,0)}(t_m, x) \frac{3}{2} (\Delta t)^2 + \mathcal{O}((\Delta t)^3), \tag{4.3d}$$

$$\mathbb{E} [h(t_{m+2}, X_{m+2}^{m,x}) \Delta \omega_{m+2}] = h_{(1)}(t_m, x) \Delta t + h_{(0,1)}(t_m, x) \frac{3}{2} (\Delta t)^2 + h_{(1,0)}(t_m, x) \frac{1}{2} (\Delta t)^2 + \mathcal{O}((\Delta t)^3), \tag{4.3e}$$

$$\frac{2}{\Delta t} \mathbb{E} [h(t_{m+2}, X_{m+2}^{m,x}) (\Delta \omega_{m+1} - \Delta \omega_{m+2})] = \frac{2}{\Delta t} [h_{(1,0)}(t_m, x) - h_{(0,1)}(t_m, x)] (\Delta t)^2 + \mathcal{O}((\Delta t)^2). \tag{4.3f}$$

This lemma can be proved with the help of [17], Chapter 5.7. The definitions of the operators L^0 and L^1 , equation (2.9), give

$$h_{(1,0)}(t_m, x) - h_{(0,1)}(t_m, x) = D_x h(t_m, x) \left[\sigma(x) \mu_x(x) - \mu(x) \sigma_x(x) - \frac{1}{2} \sigma^2(x) \sigma_{xx}(x) \right]. \tag{4.4}$$

4.2. Local error Δ -time-discretization FSDE

The weak convergence rate of the 2.0-weak-Taylor scheme is denoted by $\gamma_w^{2T} = 2$.

Lemma 4.2. *For a sufficiently smooth function $h(t, x)$ we have the following local weak errors of the 2.0-weak-Taylor scheme*

$$\mathbb{E} [h(t_{m+1}, X_{m+1}^{m,x}) - h(t_{m+1}, X_{m+1}^{\Delta, m,x})] = \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}), \tag{4.5a}$$

$$\mathbb{E} \left[\left(h(t_{m+1}, X_{m+1}^{m,x}) - h(t_{m+1}, X_{m+1}^{\Delta, m,x}) \right) \Delta \omega_{m+1} \right] = \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}), \tag{4.5b}$$

$$\mathbb{E} [h(t_{m+2}, X_{m+2}^{m,x}) - h(t_{m+2}, X_{m+2}^{\Delta, m,x})] = \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}), \tag{4.5c}$$

$$\mathbb{E} \left[\left(h(t_{m+2}, X_{m+2}^{m,x}) - h(t_{m+2}, X_{m+2}^{\Delta, m,x}) \right) \Delta \omega_{m+1} \right] = \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}), \tag{4.5d}$$

$$\mathbb{E} \left[\left(h(t_{m+2}, X_{m+2}^{m,x}) - h(t_{m+2}, X_{m+2}^{\Delta, m,x}) \right) \Delta \omega_{m+2} \right] = \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}), \tag{4.5e}$$

$$\frac{2}{\Delta t} \mathbb{E} \left[\left(h(t_{m+2}, X_{m+2}^{m,x}) - h(t_{m+2}, X_{m+2}^{\Delta, m,x}) \right) (\Delta \omega_{m+1} - \Delta \omega_{m+2}) \right] = \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}). \tag{4.5f}$$

For a proof we refer to Appendix A.

4.3. Local error θ -time-discretization scheme

The equation for $y_m(x)$ is given by (see (3.19))

$$y_m(x) = \mathbb{E} [y_{m+1}(X_{m+1}^{m,x})] + \Delta t \frac{1}{2} f(t_m, \mathbf{\Lambda}_m(x)) + \Delta t \frac{1}{2} \mathbb{E} [f(t_{m+1}, \mathbf{\Lambda}_{m+1}(X_{m+1}^{m,x}))] + \mathcal{R}_m^y(x), \tag{4.6}$$

and the equation for $z_m(x)$ reads (see (3.20))

$$z_m(x) = \frac{2}{\Delta t} \mathbb{E} [y_{m+1}(X_{m+1}^{m,x}) \Delta \omega_{m+1}] + \mathbb{E} [f(t_{m+1}, \mathbf{\Lambda}_{m+1}(X_{m+1}^{m,x})) \Delta \omega_{m+1}] - \mathbb{E} [z_{m+1}(X_{m+1}^{m,x})] + \frac{2}{\Delta t} \mathcal{R}_m^z(x), \tag{4.7}$$

with θ -discretization errors

$$\begin{aligned} \mathcal{R}_m^y(x) &= \int_{t_m}^{t_{m+1}} \mathbb{E} [f(s, \mathbf{\Lambda}_s(X_s^{m,x}))] ds - \left\{ \Delta t \frac{1}{2} f(t_m, \mathbf{\Lambda}_m(x)) + \Delta t \frac{1}{2} \mathbb{E} [f(t_{m+1}, \mathbf{\Lambda}_{m+1}(X_{m+1}^{m,x}))] \right\}, \tag{4.8} \\ \mathcal{R}_m^z(x) &= \int_{t_m}^{t_{m+1}} \mathbb{E} [f(s, \mathbf{\Lambda}_s(X_s^{m,x})) (\omega_s - \omega_{t_m})] ds - \left\{ \Delta t \frac{1}{2} \mathbb{E} [f(t_{m+1}, \mathbf{\Lambda}_{m+1}(X_{m+1}^{m,x})) \Delta \omega_{m+1}] \right\} \\ &\quad - \int_{t_m}^{t_{m+1}} \mathbb{E} [z_s(X_s^{m,x})] ds + \left\{ \Delta t \frac{1}{2} z_m(x) + \Delta t \frac{1}{2} \mathbb{E} [z_{m+1}(X_{m+1}^{m,x})] \right\}. \tag{4.9} \end{aligned}$$

Here x denotes the exact solution of the FSDE (2.4) at time t_m , in other words, $x = X_m$.

Lemma 4.3. For sufficiently smooth functions $f(\cdot)$ and $g(\cdot)$ the θ -discretization errors are of order

$$\mathcal{R}_m^y(x) = \mathcal{O}((\Delta t)^3), \tag{4.10a}$$

$$\mathcal{R}_m^z(x) = \mathcal{O}((\Delta t)^3). \tag{4.10b}$$

Besides,

$$\mathbb{E} [\mathcal{R}_{m+1}^y(X_{m+1}^{m,x}) \Delta \omega_{m+1}] = \mathcal{O}((\Delta t)^4), \tag{4.11a}$$

$$\mathbb{E} [\mathcal{R}_{m+1}^z(X_{m+1}^{m,x})] - \mathcal{R}_m^z(x) = \mathcal{O}((\Delta t)^4). \tag{4.11b}$$

Proof. For equations (4.10a) and (4.10b) we use that for a general sufficiently smooth function $h(t, x)$, with bounded derivatives, it holds that

$$\begin{aligned} &\mathbb{E} \left[\int_{t_m}^{t_{m+1}} h(s, X_s^{m,x}) ds - \Delta t \frac{1}{2} (h(t_m, x) + h(t_{m+1}, X_{m+1}^{m,x})) \right] \\ &= h(t_m, x) \Delta t + h_{(0)}(t_m, x) \frac{1}{2} (\Delta t)^2 + h_{(0,0)}(t_m, x) \frac{1}{3!} (\Delta t)^3 + \mathcal{O}((\Delta t)^4) \\ &\quad - \Delta t \frac{1}{2} (h(t_m, x) + h(t_m, x) + h_{(0)}(t_m, x) \Delta t + h_{(0,0)}(t_m, x) \frac{1}{2} (\Delta t)^2 + \mathcal{O}((\Delta t)^3)) \\ &= h_{(0,0)}(t_m, x) \frac{-1}{12} (\Delta t)^3 + \mathcal{O}((\Delta t)^4) = \mathcal{O}((\Delta t)^3). \tag{4.12} \end{aligned}$$

For equation (4.10b) we use:

$$\begin{aligned} &\mathbb{E} \left[\int_{t_m}^{t_{m+1}} h(s, X_s^{m,x}) \Delta \omega_{m+1} ds - \Delta t \frac{1}{2} h(t_{m+1}, X_{m+1}^{m,x}) \Delta \omega_{m+1} \right] \\ &= h_{(1)}(t_m, x) \frac{1}{2} (\Delta t)^2 + [h_{(1,0)}(t_m, x) + h_{(0,1)}(t_m, x)] \frac{1}{3!} (\Delta t)^3 + \mathcal{O}((\Delta t)^4) \\ &\quad - \Delta t \frac{1}{2} (h_{(1)}(t_m, x) \Delta t + [h_{(1,0)}(t_m, x) + h_{(0,1)}(t_m, x)] \frac{1}{2} (\Delta t)^2 + \mathcal{O}((\Delta t)^3)) \\ &= [h_{(1,0)}(t_m, x) + h_{(0,1)}(t_m, x)] \frac{-1}{12} (\Delta t)^3 + \mathcal{O}((\Delta t)^4) = \mathcal{O}((\Delta t)^3). \tag{4.13} \end{aligned}$$

For equation (4.11a) we use:

$$\begin{aligned}
 & \mathbb{E} \left[\Delta\omega_{m+1} \left(\int_{t_{m+1}}^{t_{m+2}} h(s, X_s^{m,x}) ds - \frac{1}{2} \Delta t (h(t_{m+1}, X_{m+1}^{m,x}) + h(t_{m+2}, X_{m+2}^{m,x})) \right) \right] \\
 &= \mathbb{E} \left[\Delta\omega_{m+1} \mathbb{E} \left[\int_{t_{m+1}}^{t_{m+2}} h(s, X_s^{m+1, X_{m+1}^{m,x}}) ds - \Delta t \frac{1}{2} (h(t_{m+1}, X_{m+1}^{m,x}) + h(t_{m+2}, X_{m+2}^{m+1, X_{m+1}^{m,x}})) \right] \right] \\
 &= \mathbb{E} \left[\Delta\omega_{m+1} h_{(0,0)}(t_{m+1}, X_{m+1}^{m,x}) \frac{1}{12} (\Delta t)^3 + \mathcal{O}((\Delta t)^4) \right] \\
 &= h_{(1,0,0)}(t_m, x) \frac{1}{12} (\Delta t)^4 + \mathcal{O}((\Delta t)^4) = \mathcal{O}((\Delta t)^4). \tag{4.14}
 \end{aligned}$$

For equation (4.11b) we use:

$$\begin{aligned}
 & \mathbb{E} \left[\mathbb{E} \left[\int_{t_{m+1}}^{t_{m+2}} h(s, X_s^{m+1, X_{m+1}^{m,x}}) ds - \Delta t \frac{1}{2} (h(t_{m+1}, X_{m+1}^{m,x}) + h(t_{m+2}, X_{m+2}^{m+1, X_{m+1}^{m,x}})) \right] \right] \\
 & - \mathbb{E} \left[\int_{t_m}^{t_{m+1}} h(s, X_s^{m,x}) ds - \Delta t \frac{1}{2} (h(t_m, x) + h(t_{m+1}, X_{m+1}^{m,x})) \right] \\
 &= \mathbb{E} \left[h_{(0,0)}(t_{m+1}, X_{m+1}^{m,x}) \frac{1}{12} (\Delta t)^3 - h_{(0,0)}(t_m, x) \frac{1}{12} (\Delta t)^3 + \mathcal{O}((\Delta t)^4) \right] \\
 &= (h_{(0,0)}(t_m, x) + h_{(0,0,0)}(t_m, x) \Delta t + \mathcal{O}((\Delta t)^2)) \frac{1}{12} (\Delta t)^3 \\
 & - h_{(0,0)}(t_m, x) \frac{1}{12} (\Delta t)^3 + \mathcal{O}((\Delta t)^4) \\
 &= h_{(0,0,0)}(t_m, x) \frac{1}{12} (\Delta t)^4 + \mathcal{O}((\Delta t)^4) = \mathcal{O}((\Delta t)^4) \tag{4.15}
 \end{aligned}$$

and

$$\begin{aligned}
 & \mathbb{E} \left[\mathbb{E} \left[\int_{t_{m+1}}^{t_{m+2}} h(s, X_s^{m+1, X_{m+1}^{m,x}}) \Delta\omega_{m+2} ds - \Delta t \frac{1}{2} h(t_{m+2}, X_{m+2}^{m+1, X_{m+1}^{m,x}}) \Delta\omega_{m+2} \right] \right] \\
 & - \mathbb{E} \left[\int_{t_m}^{t_{m+1}} h(s, X_s^{m,x}) \Delta\omega_{m+1} ds - \Delta t \frac{1}{2} h(t_{m+1}, X_{m+1}^{m,x}) \Delta\omega_{m+1} \right] \\
 &= \mathbb{E} \left[[h_{(1,0)}(t_{m+1}, X_{m+1}^{m,x}) + h_{(0,1)}(t_{m+1}, X_{m+1}^{m,x})] \frac{1}{12} (\Delta t)^3 - [h_{(1,0)}(t_m, x) + h_{(0,1)}(t_m, x)] \frac{1}{12} (\Delta t)^3 + \mathcal{O}((\Delta t)^4) \right] \\
 &= ([h_{(1,0)}(t_m, x) + h_{(0,1)}(t_m, x)] + [h_{(0,1,0)}(t_m, x) + h_{(0,0,1)}(t_m, x)] \Delta t + \mathcal{O}((\Delta t)^2)) \frac{1}{12} (\Delta t)^3 \\
 & - [h_{(1,0)}(t_m, x) + h_{(0,1)}(t_m, x)] \frac{1}{12} (\Delta t)^3 + \mathcal{O}((\Delta t)^4) \\
 &= [h_{(0,1,0)}(t_m, x) + h_{(0,0,1)}(t_m, x)] \frac{1}{12} (\Delta t)^4 + \mathcal{O}((\Delta t)^4) = \mathcal{O}((\Delta t)^4). \tag{4.16}
 \end{aligned}$$

4.4. Global error Δ -time-discretization scheme FBSDE

The equation for $y_m^\Delta(x)$ is now given by

$$y_m^\Delta(x) = \mathbb{E} \left[y_{m+1}^\Delta(X_{m+1}^{\Delta, m, x}) \right] + \Delta t \frac{1}{2} f(t_m, \mathbf{\Lambda}_m^\Delta(x)) + \Delta t \frac{1}{2} \mathbb{E} \left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}^\Delta(X_{m+1}^{\Delta, m, x})) \right], \tag{4.17}$$

and the equation for $z_m^\Delta(x)$ reads

$$\begin{aligned}
 z_m^\Delta(x) &= \frac{2}{\Delta t} \mathbb{E} \left[y_{m+1}^\Delta(X_{m+1}^{\Delta, m, x}) \Delta\omega_{m+1} \right] + \mathbb{E} \left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}^\Delta(X_{m+1}^{\Delta, m, x})) \Delta\omega_{m+1} \right] \\
 & - \mathbb{E} \left[z_{m+1}^\Delta(X_{m+1}^{\Delta, m, x}) \right]. \tag{4.18}
 \end{aligned}$$

We define the following global errors,

$$\epsilon_m^y(X_m, X_m^\Delta) := y_m(X_m) - y_m^\Delta(X_m^\Delta), \quad (4.19a)$$

$$\epsilon_m^z(X_m, X_m^\Delta) := z_m(X_m) - z_m^\Delta(X_m^\Delta), \quad (4.19b)$$

$$\epsilon_m^f(X_m, X_m^\Delta) := f(t_m, \mathbf{\Lambda}_m(X_m)) - f(t_m, \mathbf{\Lambda}_m^\Delta(X_m^\Delta)), \quad (4.19c)$$

and

$$\epsilon_m^y(x) := y_m(x) - y_m^\Delta(x), \quad (4.20a)$$

$$\epsilon_m^z(x) := z_m(x) - z_m^\Delta(x), \quad (4.20b)$$

$$\epsilon_m^f(x) := f(t_m, \mathbf{\Lambda}_m(x)) - f(t_m, \mathbf{\Lambda}_m^\Delta(x)). \quad (4.20c)$$

For the error at time t_m we find

$$\begin{aligned} \epsilon_m^y(X_m, X_m^\Delta) &= y_m(X_m) - y_m^\Delta(X_m) + y_m^\Delta(X_m) - y_m^\Delta(X_m^\Delta) \\ &= \epsilon_m^y(X_m) + y_m^\Delta(X_m) - y_m^\Delta(X_m^\Delta), \end{aligned} \quad (4.21a)$$

$$|\mathbb{E}_0[\epsilon_m^y(X_m, X_m^\Delta)]| \leq |\mathbb{E}_0[\epsilon_m^y(X_m)]| + |\mathbb{E}_0[y_m^\Delta(X_m) - y_m^\Delta(X_m^\Delta)]|. \quad (4.21b)$$

The second part is the weak error and, assuming that $y_m^\Delta(x)$ and $z_m^\Delta(x)$ are $2(\gamma_w^{2T} + 1)$ -times continuously differentiable with polynomial growth, we have

$$|\mathbb{E}_0[\epsilon_m^y(X_m, X_m^\Delta)]| \leq |\mathbb{E}_0[\epsilon_m^y(X_m)]| + \mathcal{O}((\Delta t)^{\gamma_w^{2T}}). \quad (4.22)$$

Similarly, we get

$$|\mathbb{E}_0[\epsilon_m^z(X_m, X_m^\Delta)]| \leq |\mathbb{E}_0[\epsilon_m^z(X_m)]| + \mathcal{O}((\Delta t)^{\gamma_w^{2T}}). \quad (4.23)$$

We find the following bounds on the errors $\epsilon_m^y(x)$ and $\epsilon_m^z(x)$.

Theorem 4.1. *Given*

$$\mathbb{E}_{M-1}[|\epsilon_M^y(X_M)|] \sim \mathcal{O}((\Delta t)^3), \quad \mathbb{E}_{M-1}[|\epsilon_M^z(X_M)|] \sim \mathcal{O}((\Delta t)^3), \quad (4.24)$$

then

$$\mathbb{E}_0[|\epsilon_m^y(X_m)| + \Delta t |\epsilon_m^z(X_m)|] \leq \mathcal{O}((\Delta t)^2). \quad (4.25)$$

Remark 4.1. [Theorem 4.1](#) states second-order convergence for Y and first-order convergence for Z . However, in our numerical experiments we also find second-order convergence for Z if we apply the 2.0-weak-Taylor scheme and $\theta = 1/2$. Higher-order schemes are developed in [\[5,4\]](#). In our case a Fourier-based method results in a very efficient numerical scheme.

Remark 4.2. For the scheme with $\theta = 1$ the θ -discretization errors are one order lower and we find first-order convergence. For the Euler and Milstein schemes, the weak convergence rate $\gamma_w^{2T} = 2$ should be replaced by their weak convergence rates $\gamma_w^E = 1$ and $\gamma_w^M = 1$, respectively, and we find, for both $\theta = 1/2$ and $\theta = 1$, first order convergence.

Proof. Error Y For the Y -component we have with [\(4.6\)](#) and [\(4.17\)](#)

$$\begin{aligned} \epsilon_m^y(x) &= y_m(x) - y_m^\Delta(x) \\ &= \mathbb{E} \left[\epsilon_{m+1}^y(X_{m+1}^{m,x}, X_{m+1}^{\Delta,m,x}) \right] + \Delta t \frac{1}{2} \epsilon_m^f(x) + \Delta t \frac{1}{2} \mathbb{E} \left[\epsilon_{m+1}^f(X_{m+1}^{m,x}, X_{m+1}^{\Delta,m,x}) \right] + \mathcal{R}_m^y(x). \end{aligned} \quad (4.26)$$

It follows, with equality [\(4.5a\)](#), that

$$\begin{aligned} \mathbb{E} \left[\epsilon_{m+1}^y(X_{m+1}^{m,x}, X_{m+1}^{\Delta,m,x}) \right] &= \mathbb{E} \left[\epsilon_{m+1}^y(X_{m+1}^{m,x}) \right] + \mathbb{E} \left[y_{m+1}^\Delta(X_{m+1}^{m,x}) - y_{m+1}^\Delta(X_{m+1}^{\Delta,m,x}) \right] \\ &= \mathbb{E} \left[\epsilon_{m+1}^y(X_{m+1}^{m,x}) \right] + \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}), \end{aligned} \quad (4.27a)$$

$$\begin{aligned} \mathbb{E} \left[\epsilon_{m+1}^f(X_{m+1}^{m,x}, X_{m+1}^{\Delta,m,x}) \right] &= \mathbb{E} \left[\epsilon_{m+1}^f(X_{m+1}^{m,x}) \right] + \mathbb{E} \left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}^\Delta(X_{m+1}^{m,x})) - f(t_{m+1}, \mathbf{\Lambda}_{m+1}^\Delta(X_{m+1}^{\Delta,m,x})) \right] \\ &= \mathbb{E} \left[\epsilon_{m+1}^f(X_{m+1}^{m,x}) \right] + \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}). \end{aligned} \quad (4.27b)$$

Driver function $f(t, x, y, z)$ is Lipschitz in $x, y,$ and $z,$ so that

$$|\epsilon_m^f(x)| \leq L_f (|\epsilon_m^y(x)| + |\epsilon_m^z(x)|), \quad (4.28a)$$

$$\mathbb{E} \left[|\epsilon_{m+1}^f(X_{m+1}^{m,x})| \right] \leq L_f \mathbb{E} [|\epsilon_{m+1}^y(X_{m+1}^{m,x})| + |\epsilon_{m+1}^z(X_{m+1}^{m,x})|]. \quad (4.28b)$$

Then,

$$\begin{aligned} |\epsilon_m^y(x)| &\leq \frac{1 + \frac{1}{2} \Delta t L_f}{1 - \frac{1}{2} \Delta t L_f} |\mathbb{E} [\epsilon_{m+1}^y(X_{m+1}^{m,x})]| + \frac{\frac{1}{2} \Delta t L_f}{1 - \frac{1}{2} \Delta t L_f} |\epsilon_m^z(x)| \\ &+ \frac{\frac{1}{2} \Delta t L_f}{1 - \frac{1}{2} \Delta t L_f} \mathbb{E} [|\epsilon_{m+1}^z(X_{m+1}^{m,x})|] + \frac{1}{1 - \frac{1}{2} \Delta t L_f} \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}) + \frac{1}{1 - \frac{1}{2} \Delta t L_f} \mathcal{O}((\Delta t)^3). \end{aligned} \quad (4.29)$$

Error Z For the Z -component we have, with (4.7) and (4.18)

$$\begin{aligned} \epsilon_m^z(x) &= z_m(x) - z_m^\Delta(x) \\ &= \frac{2}{\Delta t} \mathbb{E} \left[\epsilon_{m+1}^y(X_{m+1}^{m,x}, X_{m+1}^{\Delta,m,x}) \Delta \omega_{m+1} \right] + \mathbb{E} \left[\epsilon_{m+1}^f(X_{m+1}^{m,x}, X_{m+1}^{\Delta,m,x}) \Delta \omega_{m+1} \right] \\ &- \mathbb{E} \left[\epsilon_{m+1}^z(X_{m+1}^{m,x}, X_{m+1}^{\Delta,m,x}) \right] + \frac{2}{\Delta t} \mathcal{R}_m^z(x). \end{aligned} \quad (4.30)$$

Substituting the similar equations for ϵ_{m+1}^y and ϵ_{m+1}^z as in (4.26) and (4.30) gives,

$$\begin{aligned} \epsilon_m^z(x) &= \frac{2}{\Delta t} \mathbb{E} \left[\epsilon_{m+2}^y(X_{m+2}^{m,x}, X_{m+2}^{\Delta,m,x}) \Delta \omega_{m+1} \right] + \mathbb{E} \left[\epsilon_{m+1}^f(X_{m+1}^{m,x}, X_{m+1}^{\Delta,m,x}) \Delta \omega_{m+1} \right] \\ &+ \mathbb{E} \left[\epsilon_{m+2}^f(X_{m+2}^{m,x}, X_{m+2}^{\Delta,m,x}) \Delta \omega_{m+1} \right] + \frac{2}{\Delta t} \mathbb{E} \left[\mathcal{R}_{m+1}^y(X_{m+1}^{m,x}) \Delta \omega_{m+1} \right] \\ &+ \mathbb{E} \left[\epsilon_{m+1}^f(X_{m+1}^{m,x}, X_{m+1}^{\Delta,m,x}) \Delta \omega_{m+1} \right] - \frac{2}{\Delta t} \mathbb{E} \left[\epsilon_{m+2}^y(X_{m+2}^{m,x}, X_{m+2}^{\Delta,m,x}) \Delta \omega_{m+2} \right] \\ &- \mathbb{E} \left[\epsilon_{m+2}^f(X_{m+2}^{m,x}, X_{m+2}^{\Delta,m,x}) \Delta \omega_{m+2} \right] + \mathbb{E} \left[\epsilon_{m+2}^z(X_{m+2}^{m,x}, X_{m+2}^{\Delta,m,x}) \right] \\ &- \frac{2}{\Delta t} \mathbb{E} \left[\mathcal{R}_{m+1}^z(X_{m+1}^{m,x}) \right] + \frac{2}{\Delta t} \mathcal{R}_m^z(x). \end{aligned} \quad (4.31)$$

The different terms in the above equation can be bounded as follows:

- With equality (4.5c),

$$\left| \mathbb{E} \left[\epsilon_{m+2}^z(X_{m+2}^{m,x}, X_{m+2}^{\Delta,m,x}) \right] \right| \leq \left| \mathbb{E} \left[\epsilon_{m+2}^z(X_{m+2}^{m,x}) \right] \right| + \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}). \quad (4.32)$$

- With equality (4.5b),

$$\left| 2 \mathbb{E} \left[\epsilon_{m+1}^f(X_{m+1}^{m,x}, X_{m+1}^{\Delta,m,x}) \Delta \omega_{m+1} \right] \right| \leq C \sqrt{\Delta t} |\epsilon_{m+1}^f(X_{m+1}^{m,x})|_\infty + \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}), \quad (4.33)$$

where $|\cdot|_\infty$ denotes the infinite norm and $C > 0$ a constant.

- With equality (4.5d),

$$\left| \mathbb{E} \left[\epsilon_{m+2}^f(X_{m+2}^{m,x}, X_{m+2}^{\Delta,m,x}) \Delta \omega_{m+1} \right] \right| \leq C \sqrt{\Delta t} |\epsilon_{m+2}^f(X_{m+2}^{m,x})|_\infty + \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}). \quad (4.34)$$

- With equality (4.5e),

$$\left| \mathbb{E} \left[\epsilon_{m+2}^f(X_{m+2}^{m,x}, X_{m+2}^{\Delta,m,x}) \Delta \omega_{m+2} \right] \right| \leq C \sqrt{\Delta t} |\epsilon_{m+2}^f(X_{m+2}^{m,x})|_\infty + \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}). \quad (4.35)$$

- With² equality (4.5f),

$$\begin{aligned} &\left| \frac{2}{\Delta t} \mathbb{E} \left[\epsilon_{m+2}^y(X_{m+2}^{m,x}, X_{m+2}^{\Delta,m,x}) (\Delta \omega_{m+1} - \Delta \omega_{m+2}) \right] \right| \\ &\leq \frac{2}{\Delta t} \left| \mathbb{E} \left[\epsilon_{m+2}^y(X_{m+2}^{m,x}) (\Delta \omega_{m+1} - \Delta \omega_{m+2}) \right] \right| + \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}). \end{aligned} \quad (4.36)$$

We use the following inequality (see equation (4.3f))

$$\frac{2}{\Delta t} \left| \mathbb{E} \left[\epsilon_{m+2}^y(X_{m+2}^{m,x}) (\Delta \omega_{m+1} - \Delta \omega_{m+2}) \right] \right| \leq C \Delta t |\sigma(x) \epsilon_m^{D_x y}(x)|, \quad (4.37)$$

with

² For constant drift $\mu(\cdot)$ and volatility $\sigma(\cdot)$ this term cancels, see [26].

$$\begin{aligned}
\sigma(x)\epsilon_m^{D_x y}(x) &:= \sigma(x)D_x y_m(x) - \sigma(x)D_x y_m^\Delta(x) \\
&= z_m(x) - z_m^\Delta(x) + z_m^\Delta(x) - L^1 y_m^\Delta(x) \\
&= \epsilon_m^z(x) + z_m^\Delta(x) - L^1 y_m^\Delta(x)
\end{aligned} \tag{4.38}$$

and

$$\begin{aligned}
z_m^\Delta(x) - L^1 y_m^\Delta(x) &= \frac{2}{\Delta t} \mathbb{E} \left[y_{m+1}^\Delta(X_{m+1}^{\Delta, m, x}) \Delta \omega_{m+1} \right] + \mathbb{E} \left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}^\Delta(X_{m+1}^{\Delta, m, x})) \Delta \omega_{m+1} \right] \\
&\quad - \mathbb{E} \left[z_{m+1}^\Delta(X_{m+1}^{\Delta, m, x}) \right] - L^1 y_m^\Delta(x) \\
&= \frac{2}{\Delta t} \mathbb{E} \left[y_{m+1}^\Delta(X_{m+1}^{m, x}) \Delta \omega_{m+1} \right] + \mathbb{E} \left[f(t_{m+1}, \mathbf{\Lambda}_{m+1}^\Delta(X_{m+1}^{m, x})) \Delta \omega_{m+1} \right] \\
&\quad - \mathbb{E} \left[z_{m+1}^\Delta(X_{m+1}^{m, x}) \right] + \mathcal{O}((\Delta t)^{\gamma_w^{2T}}) - L^1 y_m^\Delta(x).
\end{aligned} \tag{4.39}$$

With equations (4.3a) and (4.3b) we find

$$\begin{aligned}
z_m^\Delta(x) - L^1 y_m^\Delta(x) &\approx \frac{2}{\Delta t} L^1 y_m^\Delta(x) \Delta t + \frac{2}{\Delta t} [L^1 L^0 y_m^\Delta(x) + L^0 L^1 y_m^\Delta(x)] \frac{1}{2} (\Delta t)^2 \\
&\quad + L^1 f(t_m, \mathbf{\Lambda}_m^\Delta(x)) \Delta t + [L^1 L^0 f(t_m, \mathbf{\Lambda}_m^\Delta(x)) + L^0 L^1 f(t_m, \mathbf{\Lambda}_m^\Delta(x))] \frac{1}{2} (\Delta t)^2 \\
&\quad - z_m^\Delta(x) - L^0 z_m^\Delta(x) \Delta t + \mathcal{O}((\Delta t)^{\gamma_w^{2T}}) - L^1 y_m^\Delta(x).
\end{aligned} \tag{4.40}$$

We deduce that $z_m^\Delta(x) - L^1 y_m^\Delta(x) = \mathcal{O}(\Delta t)$ and

$$\frac{2}{\Delta t} |\mathbb{E} [\epsilon_{m+2}^y(X_{m+2}^{m, x}) (\Delta \omega_{m+1} - \Delta \omega_{m+2})]| \leq C \Delta t |\epsilon_m^z(x)| + \mathcal{O}((\Delta t)^2). \tag{4.41}$$

We can now bound the absolute error by

$$\begin{aligned}
(1 - C \Delta t) |\epsilon_m^z(x)| &\leq |\mathbb{E} [\epsilon_{m+2}^z(X_{m+2}^{m, x})]| + c \sqrt{\Delta t} |\epsilon_{m+2}^f(X_{m+2}^{m, x})|_\infty + c \sqrt{\Delta t} |\epsilon_{m+1}^f(X_{m+1}^{m, x})|_\infty \\
&\quad + \frac{2}{\Delta t} |\mathbb{E} [\epsilon_{m+2}^y(X_{m+2}^{m, x}) (\Delta \omega_{m+1} - \Delta \omega_{m+2})]| + \mathcal{O}((\Delta t)^2) + \mathcal{O}((\Delta t)^{\gamma_w^{2T} + 1}).
\end{aligned} \tag{4.42}$$

Total error Summing up the errors gives us, for $\Delta t \leq 1$, ($m \leq M - 3$)

$$\begin{aligned}
&\mathbb{E} [|\epsilon_m^y(x)|_\infty + \Delta t |\epsilon_m^z(x)|_\infty + \Delta t |\epsilon_{m+1}^z(X_{m+1}^{m, x})|_\infty] \\
&\leq A \mathbb{E} [|\epsilon_{m+2}^y(X_{m+2}^{m, x})|_\infty + \Delta t |\epsilon_{m+2}^z(X_{m+2}^{m, x})|_\infty + \Delta t |\epsilon_{m+3}^z(X_{m+3}^{m, x})|_\infty] + B \quad \text{with} \\
A &= 1 + Q \Delta t, \\
B &= \mathcal{O}((\Delta t)^3) + \mathcal{O}((\Delta t)^{\gamma_w^{2T} + 1}),
\end{aligned} \tag{4.43}$$

with $Q > 0$ a constant. Iterating this equality results in [Theorem 4.1](#).

5. Numerical experiments FBSDE

In this section we discuss numerical experiments. MATLAB 7.11.0 is used for the computations. In the experiments we use $\theta = 1/2$ and $\theta = 1$. Similar as in [\[26\]](#), we prescribe a computational domain $[a, b]$ by

$$[a, b] = [\kappa_1 - L\sqrt{\kappa_2}, \kappa_1 + L\sqrt{\kappa_2}], \tag{5.1}$$

with cumulants of one Euler step $\kappa_1 = X_0 + \mu(X_0)T$ and $\kappa_2 = \sigma^2(X_0)T$, and $L = 10$. Furthermore, we set the number of terms in the Fourier cosine series expansions equal to $N = 2^9$. For these values the BCOS method has converged in N to machine precision.

5.1. Example 1

The first example is derived from [\[21,20\]](#). We take drift and diffusion term

$$\mu(x) = \frac{x(1+x^2)}{(2+x^2)^3}, \quad \sigma(x) = \frac{1+x^2}{2+x^2}. \tag{5.2}$$

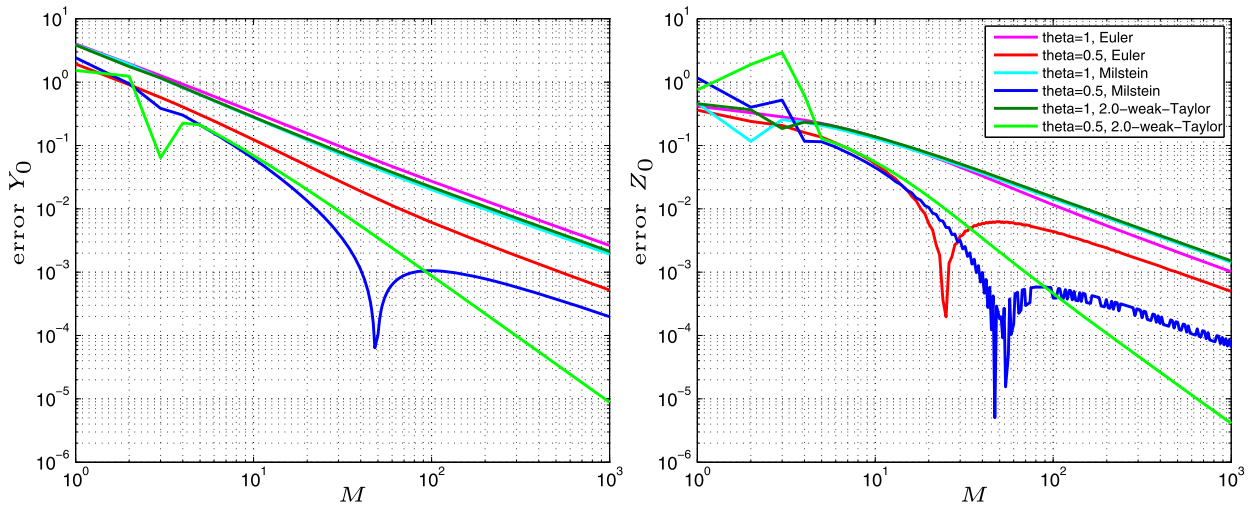


Fig. 5.1. Results example 1, left: error in $\hat{Y}_0^\Delta(x_0)$, right: error in $\hat{Z}_0^\Delta(x_0)$.

Table 2
CPU time (s).

$(N = 2^9) M$	4	8	16	32	64	128	256	512
Euler $\theta = 1$	0.0376	0.0507	0.0563	0.0852	0.1442	0.2620	0.5005	0.9958
2.0-weak-Taylor $\theta = \frac{1}{2}$	0.0976	0.1247	0.1249	0.1738	0.2738	0.4698	0.8668	1.6545
$(M = 256) N$	2^6	2^7	2^8	2^9				
Euler $\theta = 1$	0.1760	0.2028	0.2661	0.5005				
2.0-weak-Taylor $\theta = \frac{1}{2}$	0.2068	0.2438	0.4711	0.8668				

The driver function and terminal condition are given by

$$f(t, x, y, z) = \frac{1}{t+1} \exp\left(-\frac{x^2}{t+1}\right) \left(\frac{4x^2(1+x^2)}{(2+x^2)^3} + \frac{(1+x^2)^2}{(2+x^2)^2} \left(1 - \frac{2x^2}{t+1}\right) - \frac{x^2}{t+1} \right) - \frac{zx}{(2+x^2)^2} \sqrt{\frac{1+y^2 + \exp\left(-\frac{2x^2}{t+1}\right)}{1+2y^2}}, \tag{5.3a}$$

$$g(x) = \exp\left(-\frac{x^2}{T+1}\right). \tag{5.3b}$$

The exact solution reads $v(t, x) = \exp\left(-\frac{x^2}{t+1}\right)$. For the experiment we use parameter values $T = 10$ and $X_0 = 1$. The results are shown in Fig. 5.1. Only with the 2.0-weak-Taylor scheme and $\theta = 1/2$ we achieve second-order convergence in M , as expected. Notice that $\kappa(x) = 0$ for $x = 0$, and then the characteristic function is given by equation (3.6). The characteristic function decreases at a lower exponential rate in u for $x \approx 0$. This gives rise to an unstable behavior when N is not sufficiently high. In the figure this is visible for values $M < 4$.

Table 2 shows CPU times, where the small values demonstrate the efficiency of the BCOS method. Computation of the characteristic function and the function f on an x -grid are the most time-consuming parts of the algorithm. The computation time is linear in the number of timesteps M and of $\mathcal{O}(N^2)$ in the number of terms in the Fourier cosine series expansions. The latter is not clearly measurable for these values of N as a significant part of the computation time is spend on discrete Fourier cosine transforms, which is of order $\mathcal{O}(N \log N)$.

5.2. Example 2: European call option – CEV – \mathbb{P} -measure

In the second example we compute the price $v(t, X_t)$ of a European call option where the underlying asset follows the CEV asset price process,

$$dX_s = \bar{\mu} X_s ds + \bar{\sigma} X_s^\gamma d\omega_s, \tag{5.4}$$

with $\gamma > 0$. The exact solution is given by the CEV price [13]. For the derivation of the corresponding semilinear PDE we set up a self-financing portfolio Y_s with a_s assets and bonds with risk-free return rate r . Markets are assumed to be complete

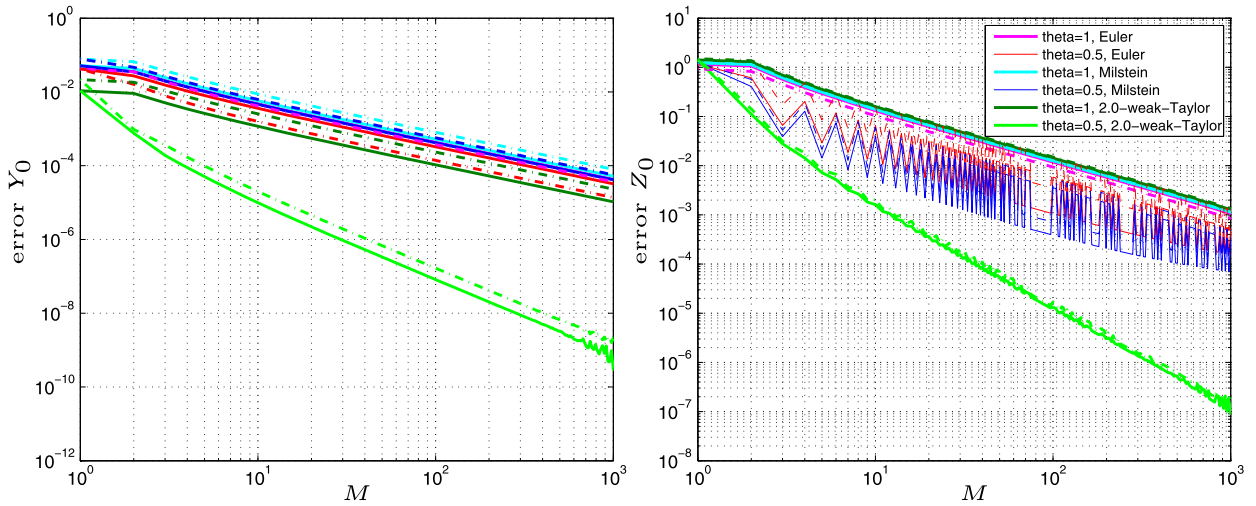


Fig. 5.2. Results example 2, left: error in $\hat{y}_0^\Delta(x_0)$, right: error in $\hat{z}_0^\Delta(x_0)$, $\gamma = 0.2$ (lines) and $\gamma = 0.8$ (dashes).

in this model, there are no trading restrictions, and the option can be exactly replicated by the hedging portfolio, that is $Y_T = \max(X_T - K, 0)$. Then, the option value at initial time should be equal to the initial value of the portfolio. The hedge portfolio evolves according to the SDE

$$dY_s = r(Y_s - a_s X_s) ds + a_s dX_s = (rY_s + (\mu(X_s) - rX_s)a_s) ds + \sigma(X_s)a_s d\omega_s. \tag{5.5}$$

If we set $Z_s = \sigma(X_s)a_s$, then (Y, Z) solve the FBSDE

$$dY_s = -f(s, X_s, Y_s, Z_s) ds + Z_s d\omega_s, \quad Y_T = \max(X_T - K, 0), \tag{5.6a}$$

$$f(t, x, y, z) = -ry - \frac{\mu(x) - rx}{\sigma(x)} z = -ry - \frac{\bar{\mu} - r}{\bar{\sigma}} x^{1-\gamma} z. \tag{5.6b}$$

Y_t corresponds to the value of the portfolio and Z_t is related to the hedging strategy. The option value is given by $v(t, X_t) = Y_t$ and $\sigma(X_t)D_x v(t, X_t) = Z_t$.

For the tests, we use the following parameter values

$$X_0 = 100, K = 100, r = 0.1, \bar{\mu} = 0.2, T = 0.1. \tag{5.7}$$

We take the elasticity of variance equal to $\gamma = 0.2$ and $\gamma = 0.8$ and choose $\bar{\sigma}$ so that $\sigma(X_0) = 25$.

As the terminal coefficients $Z_k^\Delta(t_M)$ (equation (3.41b)) and $\mathcal{F}_k^\Delta(t_M)$ (equation (3.41c)) are not known analytically and the corresponding functions are not smooth we take $\theta = 1$ in the first iteration with time step $(\Delta t)^2$. With this choice we do not need to compute these terminal coefficients but keep second-order convergence in the first iteration. We use a very large number of timesteps $M = 10^4$ to get reference values Z_0 . The results for the Euler, Milstein, and 2.0-weak-Taylor schemes are shown in Fig. 5.2. Again we achieve second-order convergence in M with the 2.0-weak-Taylor scheme and $\theta = 1/2$.

5.3. Example 3: bond price – CIR

In this section we consider the CIR interest rate process

$$dX_s = \kappa(\bar{x} - X_s) ds + \eta\sqrt{X_s} d\omega_s. \tag{5.8}$$

The PDE for the zero-coupon bond price is given by [27, p. 275]

$$\frac{\partial v}{\partial t}(t, x) + \kappa(\bar{x} - x)D_x v(t, x) + \frac{1}{2}\eta^2 x D_x^2 v(t, x) - xv(t, x) = 0, \quad (t, x) \in [0, T] \times \mathbb{R}_+, \tag{5.9a}$$

$$v(T, x) = 1, \quad x \in \mathbb{R}_+. \tag{5.9b}$$

This problem is related to the FBSDE

$$dY_s = -f(s, X_s, Y_s, Z_s) ds + Z_s d\omega_s, \quad Y_T = 1, \tag{5.10a}$$

$$f(t, x, y, z) = -xy. \tag{5.10b}$$

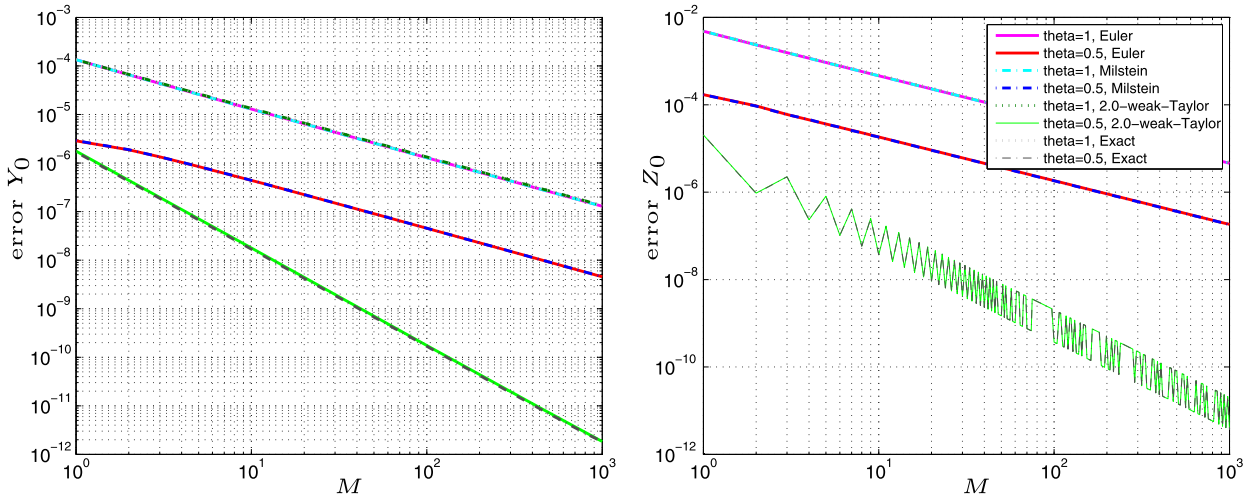


Fig. 5.3. Results example 3, left: error in $\hat{y}_0^\Delta(x_0)$, right: error in $\hat{z}_0^\Delta(x_0)$.

The exact solution is given by

$$h = \sqrt{\varkappa^2 + 2\eta^2}, \tag{5.11a}$$

$$A(t, T) = \left(\frac{2he^{\frac{1}{2}(\varkappa+h)(T-t)}}{2h + (\varkappa + h)(e^{h(T-t)} - 1)} \right)^{2\varkappa\bar{x}/\eta^2}, \tag{5.11b}$$

$$B(t, T) = \frac{2(e^{(T-t)h} - 1)}{2h + (\varkappa + h)(e^{h(T-t)} - 1)}, \tag{5.11c}$$

$$v(t, x) = A(t, T) \exp(-B(t, T)x). \tag{5.11d}$$

For the tests, we use the following parameter values

$$X_0 = 0.04, \bar{x} = 0.01, \varkappa = 0.2, \eta = 0.1, T = 0.25. \tag{5.12}$$

The Feller condition is not satisfied and the process may reach zero. This is a nontrivial situation for Monte Carlo methods but for BCOS it is no problem, as we take the left-hand side of the computational domain equal to $a = 0$. For the exact simulation approach (Section 3.3.1) we use the characteristic function in equation (2.24). The results for the Euler, Milstein, 2.0-weak-Taylor and exact simulation schemes are shown in Fig. 5.3. The 2.0-weak-Taylor and exact simulation scheme give almost the same results, with second-order convergence for $\theta = 1/2$.

5.4. Example 4: time-dependent drift and diffusion

The drift and diffusion terms of the FSDE, see equation (2.4), did not depend on time. In this last example we generalize the method by a time-dependent drift $\mu(t, x)$ and diffusion $\sigma(t, x)$. The numerical schemes in Sections 2 and 3 remain valid and the only change is due to the time-derivative in the diffusion operator L^0 , equation (2.9). The functions $m(x)$ and $s(x)$ in Section 3.1 for the 2.0-weak-Taylor scheme become time-dependent:

$$m(t, x) = \mu(t, x) - \frac{1}{2}\sigma(t, x)\sigma_x(t, x) + \frac{1}{2} \left(\mu_t(t, x) + \mu(t, x)\mu_x(t, x) + \frac{1}{2}\mu_{xx}(t, x)\sigma^2(t, x) \right) \Delta t, \tag{5.13a}$$

$$s(t, x) = \sigma(t, x) + \frac{1}{2} \left(\mu_x(t, x)\sigma(t, x) + \sigma_t(t, x) + \mu(t, x)\sigma_x(t, x) + \frac{1}{2}\sigma_{xx}(t, x)\sigma^2(t, x) \right) \Delta t. \tag{5.13b}$$

We perform a test for the price $v(t, X_t)$ of a call option where the underlying asset follows a geometric Brownian motion with time-dependent drift and volatility. We choose the following periodic functionals:

$$\mu(t, x) := \bar{\mu}(t)x, \quad \bar{\mu}(t) := \bar{\mu}_0 + \bar{\mu}_1 \sin\left(\frac{2\pi t}{\zeta_1}\right) + \bar{\mu}_2 \sin\left(\frac{2\pi t}{\zeta_2}\right), \tag{5.14a}$$

$$\sigma(t, x) := \bar{\sigma}(t)x, \quad \bar{\sigma}(t) := \bar{\sigma}_0 + \bar{\sigma}_1 \sin\left(\frac{2\pi t}{\zeta_1}\right) + \bar{\sigma}_2 \sin\left(\frac{2\pi t}{\zeta_2}\right). \tag{5.14b}$$

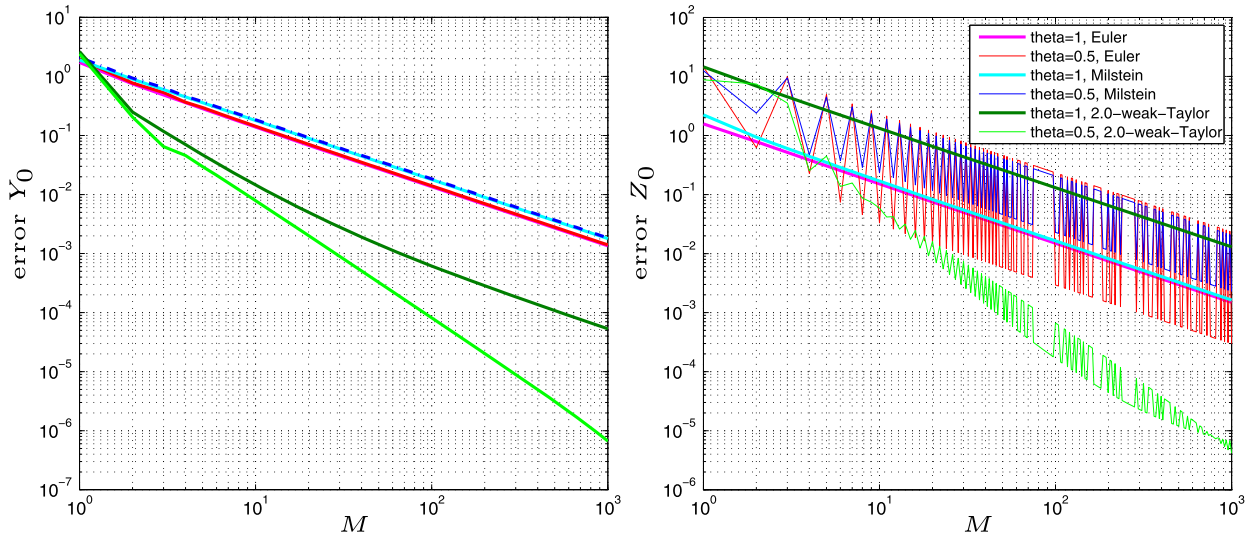


Fig. 5.4. Results example 4, left: error in $\hat{y}_0^\Delta(x_0)$, right: error in $\hat{z}_0^\Delta(x_0)$.

The hedge portfolio, see Section 5.2, evolves according to the SDE

$$dY_s = \left(rY_s + \frac{\bar{\mu}(s) - r}{\bar{\sigma}(s)} \bar{\sigma}(s) a_s X_s \right) ds + \sigma(s, X_s) a_s d\omega_s. \tag{5.15}$$

If we set $Z_s = \sigma(s, X_s) a_s$, then (Y, Z) solve the FBSDE

$$dY_s = -f(s, X_s, Y_s, Z_s) ds + Z_s d\omega_s, \quad Y_T = \max(X_T - K, 0), \tag{5.16a}$$

$$f(t, x, y, z) = -ry - \frac{\bar{\mu}(t) - r}{\bar{\sigma}(t)} z. \tag{5.16b}$$

Again the option value is given by $v(t, X_t) = Y_t$ and $\sigma(t, X_t) D_x v(t, X_t) = Z_t$. The exact solution of this local volatility model is given by the Black–Scholes price and Delta with volatility parameter $\sqrt{\frac{1}{T-t} \int_t^T \bar{\sigma}^2(s) ds}$ [24]. For the tests, we use the following parameter values

$$\begin{aligned} X_0 = 100, K = 100, r = 0.1, \bar{\mu}_0 = 0.2, \bar{\sigma}_0 = 0.25, T = 0.25, \\ \bar{\mu}_1 = 0.1, \bar{\mu}_2 = 0.02, \bar{\sigma}_1 = 0.125, \bar{\sigma}_2 = 0.025, \zeta_1 = 1, \zeta_2 = \frac{1}{4}, \end{aligned} \tag{5.17}$$

with the exact solutions $Y_0 = v(t_0, X_0) = 7.8159$ and $Z_0 = \sigma X_0 D_x v(t_0, X_0) = 14.8115$. The results are shown in Fig. 5.4. The computation time has increased because the characteristic function (3.5) changes with each timestep. The computation time for $M = 1000$ is about eight times longer.

6. Conclusion

In this paper we extended the probabilistic numerical BCOS method [26], for solving decoupled forward–backward stochastic differential equations. The underlying forward stochastic differential equation is now approximated by different Taylor discretization schemes, such as the Euler, Milstein, and Order 2.0 weak Taylor schemes, or by exact simulation. The discretization of the FBSDE with the θ -method results in a backward induction scheme with conditional expectations. The expected values are approximated by a Fourier cosine method and relies on the availability of the characteristic function for these discrete Taylor schemes. In this way we generalize the applicability of the BCOS method to FSDEs for which the ‘continuous’ characteristic function is not available. The Fourier cosine coefficients are recovered recursively in an efficient way by using discrete Fourier cosine transforms and an FFT algorithm. Numerical tests demonstrate the applicability of the BCOS method for BSDEs in financial problems. In the tests we observed different convergence results for Z_0 and Y_0 . The 2.0-weak-Taylor and exact simulation scheme with $\theta = 1/2$ result in second-order convergence in the number of timesteps of the BCOS method, as expected from the error analysis.

Appendix A. Proof Lemma 4.2

The Itô–Taylor expansion of FSDE X_t reads (with $c(x) = x$)

$$X_{m+1}^{m,x} = \sum_{\alpha \in \mathcal{A}} c_\alpha(x) I_{\alpha,m+1} + \sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_\alpha [c_\alpha(X_{m+1}^{m,x})]_{m+1}, \quad (\text{A.1a})$$

$$X_{m+2}^{m,x} = \sum_{\beta \in \mathcal{A}} c_\beta(X_{m+1}^{m,x}) I_{\beta,m+2} + \sum_{\beta \in \mathcal{B}(\mathcal{A})} I_\beta [c_\beta(X_{m+1}^{m,x})]_{m+2}. \quad (\text{A.1b})$$

For the Order 2.0 weak Taylor scheme we have

$$\mathcal{A} = \{v, (0), (1), (1, 1), (1, 0), (0, 1), (0, 0)\} \quad \text{and}$$

$$\mathcal{B}(\mathcal{A}) = \{(1, 1, 1), (0, 0, 1), (0, 1, 0), (1, 0, 0), (0, 1, 1), (1, 0, 1), (1, 1, 0), (0, 0, 0)\}. \quad (\text{A.2})$$

The discrete approximations read

$$X_{m+1}^{\Delta,m,x} = \sum_{\alpha \in \mathcal{A}} c_\alpha(x) I_{\alpha,m+1}, \quad X_{m+2}^{\Delta,m,x} = \sum_{\beta \in \mathcal{A}} c_\beta(X_{m+1}^{\Delta,m,x}) I_{\beta,m+2}. \quad (\text{A.3})$$

With the theory in Chapter 5.7 of [17] we find

$$\mathbb{E}_m \left[\sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_\alpha [h(\cdot, X)]_{m+1} \right] = \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}), \quad (\text{A.4a})$$

$$\mathbb{E}_m \left[\sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_\alpha [h(\cdot, X)]_{m+1} \Delta \omega_{m+1} \right] = \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}), \quad (\text{A.4b})$$

$$\frac{2}{\Delta t} \mathbb{E}_m \left[\sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_{(0,\alpha)} [h_{(0,\alpha)}(\cdot, X)]_{m+1} \Delta \omega_{m+1} \right] = 0, \quad (\text{A.4c})$$

$$\frac{2}{\Delta t} \mathbb{E}_m \left[\sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_{(1,\alpha)} [h_{(1,\alpha)}(\cdot, X)]_{m+1} \Delta \omega_{m+1} \right] = \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}), \quad (\text{A.4d})$$

$$\frac{2}{\Delta t} \mathbb{E}_m \left[\sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_0 [h_{(0,\alpha)}(\cdot, X)]_{m+1} \right] = \mathcal{O}((\Delta t)^0), \quad (\text{A.4e})$$

$$\frac{2}{\Delta t} \mathbb{E}_m \left[\sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_1 [h_{(1,\alpha)}(\cdot, X)]_{m+1} \right] = 0, \quad (\text{A.4f})$$

$$\frac{2}{\Delta t} \mathbb{E}_m \left[\sum_{\alpha \in \mathcal{A} \setminus \alpha_\emptyset} I_{\alpha,m+1} \right] + \frac{2}{\Delta t} \mathbb{E}_{m+1} \left[\sum_{\alpha \in \mathcal{A} \setminus \alpha_\emptyset} I_{\alpha,m+1} \Delta \omega_{m+1} \right] = \mathcal{O}((\Delta t)^0). \quad (\text{A.4g})$$

With the above equations we find

$$\mathbb{E} \left[X_{m+1}^{m,x} - X_{m+1}^{\Delta,m,x} \right] = \mathbb{E} \left[\sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_\alpha [c_\alpha(X_{m+1}^{m,x})]_{m+1} \right] = \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}), \quad (\text{A.5a})$$

$$\mathbb{E} \left[\left(X_{m+1}^{m,x} - X_{m+1}^{\Delta,m,x} \right) \Delta \omega_{m+1} \right] = \mathbb{E} \left[\sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_\alpha [c_\alpha(X_{m+1}^{m,x})]_{m+1} \Delta \omega_{m+1} \right] = \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}). \quad (\text{A.5b})$$

The Taylor series of function $h(t_{m+1}, x)$ in x gives

$$h(t_{m+1}, X_{m+1}^{m,x}) - h(t_{m+1}, X_{m+1}^{\Delta,m,x}) = (-1)^{\ell+1} \sum_{\ell=1}^{\infty} \frac{1}{\ell!} h^{(\ell)}(t_{m+1}, X_{m+1}^{m,x}) \left(X_{m+1}^{m,x} - X_{m+1}^{\Delta,m,x} \right)^\ell. \quad (\text{A.6})$$

By using the strong error result ([17], Chapter 5.9)

$$\mathbb{E} \left[\left| X_{m+1}^{m,x} - X_{m+1}^{\Delta,m,x} \right|^\ell \right] = \mathcal{O}((\Delta t)^{1.5\ell}) \quad (\text{A.7})$$

for $\ell \geq 2$ and formula (4.1) for $\ell = 1$, we find

$$\mathbb{E} \left[h(t_{m+1}, X_{m+1}^{m,x}) - h(t_{m+1}, X_{m+1}^{\Delta,m,x}) \right] = \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}), \quad (\text{A.8a})$$

$$\mathbb{E} \left[\left(h(t_{m+1}, X_{m+1}^{m,x}) - h(t_{m+1}, X_{m+1}^{\Delta,m,x}) \right) \Delta \omega_{m+1} \right] = \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}). \quad (\text{A.8b})$$

We can rewrite the difference between $X_{m+2}^{m,x}$ and $X_{m+2}^{\Delta,m,x}$ as

$$\begin{aligned}
X_{m+2}^{m,x} - X_{m+2}^{\Delta,m,x} &= \left(\sum_{\beta \in \mathcal{A}} c_{\beta}(X_{m+1}^{m,x}) I_{\beta,m+2} + \sum_{\beta \in \mathcal{B}(\mathcal{A})} I_{\beta}[c_{\beta}(X_{m+1}^{m,x})]_{m+2} \right) - \sum_{\beta \in \mathcal{A}} c_{\beta}(X_{m+1}^{m,x}) I_{\beta,m+2} \\
&\quad + \sum_{\beta \in \mathcal{A}} c_{\beta}(X_{m+1}^{m,x}) I_{\beta,m+2} - \sum_{\beta \in \mathcal{A}} c_{\beta}(X_{m+1}^{\Delta,m,x}) I_{\beta,m+2} \\
&= \sum_{\beta \in \mathcal{B}(\mathcal{A})} I_{\beta}[c_{\beta}(X_{m+1}^{m,x})]_{m+2} \\
&\quad + \sum_{\beta \in \mathcal{A}} \left(c_{\beta}(X_{m+1}^{m,x}) - c_{\beta}(X_{m+1}^{\Delta,m,x}) \right) I_{\beta,m+2} \\
&= \sum_{\beta \in \mathcal{B}(\mathcal{A})} I_{\beta}[c_{\beta}(X_{m+1}^{m,x})]_{m+2} \\
&\quad + \sum_{\beta \in \mathcal{A} \setminus \alpha_{\emptyset}} \left(c_{\beta}(X_{m+1}^{m,x}) - c_{\beta}(X_{m+1}^{\Delta,m,x}) \right) I_{\beta,m+2} + \sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_{\alpha}[c_{\alpha}(X_{m+1}^{m,x})]_{m+1}. \tag{A.9}
\end{aligned}$$

Then we find

$$\begin{aligned}
\mathbb{E} \left[X_{m+2}^{m,x} - X_{m+2}^{\Delta,m,x} \right] &= \mathbb{E} \left[\sum_{\beta \in \mathcal{B}(\mathcal{A})} I_{\beta}[c_{\beta}(X_{m+1}^{m,x})]_{m+2} \right] \\
&\quad + \mathbb{E} \left[\sum_{\beta \in \mathcal{A}} \left(c_{\beta}(X_{m+1}^{m,x}) - c_{\beta}(X_{m+1}^{\Delta,m,x}) \right) I_{\beta,m+2} \right] = \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}), \tag{A.10a}
\end{aligned}$$

$$\begin{aligned}
\mathbb{E} \left[\left(X_{m+2}^{m,x} - X_{m+2}^{\Delta,m,x} \right) \Delta \omega_{m+1} \right] &= \mathbb{E} \left[\sum_{\beta \in \mathcal{B}(\mathcal{A})} I_{\beta}[c_{\beta}(X_{m+1}^{m,x})]_{m+2} \Delta \omega_{m+1} \right] \\
&\quad + \mathbb{E} \left[\sum_{\beta \in \mathcal{A}} \left(c_{\beta}(X_{m+1}^{m,x}) - c_{\beta}(X_{m+1}^{\Delta,m,x}) \right) I_{\beta,m+2} \Delta \omega_{m+1} \right] = \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}), \tag{A.10b}
\end{aligned}$$

$$\begin{aligned}
\mathbb{E} \left[\left(X_{m+2}^{m,x} - X_{m+2}^{\Delta,m,x} \right) \Delta \omega_{m+2} \right] &= \mathbb{E} \left[\sum_{\beta \in \mathcal{B}(\mathcal{A})} I_{\beta}[c_{\beta}(X_{m+1}^{m,x})]_{m+2} \Delta \omega_{m+2} \right] \\
&\quad + \mathbb{E} \left[\sum_{\beta \in \mathcal{A}} \left(c_{\beta}(X_{m+1}^{m,x}) - c_{\beta}(X_{m+1}^{\Delta,m,x}) \right) I_{\beta,m+2} \Delta \omega_{m+2} \right] = \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}). \tag{A.10c}
\end{aligned}$$

Furthermore,

$$\frac{2}{\Delta t} \mathbb{E} \left[\left(X_{m+2}^{m,x} - X_{m+2}^{\Delta,m,x} \right) (\Delta \omega_{m+1} - \Delta \omega_{m+2}) \right] = \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}). \tag{A.11}$$

For the above result we divide the expected value into three parts with equation (A.9). Firstly

$$\begin{aligned}
&\frac{2}{\Delta t} \mathbb{E} \left[\sum_{\beta \in \mathcal{B}(\mathcal{A})} I_{\beta}[c_{\beta}(X_{m+1}^{m,x})]_{m+2} \Delta \omega_{m+1} \right] \\
&= \frac{2}{\Delta t} \mathbb{E} \left[\left(\sum_{\beta \in \mathcal{B}(\mathcal{A})} c_{\beta}(X_{m+1}^{m,x}) I_{\beta,m+2} + \sum_{\beta \in \mathcal{B}(\mathcal{A})} I_{(0,\beta)}[c_{(0,\beta)}(X_{m+1}^{m,x})]_{m+2} + \sum_{\beta \in \mathcal{B}(\mathcal{A})} I_{(1,\beta)}[c_{(1,\beta)}(X_{m+1}^{m,x})]_{m+2} \right) \Delta \omega_{m+1} \right] \\
&= \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}), \tag{A.12}
\end{aligned}$$

secondly

$$\frac{2}{\Delta t} \mathbb{E} \left[\sum_{\beta \in \mathcal{A} \setminus \alpha_{\emptyset}} \left(c_{\beta}(X_{m+1}^{m,x}) - c_{\beta}(X_{m+1}^{\Delta,m,x}) \right) I_{\beta,m+2} (\Delta \omega_{m+1} - \Delta \omega_{m+2}) \right] = \mathcal{O}((\Delta t)^{\gamma_w^{2T}+1}), \tag{A.13}$$

and thirdly

$$\begin{aligned}
&\frac{2}{\Delta t} \mathbb{E} \left[\sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_{\alpha}[c_{\alpha}(X_{m+1}^{m,x})]_{m+1} \Delta \omega_{m+1} \right] - \frac{2}{\Delta t} \mathbb{E} \left[\sum_{\beta \in \mathcal{B}(\mathcal{A})} I_{\beta}[c_{\beta}(X_{m+1}^{m,x})]_{m+2} \Delta \omega_{m+2} \right] \\
&= \frac{2}{\Delta t} \mathbb{E} \left[\left(\sum_{\alpha \in \mathcal{B}(\mathcal{A})} c_{\alpha}(X) I_{\alpha,m+1} + I_{(0,\alpha)}[c_{(0,\alpha)}(X_{m+1}^{m,x})]_{m+1} + I_{(1,\alpha)}[c_{(1,\alpha)}(X_{m+1}^{m,x})]_{m+1} \right) \Delta \omega_{m+1} \right]
\end{aligned}$$

$$\begin{aligned}
& - \frac{2}{\Delta t} \mathbb{E} \left[\left(\sum_{\beta \in \mathcal{B}(\mathcal{A})} c_{\beta}(x) I_{\beta, m+2} + I_0 [c_{(0, \beta)}(X_{m+1}^{m, x})]_{m+1} I_{\beta, m+2} + I_1 [c_{(1, \beta)}(X_{m+1}^{m, x})]_{m+1} I_{\beta, m+2} \right. \right. \\
& \left. \left. + I_{(0, \beta)} [c_{(0, \beta)}(X_{m+2}^{m, x})]_{m+2} + I_{(1, \beta)} [c_{(1, \beta)}(X_{m+2}^{m, x})]_{m+2} \right) \Delta \omega_{m+2} \right] \\
& = \mathcal{O}((\Delta t)^{\gamma_w^{2T} + 1}).
\end{aligned} \tag{A.14}$$

Note that in the above equation the terms $c_{\alpha}(x) I_{\alpha, m+1}$ and $c_{\beta}(x) I_{\beta, m+2}$ cancel out.

The Taylor series of function $h(t_{m+2}, x)$ in x gives

$$h(t_{m+2}, X_{m+2}^{m, x}) - h(t_{m+2}, X_{m+2}^{\Delta, m, x}) = (-1)^{\ell+1} \sum_{\ell=1}^{\infty} \frac{1}{\ell!} h^{(\ell)}(t_{m+2}, X_{m+2}^{m, x}) \left(X_{m+2}^{m, x} - X_{m+2}^{\Delta, m, x} \right)^{\ell}. \tag{A.15}$$

By using the strong error result

$$\mathbb{E} \left[\left| X_{m+2}^{m, x} - X_{m+2}^{\Delta, m, x} \right|^{\ell} \right] = \mathcal{O}((\Delta t)^{1.5\ell}) \tag{A.16}$$

for $\ell \geq 2$ and formula (4.2) for $\ell = 1$, we find

$$\mathbb{E} \left[h(t_{m+2}, X_{m+2}^{m, x}) - h(t_{m+2}, X_{m+2}^{\Delta, m, x}) \right] = \mathcal{O}((\Delta t)^{\gamma_w^{2T} + 1}), \tag{A.17a}$$

$$\mathbb{E} \left[\left(h(t_{m+2}, X_{m+2}^{m, x}) - h(t_{m+2}, X_{m+2}^{\Delta, m, x}) \right) \Delta \omega_{m+1} \right] = \mathcal{O}((\Delta t)^{\gamma_w^{2T} + 1}), \tag{A.17b}$$

$$\mathbb{E} \left[\left(h(t_{m+2}, X_{m+2}^{m, x}) - h(t_{m+2}, X_{m+2}^{\Delta, m, x}) \right) \Delta \omega_{m+2} \right] = \mathcal{O}((\Delta t)^{\gamma_w^{2T} + 1}), \tag{A.17c}$$

$$\frac{2}{\Delta t} \mathbb{E} \left[\left(h(t_{m+2}, X_{m+2}^{m, x}) - h(t_{m+2}, X_{m+2}^{\Delta, m, x}) \right) (\Delta \omega_{m+1} - \Delta \omega_{m+2}) \right] = \mathcal{O}((\Delta t)^{\gamma_w^{2T} + 1}). \tag{A.17d}$$

Remark A.1. For the Euler scheme we have

$$\mathcal{A} = \{v, (0), (1)\} \quad \text{and} \quad \mathcal{B}(\mathcal{A}) = \{(1, 1), (1, 0), (0, 1), (0, 0)\} \tag{A.18}$$

and for the Milstein scheme

$$\mathcal{A} = \{v, (0), (1), (1, 1)\} \quad \text{and} \quad \mathcal{B}(\mathcal{A}) = \{(1, 0), (0, 1), (0, 0), (1, 1, 1), (0, 1, 1)\}. \tag{A.19}$$

For the Euler and Milstein scheme the weak convergence rate $\gamma_w^{2T} = 2$ is replaced by their weak convergence rates $\gamma_w^E = 1$ and $\gamma_w^M = 1$, respectively.

References

- [1] C. Bender, J. Steiner, Least-squares Monte Carlo for backward SDEs, in: Numerical Methods in Finance, in: Springer Proceedings in Mathematics, vol. 12, 2012, pp. 257–289.
- [2] B. Bouchard, R. Elie, Discrete-time approximation of decoupled forward–backward SDE with jumps, Stoch. Process. Appl. 118 (1) (2008) 53–75.
- [3] P. Briand, C. Labart, Simulation of BSDEs by Wiener chaos expansion, HAL: hal-00688523, version 2 (2013).
- [4] J.-F. Chassagneux, Linear multistep schemes for BSDEs, SIAM J. Numer. Anal. 52 (6) (2014) 2815–2836.
- [5] J.-F. Chassagneux, D. Crisan, Runge–Kutta schemes for backward stochastic differential equations, Ann. Appl. Probab. 24 (2) (2014) 679–720.
- [6] S. Crépey, Bilateral counterparty risk under funding constraints part I: Pricing, mathematical finance.
- [7] D. Duffie, D. Filipović, W. Schachermayer, Affine processes and applications in finance, Ann. Appl. Probab. 13 (3) (2003) 984–1053.
- [8] N. El Karoui, S. Peng, M.C. Quenez, Backward stochastic differential equations in finance, Math. Finance 7 (1) (1997) 1–71.
- [9] F. Fang, C.W. Oosterlee, A novel pricing method for European options based on Fourier-cosine series expansions, SIAM J. Sci. Comput. 31 (2) (2008) 826–848.
- [10] F. Fang, C.W. Oosterlee, Pricing early-exercise and discrete barrier options by Fourier-cosine series expansions, Numer. Math. 114 (1) (2009) 27–62.
- [11] F. Fang, C.W. Oosterlee, A Fourier-based valuation method for Bermudan and barrier options under Heston’s model, SIAM J. Financ. Math. 2 (1) (2011) 439–463.
- [12] E. Gobet, J.P. Lemor, X. Warin, A regression-based Monte Carlo method to solve backward stochastic differential equations, Ann. Appl. Probab. 15 (3) (2005) 2172–2202.
- [13] J.C. Hull, Options, Futures and Other Derivatives, Pearson/Prentice-Hall, New Jersey, 2009.
- [14] C.B. Hyndman, P. Oyono Ngou, A convolution method for numerical solution of backward stochastic differential equations, arXiv:1304.1783, 2014.
- [15] A. Iserles, A First Course in the Numerical Analysis of Differential Equations, Cambridge Texts in Applied Mathematics, Cambridge University Press, 2009.
- [16] M. Jeanblanc, M. Yor, M. Chesney, Mathematical Methods for Financial Markets, Springer Finance, Springer, Berlin, 2009.
- [17] P.E. Kloeden, E. Platen, Numerical Solution of Stochastic Differential Equations, Applications of Mathematics Series, Springer, 1992.
- [18] J.P. Lemor, E. Gobet, X. Warin, Rate of convergence of an empirical regression method for solving generalized backward stochastic differential equations, Bernoulli 12 (5) (2006) 889–916.
- [19] A.L. Lewis, Option Valuation Under Stochastic Volatility: With Mathematica Code, Finance Press, 2000.
- [20] J. Ma, J. Shen, Y. Zhao, On numerical approximations of forward–backward stochastic differential equations, SIAM J. Numer. Anal. 46 (5) (2008) 2636–2661.
- [21] G.N. Milstein, M.V. Tretyakov, Numerical algorithms for forward–backward stochastic differential equations, SIAM J. Sci. Comput. 28 (2) (2006) 561–582.
- [22] E. Pardoux, S.G. Peng, Adapted solution of a backward stochastic differential equation, Syst. Control Lett. 14 (1) (1990) 55–61.

- [23] E. Pardoux, S. Peng, Backward stochastic differential equations and quasilinear parabolic partial differential equations, in: *Stochastic Partial Differential Equations and Their Applications*, in: *Lecture Notes in Control and Information Sciences*, vol. 176, Springer, Berlin, 1992, pp. 200–217.
- [24] A. Pascucci, *PDE and Martingale Methods in Option Pricing*, Bocconi & Springer Series, Springer, 2011.
- [25] H. Pham, *Continuous-Time Stochastic Control and Optimization with Financial Applications*, Springer-Verlag, Berlin, 2009.
- [26] M.J. Ruijter, C.W. Oosterlee, A Fourier cosine method for an efficient computation of solutions to BSDEs, *SIAM J. Sci. Comput.* 37 (2) (2015) A859–A889.
- [27] S.E. Shreve, *Stochastic Calculus for Finance II: Continuous-Time Models*, Springer Finance, Springer, Berlin, 2010.
- [28] 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.
- [29] W. Zhao, W. Zhang, L. Ju, A numerical method and its error estimates for the decoupled forward–backward stochastic differential equations, *Commun. Comput. Phys.* 15 (3) (2014) 618–646.