Bergische Universität Wuppertal

Fakultät für Mathematik und Naturwissenschaften

Institute of Mathematical Modelling, Analysis and Computational Mathematics (IMACM)

Preprint BUW-IMACM 20/35

Long Teng and Weidong Zhao

High-order combined Multi-step Scheme for solving forward Backward Stochastic Differential Equations

http://www.math.uni-wuppertal.de

High-order combined Multi-step Scheme for solving forward Backward Stochastic Differential Equations

Long Teng[1,*] and Weidong Zhao[2]

[1]Lehrstuhl für Angewandte Mathematik und Numerische Analysis,
Fakultät für Mathematik und Naturwissenschaften,
Bergische Universität Wuppertal, Gaußstr. 20, 42119 Wuppertal, Germany
[2]School of Mathematics & Finance Institute,
Shandong University, Jinan 250100, China(wdzhao@sdu.edu.cn)

Abstract

In this work, in order to obtain higher-order schemes for solving forward backward stochastic differential equations, we adopt the high-order multi-step method in [W. Zhao, Y. Fu and T. Zhou, SIAM J. Sci. Comput., 36(4) (2014), pp.A1731-A1751] by combining multi-steps. Two reference ordinary differential equations containing the conditional expectations and their derivatives are derived from the backward component. These derivatives are approximated by finite difference methods with multi-step combinations. The resulting scheme is a semi-discretization in the time direction involving conditional expectations, which are solved by using the Gaussian quadrature rules and polynomial interpolations on the spatial grids. Our new proposed multi-step scheme allows for higher convergence rate up to ninth order, and are more efficient. Finally, we provide a numerical illustration of the convergence of the proposed method.

Keywords *forward backward stochastic differential equations, multi-step scheme, finite difference method, time-space grid, Gauss-Hermite quadrature rule*

1 Introduction

Recently, the forward-backward stochastic differential equation (FBSDE) becomes an important tool for formulating many problems in various areas including physics and financial mathematics. We are interested in the numerical approximation of the general FBSDEs

$$\begin{cases} dX_t = a(t, X_t, Y_t, Z_t)\,dt + b(t, X_t, Y_t, Z_t)\,dW_t, \quad X_0 = x_0, & \text{forward component} \\ -dY_t = f(t, X_t, Y_t, Z_t)\,dt - Z_t\,dW_t, & \text{backward component} \\ Y_T = \xi = g(X_T) \end{cases} \tag{1}$$

on a filtered complete probability space $(\Omega, \mathcal{F}, P)$ with the natural filtration $(\mathcal{F}_t)_{0\le t\le T}$, where $a : [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^n$ and $b : [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^{n\times d}$, are drift and diffusion coefficients in the forward component, respectively; $W_t = (W_t^1, \cdots, W_t^d)^T$ is a d-dimensional Brownian motion (all Brownian motions are independent with each other); $f(t, X_t, Y_t, Z_t) : [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^m$ is the driver function and ξ is the square-integrable terminal condition. We see that the terminal condition Y_T depends on final value of the forward component. Note that a, b and f are all $\mathcal{F}_t$-adapted, and a triple (X_t, Y_t, Z_t) is called an L^2-adapted solution of (1) if it is $\mathcal{F}_t$-adapted, square integrable, and satisfies

$$\begin{cases} X_t = X_0 + \int_0^t a(s, X_s, Y_s, Z_s)\,ds + \int_0^t b(s, X_s, Y_s, Z_s)\,dW_s, \\ Y_t = \xi + \int_t^T f(s, X_s, Y_s, Z_s)\,ds - \int_t^T Z_s\,dW_s. \end{cases} \tag{2}$$

*Corresponding author (teng@math.uni-wuppertal.de)

One obtains decoupled FBSDEs if a and b are independent with Y_t and Z_t in (1), which become backward stochastic differential equations (BSDEs) when $a = 0$ and $b = 1$.
The existence and uniqueness of solution of the BSDEs assuming the Lipschitz conditions on f, a, b and g are proven by Pardoux and Peng [Pardoux and Peng, 1990, Pardoux and Peng, 1992]. The uniqueness of solution is extended under more general assumptions for f in [Lepeltier and Martin, 1997], but only in the one-dimensional case. The existence and uniqueness of solution of FBSDEs have been studied in [Ma et al., 1994, Peng and Wu, 1999].
In recent years, many numerical methods have been proposed for the BSDEs and FBSDEs. We list some of them here: [Bender and Zhang, 2008, Bender and Steiner, 2012, Bouchard and Touzi, 2004, Crisan and Manolarakis, 2010, Cvitanic and Zhang, 2006, Delarue and Menozzi, 2006, Douglas et al., 1996, Fu et al., 2017, Gobet et al., 2005, Lemor et al., 2006, Ma and Zhang, 2005, Ma et al., 2008, Milsetin and Tretyakov, 2006, Ruijter and Oosterlee, 2015, Teng, 2019, Teng et al., 2020, Zhao et al., 2006, Zhao et al., 2009, Zhao et al., 2010, Zhao et al., 2012, Zhao et al., 2013, Zhao et al., 2014b, Zhang, 2004, Zhang et al., 2013], and many others. In this literature, the high-order methods rely on the high-order approaches for both the forward and backward components, where are clearly difficult and computationally expensive to achieve.
Moreover, Zhao et al. proposed in [Zhao et al., 2014a] new kinds of high-order multi-step schemes for FBSDEs, which can keep high-order accuracy while using the Euler method to solve the forward component. This is of great interesting since the use of Euler method can dramatically simplify the entire computations. However, the convergence rate is restricted to sixth order, since the stability condition cannot be satisfied for a higher order. For this reason, we adopt in this work this method by combining some multi-steps to achieve higher rate of convergence. More precisely, we derive two reference ordinary differential equations (ODEs) which contain the conditional expectations and their derivatives. We approximate these derivatives using finite difference methods with the combination of multi-steps for a better stability. The resulting conditional expectations are solved using the Gaussian quadrature rules, whereas the first component, X_t is solved using the Euler method thanks to the local property of the generator of diffusion processes. FBSDEs are numerically solved on the time-space grids. Numerical experiments are presented to demonstrate the improvement in the rate of convergence.
In the next section, we start with preliminaries on FBSDEs and derive in Section 3 the approximations of derivatives using finite difference method with combined multi-steps. In Section 4, we derive the reference ODEs, based on which the semi-discrete higher-order multi-step schemes are introduced for solving decoupled FBSDEs. Section 5 is devoted to the fully discrete higher-order schemes. In Section 6, these methods are extended to solve a coupled FBSDE. In Section 7, several numerical experiments on the decoupled and coupled FBSDEs including two-dimensional applications are provided to show the higher efficiency and accuracy. Finally, Section 8 concludes this work.

2 Preliminaries

As mentioned before, throughout the paper we assume that $(\Omega, \mathcal{F}, P)$ is a complete, filtered probability space. A standard d-dimensional Brownian motion W_t with a finite terminal time T is defined, and the first component, X_t generates the filtration $\mathcal{F}_t = \sigma\{X_s, 0 \le s \le t\}$. And the usual hypotheses should be satisfied. We denote the set of all $\mathcal{F}_t$-adapted and square integrable processes in $\mathbb{R}^d$ with $L^2 = L^2(0, T; \mathbb{R}^d)$, and list following notation to be used:

- $|\cdot|$: the Euclidean norm in $\mathbb{R}$, $\mathbb{R}^n$ and $\mathbb{R}^{n\times d}$;
- $\mathcal{F}_t^{s,x}$: σ-algebra generated by the diffusion process $\{X_r, s \le r \le t, X_s = x\}$;

- $\mathbb{E}_t^{s,x}[\cdot]$: conditional expectation under $\mathcal{F}_t^{s,x}$, i.e., $\mathbb{E}_t^{s,x}[\cdot|\mathcal{F}_t^{s,x}]$;
- C_b^k : the set of continuous functions with uniformly bounded derivatives up to order k;
- C^{k_1,k_2} : the set of functions with continuous partial derivatives $\frac{\partial}{\partial t}$ and $\frac{\partial}{\partial x}$ up to k_1 and k_2, respectively;
- C_L : the set of uniformly Lipschitz continuous function with respect to the spatial variables;
- $C_L^{\frac{1}{2}}$: the subset of C_L such that its element is Hölder-$\frac{1}{2}$ continuous with respect to time, with uniformly bounded Lipschitz and Hölder constants.

Let X_t be a diffusion process

$$X_t = x_0 + \int_0^t a(s, X_s)\, ds + \int_0^t b(s, X_s)\, dW_s \tag{3}$$

starting at (t_0, x_0) and $t \in [t_0, T]$, which has a unique solution. Note that $\mathbb{E}_s^x[X_t] := \mathbb{E}_s^{s,x}[X_t]$ is equal to $\mathbb{E}[X_t|X_s = x]$ for all $s \le t$ with the Markov property of the diffusion process. Given a measurable function $g : [0,T] \times \mathbb{R}^n \to \mathbb{R}$, $\mathbb{E}_s^x[g(t, X_t)]$ is a function of (t, s, x) whose partial derivative with respect to t reads

$$\frac{\partial \mathbb{E}_s^x[g(t, X_t)]}{\partial t} = \lim_{\tau \to 0^+} \frac{\mathbb{E}_s^x[g(t+\tau, X_{t+\tau})] - \mathbb{E}_s^x[g(t, X_t)]}{\tau}$$

provided that the limit exists and is finite.

Definition 2.1 (Generator)**.** *The generator $\mathcal{A}_t^x$ of X_t satisfying* (3) *on a measurable function $g : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ is defined by*

$$\mathcal{A}_t^x g(t,x) = \lim_{h \to 0^+} \frac{\mathbb{E}_t^x[g(t+h, X_{t+h})] - g(t,x)}{h}, \quad x \in \mathbb{R}^n.$$

Theorem 2.1. *Let X_t be the diffusion process defined by* (3)*, then it holds*

$$\mathcal{A}_t^x f(t,x) = \mathcal{L}_{t,x} f(t,x) \tag{4}$$

for $f \in C^{1,2}([0,T] \times \mathbb{R}^n)$ with

$$\mathcal{L}_{t,x} = \frac{\partial}{\partial t} + \sum_i a_i(t,x) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j} (bb^\top)_{i,j}(t,x) \frac{\partial^2}{\partial x_i \partial x_j}.$$

The proof can be simply completed by using the Itô's lemma and the dominated convergence theorem.

Remark 2.1. *From* (4) *one can straightforwardly deduce that*

$$\mathcal{A}_t^{X_t} f(t, X_t) = \mathcal{L}_{t,X_t} f(t, X_t),$$

which is a stochastic process.

By using the Itô's lemma and Theorem 2.1 we calculate

$$\left.\frac{d\, \mathbb{E}_{t_0}^{x_0}[f(t, X_t)]}{dt}\right|_{t=t_0} = \lim_{t \to t_0^+} \frac{\mathbb{E}_{t_0}^{x_0}[f(t, X_t)] - g(t_0, x_0)}{t - t_0} = \mathcal{L}_{t,x} f(t_0, x_0) = \mathcal{A}_t^x f(t_0, x_0),$$

from which we deduce Theorem (2.2) as follows.

Theorem 2.2. *Assume that $f \in C^{1,2}([0,T]\times\mathbb{R}^n)$ and $\mathbb{E}_{t_0}^{x_0}[|\mathcal{L}_{t,X_t}f(t,X_t)|]<\infty$, let $t_0<t$ be a fixed time, and $x_0\in\mathbb{R}^n$ be a fixed space point, it holds that*

$$\frac{d\,\mathbb{E}_{t_0}^{x_0}[f(t,X_t)]}{dt}=\mathbb{E}_{t_0}^{x_0}\left[\mathcal{A}_t^{X_t}f(t,X_t)\right],\quad t\geq t_0.$$

Furthermore, one has the following identity

$$\left.\frac{d\,\mathbb{E}_{t_0}^{x_0}[f(t,X_t)]}{dt}\right|_{t=t_0}=\left.\frac{d\,\mathbb{E}_{t_0}^{x_0}[f(t,\tilde{X}_t)]}{dt}\right|_{t=t_0},\tag{5}$$

where $\tilde{X}_t$ is an approximating diffusion process defined by

$$\tilde{X}_t=x_0+\int_0^t\tilde{a}\,ds+\int_0^t\tilde{b}\,dW_s$$

with the smooth functions $\tilde{a}_t=\tilde{a}(t,\tilde{X}_t;t_0,x_0)$ and $\tilde{b}_t=\tilde{b}(t,\tilde{X}_t;t_0,x_0)$ of $(t,\tilde{X}_t)$ with the parameter (t_0,x_0) satisfying

$$\tilde{a}(t_0,\tilde{X}_{t_0};t_0,x_0)=a(t_0,x_0)\ \text{and}\ \tilde{b}(t_0,\tilde{X}_{t_0};t_0,x_0)=b(t_0,x_0)$$

It has been noted in [Zhao et al., 2014a] that the different approximations of (5) can be obtained by choosing different $\tilde{a}_t$'s and $\tilde{b}_t$'s. One can simply e.g., choose $\tilde{a}(s,\tilde{X}_s;t_0,x_0)=a(t_0,x_0)$ and $\tilde{b}(s,\tilde{X}_s;t_0,x_0)=b(t_0,x_0)$ for all $s\in[t_0,t]$.
For existence, regularity and representation of solutions of decoupled FBSDEs we refer to [Ma and Zhang, 2005, Peng, 1991, Zhang, 2001]. In the following of this section we will present some of those. We denote the forward stochastic differential equation (SDE) starting from (s,x) with $X_t^{s,x}$ and consider the decoupled FBSDEs

$$\begin{cases}X_t^{s,x}=x+\int_s^t a(r,X_r^{s,x})\,ds+\int_s^t b(r,X_r^{s,x})\,dW_r,\\ Y_t^{s,x}=g(X_T^{s,x})+\int_t^T f(r,X_r^{s,x},Y_r^{s,x},Z_r^{s,x})\,dr-\int_t^T Z_r^{s,x}\,dW_r,\end{cases}\tag{6}$$

where $t\in[s,T]$, and the superscript ^{s,x} will be omitted when the context is clear.
Throughout the paper, we shall often make use of the following standing assumptions:

1. The functions $a,b\in C_b^1$, and assume

$$\sup_{0\leq t\leq T}\{|a(t,0)|+|b(t,0)|\}\leq L,$$

   where the common constant $L>0$ denotes all the Lipschitz constants.

2. $n=d$ and we assume that b satisfies

$$b(t,x)b^\top(t,x)\geq\frac{1}{L}I_n,\quad\forall(t,x)\in[0,T]\times\mathbb{R}^n.$$

3. $a,b,f,g\in C_L$, and assume that

$$\sup_{0\leq t\leq T}|f(t,0,0,0)|+|g(0)|\leq L,$$

   where L denotes all the Lipschitz constants.

4. $a,b,f\in C_L^{\frac{1}{2}}$.

Under the above conditions, it is clear that (6) is well-posed; the resulting integrands by taking conditional expectation on both side of the backward component is continuous with respect to time; the nonlinear Feynman-Kac formula [Ma and Zhang, 2005, Peng, 1991] can be given as follows.

Theorem 2.3. *Let $u \in C^{1,2}([0,T] \times \mathbb{R}^n)$ be a classical solution to the following PDE*

$$\mathcal{L}_{t,x} u(t,x) + f(t,x,u(t,x),\nabla u(t,x) b(t,x)) = 0, \quad u(T,x) = g(x),$$

then $Y_t^{s,x} = u(t, X_t^{s,x})$, $Z_t^{s,x} = \nabla_x u(t, X_t^{s,x}) b(t, X_t^{s,x})$, $\forall t \in (s,T]$ is the unique solution to (6).

3 Calculation of the weights in the FDM for approximating derivative

In this section we calculate the weights in the FDM for approximating the function derivatives, e.g., $\frac{du(t)}{dt}$. Let $u(t) \in C_b^{k+1}$, k is a positive integer, and $t_i = i\Delta t$, i.e., $t_0 < t_1 < \cdots < t_k$.

3.1 Combination of two time points

We consider the Taylor's expansions of $u(t_i)$ and $u(t_{i+1}), i = 0, \cdots, k$

$$\begin{cases} u(t_i) = \sum_{j=0}^{k} \frac{(\Delta t_i)^j}{j!} \frac{d^j u}{dt^j}(t_0) + \mathcal{O}(\Delta t_i)^{k+1} \\ u(t_{i+1}) = \sum_{j=0}^{k} \frac{(\Delta t_{i+1})^j}{j!} \frac{d^j u}{dt^j}(t_0) + \mathcal{O}(\Delta t_{i+1})^{k+1}, \end{cases}$$

from which we can deduce

$$\begin{cases} \sum_{i=0}^{k} \alpha_{k,i} u(t_i) = \sum_{j=0}^{k} \frac{\sum_{i=0}^{k} \alpha_{k,i} (\Delta t_i)^j}{j!} \frac{d^j u}{dt^j}(t_0) + \mathcal{O}\left(\sum_{i=0}^{k} \alpha_{k,i} (\Delta t_i)^{k+1} \right) \\ \sum_{i=0}^{k} \alpha_{k,i} u(t_{i+1}) = \sum_{j=0}^{k} \frac{\sum_{i=0}^{k} \alpha_{k,i} (\Delta t_{i+1})^j}{j!} \frac{d^j u}{dt^j}(t_0) + \mathcal{O}\left(\sum_{i=0}^{k} \alpha_{k,i} (\Delta t_{i+1})^{k+1} \right), \end{cases}$$

where $\alpha_{k,i}, i = 0, 1, \cdots, k$ are real numbers. Clearly, we obtain

$$\sum_{i=0}^{k} \alpha_{k,i} (u(t_i) + u(t_{i+1})) = \sum_{j=0}^{k} \frac{\sum_{i=0}^{k} \alpha_{k,i} \left((\Delta t_i)^j + (\Delta t_{i+1})^j \right)}{j!} \frac{d^j u}{dt^j}(t_0) + \mathcal{O}\left(\sum_{i=0}^{k} \alpha_{k,i} \left((\Delta t_i)^{k+1} + (\Delta t_{i+1})^{k+1} \right) \right)$$

and thus

$$\frac{du}{dt}(t_0) = \sum_{i=0}^{k} \alpha_{k,i} \left(u(t_i) + u(t_{i+1}) \right) + \mathcal{O}\left(\sum_{i=0}^{k} \alpha_{k,i} \left((\Delta t_i)^{k+1} + (\Delta t_{i+1})^{k+1} \right) \right) \tag{7}$$

by choosing

$$\frac{\sum_{i=0}^{k} \alpha_{k,i} \left((\Delta t_i)^j + (\Delta t_{i+1})^j \right)}{j!} = \begin{cases} 1, & j = 1, \\ 0, & j \neq 1. \end{cases} \tag{8}$$

Due to $t_i = i\Delta t$ and $t_{i+1} = (i+1)\Delta t$ the conditions in (8) are equivalent to the following system:

$$\begin{bmatrix} 2 & 2 & 2 & \dots & 2 \\ 1 & 3 & 5 & \dots & k+(k+1) \\ 1 & 5 & 13 & \dots & k^2+(k+1)^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1^k+2^k & 2^k+3^k & \dots & k^k+(k+1)^k \end{bmatrix} \times \begin{bmatrix} \alpha_{k,0}\Delta t \\ \alpha_{k,1}\Delta t \\ \alpha_{k,2}\Delta t \\ \vdots \\ \alpha_{k,k}\Delta t \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

which can be solved for $\alpha_{k,i}\Delta t, i = 0, \cdots, k$. We refer to the algorithm proposed in [Fornberg, 1988] for those solutions. We report $\alpha_{k,i}\Delta t$ for $k = 1, 2, \cdots, 7$ in Table 1, since the related multi-step schemes proposed in this paper is unstable from $k = 8$, which will be explained below. The multi-step schemes (combining two time points) can be constructed by

$\alpha_{k,i}\Delta t$	$i=0$	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$
$k=1$	$-\frac{1}{2}$	$\frac{1}{2}$						
$k=2$	-1	$\frac{3}{2}$	$-\frac{1}{2}$					
$k=3$	$-\frac{17}{12}$	$\frac{11}{4}$	$-\frac{7}{4}$	$\frac{5}{12}$				
$k=4$	$-\frac{7}{4}$	$\frac{49}{12}$	$-\frac{15}{4}$	$\frac{7}{4}$	$-\frac{1}{3}$			
$k=5$	$-\frac{121}{60}$	$\frac{65}{12}$	$-\frac{77}{12}$	$\frac{53}{12}$	$-\frac{5}{3}$	$\frac{4}{15}$		
$k=6$	$-\frac{67}{30}$	$\frac{403}{60}$	$-\frac{29}{3}$	$\frac{35}{4}$	$-\frac{59}{12}$	$\frac{47}{30}$	$-\frac{13}{60}$	
$k=7$	$-\frac{2027}{840}$	$\frac{319}{40}$	$-\frac{1613}{120}$	$\frac{361}{24}$	$-\frac{269}{24}$	$\frac{641}{120}$	$-\frac{59}{40}$	$\frac{151}{840}$

Table 1: The values of $\alpha_{k,i}\Delta t$ for combining two time points.

approximating the reference ODEs (see Sec. 4.1) using (7). Therefore, we consider the following ODE

$$\frac{Y(t)}{dt} = f(t, Y(t)), \quad t \in [0, T) \tag{9}$$

with the known terminal condition $Y(T)$ for studying stability, see also [Zhao et al., 2014a]. Applying (7) to (9) one obtain the multi-step scheme as

$$\alpha_{k,0} Y^n + \sum_{j=1}^{k} \left(\alpha_{k,j-1} + \alpha_{k,j} \right) Y^{n+j} + \alpha_{k,k} Y^{n+k+1} = f(t_n, Y^n) \tag{10}$$

under the uniform time partition $0 = t_0 < t_1 < \cdots < t_N = T$. (10) is stable if the roots $\{\lambda_{k,j}\}_{j=1}^{k}$ of the characteristic equation

$$P(\lambda) = \alpha_{k,0} \lambda^{k+1} + \sum_{j=1}^{k} \left(\alpha_{k,j-1} + \alpha_{k,j} \right) \lambda^{k+1-j} + \alpha_{k,k} \lambda^0 \tag{11}$$

satisfies the following root conditions [Butcher, 2008]

- $|\lambda_{k,j}| \leq 1$,
- $P'(\lambda_{k,j}) \neq 0$ if $|\lambda_{k,j}| = 1$ (simple roots).

k	2	3	4	5	6	7	8
$\max(\lvert\lambda_{k,j}\rvert)$	0.5000	0.5424	0.6344	0.7438	0.8636	0.9915	1.1264

Table 2: The maximum absolute root of (11) except the simple roots

With the $\alpha_{k,j}$ in Tabel 1, 1 is the simple root of the latter characteristic function for each k, except which we list the maximum absolute values of the roots for $k = 2, \cdots, 8$ in Table 2, from which we see that the multi-step scheme (10) is unstable for $k \geq 8$. However, compared to the multi-step scheme proposed in [Zhao et al., 2014a](unstable ≥ 7), stability for $k = 7$ has been achieved, i.e., 1-order higher convergence rate is obtained. Combination of more time points can be done similarly, and provide other multi-step schemes, which have different instabilities. In our investigation we find that the multi-step scheme resulted by combining four time points are stable for $k \leq 9$, which is the best. Thus, we show its detailed derivation in next subsection and will consider it in the numerical experiments.

3.2 Combination of four time points

Similarly but slightly different to the multi-step scheme in Section 3.1, we need to consider the Taylor's expansions of $u(t_i)$, $u(t_{i+1})$, $u(t_{i+2})$ and $u(t_{i+3}), i = 0, \cdots, k$,

$$\begin{cases} u(t_i) = \sum_{j=0}^{k} \frac{(\Delta t_i)^j}{j!} \frac{d^j u}{dt^j}(t_0) + \mathcal{O}(\Delta t_i)^{k+1}, \\ u(t_{i+1}) = \sum_{j=0}^{k} \frac{(\Delta t_{i+1})^j}{j!} \frac{d^j u}{dt^j}(t_0) + \mathcal{O}(\Delta t_{i+1})^{k+1}, \\ u(t_{i+2}) = \sum_{j=0}^{k} \frac{(\Delta t_{i+2})^j}{j!} \frac{d^j u}{dt^j}(t_0) + \mathcal{O}(\Delta t_{i+2})^{k+1}, \\ u(t_{i+3}) = \sum_{j=0}^{k} \frac{(\Delta t_{i+3})^j}{j!} \frac{d^j u}{dt^j}(t_0) + \mathcal{O}(\Delta t_{i+3})^{k+1}, \end{cases}$$

from which we can deduce

$$\begin{cases} \sum_{i=0}^{k} \alpha_{k,i} u(t_i) = \sum_{j=0}^{k} \frac{\sum_{i=0}^{k} \alpha_{k,i} (\Delta t_i)^j}{j!} \frac{d^j u}{dt^j}(t_0) + \mathcal{O}\left(\sum_{i=0}^{k} \alpha_{k,i} (\Delta t_i)^{k+1} \right), \\ \sum_{i=0}^{k} \alpha_{k,i} u(t_{i+1}) = \sum_{j=0}^{k} \frac{\sum_{i=0}^{k} \alpha_{k,i} (\Delta t_{i+1})^j}{j!} \frac{d^j u}{dt^j}(t_0) + \mathcal{O}\left(\sum_{i=0}^{k} \alpha_{k,i} (\Delta t_{i+1})^{k+1} \right), \\ \sum_{i=0}^{k} \alpha_{k,i} u(t_{i+2}) = \sum_{j=0}^{k} \frac{\sum_{i=0}^{k} \alpha_{k,i} (\Delta t_{i+2})^j}{j!} \frac{d^j u}{dt^j}(t_0) + \mathcal{O}\left(\sum_{i=0}^{k} \alpha_{k,i} (\Delta t_{i+2})^{k+1} \right), \\ \sum_{i=0}^{k} \alpha_{k,i} u(t_{i+3}) = \sum_{j=0}^{k} \frac{\sum_{i=0}^{k} \alpha_{k,i} (\Delta t_{i+3})^j}{j!} \frac{d^j u}{dt^j}(t_0) + \mathcal{O}\left(\sum_{i=0}^{k} \alpha_{k,i} (\Delta t_{i+3})^{k+1} \right), \end{cases}$$

where $\alpha_{k,i}, i = 0, 1, \cdots, k$ are real numbers as well. Straightforwardly, we obtain

$$\begin{aligned}\sum_{i=0}^{k} \alpha_{k,i}(u(t_i) + u(t_{i+1}) + u(t_{i+2}) + u(t_{i+3})) = \\ \sum_{j=0}^{k} \frac{\sum_{i=0}^{k} \alpha_{k,i}\left((\Delta t_i)^j + (\Delta t_{i+1})^j + (\Delta t_{i+2})^j + (\Delta t_{i+3})^j\right)}{j!} \frac{d^j u}{dt^j}(t_0) \\ + \underbrace{\mathcal{O}\left(\sum_{i=0}^{k} \alpha_{k,i}\left((\Delta t_i)^{k+1} + (\Delta t_{i+1})^{k+1} + (\Delta t_{i+2})^{k+1} + (\Delta t_{i+3})^{k+1}\right)\right)}_{:=\epsilon}\end{aligned} \tag{12}$$

and thus

$$\frac{du}{dt}(t_0) = \sum_{i=0}^{k} \alpha_{k,i}\left(u(t_i) + u(t_{i+1}) + u(t_{i+2}) + u(t_{i+3})\right) + \epsilon \tag{13}$$

by choosing

$$\frac{\sum_{i=0}^{k} \alpha_{k,i}\left((\Delta t_i)^j + (\Delta t_{i+1})^j + (\Delta t_{i+2})^j + (\Delta t_{i+3})^j\right)}{j!} = \begin{cases} 1, & j = 1, \\ 0, & j \neq 1 \end{cases}$$

which are equivalent to the following system:

$$\begin{bmatrix} 4 & 4 & \dots & 4 \\ 6 & 10 & \dots & k+(k+1)+(k+2)+(k+3) \\ 14 & 30 & \dots & k^2+(k+1)^2+(k+2)^2+(k+3)^2 \\ \vdots & \vdots & \vdots & \vdots \\ 1^k+2^k+3^k & 1^k+2^k+3^k+4^k & \dots & k^k+(k+1)^k+(k+2)^k+(k+3)^k \end{bmatrix} \times \begin{bmatrix} \alpha_{k,0}\Delta t \\ \alpha_{k,1}\Delta t \\ \alpha_{k,2}\Delta t \\ \vdots \\ \alpha_{k,k}\Delta t \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

In Table 3 we report solutions of the latter system for $k = 1, \cdots, 9$. Applying (13) to (9) one

$\alpha_{k,i}\Delta t$	$i=0$	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$	$i=8$	$i=9$
$k=1$	$-\frac{1}{4}$	$\frac{1}{4}$								
$k=2$	$-\frac{3}{4}$	$\frac{5}{4}$	$-\frac{1}{2}$							
$k=3$	$-\frac{4}{3}$	3	$-\frac{9}{4}$	$\frac{7}{12}$						
$k=4$	$-\frac{11}{6}$	5	$-\frac{21}{4}$	$\frac{31}{12}$	$-\frac{1}{2}$					
$k=5$	$-\frac{87}{40}$	$\frac{161}{24}$	$-\frac{26}{3}$	6	$-\frac{53}{24}$	$\frac{41}{120}$				
$k=6$	$-\frac{19}{8}$	$\frac{949}{120}$	$-\frac{35}{3}$	10	$-\frac{125}{24}$	$\frac{37}{24}$	$-\frac{1}{5}$			
$k=7$	$-\frac{419}{168}$	$\frac{1049}{120}$	$-\frac{85}{6}$	$\frac{85}{6}$	$-\frac{75}{8}$	$\frac{97}{24}$	$-\frac{31}{30}$	$\frac{5}{42}$		
$k=8$	$-\frac{145}{56}$	$\frac{2661}{280}$	$-\frac{101}{6}$	$\frac{39}{2}$	$-\frac{385}{24}$	$\frac{75}{8}$	$-\frac{37}{10}$	$\frac{37}{42}$	$-\frac{2}{21}$	
$k=9$	$-\frac{6781}{2520}$	$\frac{2917}{280}$	$-\frac{4303}{210}$	$\frac{841}{30}$	$-\frac{3461}{120}$	$\frac{887}{40}$	$-\frac{367}{30}$	$\frac{953}{210}$	$-\frac{106}{105}$	$\frac{32}{315}$

Table 3: The values of $\alpha_{k,i}\Delta t$ for combining four time points.

obtain the multi-step scheme as

$$\begin{aligned}&\alpha_{k,0}Y^n + (\alpha_{k,0} + \alpha_{k,1})Y^{n+1} + (\alpha_{k,0} + \alpha_{k,1} + \alpha_{k,2})Y^{n+2} \\ &\quad + \sum_{j=3}^{k} \left(\alpha_{k,j-3} + \alpha_{k,j-2} + \alpha_{k,j-1} + \alpha_{k,j}\right) Y^{n+j} + (\alpha_{k,k-2} + \alpha_{k,k-1} + \alpha_{k,k})Y^{n+k+1} \\ &\quad + (\alpha_{k,k-1} + \alpha_{k,k})Y^{n+k+2} + \alpha_{k,k}Y^{n+k+3} = f(t_n, Y^n)\end{aligned} \tag{14}$$

whose characteristic equation reads

$$\begin{aligned}\alpha_{k,0}\lambda^{k+3} &+ (\alpha_{k,0}+\alpha_{k,1})\lambda^{k+2} + (\alpha_{k,0}+\alpha_{k,1}+\alpha_{k,2})\lambda^{k+1} \\ &+ \sum_{j=3}^{k} \left(\alpha_{k,j-3}+\alpha_{k,j-2}+\alpha_{k,j-1}+\alpha_{k,j}\right)\lambda^{k+3-j} \\ &+ (\alpha_{k,k-2}+\alpha_{k,k-1}+\alpha_{k,k})\lambda^2 + (\alpha_{k,k-1}+\alpha_{k,k})\lambda^1 + \alpha_{k,k}\lambda^0 = 0.\end{aligned} \tag{15}$$

With the $\alpha_{k,j}$ in Table 3, 1 is the simple root of the latter characteristic function for each k. The maximum absolute values of the roots for $k = 2, \cdots, 10$ expect the simple roots are listed in Table 4, also the multi-step scheme (14) is stable for $k \leq 9$. We remark that the stability cannot

k	2	3	4	5	6	7	8	9	10
$\max(\lvert\lambda_{k,j}\rvert)$	0.6667	0.6614	0.6875	0.7104	0.7224	0.7376	0.8134	0.9931	1.2286

Table 4: The maximum absolute root of (15) except the simple roots

be guaranteed for $k > 9$ by combining more time points, e.g., the multi-step scheme constructed by combining five time points is stable for $k \leq 8$.

4 The semi-discrete multi-step scheme for decoupled FBSDEs

Following the idea in [Zhao et al., 2014a] we derive the semi-discrete scheme for (1) in the decoupled case. We consider the time interval $[0, T]$ with the following partition

$$0 = t_0 < t_1 < t_2 < \cdots t_{N_T} = T.$$

We denote $t_{n+k} - t_n$ by $\Delta t_{n,k}$ and $W_{t_{n+k}} - W_{t_n}$ by $\Delta W_{n,k}$, i.e., $\Delta t_{t_n,t} = t - t_n$ and $\Delta W_{t_n,t} = W_t - W_{t_n}$ for $t \geq t_n$.

4.1 Two reference ODEs

Let (X_t, Y_t, Z_t) be the solution of the decoupled FBSDEs (1). By taking conditional expectation $\mathbb{E}^x_{t_n}[\cdot]$ on both sides of the backward component in (1) one obtains the integral equation

$$\mathbb{E}^x_{t_n}[Y_t] = \mathbb{E}^x_{t_n}[\xi] + \int_t^T \mathbb{E}^x_{t_n}[f(s, X_s, Y_s, Z_s)]\, ds, \quad \forall t \in [t_n, T].$$

As explained in Sec. 2, the integrand in the latter integral equation is continuous with respect to the time. By taking the derivative with respect to t on both sides one thus obtain the first reference ODE:

$$\frac{d\,\mathbb{E}^x_{t_n}[Y_t]}{dt} = -\mathbb{E}^x_{t_n}[f(t, X_t, Y_t, Z_t)], \quad \forall t \in [t_n, T]. \tag{16}$$

Furthermore, we have

$$Y_{t_n} = Y_t + \int_{t_n}^t f(s, X_s, Y_s, Z_s)\, ds - \int_{t_n}^t Z_s\, dW_s, \quad t \in [t_n, T].$$

By multiplying both sides of the latter equation by $\Delta W^\top_{t_n,t}$ and again taking the conditional expectation $\mathbb{E}^x_{t_n}[\cdot]$ on its both sides we obtain

$$0 = \mathbb{E}^x_{t_n}\left[Y_t \Delta W^\top_{t_n,t}\right] + \int_{t_n}^t \mathbb{E}^x_{t_n}\left[f(s, X_s, Y_s, Z_s)\Delta W^\top_{t_n,s}\right] ds - \int_{t_n}^t \mathbb{E}^x_{t_n}[Z_s]\, ds, \quad t \in [t_n, T].$$

Similarly, we obtain the second reference ODE:

$$\frac{d\,\mathbb{E}_{t_n}^x\left[Y_t\Delta W_{t_n,t}^\top\right]}{dt} = -\mathbb{E}_{t_n}^x\left[f(t,X_t,Y_t,Z_t)\Delta W_{t_n,t}^\top\right] + \mathbb{E}_{t_n}^x\left[Z_t\right], \quad t\in[t_n,T]. \tag{17}$$

by taking the derivative with respect to $t\in[t_n,T]$.

4.2 The semi-discrete scheme

Let $\bar{a}(t,x)$ and $\bar{b}(t,x)$ be smooth functions for $t\in[t_n,T]$ and $x\in\mathbb{R}^n$ satisfying $\bar{b}(t,x)=b(t,x)$ and $\bar{b}(t,x)=b(t,x)$, and thus define the diffusion process

$$\bar{X}_t^{t_n,x} = x + \int_{t_n}^t \bar{b}(s,\bar{X}_s^{t_n,x})\,ds + \int_{t_n}^t \bar{b}(s,\bar{X}_s^{t_n,x})\,dW_s. \tag{18}$$

Let $(X_t^{t_n,x},Y_t^{t_n,x},Z_t^{t_n,x})$ be the solution of the decoupled FBSDEs, i.e., $Y_t^{t_n,x}$ and $Z_t^{t_n,x}$ can be represented by $u(t,X_t^{t_n,x})$ and $\nabla_x u(t,X_t^{t_n,x})b(s,X_s^{t_n,x})$, respectively, see Theorem 2.3. Therefore, we set $\bar{Y}_t^{t_n,x}=u(t,\bar{X}_t^{t_n,x})$ and $\bar{Z}_t^{t_n,x}=\nabla_x u(t,\bar{X}_t^{t_n,x})b(s,\bar{X}_s^{t_n,x})$ to have

$$\left.\frac{d\,\mathbb{E}_{t_n}^x[Y_t^{t_n,x}]}{dt}\right|_{t=t_n} = \left.\frac{d\,\mathbb{E}_{t_n}^x[\bar{Y}_t^{t_n,x}]}{dt}\right|_{t=t_n}$$

and

$$\left.\frac{d\,\mathbb{E}_{t_n}^x[Y_t^{t_n,x}\Delta W_{t_n,t}^\top]}{dt}\right|_{t=t_n} = \left.\frac{d\,\mathbb{E}_{t_n}^x[\bar{Y}_t^{t_n,x}\Delta W_{t_n,t}^\top]}{dt}\right|_{t=t_n}$$

by Theorem 2.2. Then, we apply (13) to terms on the right hand side of both the latter equations to obtain

$$\begin{aligned}
\left.\frac{d\,\mathbb{E}_{t_n}^x[Y_t^{t_n,x}]}{dt}\right|_{t=t_n} &= \sum_{i=0}^{k}\alpha_{k,i}\mathbb{E}_{t_n}^x\left[\bar{Y}_{t_{n+i}}^{t_n,x}+\bar{Y}_{t_{n+i+1}}^{t_n,x}+\bar{Y}_{t_{n+i+2}}^{t_n,x}+\bar{Y}_{t_{n+i+3}}^{t_n,x}\right]+\bar{R}_{y,n}^k\\
&= \alpha_{k,0}\mathbb{E}_{t_n}^x\left[\bar{Y}_{t_n}^{t_n,x}\right]+(\alpha_{k,0}+\alpha_{k,1})\mathbb{E}_{t_n}^x\left[\bar{Y}_{t_{n+1}}^{t_n,x}\right]+(\alpha_{k,0}+\alpha_{k,1}+\alpha_{k,2})\mathbb{E}_{t_n}^x\left[\bar{Y}_{t_{n+2}}^{t_n,x}\right]\\
&\quad+\sum_{j=3}^{k}\left(\alpha_{k,j-3}+\alpha_{k,j-2}+\alpha_{k,j-1}+\alpha_{k,j}\right)\mathbb{E}_{t_n}^x\left[\bar{Y}_{t_{n+j}}^{t_n,x}\right]\\
&\quad+(\alpha_{k,k-2}+\alpha_{k,k-1}+\alpha_{k,k})\mathbb{E}_{t_n}^x\left[\bar{Y}_{t_{n+k+1}}^{t_n,x}\right]\\
&\quad+(\alpha_{k,k-1}+\alpha_{k,k})\mathbb{E}_{t_n}^x\left[\bar{Y}_{t_{n+k+2}}^{t_n,x}\right]+\alpha_{k,k}\mathbb{E}_{t_n}^x\left[\bar{Y}_{t_{n+k+3}}^{t_n,x}\right]+\bar{R}_{y,n}^k
\end{aligned} \tag{19}$$

and

$$\begin{aligned}
\left.\frac{d\,\mathbb{E}_{t_n}^x[Y_t^{t_n,x}\Delta W_{t_n,t}^\top]}{dt}\right|_{t=t_n} &= \sum_{i=1}^{k}\alpha_{k,i}\mathbb{E}_{t_n}^x\left[\bar{Y}_{t_{n+i}}^{t_n,x}\Delta W_{n,i}^\top+\bar{Y}_{t_{n+i+1}}^{t_n,x}\Delta W_{n,i+1}^\top\right.\\
&\quad\left.+\bar{Y}_{t_{n+i+2}}^{t_n,x}\Delta W_{n,i+2}^\top+\bar{Y}_{t_{n+i+3}}^{t_n,x}\Delta W_{n,i+3}^\top\right]+\bar{R}_{z,n}^k\\
&= (\alpha_{k,0}+\alpha_{k,1})\mathbb{E}_{t_n}^x\left[\bar{Y}_{t_{n+1}}^{t_n,x}\Delta W_{n,1}^\top\right]+(\alpha_{k,0}+\alpha_{k,1}+\alpha_{k,2})\mathbb{E}_{t_n}^x\left[\bar{Y}_{t_{n+2}}^{t_n,x}\Delta W_{n,2}^\top\right]\\
&\quad+\sum_{j=3}^{k}\left(\alpha_{k,j-3}+\alpha_{k,j-2}+\alpha_{k,j-1}+\alpha_{k,j}\right)\mathbb{E}_{t_n}^x\left[\bar{Y}_{t_{n+j}}^{t_n,x}\Delta W_{n,j}^\top\right]\\
&\quad+(\alpha_{k,k-2}+\alpha_{k,k-1}+\alpha_{k,k})\mathbb{E}_{t_n}^x\left[\bar{Y}_{t_{n+k+1}}^{t_n,x}\Delta W_{n,k+1}^\top\right]\\
&\quad+(\alpha_{k,k-1}+\alpha_{k,k})\mathbb{E}_{t_n}^x\left[\bar{Y}_{t_{n+k+2}}^{t_n,x}\Delta W_{n,k+2}^\top\right]+\alpha_{k,k}\mathbb{E}_{t_n}^x\left[\bar{Y}_{t_{n+k+3}}^{t_n,x}\Delta W_{n,k+3}^\top\right]+\bar{R}_{z,n}^k
\end{aligned} \tag{20}$$

where $\alpha_{k,i}$ are given in Table 3, $\bar{R}^k_{y,n}$ and $\bar{R}^k_{z,n}$ are truncation errors. We insert respectively (19) and (20) into (16) and (17), and obtain

$$\begin{aligned}
&\alpha_{k,0}\mathbb{E}^x_{t_n}\left[\bar{Y}^{t_n,x}_{t_n}\right] + (\alpha_{k,0}+\alpha_{k,1})\mathbb{E}^x_{t_n}\left[\bar{Y}^{t_n,x}_{t_{n+1}}\right] + (\alpha_{k,0}+\alpha_{k,1}+\alpha_{k,2})\mathbb{E}^x_{t_n}\left[\bar{Y}^{t_n,x}_{t_{n+2}}\right] \\
&+ \sum_{j=3}^{k}\left(\alpha_{k,j-3}+\alpha_{k,j-2}+\alpha_{k,j-1}+\alpha_{k,j}\right)\mathbb{E}^x_{t_n}\left[\bar{Y}^{t_n,x}_{t_{n+j}}\right] \\
&+ (\alpha_{k,k-2}+\alpha_{k,k-1}+\alpha_{k,k})\mathbb{E}^x_{t_n}\left[\bar{Y}^{t_n,x}_{t_{n+k+1}}\right] + (\alpha_{k,k-1}+\alpha_{k,k})\mathbb{E}^x_{t_n}\left[\bar{Y}^{t_n,x}_{t_{n+k+2}}\right] \\
&+ \alpha_{k,k}\mathbb{E}^x_{t_n}\left[\bar{Y}^{t_n,x}_{t_{n+k+3}}\right] = -f(t_n,x,Y_{t_n},Z_{t_n}) + R^k_{y,n}
\end{aligned} \tag{21}$$

and

$$\begin{aligned}
&(\alpha_{k,0}+\alpha_{k,1})\mathbb{E}^x_{t_n}\left[\bar{Y}^{t_n,x}_{t_{n+1}}\Delta W^\top_{n,1}\right] + (\alpha_{k,0}+\alpha_{k,1}+\alpha_{k,2})\mathbb{E}^x_{t_n}\left[\bar{Y}^{t_n,x}_{t_{n+2}}\Delta W^\top_{n,2}\right] \\
&+ \sum_{j=3}^{k}\left(\alpha_{k,j-3}+\alpha_{k,j-2}+\alpha_{k,j-1}+\alpha_{k,j}\right)\mathbb{E}^x_{t_n}\left[\bar{Y}^{t_n,x}_{t_{n+j}}\Delta W^\top_{n,j}\right] \\
&+ (\alpha_{k,k-2}+\alpha_{k,k-1}+\alpha_{k,k})\mathbb{E}^x_{t_n}\left[\bar{Y}^{t_n,x}_{t_{n+k+1}}\Delta W^\top_{n,k+1}\right] \\
&+ (\alpha_{k,k-1}+\alpha_{k,k})\mathbb{E}^x_{t_n}\left[\bar{Y}^{t_n,x}_{t_{n+k+2}}\Delta W^\top_{n,k+2}\right] + \alpha_{k,k}\mathbb{E}^x_{t_n}\left[\bar{Y}^{t_n,x}_{t_{n+k+3}}\Delta W^\top_{n,k+3}\right] = Z_{t_n} + R^k_{z,n}
\end{aligned} \tag{22}$$

with $R^k_{y,n} = -\bar{R}^k_{y,n}$ and $R^k_{z,n} = -\bar{R}^k_{z,n}$.
We denote the numerical approximations of Y_t and Z_t at t_n by Y^n and Z^n, respectively. Furthermore, for $\bar{a}$ and $\bar{b}$ in (18) we choose $\bar{a}(t,\bar{X}^{t_n,x}_t) = a(t_n,x)$ and $\bar{b}(t,\bar{X}^{t_n,x}_t) = b(t_n,x)$ for $t\in[t_n,T]$. Finally, from (21) and (22), the semi-discrete scheme can be obtained as

Scheme 1. *Assume that Y^{N_T-i} and Z^{N_T-i} are known for $i = 0,1,\cdots,k+2$. For $n = N_T-k-3,\cdots,0$, $X^{n,j}$, $Y^n = Y^n(X^n)$ and $Z^n = Z^n(X^n)$ can be solved by*

$$X^{n,j} = X^n + a(t_n,X^n)\Delta t_{n,j} + b(t_n,X^n)\Delta W_{n,j}, \quad j = 1,\cdots,k+3, \tag{23}$$

$$\begin{aligned}
Z^n &= (\alpha_{k,0}+\alpha_{k,1})\mathbb{E}^{X^n}_{t_n}\left[\bar{Y}^{n+1}\Delta W^\top_{n,1}\right] + (\alpha_{k,0}+\alpha_{k,1}+\alpha_{k,2})\mathbb{E}^{X^n}_{t_n}\left[\bar{Y}^{n+2}\Delta W^\top_{n,2}\right] \\
&+ \sum_{j=3}^{k}\left(\alpha_{k,j-3}+\alpha_{k,j-2}+\alpha_{k,j-1}+\alpha_{k,j}\right)\mathbb{E}^{X^n}_{t_n}\left[\bar{Y}^{n+j}\Delta W^\top_{n,j}\right] \\
&+ (\alpha_{k,k-2}+\alpha_{k,k-1}+\alpha_{k,k})\mathbb{E}^{X^n}_{t_n}\left[\bar{Y}^{n+k+1}\Delta W^\top_{n,k+1}\right] \\
&+ (\alpha_{k,k-1}+\alpha_{k,k})\mathbb{E}^{X^n}_{t_n}\left[\bar{Y}^{n+k+2}\Delta W^\top_{n,k+2}\right] + \alpha_{k,k}\mathbb{E}^{X^n}_{t_n}\left[\bar{Y}^{n+k+3}\Delta W^\top_{n,k+3}\right],
\end{aligned} \tag{24}$$

$$\begin{aligned}
\alpha_{k,0}Y^n &= -(\alpha_{k,0}+\alpha_{k,1})\mathbb{E}^{X^n}_{t_n}\left[\bar{Y}^{n+1}\right] - (\alpha_{k,0}+\alpha_{k,1}+\alpha_{k,2})\mathbb{E}^{X^n}_{t_n}\left[\bar{Y}^{n+2}\right] \\
&- \sum_{j=3}^{k}(\alpha_{k,j-3}+\alpha_{k,j-2}+\alpha_{k,j-1}+\alpha_{k,j})\mathbb{E}^{X^n}_{t_n}\left[\bar{Y}^{n+j}\right] \\
&- (\alpha_{k,k-2}+\alpha_{k,k-1}+\alpha_{k,k})\mathbb{E}^{X^n}_{t_n}\left[\bar{Y}^{n+k+1}\right] - (\alpha_{k,k-1}+\alpha_{k,k})\mathbb{E}^{X^n}_{t_n}\left[\bar{Y}^{n+k+2}\right] \\
&- \alpha_{k,k}\mathbb{E}^{X^n}_{t_n}\left[\bar{Y}^{n+k+3}\right] - f(t_n,X^n,Y^n,Z^n).
\end{aligned} \tag{25}$$

Remark 4.1. *1. $\bar{Y}^{n+j}$ is the value of Y^{n+j} at the space point $X^{n,j}$ for $j = 1,\cdots,k+3$.*

2. The latter implicit equation can be solved by using iterative methods, e.g., Newton's method or Picard scheme.

3. *By Theorem 2.2 and* (12) *it holds [Butcher, 2008]*

$$\bar{R}^k_{y,n} = \mathcal{O}(\Delta t)^k \text{ and } \bar{R}^k_{z,n} = \mathcal{O}(\Delta t)^k$$

provided that $\mathcal{L}^{k+4}_{t,x}u(t,x)$ *is bounded, where* $\bar{R}^k_{y,n}$ *and* $\bar{R}^k_{z,n}$ *are defined in* (19) *and* (20)*, respectively.*

4. *Similar to the scheme proposed in [Zhao et al., 2014a], one can obtain high-order accurate numerical solutions for* (24) *and* (25)*, although the Euler scheme is used for* (23)*. This is the main advantages because the usage of the Euler scheme reduces dramatically the total computational complexity, and one is only interested in the solution of* (24) *and* (25) *in many applications.*

5 The fully discrete multi-step scheme for decoupled FBSDEs

To solve (X^n, Y^n, Z^n) numerically, next we consider the space discretization. We define firstly the partition of the real space as $\mathcal{R}^n_h = \{x_i | x_i \in \mathbb{R}^n\}$ with

$$h^n = \max_{x\in\mathbb{R}^n} dist(x, \mathcal{R}^n_h),$$

where $dist(x, \mathcal{R}^n_h)$ is the distance from x to $\mathcal{R}^x_h$. Furthermore, for each x we define the neighbor grid set (local subset) $\mathcal{R}^n_{h,x}$ satisfying

1. $dist(x, \mathcal{R}^n_h) < dist(x, \mathcal{R}^n_h)/\ \mathcal{R}^n_{h,x}$,
2. the number of elements in $\mathcal{R}^n_{h,x}$ is finite and uniformly bounded.

Based on the space discretization, we can solve $Y^n(x)$ and $Z^n(x)$ for each grid point $x \in \mathcal{R}^n_h$, $n = N_t - k - 3, \cdots, 0$, by

$$\begin{aligned}
Z^n =& (\alpha_{k,0} + \alpha_{k,1})\mathbb{E}^x_{t_n}\left[\bar{Y}^{n+1}\Delta W^\top_{n,1}\right] + (\alpha_{k,0} + \alpha_{k,1} + \alpha_{k,2})\mathbb{E}^x_{t_n}\left[\bar{Y}^{n+2}\Delta W^\top_{n,2}\right] \\
&+ \sum_{j=3}^{k}(\alpha_{k,j-3} + \alpha_{k,j-2} + \alpha_{k,j-1} + \alpha_{k,j})\,\mathbb{E}^x_{t_n}\left[\bar{Y}^{n+j}\Delta W^\top_{n,j}\right] \\
&+ (\alpha_{k,k-2} + \alpha_{k,k-1} + \alpha_{k,k})\mathbb{E}^x_{t_n}\left[\bar{Y}^{n+k+1}\Delta W^\top_{n,k+1}\right] \\
&+ (\alpha_{k,k-1} + \alpha_{k,k})\mathbb{E}^x_{t_n}\left[\bar{Y}^{n+k+2}\Delta W^\top_{n,k+2}\right] + \alpha_{k,k}\mathbb{E}^x_{t_n}\left[\bar{Y}^{n+k+3}\Delta W^\top_{n,k+3}\right],
\end{aligned} \tag{26}$$

$$\begin{aligned}
\alpha_{k,0}Y^n =& -(\alpha_{k,0} + \alpha_{k,1})\mathbb{E}^x_{t_n}\left[\bar{Y}^{n+1}\right] - (\alpha_{k,0} + \alpha_{k,1} + \alpha_{k,2})\mathbb{E}^x_{t_n}\left[\bar{Y}^{n+2}\right] \\
&- \sum_{j=3}^{k}(\alpha_{k,j-3} + \alpha_{k,j-2} + \alpha_{k,j-1} + \alpha_{k,j})\mathbb{E}^x_{t_n}\left[\bar{Y}^{n+j}\right] \\
&- (\alpha_{k,k-2} + \alpha_{k,k-1} + \alpha_{k,k})\mathbb{E}^x_{t_n}\left[\bar{Y}^{n+k+1}\right] - (\alpha_{k,k-1} + \alpha_{k,k})\mathbb{E}^x_{t_n}\left[\bar{Y}^{n+k+2}\right] \\
&- \alpha_{k,k}\mathbb{E}^x_{t_n}\left[\bar{Y}^{n+k+3}\right] - f(t_n, x, Y^n, Z^n).
\end{aligned} \tag{27}$$

Note that $\bar{Y}^{n+j}$ is the value of Y^{n+j} at the space point $X^{n,j}$ generated by

$$X^{n,j} = x + a(t_n, x)\Delta t_{n,j} + b(t_n, x)\Delta W_{n,j}, \quad j = 1, \cdots, k+3.$$

However, $X^{n,j}$ does not belong to $\mathcal{R}^{n+j}_h$. This is to say that the value of Y^{n+j} at $X^{n,j}$ needs to be approximated based on the values of Y^{n+j} on $\mathcal{R}^{n+j}_h$, this can be done using a local interpolation.

By $LI^n_{h,X}F$ we denote the interpolated value of the function F at space point $X \in \mathbb{R}^n$ by using the values of F only in the neighbor grid set, namely $\mathcal{R}^n_{h,X}$. Including the interpolations, (26) and (27) become

$$\begin{aligned} Z^n = &(\alpha_{k,0}+\alpha_{k,1})\mathbb{E}^x_{t_n}\left[LI^{n+1}_{h,X^{n,j}}Y^{n+1}\Delta W^\top_{n,1}\right] \\ &+(\alpha_{k,0}+\alpha_{k,1}+\alpha_{k,2})\mathbb{E}^x_{t_n}\left[LI^{n+2}_{h,X^{n,j}}Y^{n+2}\Delta W^\top_{n,2}\right] \\ &+\sum_{j=3}^{k}\left(\alpha_{k,j-3}+\alpha_{k,j-2}+\alpha_{k,j-1}+\alpha_{k,j}\right)\mathbb{E}^x_{t_n}\left[LI^{n+j}_{h,X^{n,j}}Y^{n+j}\Delta W^\top_{n,j}\right] \\ &+(\alpha_{k,k-2}+\alpha_{k,k-1}+\alpha_{k,k})\mathbb{E}^x_{t_n}\left[LI^{n+k+1}_{h,X^{n,j}}Y^{n+k+1}\Delta W^\top_{n,k+1}\right] \\ &+(\alpha_{k,k-1}+\alpha_{k,k})\mathbb{E}^x_{t_n}\left[LI^{n+k+2}_{h,X^{n,j}}Y^{n+k+2}\Delta W^\top_{n,k+2}\right] \\ &+\alpha_{k,k}\mathbb{E}^x_{t_n}\left[LI^{n+k+3}_{h,X^{n,j}}Y^{n+k+3}\Delta W^\top_{n,k+3}\right]+R^{k,LI_h}_{z,n}, \end{aligned} \tag{28}$$

$$\begin{aligned} \alpha_{k,0}Y^n = &-(\alpha_{k,0}+\alpha_{k,1})\mathbb{E}^x_{t_n}\left[LI^{n+1}_{h,X^{n,j}}Y^{n+1}\right]-(\alpha_{k,0}+\alpha_{k,1}+\alpha_{k,2})\mathbb{E}^x_{t_n}\left[LI^{n+2}_{h,X^{n,j}}Y^{n+2}\right] \\ &-\sum_{j=3}^{k}(\alpha_{k,j-3}+\alpha_{k,j-2}+\alpha_{k,j-1}+\alpha_{k,j})\mathbb{E}^x_{t_n}\left[LI^{n+j}_{h,X^{n,j}}Y^{n+j}\right] \\ &-(\alpha_{k,k-2}+\alpha_{k,k-1}+\alpha_{k,k})\mathbb{E}^x_{t_n}\left[LI^{n+k+1}_{h,X^{n,j}}Y^{n+k+1}\right] \\ &-(\alpha_{k,k-1}+\alpha_{k,k})\mathbb{E}^x_{t_n}\left[LI^{n+k+2}_{h,X^{n,j}}Y^{n+k+2}\right] \\ &-\alpha_{k,k}\mathbb{E}^x_{t_n}\left[LI^{n+k+3}_{h,X^{n,j}}Y^{n+k+3}\right]-f(t_n,x,Y^n,Z^n)+R^{k,LI_h}_{y,n}. \end{aligned} \tag{29}$$

Furthermore, to approximate the conditional expectations in (28) and (29) we employ the Gauss-Hermite quadrature rule which is an extension of the Gaussian quadrature method for approximating the value of integrals of the form $\int_{-\infty}^{\infty}\exp(-\mathbf{x}^2)g(\mathbf{x})\,d\mathbf{x}$ by

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}g(\mathbf{x})\exp(-\mathbf{x}^\top\mathbf{x})\,d\mathbf{x}\approx\sum_{\mathbf{j}=1}^{L}w_{\mathbf{j}}g(\mathbf{a_j}), \tag{30}$$

where $\mathbf{x}=(x_1,\cdots,x_n)^\top$, $\mathbf{x}^\top\mathbf{x}=\sum_{j=1}^{n}x_j^2$, L is the number of used sample points, $\mathbf{j}=(j_1,j_2,\cdots,j_n)$, $\sum_{\mathbf{j}=\mathbf{1}}^{L}=\sum_{j_1=1,\cdots,j_n=1}^{L,\cdots,L}$, $\mathbf{a_j}=(a_{j_1},\cdots,a_{j_n})$ and $\omega_{\mathbf{j}}=\prod_{i=1}^{n}\omega_{j_i}$, $\{a_{j_i}\}_{j_i=1}^{L}$ are the roots of the Hermite polynomial $H_L(x)$ of degree L and $\{\omega_{j_i}\}_{j_i=1}^{L}$ are corresponding weights [Abramowitz and Stegun, 1972]. For a standard n-dimensional standard normal distributed random variable X we know that

$$\begin{aligned} \mathbb{E}\left[g(X)\right] &= \frac{1}{(2\pi)^{\frac{d}{2}}}\int_{-\infty}^{\infty}g(\mathbf{x})\exp\left(-\frac{\mathbf{x}^\top\mathbf{x}}{2}\right)dx \\ &= \frac{1}{(\pi)^{\frac{d}{2}}}\int_{-\infty}^{\infty}g(\sqrt{2}\mathbf{x})\exp\left(-\mathbf{x}^\top\mathbf{x}\right)dx \\ &\overset{30}{=} \frac{1}{(\pi)^{\frac{d}{2}}}\sum_{\mathbf{j}=\mathbf{1}}^{L}\omega_{\mathbf{j}}g(\mathbf{a_j})+R^{GH}_L, \end{aligned}$$

where R_L^{GH} is the truncation error of the Gauss-Hermite quadrature rule for g.
Now we consider the conditional expectations of the form $\mathbb{E}_{t_n}^x\left[LI_{h,X^{n,j}}^{n+j}Y^{n+j}\Delta W_{n,j}^\top\right]$ and $\mathbb{E}_{t_n}^x\left[LI_{h,X^{n,j}}^{n+j}Y^{n+j}\right]$ in (28) and (29). We know that $LI_{h,X^{n,j}}^{n+j}Y^{n+j}$ is the interpolated value of $\bar{Y}^{n+j}$, which is a function of $X^{n,j}$ and can be represented by (Theorem 2.3)

$$\bar{Y}^{n+j} = Y^{n+j}\left(X^{n+j}\right) = Y^{n+j}\left(X^{n+j} + a(t_n, X^n)\Delta t_{n,j} + b(t_n, X^n)\Delta W_{n,j}\right),$$

with $\Delta W_{n,j} \sim \sqrt{\Delta t_{n,j}}N(0, I_n)$. Straightforwardly, we can approximate those conditional expectations as

$$\mathbb{E}_{t_n}^{\mathbf{x},h}\left[\bar{Y}^{n+j}\right] = \frac{1}{\pi^{\frac{d}{2}}}\sum_{\mathbf{j}=\mathbf{1}}^{L}\omega_{\mathbf{j}}Y^{n+j}\left(\mathbf{x} + a(t_n,\mathbf{x})\Delta t_{n,j} + b(t_n,\mathbf{x})\Delta t_{n,j}\sqrt{2\Delta t_{n,j}}\mathbf{a_j}\right) + R_L^{GH}(Y)$$

and

$$\mathbb{E}_{t_n}^{\mathbf{x},h}\left[\bar{Y}^{n+j}\Delta W_{t_n,j}^\top\right] = \frac{1}{\pi^{\frac{d}{2}}}\sum_{\mathbf{j}=\mathbf{1}}^{L}\omega_{\mathbf{j}}Y^{n+j}\left(\mathbf{x} + a(t_n,\mathbf{x})\Delta t_{n,j} + b(t_n,\mathbf{x})\Delta t_{n,j}\sqrt{2\Delta t_{n,j}}\mathbf{a_j}\right)\mathbf{a_j} + R_L^{GH}(YW),$$

where $\mathbb{E}_{t_n}^{\mathbf{x},h}[\cdot]$ denotes the approximation of $\mathbb{E}_{t_n}^{\mathbf{x}}[\cdot]$. Finally, by inserting these approximations into (28) and (29) we obtain

$$\begin{aligned}
Z^n = &(\alpha_{k,0}+\alpha_{k,1})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+1}Y^{n+1}\Delta W_{n,1}^\top\right]\\
&+(\alpha_{k,0}+\alpha_{k,1}+\alpha_{k,2})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+2}Y^{n+2}\Delta W_{n,2}^\top\right]\\
&+\sum_{j=3}^{k}\left(\alpha_{k,j-3}+\alpha_{k,j-2}+\alpha_{k,j-1}+\alpha_{k,j}\right)\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+j}Y^{n+j}\Delta W_{n,j}^\top\right]\\
&+(\alpha_{k,k-2}+\alpha_{k,k-1}+\alpha_{k,k})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+k+1}Y^{n+k+1}\Delta W_{n,k+1}^\top\right]\\
&+(\alpha_{k,k-1}+\alpha_{k,k})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+k+2}Y^{n+k+2}\Delta W_{n,k+2}^\top\right]\\
&+\alpha_{k,k}\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+k+3}Y^{n+k+3}\Delta W_{n,k+3}^\top\right] + R_{z,n}^{k,LI_h} + R_{z,n}^{k,\mathbb{E}},
\end{aligned} \tag{31}$$

$$\begin{aligned}
\alpha_{k,0}Y^n = &-(\alpha_{k,0}+\alpha_{k,1})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+1}Y^{n+1}\right] - (\alpha_{k,0}+\alpha_{k,1}+\alpha_{k,2})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+2}Y^{n+2}\right]\\
&-\sum_{j=3}^{k}(\alpha_{k,j-3}+\alpha_{k,j-2}+\alpha_{k,j-1}+\alpha_{k,j})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+j}Y^{n+j}\right]\\
&-(\alpha_{k,k-2}+\alpha_{k,k-1}+\alpha_{k,k})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+k+1}Y^{n+k+1}\right]\\
&-(\alpha_{k,k-1}+\alpha_{k,k})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+k+2}Y^{n+k+2}\right]\\
&-\alpha_{k,k}\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+k+3}Y^{n+k+3}\right] - f(t_n,x,Y^n,Z^n) + R_{y,n}^{k,LI_h} + R_{y,n}^{k,\mathbb{E}}.
\end{aligned} \tag{32}$$

Remark 5.1. *1. The estimate of $R_{y,n}^{k,\mathbb{E}}$ or $R_{z,n}^{k,\mathbb{E}}$ reads [Abramowitz and Stegun, 1972, Shen et al., 2011, Zhao et al., 2014b]*

$$\mathcal{O}\left(\frac{L!}{2^L(2L)!}\right).$$

2. For the local interpolation errors $R_{y,n}^{k,LI_h}$ or $R_{z,n}^{k,LI_h}$ the following estimate holds

$$\mathcal{O}\left(h^{r+1}\right) \tag{33}$$

when using r-degree polynomial interpolation in k-step scheme, and provided that a, b, f and g are sufficiently smooth such that $\mathcal{L}_{t,x}^{k+4}u(t,x)$ is bounded and $u(t,\cdot) \in C_b^{r+1}$, see [Abramowitz and Stegun, 1972, Burden and Faires, 2001, Butcher, 2008, Zhao et al., 2014b].
3. To balance the time discretization error $R_{y,n}^k = \mathcal{O}(\Delta t)^k$ and $R_{z,n}^k = \mathcal{O}(\Delta t)^k$, one needs to control well both the interpolation and integration error mentioned in last two points.
4. For a k-step scheme we need to know the support values of Y^{N_T-i} and $Z^{N_T-i}, i = 0, \cdots, k+2$. One can use the following three ways to deal with this problem: before running the multi-step scheme, we choose a quite smaller Δt and run one-step scheme until $N_T - k - 2$; Alternatively, one can prepare these initial values "iteratively", namely we compute Y^{N_T-1} and Z^{N_T-1} based on Y^{N_T} and Z^{N_T} with $k = 1$, and the compute Y^{N_T-2} and Z^{N_T-2} based on $Y^{N_T}, Y^{N_T-1}, Z^{N_T}, Z^{N_T-1}$ with $k = 2$ and so on; Finally, one can use the Runge-Kutta scheme proposed in [Crisan and Chassagneux, 2014] with small Δt to initialize our proposed multi-step scheme.

By removing all the error terms, from (31) and (32) we obtain our fully discrete scheme as follows.

Scheme 2. *Assume that Y^{N_T-i} and Z^{N_T-i} on $\mathcal{R}_h^{N_T-i}$ are known for $i = 0, 1, \cdots, k+2$. For $n = N_T - k - 3, \cdots, 0$ and $x \in \mathcal{R}_h^n$, $Y^n = Y^n(x)$ and $Z^n = Z^n(x)$ can be solved by*

$$
\begin{aligned}
X^{n,j} &= x + a(t_n, x)\Delta t_{n,j} + b(t_n, x)\Delta W_{n,j}, \quad j = 1, \cdots, k+3,\\
Z^n &= (\alpha_{k,0} + \alpha_{k,1})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+1} Y^{n+1}\Delta W_{n,1}^\top\right]\\
&\quad + (\alpha_{k,0} + \alpha_{k,1} + \alpha_{k,2})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+2} Y^{n+2}\Delta W_{n,2}^\top\right]\\
&\quad + \sum_{j=3}^{k}(\alpha_{k,j-3} + \alpha_{k,j-2} + \alpha_{k,j-1} + \alpha_{k,j})\,\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+j} Y^{n+j}\Delta W_{n,j}^\top\right]\\
&\quad + (\alpha_{k,k-2} + \alpha_{k,k-1} + \alpha_{k,k})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+k+1} Y^{n+k+1}\Delta W_{n,k+1}^\top\right]\\
&\quad + (\alpha_{k,k-1} + \alpha_{k,k})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+k+2} Y^{n+k+2}\Delta W_{n,k+2}^\top\right]\\
&\quad + \alpha_{k,k}\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+k+3} Y^{n+k+3}\Delta W_{n,k+3}^\top\right],\\
\alpha_{k,0}Y^n &= -(\alpha_{k,0} + \alpha_{k,1})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+1} Y^{n+1}\right] - (\alpha_{k,0} + \alpha_{k,1} + \alpha_{k,2})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+2} Y^{n+2}\right]\\
&\quad - \sum_{j=3}^{k}(\alpha_{k,j-3} + \alpha_{k,j-2} + \alpha_{k,j-1} + \alpha_{k,j})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+j} Y^{n+j}\right]\\
&\quad - (\alpha_{k,k-2} + \alpha_{k,k-1} + \alpha_{k,k})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+k+1} Y^{n+k+1}\right]\\
&\quad - (\alpha_{k,k-1} + \alpha_{k,k})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+k+2} Y^{n+k+2}\right]\\
&\quad - \alpha_{k,k}\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+k+3} Y^{n+k+3}\right] - f(t_n, x, Y^n, Z^n).
\end{aligned}
$$

6 Numerical schemes for coupled FBSDEs

The authors in [Zhao et al., 2014b] extended their scheme proposed for solving decoupled FBSDEs to the one which can solve fully coupled FBSDEs. Similarly, our Scheme 2 can be extended to solve (1) in a fully coupled case.

Scheme 3. *Assume that Y^{N_T-i} and Z^{N_T-i} on $\mathcal{R}_h^{N_T-i}$ are known for $i = 0, 1, \cdots, k+2$. For $n = N_T - k - 3, \cdots, 0$ and $x \in \mathcal{R}_h^n$, $Y^n = Y^n(x)$ and $Z^n = Z^n(x)$ can be solved by*

1. set $Y^{n,0} = Y^{n+1}(x)$ *and* $Z^{n,0} = Z^{n+1}(x)$, *and set* $l = 0$;
2. for $l = 0, 1, \cdots$, *solve* $Y^{n,l+1} = Y^{n,l+1}(x)$ *and* $Z^{n,l+1} = Z^{n,l+1}(x)$ *by*

$$X^{n,j} = x + a(t_n, x, Y^{n,l}(x), Z^{n,l}(x))\Delta t_{n,j} + b(t_n, x, Y^{n,l}(x), Z^{n,l}(x))\Delta W_{n,j}, \quad j = 1, \cdots, k+3,$$

$$\begin{aligned}
Z^n = {} & (\alpha_{k,0} + \alpha_{k,1})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+1} Y^{n+1}\Delta W_{n,1}^\top\right] + (\alpha_{k,0} + \alpha_{k,1} + \alpha_{k,2})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+2} Y^{n+2}\Delta W_{n,2}^\top\right] \\
& + \sum_{j=3}^{k} \left(\alpha_{k,j-3} + \alpha_{k,j-2} + \alpha_{k,j-1} + \alpha_{k,j}\right) \mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+j} Y^{n+j}\Delta W_{n,j}^\top\right] \\
& + (\alpha_{k,k-2} + \alpha_{k,k-1} + \alpha_{k,k})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+k+1} Y^{n+k+1}\Delta W_{n,k+1}^\top\right] \\
& + (\alpha_{k,k-1} + \alpha_{k,k})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+k+2} Y^{n+k+2}\Delta W_{n,k+2}^\top\right] + \alpha_{k,k}\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+k+3} Y^{n+k+3}\Delta W_{n,k+3}^\top\right],
\end{aligned}$$

$$\begin{aligned}
\alpha_{k,0} Y^n = {} & -(\alpha_{k,0} + \alpha_{k,1})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+1} Y^{n+1}\right] - (\alpha_{k,0} + \alpha_{k,1} + \alpha_{k,2})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+2} Y^{n+2}\right] \\
& - \sum_{j=3}^{k}(\alpha_{k,j-3} + \alpha_{k,j-2} + \alpha_{k,j-1} + \alpha_{k,j})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+j} Y^{n+j}\right] \\
& - (\alpha_{k,k-2} + \alpha_{k,k-1} + \alpha_{k,k})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+k+1} Y^{n+k+1}\right] - (\alpha_{k,k-1} + \alpha_{k,k})\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+k+2} Y^{n+k+2}\right] \\
& - \alpha_{k,k}\mathbb{E}_{t_n}^{x,h}\left[LI_{h,X^{n,j}}^{n+k+3} Y^{n+k+3}\right] - f(t_n, x, Y^n, Z^n).
\end{aligned}$$

until $\max\left(\left|Y^{n,l+1} - Y^{n,l}\right|, \left|Z^{n,l+1} - Z^{n,l}\right|\right) < \epsilon_0$,
3. let $Y^n = Y^{n,l+1}$ *and* $Z^n = Z^{n,l+1}$.

Remark 6.1. *1. Scheme 3 coincides with Scheme 2 if* a *and* b *do not depend on* Y *and* Z.
2. We only assume that the coupled FBSDEs are uniquely solvable, the lacking analysis will be the task of future work.

7 Numerical experiments

In this section we use some numerical examples to show that our Schemes 2 and 3 can reach ninth-order convergence rate for solving FBSDEs. The uniform partitions in both time and space will be used, that is, the time interval $[0, T]$ will be uniformly divided into N_T parts with $\Delta t = \frac{T}{N_T}$ such that $t_n = n\Delta t, n = 0, 1, \cdots, N_T$; the space partition is $\mathcal{R}_h^n = R_h$ for all n with

$$\mathcal{R}_h = \mathcal{R}_{1,h} \times \mathcal{R}_{2,h} \times \cdots \mathcal{R}_{n,h},$$

where $\mathcal{R}_{j,h}$ is the partition of $\mathbb{R}$

$$\mathcal{R}_{j,h} = \left\{x_i^j : x_i^j = ih, i = 0, \pm 1, \cdots, \pm\infty\right\}, \quad j = 1, 2, \cdots, n.$$

In our numerical experiments we choose the local Lagrange interpolation for $LI_{h,x}^n$ based on the set of some neighbor grids near x, i.e., $\mathcal{R}_{h,x} \subset \mathcal{R}_h$ such that (33) holds. Following [Zhao et al., 2014a], we set sufficiently many Gauss-Hermite quadrature points such that the quadrature error could be negligible. Note that the truncation error is defined in (12), in order to thus balance the time and space truncation error in our numerical examples, we force $h^{r+1} = (\Delta t)^{k+1}$, where r is the degree of the Lagrangian interpolation polynomials. For example, one can firstly specify a value of r, and then adjust the value of h such that $h = \Delta t^{\frac{k+1}{r+1}}$. For the numerical results in this paper, r is set to be a value from the set $\{10, 11, \cdots, 21\}$ to control the errors. Furthermore, we will consider k from 3 such that at least one combination of four

$\alpha_{k,i}$ s is included, but until $k = 9$ due to the stability condition, see Scheme 2 and 3. Finally, CR and RT are used to denote the convergence rate and the running time in second, respectively. For the comparison purpose, we will directly take examples considered in [Zhao et al., 2014a]. Numerical experiment were performed in MATLAB with an Intel(R) Core(TM) i5-8350 CPU @ 1.70 GHz and 15 G RAM.

Example 1 The first example reads

$$\begin{cases} dX_t = \frac{1}{1+2\exp(t+X_t)}\,dt + \frac{\exp(t+X_t)}{1+\exp(t+X_t)}\,dW_t, \quad X_0 = 1, \\ -dY_t = \left(-\frac{2Y_t}{1+2\exp(t+X_t)} - \frac{1}{2}\left(\frac{Y_t Z_t}{1+\exp(t+X_t)} - Y_t^2 Z_t\right)\right)\,dt - Z_t\,dW_t, \\ Y_T = \frac{\exp(T+X_T)}{1+\exp(T+X_T)}, \end{cases}$$

with the analytic solution

$$\begin{cases} Y_t = \frac{\exp(t+X_t)}{1+\exp(t+X_t)}, \\ Z_t = \frac{(\exp(t+X_t))^2}{(1+\exp(t+X_t))^3}. \end{cases}$$

Obviously, in this example, the generator a and b does not depend on Y_t and Z_t, i.e., a decoupled FBSDE. All the convergence rates, running time and absolute errors are reported in Table 5. In the proposed scheme, $k + 3$ points are needed for the iterations, i.e., one need 12 points when $k = 9$. From the other side we can not choose a large value for N_T due to the accuracy of double precision. Therefore, to show the convergence rate up to ninth order we consider $N_T = \{16, 20, 24, 28, 32\}$ in this example. From Table 5 we see that the quite high accuracy of

Scheme 2		$N_T = 16$	$N_T = 20$	$N_T = 24$	$N_T = 28$	$N_T = 32$	CR
$k = 3$	$\lvert Y^0 - Y_0\rvert$	4.717e-06	2.613e-06	1.569e-06	1.015e-06	6.905e-07	2.78
	$\lvert Z^0 - Z_0\rvert$	2.547e-05	1.552e-05	1.009e-05	6.889e-06	4.903e-06	2.39
	RT	0.37	0.49	0.65	0.91	1.19	
$k = 4$	$\lvert Y^0 - Y_0\rvert$	6.871e-07	3.152e-07	1.629e-07	9.240e-08	5.618e-08	3.61
	$\lvert Z^0 - Z_0\rvert$	6.879e-06	3.097e-06	1.595e-06	9.027e-07	5.488e-07	3.65
	RT	0.39	0.61	0.85	1.12	1.41	
$k = 5$	$\lvert Y^0 - Y_0\rvert$	5.623e-08	2.077e-08	9.011e-09	4.355e-09	2.343e-09	4.59
	$\lvert Z^0 - Z_0\rvert$	6.522e-07	2.427e-07	1.047e-07	5.016e-08	2.704e-08	4.60
	RT	0.46	0.69	0.95	1.25	1.58	
$k = 6$	$\lvert Y^0 - Y_0\rvert$	3.549e-09	1.073e-09	3.929e-10	1.623e-10	7.549e-11	5.56
	$\lvert Z^0 - Z_0\rvert$	5.658e-08	1.632e-08	6.000e-09	2.519e-09	1.168e-09	5.59
	RT	0.52	0.85	1.26	1.74	2.10	
$k = 7$	$\lvert Y^0 - Y_0\rvert$	2.156e-10	4.809e-11	1.457e-11	5.075e-12	2.019e-12	6.73
	$\lvert Z^0 - Z_0\rvert$	6.349e-09	1.556e-09	4.796e-10	1.749e-10	7.147e-11	6.47
	RT	0.60	1.10	1.65	2.15	2.79	
$k = 8$	$\lvert Y^0 - Y_0\rvert$	6.025e-11	8.573e-12	3.292e-12	7.027e-13	4.868e-13	7.10
	$\lvert Z^0 - Z_0\rvert$	1.029e-09	1.934e-10	5.811e-11	1.459e-11	6.696e-12	7.35
	RT	0.62	1.17	1.73	2.37	3.11	
$k = 9$	$\lvert Y^0 - Y_0\rvert$	2.315e-11	4.131e-12	9.073e-13	2.169e-13	2.398e-14	9.55
	$\lvert Z^0 - Z_0\rvert$	3.672e-10	5.073e-11	1.184e-11	2.528e-12	5.760e-13	9.19
	RT	0.69	1.32	2.04	2.83	3.74	

Table 5: Errors, running time and convergence rates for Example 1, $T = 1$

Scheme 2 for solving decoupled FBSDEs. Scheme 2 is a k-order scheme up to $k = 9$, and more efficient for taking a larger value for k, which is consistent with the theory [Butcher, 2008], see also Table 4

For the second example we consider the coupled FBSDE (taken from [Zhao et al., 2014a]) to test Scheme 3, in which an iterative process is required with longer computational time.

Example 2

$$\begin{cases} dX_t = -\frac{1}{2}\sin(t+X_t)\cos(t+X_t)(Y_t^2+Z_t)\,ds \\ \qquad +\frac{1}{2}\cos(t+X_t) + (Y_t\sin(t+X_t)+Z_t+1)\,dW_s, \quad X_0 = 1.5, \\ -dY_t = (Y_tZ_t - \cos(t+X_t))\;dt - Z_t\,dW_t, \\ Y_T = \sin(T+X_T). \end{cases}$$

has the analytic solution

$$\begin{cases} Y_t = \sin(t+X_t), \\ Z_t = \cos^2(t+X_t). \end{cases}$$

In this coupled FBSDE, the diffusion coefficient b depends on X, Y and Z, i.e., quite general. Due to the same reasons as those explained for Example 1, we set $N_T = \{13, 15, 17, 19, 21\}$ in order to show the convergence rate up to ninth order. From the results listed in Table 6 one can

Scheme 3		$N_T = 13$	$N_T = 15$	$N_T = 17$	$N_T = 19$	$N_T = 21$	CR
$k = 3$	$\lvert Y^0 - Y_0\rvert$	2.269e-04	1.398e-04	9.186e-05	6.025e-05	4.336e-05	3.47
	$\lvert Z^0 - Z_0\rvert$	1.562e-04	1.143e-04	6.555e-05	4.431e-05	3.359e-05	3.36
	RT	3.89	4.79	5.78	6.53	7.84	
$k = 4$	$\lvert Y^0 - Y_0\rvert$	9.569e-06	7.654e-06	4.799e-06	2.734e-06	1.571e-06	3.83
	$\lvert Z^0 - Z_0\rvert$	1.447e-04	1.126e-04	6.268e-05	3.009e-05	1.223e-05	5.14
	RT	4.51	5.73	6.79	8.12	10.43	
$k = 5$	$\lvert Y^0 - Y_0\rvert$	4.773e-07	1.740e-07	3.835e-08	6.464e-08	2.325e-08	5.96
	$\lvert Z^0 - Z_0\rvert$	2.433e-06	6.129e-07	2.215e-08	1.988e-07	2.968e-07	4.89
	RT	21.07	28.68	35.43	41.14	53.55	
$k = 6$	$\lvert Y^0 - Y_0\rvert$	4.469e-08	2.509e-08	1.361e-08	6.572e-09	3.257e-09	5.47
	$\lvert Z^0 - Z_0\rvert$	4.257e-07	3.145e-07	1.629e-07	7.017e-08	2.121e-08	6.15
	RT	33.49	48.76	63.83	80.93	99.07	
$k = 7$	$\lvert Y^0 - Y_0\rvert$	6.510e-10	4.904e-10	1.536e-11	4.256e-11	3.250e-11	7.20
	$\lvert Z^0 - Z_0\rvert$	1.218e-08	1.207e-08	8.072e-10	3.704e-09	2.249e-10	7.56
	RT	34.39	53.76	75.34	97.09	120.69	
$k = 8$	$\lvert Y^0 - Y_0\rvert$	1.876e-10	8.477e-11	3.098e-11	1.028e-11	5.961e-12	7.51
	$\lvert Z^0 - Z_0\rvert$	8.841e-09	2.877e-09	1.727e-10	7.179e-11	5.141e-10	8.19
	RT	35.76	90.17	149.87	172.70	247.68	
$k = 9$	$\lvert Y^0 - Y_0\rvert$	1.926e-11	5.744e-12	5.828e-13	9.910e-13	3.098e-14	12.10
	$\lvert Z^0 - Z_0\rvert$	2.502e-10	1.526e-10	3.190e-11	3.039e-11	2.234e-12	9.07
	RT	55.19	173.80	275.15	384.02	503.85	

Table 6: Errors, running time and convergence rates for Example 2, $T = 1$

clearly draw same conclusions as those having been for Example 1.
Finally, we illustrate the accuracy of the proposed scheme for a two-dimensional example, which is also taken from [Zhao et al., 2014a] and reads

Example 3

$$\begin{cases} \begin{pmatrix} dX_t^1 \\ dX_t^2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\sin^2(t+X_t^1) \\ \frac{1}{2}\sin^2(t+X_t^2) \end{pmatrix} dt + \begin{pmatrix} \frac{1}{2}\cos^2(t+X_t^1) \\ \frac{1}{2}\cos^2(t+X_t^2) \end{pmatrix} dW_t, \quad \begin{pmatrix} X_0^1 \\ X_0^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \\ \begin{pmatrix} dY_t^1 \\ dY_t^2 \end{pmatrix} = \begin{pmatrix} -\frac{3}{2}\cos(t+X_t^1)\sin(t+X_t^2) - \frac{3}{2}\sin(t+X_t^1)\cos(t+X_t^2) - Z_t^2 \\ +\frac{1}{2}Y_t^1\left(\frac{1}{4}\cos^4(t+X_t^2) + \frac{1}{4}\cos^4(t+X_t^1)\right) - \frac{1}{4}(Y_t^2)^3 \\ \frac{3}{2}\sin(t+X_t^1)\cos(t+X_t^2) + \frac{3}{2}\cos(t+X_t^1)\sin(t+X_t^2) - Z_t^1 \\ +\frac{1}{2}Y_t^2\left(\frac{1}{4}\cos^4(t+X_t^2) + \frac{1}{4}\cos^4(t+X_t^1)\right) - \frac{1}{4}Y_t^1(Y_t^2)^2 \end{pmatrix} dt - \begin{pmatrix} Z_t^1 \\ Z_t^2 \end{pmatrix} dW_t, \\ \begin{pmatrix} Y_T^1 \\ Y_T^2 \end{pmatrix} = \begin{pmatrix} \sin(T+X_T^1)\sin(T+X_T^2) \\ \cos(T+X_T^1)\cos(T+X_T^2) \end{pmatrix} \end{cases}$$

with the analytic solution

$$\begin{cases} \begin{pmatrix} Y_t^1 \\ Y_t^2 \end{pmatrix} = \begin{pmatrix} \sin(t+X_t^1)\sin(t+X_t^2) \\ \cos(t+X_t^1)\cos(t+X_t^2) \end{pmatrix}, \\ \begin{pmatrix} Z_t^1 \\ Z_t^2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\cos(t+X_t^1)\sin(t+X_t^2)\cos^2(t+X_t^2) \\ +\frac{1}{2}\sin(t+X_t^1)\cos(t+X_t^2)\cos^2(t+X_t^1) \\ -\frac{1}{2}\sin(t+X_t^1)\cos^3(t+X_t^2) - \frac{1}{2}\cos^3(t+X_t^1)\sin(t+X_t^2) \end{pmatrix}. \end{cases}$$

The numerical approximations are reported in Table 7, which show that our multi-step scheme is still quite highly accurate for solving a two-dimensional FBSDE. We observe that the convergence rates are roughly consistent with the theoretical results, the slight deviation comes from the quadratures and especially the two-dimensional interpolations. Obviously, the high efficiency and accuracy have been shown in this example. Note that the parallel computing toolbox in MATLAB has been used in this example, more precisely, the parallel for-Loops (parfor) is used for the two-dimensional interpolation on the grid points.

8 Conclusion

In this work, by combining the multi-steps we have adopted the high-order multi-step method in [W. Zhao, Y. Fu and T. Zhou, SIAM J. Sci. Comput., 36(4) (2014), pp.A1731-A1751] for solving FBSDEs. First of all, our new schemes allow for higher convergence rate up to ninth order, and are more efficient. Secondly, they keep the key feature of the method in [W. Zhao, Y. Fu and T. Zhou, SIAM J. Sci. Comput. 36(4), pp.A1731-A1751], that is the numerical solution of backward component maintains the higher-order accuracy by using the Euler method to the forward component. This feature makes our schemes be promising in solving problems in practice. The effectiveness and higher-order accuracy have been confirmed by the numerical experiments. A rigorous stability analysis for the proposed schemes is the task of future work.

Acknowledge

We thank Suman Kumar from the University of Wuppertal for his assistance with Matlab programming partially used for the numerical experiments.

Scheme 2		$N_T = 16$	$N_T = 20$	$N_T = 24$	$N_T = 28$	$N_T = 32$	CR
$k = 3$	$\lvert Y^{0,1} - Y_0^1\rvert$	4.308e-03	2.447e-03	1.495e-03	9.726e-04	6.651e-04	2.70
	$\lvert Y^{0,2} - Y_0^2\rvert$	3.896e-03	2.200e-03	1.340e-03	8.705e-04	5.949e-04	2.71
	$\lvert Z^{0,1} - Z_0^1\rvert$	3.887e-03	2.041e-03	1.201e-03	7.602e-04	5.110e-04	2.93
	$\lvert Z^{0,2} - Z_0^2\rvert$	2.980e-03	1.491e-03	8.465e-04	5.243e-04	3.451e-04	3.11
	RT	34.65	53.68	74.87	92.42	130.14	
$k = 4$	$\lvert Y^{0,1} - Y_0^1\rvert$	8.829e-04	4.437e-04	2.417e-04	1.413e-04	8.759e-05	3.34
	$\lvert Y^{0,2} - Y_0^2\rvert$	7.776e-04	3.845e-04	2.072e-04	1.202e-04	7.413e-05	3.39
	$\lvert Z^{0,1} - Z_0^1\rvert$	6.653e-04	2.556e-04	1.161e-04	5.899e-05	3.270e-05	4.35
	$\lvert Z^{0,2} - Z_0^2\rvert$	6.845e-04	2.659e-04	1.218e-04	6.268e-05	3.521e-05	4.29
	RT	43.52	70.30	100.41	142.95	195.94	
$k = 5$	$\lvert Y^{0,1} - Y_0^1\rvert$	5.686e-05	2.367e-05	1.084e-05	5.445e-06	2.952e-06	4.27
	$\lvert Y^{0,2} - Y_0^2\rvert$	5.131e-05	2.112e-05	9.605e-06	4.801e-06	2.595e-06	4.31
	$\lvert Z^{0,1} - Z_0^1\rvert$	7.703e-05	2.621e-05	1.067e-05	4.923e-06	2.517e-06	4.94
	$\lvert Z^{0,2} - Z_0^2\rvert$	7.237e-05	2.402e-05	9.582e-06	4.363e-06	2.201e-06	5.04
	RT	50.72	79.97	132.00	186.09	231.84	
$k = 6$	$\lvert Y^{0,1} - Y_0^1\rvert$	4.277e-06	1.706e-06	7.258e-07	3.354e-07	1.674e-07	4.68
	$\lvert Y^{0,2} - Y_0^2\rvert$	3.857e-06	1.509e-06	6.336e-07	2.901e-07	1.437e-07	4.75
	$\lvert Z^{0,1} - Z_0^1\rvert$	5.569e-06	1.667e-06	5.803e-07	2.309e-07	1.016e-07	5.78
	$\lvert Z^{0,2} - Z_0^2\rvert$	5.531e-06	1.655e-06	5.791e-07	2.313e-07	1.026e-07	5.76
	RT	64.97	110.97	173.50	246.96	337.71	
$k = 7$	$\lvert Y^{0,1} - Y_0^1\rvert$	5.753e-07	1.843e-07	6.513e-08	2.568e-08	1.115e-08	5.70
	$\lvert Y^{0,2} - Y_0^2\rvert$	5.312e-07	1.669e-07	5.836e-08	2.283e-08	9.864e-09	5.76
	$\lvert Z^{0,1} - Z_0^1\rvert$	9.466e-07	2.291e-07	7.027e-08	2.528e-08	1.027e-08	6.52
	$\lvert Z^{0,2} - Z_0^2\rvert$	9.091e-07	2.157e-07	6.529e-08	2.326e-08	9.404e-09	6.59
	RT	73.89	139.46	234.34	339.44	453.70	
$k = 8$	$\lvert Y^{0,1} - Y_0^1\rvert$	2.586e-08	1.080e-08	4.077e-09	1.592e-09	6.652e-10	5.30
	$\lvert Y^{0,2} - Y_0^2\rvert$	2.384e-08	9.742e-09	3.623e-09	1.400e-09	5.802e-10	5.39
	$\lvert Z^{0,1} - Z_0^1\rvert$	6.238e-08	1.548e-08	4.186e-09	1.299e-09	4.791e-10	7.06
	$\lvert Z^{0,2} - Z_0^2\rvert$	6.290e-08	1.534e-08	4.077e-09	1.255e-09	4.579e-10	7.14
	RT	84.65	177.96	299.48	436.73	592.02	
$k = 9$	$\lvert Y^{0,1} - Y_0^1\rvert$	8.294e-09	2.371e-09	6.964e-10	2.250e-10	8.045e-11	6.70
	$\lvert Y^{0,2} - Y_0^2\rvert$	7.739e-09	2.153e-09	6.229e-10	1.991e-10	6.989e-11	6.80
	$\lvert Z^{0,1} - Z_0^1\rvert$	2.771e-08	5.329e-09	1.262e-09	3.491e-10	1.245e-10	7.84
	$\lvert Z^{0,2} - Z_0^2\rvert$	2.748e-08	5.202e-09	1.217e-09	3.350e-10	9.998e-11	8.08
	RT	91.19	210.99	356.54	534.76	760.94	

Table 7: Errors, running time and convergence rates for Example 3, $T = 1$

References

[Abramowitz and Stegun, 1972] Abramowitz, M. and Stegun, I. (1972). *Handbook of Mathematical Functions*. Dover Publications. Dover Books on Mathematics.

[Bender and Steiner, 2012] Bender, C. and Steiner, J. (2012). Least-squares monte carlo for backward sdes. *Numer. Methods Finance*, 12:257–289.

[Bender and Zhang, 2008] Bender, C. and Zhang, J. (2008). Time discretization and markovian iteration for coupled fbsdes. *Ann. Appl. Probab.*, 18:143–177.

[Bouchard and Touzi, 2004] Bouchard, B. and Touzi, N. (2004). Discrete-time approximation and monte-carlo simulation of backward stochastic differential equations. *Stoch. Proc. Appl.*, 111:175–206.

[Burden and Faires, 2001] Burden, R. L. and Faires, J. D. (2001). *Numerical Analysis*. Higher Education Press/Cengage Learning. 7th ed.

[Butcher, 2008] Butcher, J. C. (2008). *Numerical methods for ordinary differential equations*. John Wiley, Chichester, UK.

[Crisan and Chassagneux, 2014] Crisan, D. and Chassagneux, J. F. (2014). Runge-kutta schemes for backward stochastic differential equations. *Ann. Appl. Probab.*, 24:679–720.

[Crisan and Manolarakis, 2010] Crisan, D. and Manolarakis, K. (2010). Solving backward stochastic differential equations using the cubature method: Application to nonlinear pricing. *SIAM J. FINAN. MATH*, 3(1):534–571.

[Cvitanic and Zhang, 2006] Cvitanic, J. and Zhang, J. (2006). The steepest descent method for forward-backward sdes. *Electron. J. Probab.*, 16:940–968.

[Delarue and Menozzi, 2006] Delarue, F. and Menozzi, S. (2006). A forward-backward stochastic algorithm for quasi-linear pdes. *Ann. Appl. Probab.*, 16(1):140–184.

[Douglas et al., 1996] Douglas, J., Ma, J., and Protter, P. (1996). Numerical methods for forward-backward stochastic differential equations. *Ann. Appl. Probab.*, 6:940–968.

[Fornberg, 1988] Fornberg, B. (1988). Generation of finite difference formulas on arbitrarily spaced grids. *Math. Comput.*, 51(184):699–706.

[Fu et al., 2017] Fu, Y., Zhao, W., and Zhou, T. (2017). Efficient spectral sparse grid approximations for solving multi-dimensional forward backward sdes. *Discrete Cont. Dyn-B.*, 22(9):3439–3458.

[Gobet et al., 2005] Gobet, E., Lemor, J. P., and Warin, X. (2005). A regression-based monte carlo method to solve backward stochastic differential equations. *Ann. Appl. Probab.*, 15:2172–2202.

[Lemor et al., 2006] Lemor, J., Gobet, E., and Warin, X. (2006). Rate of convergence of an empirical regression method for solving generalized backward stochastic differential equations. *Bernoulli*, 12:889–916.

[Lepeltier and Martin, 1997] Lepeltier, J. P. and Martin, J. S. (1997). Backward stochastic differential equations with continuous generator. *Statist. Probab. Lett.*, 32(425–430).

[Ma et al., 1994] Ma, J., Protter, P., and Yong, J. (1994). Solving forward-backward stochastic differential equations explicity-a four step scheme. *Probab. Theory Related Fields*, 98(3):339–359.

[Ma et al., 2008] Ma, J., Shen, J., and Zhao, Y. (2008). On numerical approximations of forward-backward stochastic differential equations. *SIAM J. Numer. Anal.*, 46(5):2636–2661.

[Ma and Zhang, 2005] Ma, J. and Zhang, J. (2005). Representations and regularities for solutions to bsdes with reflections. *Stoch. Proc. Appl.*, 115:539–569.

[Milsetin and Tretyakov, 2006] Milsetin, G. N. and Tretyakov, M. V. (2006). Numerical algorithms for forward-backward stochastic differential equations. *SIAM J. SCI. COMPUT.*, 28:561–582.

[Pardoux and Peng, 1990] Pardoux, E. and Peng, S. (1990). Adapted solution of a backward stochastic differential equations. *System and Control Letters*, 14:55–61.

[Pardoux and Peng, 1992] Pardoux, E. and Peng, S. (1992). Backward stochastic differential equation and quasilinear parabolic partial differential equations. *Lectures Notes in CSI.*, 176:200–217.

[Peng, 1991] Peng, S. (1991). Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. *Stochastics and Stochastic Reports*, 37(1–2):61–74.

[Peng and Wu, 1999] Peng, S. and Wu, Z. (1999). Fully coupled forward-backward stochastic differential equations and applications to optimal control. *SIAM J. Control Optim.*, 37:825–843.

[Ruijter and Oosterlee, 2015] Ruijter, M. J. and Oosterlee, C. W. (2015). A fourier cosine method for an efficient computation of solutions to bsdes. *SIAM J. SCI. COMPUT.*, 37(2):A859–A889.

[Shen et al., 2011] Shen, J., Tang, T., and Wang, L. (2011). *Spectral Methods: Algorithms, Analysis and Applications*. Springer-Verlag, Berlin.

[Teng, 2019] Teng, L. (2019). A review of tree-based approaches to solve forward-backward stochastic differential equations. *arXiv:1809.00325v4, available on webpage at* `https://arxiv.org/pdf/1809.00325v4.pdf`.

[Teng et al., 2020] Teng, L., Lapitckii, A., and Günther, M. (2020). A multi-step scheme based on cubic spline for solving backward stochastic differential equations. *Appl. Numer. Math.*, 150.

[Zhang et al., 2013] Zhang, G., Gunzburger, M., and Zhao, W. (2013). A sparse-grid method for multi-dimensional backward stochastic differential equations. *J. Comput. Math.*, 31(3):221–248.

[Zhang, 2001] Zhang, J. (2001). *Some fine properties of backward stochastic differential equations*. PhD thesis, Purdue University, West Lafayette, IN.

[Zhang, 2004] Zhang, J. (2004). A numerical scheme for bsdes. *Ann. Appl. Probab.*, 14:459–488.

[Zhao et al., 2006] Zhao, W., Chen, L., and Peng, S. (2006). A new kind of accurate numerical method for backward stochastic differential equations. *SIAM J. SCI. COMPUT.*, 28(4):1563–1581.

[Zhao et al., 2014a] Zhao, W., Fu, Y., and Zhou, T. (2014a). New kinds of high-order multi-step schemes for coupled forward backward stochastic differential equations. *SIAM J. SCI. COMPUT.*, 36(4):A1731–A1751.

[Zhao et al., 2013] Zhao, W., Li, Y., and Ju, L. (2013). Error estimates of the crank-nicolson scheme for solving backward stochastic differential equations. *Int. J. Numer. Anal. Mode.l*, 10(4):876–898.

[Zhao et al., 2012] Zhao, W., Li, Y., and Zhang, G. (2012). A generalized θ-scheme for solving backward stochastic differential equations. *Discrete Cont. Dyn-B.*, 17(5):1585–1603.

[Zhao et al., 2009] Zhao, W., Wang, J., and Peng, S. (2009). Error estimates of the theta-scheme for backward stochastic differential equations. *Discrete Contin. Dyn. Syst. Ser. B*, 12:905–924.

[Zhao et al., 2010] Zhao, W., Zhang, G., and Ju, L. (2010). A stable multistep scheme for solving backward stochastic differential equations. *SIAM J. NUMER. ANAL.*, 48:1369–1394.

[Zhao et al., 2014b] Zhao, W., Zhang, W., and Ju, L. (2014b). A numerical method and its error estimates for the decoupled forward-backward stochastic differential equations. *Commun. Comput. Phys.*, 15:618–646.