East Asian Journal on Applied Mathematics
doi: 10.4208/eajam.280515.211015a

Vol. **5**, No. 4, pp. 387-404
November 2015

Convergence of Recent Multistep Schemes for a Forward-Backward Stochastic Differential Equation

Jie Yang and Weidong Zhao*

School of Mathematics & Finance Institute, Shandong University, Jinan 250100, China.

Received 28 May 2015; Accepted (in revised version) 21 October 2015.

Abstract. Convergence analysis is presented for recently proposed multistep schemes, when applied to a special type of forward-backward stochastic differential equations (FBSDEs) that arises in finance and stochastic control. The corresponding k-step scheme admits a k-order convergence rate in time, when the exact solution of the forward stochastic differential equation (SDE) is given. Our analysis assumes that the terminal conditions and the FBSDE coefficients are sufficiently regular.

AMS subject classifications: 60H35, 65C20, 60H10

Key words: Convergence analysis, multistep schemes, forward-backward stochastic differential equations.

1. Introduction

Let $(\Omega, \mathscr{F}, P; \{\mathscr{F}_t\}_{0\le t\le T})$ be a complete filtered probability space, where $\{\mathscr{F}_t\}_{0\le t\le T}$ represents the filtration generated by a standard d-dimensional Brownian motion defined by the vector $W_t = (W_t^1, W_t^2, \cdots, W_t^d)^T$ in which the superscript T denotes the transpose, and all the P-null sets are augmented to each σ-algebra $\mathscr{F}_t$. We consider the $\{\mathscr{F}_t\}_{0\le t\le T}-$adapted triplet $(X,Y,Z) = \{(X_t, Y_t, Z_t) : t \in [0,T]\}$ that satisfies the following forward backward stochastic differential equations (FBSDEs):

$$\begin{cases} X_t = X_0 + \displaystyle\int_0^t b(s, X_s)\mathrm{d}s + \int_0^t \sigma(s, X_s)\mathrm{d}W_s \quad \text{(SDE)}, \\ Y_t = \xi + \displaystyle\int_t^T f(s, X_s, Y_s, Z_s)\mathrm{d}s - \int_t^T Z_s \mathrm{d}W_s \quad \text{(BSDE)}, \end{cases} \tag{1.1}$$

where T is a fixed positive number, $b : \Omega \times [0,T] \times \mathbb{R}^q \to \mathbb{R}^q$ and $\sigma : \Omega \times [0,T] \times \mathbb{R}^q \to \mathbb{R}^{q\times d}$ are appropriate measurable functions called the drift and diffusion coefficients of the stochastic differential equation (SDE) respectively, $f : \Omega \times [0,T] \times \mathbb{R}^q \times \mathbb{R}^p \times \mathbb{R}^{p\times d} \to \mathbb{R}^p$

*Corresponding author. *Email addresses:* yangjie218@yeah.net (J. Yang), wdzhao@sdu.edu.cn (W. Zhao)

 ©2015 Global-Science Press

https://doi.org/10.4208/eajam.280515.211015a Published online by Cambridge University Press

is some appropriate measurable function called the generator of the backward stochastic differential equation (BSDE), and $\xi \in \mathscr{F}_T$ is its terminal condition.

In 1990, Pardoux & Peng [13] proved the existence and uniqueness of the solutions of the above decoupled FBSDEs. Since then, FBSDEs have been extensively studied, and applications of FBSDEs have been found in many important fields such as mathematical finance, stochastic control, risk measure, partial differential equations (PDEs) and nonlinear expectations and so on [6, 9, 10, 14–16]. There have been many articles on FBSDEs — cf. [8–10, 15] and references therein. In particular, under some standard assumptions Peng [15] proved the nonlinear Feynman-Kac formula

$$Y_s = u(s,X_s), \quad Z_s = \nabla_x u(s,X_s)\sigma(s,X_s), \quad \forall s \in [t,T)\,, \tag{1.2}$$

where $u \in C^{1,2}$ is the solution of the partial differential equation (PDE)

$$L^0_{t,x}u(t,x) + f\big(t,x,u(t,x),\nabla_x u(t,x)\sigma(t,x)\big) = 0 \tag{1.3}$$

with the terminal condition $u(T,x) = \varphi(x)$. Here $L^0_{t,x}$ is the second order differential operator defined by

$$L^0_{t,x} = \frac{\partial}{\partial t} + \sum_i b_i(t,x)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j}(\sigma\sigma^T)_{i,j}(t,x)\frac{\partial^2}{\partial x_i \partial x_j}\,. \tag{1.4}$$

These results give a precise link between the solutions of the FBSDEs and partial differential equations. In Ref. [17], some properties of the solutions of FBSDEs under weaker conditions were studied using the Malliavin derivative.

It is well known that analytic solutions (X_t, Y_t, Z_t) of FBSDEs can seldom be expressed in explicit closed-form, so great effort has been devoted to their numerical solution. Some are numerical schemes with low order convergence rates [1, 2, 5, 7, 11, 18] under lower regularity assumptions, while others are high-order numerical methods [4, 19, 21–23]. It is notable that most of the numerical methods are designed for BSDE, or decoupled FBSDEs. It is very difficult to design high-order numerical schemes for coupled FBSDEs as one usually needs to solve the forward SDE with high-order accuracy, which is prevented by the multiple stochastic integral. Recently, a novel high-order multistep numerical scheme for FBSDEs has been proposed [20], where the solution pair (Y,Z) admits high-order convergence rates even when the Euler scheme (a typical low order scheme) is used to solve the SDE.

The main purpose of this work is to prove rigorously the high-order convergence rates of the multistep schemes in Ref. [20]. To demonstrate the main idea, we narrow our discussion to a special type of decoupled FBSDEs which have applications in finance and stochastic contorl, and prove that the k-step scheme admits a k-order convergence rate in time for the solution pair (Y,Z). The analysis is carried out under sufficient regularity assumptions on the terminal conditions and on the coefficients of the FBSDEs. Our analysis invokes numerical ODE theory, a variational formula and Malliavin derivative theory. In Section 2, the multistep schemes introduced in Ref, [20] are briefly reviewed, together with some useful lemmas. Section 3 presents our main result — viz. that the k-step scheme in [20] admits a k-order convergence rate in time. Some concluding remarks are made in Section 4.

https://doi.org/10.4208/eajam.280515.211015a Published online by Cambridge University Press

2. Preliminaries

Let us first summarise some notation that we use.

- $|\cdot|$: the Euclidean norm in the Euclidean space $\mathbb{R}$.
- $\mathscr{F}_s^{t,x}(t \le s \le T)$: the σ-algebra generated by the diffusion process $\{X_r, t \le r \le s, X_t = x\}$.
- $\mathbb{E}_s^{t,x}[\eta]$: the conditional expectation of the random variable η under $\mathscr{F}_s^{t,x}$, i.e., $\mathbb{E}_s^{t,x}[\eta] = \mathbb{E}[\eta|\mathscr{F}_s^{t,x}]$. When $s = t$, we use $\mathbb{E}_t^x[\eta]$ to denote $\mathbb{E}[\eta|\mathscr{F}_t^{t,x}]$.
- $\|\eta\|_\infty := \sup_\omega |\eta(\omega)|$, $\|\eta\|_{L^p} := (\mathbb{E}[|\eta|^p])^{1/p}, \quad 1 \le p < \infty$.
- C_b^k: the set of function $\varphi(x)$ with continuous bounded derivatives up to and including the order k.
- C^{k_1,k_2}: the set of function $f(t,x)$ with continuous partial derivatives up to and including k_1 with respect to t, and up to and including k_2 with respect to x.

As mentioned, we narrow our attention to a special type of FBSDEs — viz.

$$\begin{cases} X_s^{t,x} = x + \int_t^s b(r, X_r^{t,x})dr + \int_t^s \sigma(r, X_r^{t,x})\mathrm{d}W_r\,, \\ Y_s^{t,x} = \varphi(X_T) + \int_s^T f(r, X_r^{t,x}, Y_r^{t,x})dr - \int_s^T Z_r^{t,x} dW_r \end{cases} \tag{2.1}$$

for $t \le s \le T$, where the generator function f is independent of Z. Thus although the numerical schemes in Ref. [20] are applicable to general FBSDEs, our theoretical analysis in this article specifically addresses the problem (2.1), and we also assume that the terminal condition ξ is a deterministic function $\varphi(X_T)$. The superscript ^{t,x} is used to signify that the forward SDE starts from time-space point $(t, X_t = x)$, but such superscripts can be omitted if no ambiguity arises. This problem is well-posed under standard assumptions [13, 18].

2.1. Reference equations

For the numerical scheme of Ref. [20], we first uniformly divide the interval $[0, T]$ into N parts with time step $\Delta t = T/N$ and grid $t_n = n\Delta t$, $n = 0, 1, \cdots, N$, and write $\forall j = 0, 1, \cdots, N-n$:

$$\Delta t_{n,j} = t_{n+j} - t_n = j\Delta t, \qquad \Delta W_{t_n,t} = W_t - W_{t_n}, \qquad \Delta W_{n,j} = W_{t_{n+j}} - W_{t_n}.$$

Let (X_t, Y_t, Z_t) be the solution of (2.1) with terminal condition $\xi = \varphi(X_T)$. Then for any $t \in [t_n, T]$,

$$Y_{t_n} = Y_t + \int_{t_n}^t f(s, X_s, Y_s)ds - \int_{t_n}^t Z_s dW_s\,. \tag{2.2}$$

https://doi.org/10.4208/eajam.280515.211015a Published online by Cambridge University Press

Taking the conditional expectation $\mathbb{E}^x_{t_n}[\cdot]$ on both sides of Eq. (2.2), we obtain

$$\mathbb{E}^x_{t_n}[Y_t] = \mathbb{E}^x_{t_n}[\xi] + \int_t^T \mathbb{E}^x_{t_n}[f(s,X_s,Y_s)]ds\,. \tag{2.3}$$

Further, on multiplying both sides of Eq. (2.2) by $(\Delta W_{t_n,t})^T$ and then taking the conditional expectation $\mathbb{E}^x_{t_n}[\cdot]$ on both sides of the resulting equation, we obtain

$$0 = \mathbb{E}^x_{t_n}[Y_t(\Delta W_{t_n,t})^T] + \int_{t_n}^t \mathbb{E}^x_{t_n}[f(s,X_s,Y_s)(\Delta W_{t_n,t})^T]ds - \int_{t_n}^t \mathbb{E}^x_{t_n}[Z_s]ds\,. \tag{2.4}$$

The integrands $\mathbb{E}^x_{t_n}[f(s,X_s,Y_s)]$, $\mathbb{E}^x_{t_n}[f(s,X_s,Y_s)\Delta(W_{t_n,t})^T]$ and $\mathbb{E}^x_{t_n}[Z_s]$ are continuous functions of s, under standard assumptions. Differentiating with respect to t, Eqs. (2.3) and (2.4) respectively yield the following two reference ordinary differential equations (ODEs):

$$\frac{\mathrm{d}\mathbb{E}^x_{t_n}[Y_t]}{\mathrm{d}t} = -\mathbb{E}^x_{t_n}[f(t,X_t,Y_t)], \tag{2.5}$$

$$\frac{\mathrm{d}\mathbb{E}^x_{t_n}[Y_t(\Delta W_{t_n,t})^T]}{\mathrm{d}t} = -\mathbb{E}^x_{t_n}[f(t,X_t,Y_t)(\Delta W_{t_n,t})^T] + \mathbb{E}^x_{t_n}[Z_t]. \tag{2.6}$$

Based on the above two reference ODEs, efficient multistep schemes may be obtained by carefully approximating the derivatives and conditional expectations [20].

2.2. Multistep numerical schemes

We now briefly review the multistep schemes proposed in Ref. [20] based on the reference equations (2.5) and (2.6), after a summary of some classical results on approximating derivatives.

Let $u \in C_b^{k+1}$ and $\{t_i | t_i \in \mathbb{R},\ t_0 < t_1 < \cdots < t_k,\ k \in \mathbb{N}\}$. By Taylor's expansion, for each t_i, $i = 0, 1, \cdots, k$ we have

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

where $\Delta t_i = t_i - t_0$. Then

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

where $\{\alpha_{k,i}\}_{0\le i\le k}$ are real numbers. If we choose $\alpha_{k,i}$ such that

$$\frac{1}{j!}\sum_{i=0}^{k} \alpha_{k,i}(\Delta t_i)^j = \delta_{j1} = \begin{cases} 1\,, & j = 1\,, \\ 0\,, & j \neq 1\,, \end{cases} \tag{2.7}$$

https://doi.org/10.4208/eajam.280515.211015a Published online by Cambridge University Press

Table 1: Values of $\alpha_{k,i}\Delta t$, $i=0,1,\cdots,k$, $k=1,\cdots,6$

$\alpha_{k,i}\Delta t$	$i=0$	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$
$k=1$	-1	1					
$k=2$	$-\frac{3}{2}$	2	$-\frac{1}{2}$				
$k=3$	$-\frac{11}{6}$	3	$-\frac{3}{2}$	$\frac{1}{3}$			
$k=4$	$-\frac{25}{12}$	4	-3	$\frac{4}{3}$	$-\frac{1}{4}$		
$k=5$	$-\frac{137}{60}$	5	-5	$\frac{10}{3}$	$-\frac{5}{4}$	$\frac{1}{5}$	
$k=6$	$-\frac{49}{20}$	6	$-\frac{15}{2}$	$\frac{20}{3}$	$-\frac{15}{4}$	$\frac{6}{5}$	$-\frac{1}{6}$

we obtain

$$\frac{du}{dt}(t_0)=\sum_{i=0}^{k}\alpha_{k,i}u(t_i)+R_e\,, \tag{2.8}$$

where the truncation error is

$$R_e=\mathcal{O}\left(\sum_{i=0}^{k}\alpha_{k,i}(\Delta t_i)^{k+1}\right).$$

In particular, when $\Delta t_i=i\Delta t$, the conditions (2.7) are equivalent to the following linear system:

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1\\ 0 & 1 & 2 & \cdots & k\\ 0 & 1^2 & 2^2 & \cdots & k^2\\ \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 1^k & 2^k & \cdots & k^k \end{bmatrix}\times\begin{bmatrix} \alpha_{k,0}\Delta t\\ \alpha_{k,1}\Delta t\\ \alpha_{k,2}\Delta t\\ \vdots\\ \alpha_{k,k}\Delta t \end{bmatrix}=\begin{bmatrix} 0\\ 1\\ 0\\ \vdots\\ 0 \end{bmatrix}. \tag{2.9}$$

In Table 1, we list the quantities $\alpha_{k,i}\Delta t$ $(i=0,1,\cdots,k)$ of the system for $1\le k\le 6$. For a fixed k-step scheme, it is notable that the quantity $\alpha_{k,i}\Delta t$ is a constant.

Under the scheme (2.8), we obtain

$$\left.\frac{\mathrm{d}\mathbb{E}^x_{t_n}[Y_t]}{\mathrm{d}t}\right|_{t=t_n}=\sum_{i=0}^{k}\alpha_{k,i}\mathbb{E}^x_{t_n}\left[Y_{t_{n+i}}\right]+R^k_{y,n}\,, \tag{2.10}$$

$$\left.\frac{\mathrm{d}\mathbb{E}^x_{t_n}[Y_t(\Delta W_{t_n,t})^T]}{\mathrm{d}t}\right|_{t=t_n}=\sum_{i=0}^{k}\alpha_{k,i}\mathbb{E}^x_{t_n}\left[Y_{t_{n+i}}(\Delta W_{t_n,t})^T\right]+R^k_{z,n}\,, \tag{2.11}$$

where $R^k_{y,n}$ and $R^k_{z,n}$ are the associated truncation errors — viz.

$$R^k_{y,n}=\left.\frac{\mathrm{d}\mathbb{E}^x_{t_n}[Y_t]}{\mathrm{d}t}\right|_{t=t_n}-\sum_{j=0}^{k}\alpha_{k,j}\mathbb{E}^x_{t_n}[Y_{t_{n+j}}], \tag{2.12}$$

$$R^k_{z,n}=\left.\frac{\mathrm{d}\mathbb{E}^x_{t_n}[Y_t\Delta W_{t_n,t}]}{\mathrm{d}t}\right|_{t=t_n}-\sum_{j=0}^{k}\alpha_{k,j}\mathbb{E}^x_{t_n}[Y_{t_{n+j}}(\Delta W_{n,j})^T]. \tag{2.13}$$

https://doi.org/10.4208/eajam.280515.211015a Published online by Cambridge University Press

Inserting Eqs. (2.10) and (2.11) into Eqs. (2.5) and (2.6) respectively, we get the following two reference equations for solving Y and Z:

$$\alpha_{k,0}Y_{t_n} = -\sum_{j=1}^{k}\alpha_{k,j}\mathbb{E}^x_{t_n}[Y_{t_{n+j}}] - f(t_n, X_{t_n}, Y_{t_n}) - R^k_{y,n}\,, \tag{2.14}$$

$$Z_{t_n} = \sum_{j=1}^{k}\alpha_{k,j}\mathbb{E}^x_{t_n}[Y_{t_{n+j}}(\Delta W_{n,j})^T] + R^k_{z,n}\,. \tag{2.15}$$

Then with Y^n and Z^n denoting the respective numerical approximations for the solutions of Y_t and Z_t at $t = t_n$, on removing the truncation error terms $R^k_{y,n}$ and $R^k_{z,n}$ from Eqs. (2.14) and (2.15) one has the following multistep scheme.

Scheme I: Assume that Y^{N-i} and Z^{N-i}, $i = 0, 1, \cdots, k-1$, are known. For $n = N-k, \cdots, 1, 0$, with $X^{t_n,x}_t$ being the solution of the SDE in (2.1), solve Y^n and Z^n by

$$\alpha_{k,0}Y^n = -\sum_{j=1}^{k}\alpha_{k,j}\mathbb{E}_{t_n}[Y^{n+j}] - f(t_n, X_{t_n}, Y^n), \tag{2.16}$$

$$Z^n = \sum_{j=1}^{k}\alpha_{k,j}\mathbb{E}_{t_n}[Y^{n+j}(\Delta W_{n,j})^T], \tag{2.17}$$

where Y^{n+j} are the approximation values of $Y_{t_{n+j}}$ at the space points $X^{t_n,x}_{t_{n+j}}$.

In Ref. [20], there are many numerical examples demonstrating that the k-step scheme admits k-order convergence rate in solving FBSDEs, where $k \le 6$. However, when $k > 6$ the scheme becomes unstable, as classical numerical ODE theory predicts [3]. Our main purpose here is to provide a theoretical justification for such high-order convergence.

2.3. Some useful lemmas

To show the convergence rate of the multistep scheme (Scheme I), we need some regularity assumptions and lemmas.

Lemma 2.1. *Suppose that the given data b, σ, f and ξ are sufficiently smooth, such that $(L^0_{t,x})^{k+1}u(t,x)$ and $(L^0_{t,x})^{k+1}\nabla u(t,x)$ (see definition (1.4)) are bounded for $(t,x) \in [0,T]\times\mathbb{R}^q$. Then we have the following local estimates:*

$$R^k_{y,n} = \mathcal{O}(\Delta t)^k\,, \qquad R^k_{z,n} = \mathcal{O}(\Delta t)^k\,. \tag{2.18}$$

Proof. Let A^x_t be the generator of a diffusion process X_t that satisfies the forward SDE in (2.1):

$$A^x_t g(t,x) = \lim_{s\downarrow t}\frac{\mathbb{E}^x_t[g(s,X_s)] - g(t,x)}{s-t}\,, \quad x \in \mathbb{R}\,.$$

https://doi.org/10.4208/eajam.280515.211015a Published online by Cambridge University Press

For $g \in C^{1,2}([0,T]\times\mathbb{R})$, we have

$$A_t^{X_t} g(t,X_t) = L^0_{t,X_t} g(t,X_t)\,.$$

Let $t_0 < t$ be fixed, and let $x_0 \in \mathbb{R}$ be a fixed point. If $\mathbb{E}^{x_0}_{t_0}[|L^0_{t,X_t} g(t,X_t)|] < +\infty$,

$$\frac{\mathrm{d}\mathbb{E}^{x_0}_{t_0}[g(t,X_t)]}{\mathrm{d}t} = \mathbb{E}^{x_0}_{t_0}\big[A_t^{X_t} g(t,X_t)\big] = \mathbb{E}^{x_0}_{t_0}[L^0_{t,X_t} g(t,X_t)], \quad t \geq t_0\,,$$

hence from the Feynman-Kac formula (1.2) we have

$$\frac{\mathrm{d}^{k+1}\mathbb{E}^{x_0}_{t_0}[Y(t,X_t)]}{\mathrm{d}t^{k+1}} = \frac{\mathrm{d}^{k+1}\mathbb{E}^{x_0}_{t_0}[u(t,X_t)]}{\mathrm{d}t^{k+1}} = \frac{\mathrm{d}^{k}\mathbb{E}^{x_0}_{t_0}[L^0_{t,X_t} u(t,X_t)]}{\mathrm{d}t^{k}} = \mathbb{E}^{x_0}_{t_0}[(L^0_{t,X_t})^{k+1} u(t,X_t)].$$

From (2.8), if $(L^0_{t,x})^{k+1}u(t,x)$ is bounded for all $(t,x)\in[0,T]\times\mathbb{R}$ it follows that

$$R^k_{y,n} = \mathscr{O}(\Delta t)^k.$$

The estimate for $R^k_{z,n} = \mathscr{O}(\Delta t)^k$ follows by using the same arguments. □

For the derivative approximation formula (2.8), we make a standard assumption as follows [3].

Assumption 2.1. The positive integers k are chosen such that the roots of the polynomial $P(\lambda) = \alpha_0\lambda^k + \cdots + \alpha_{k-1}\lambda + \alpha_k$ satisfy the root conditions — i.e. the roots of $P(\lambda)$ lie in the closed unit disk $|\lambda| \leq 1$ and those which lie on the unit circle $|\lambda| = 1$ are all simple roots.

Under assumption 2.1, we have the following classical result [3].

Lemma 2.2. *Suppose that all roots of the polynomial $P(\lambda) = \alpha_0\lambda^k + \cdots + \alpha_{k-1}\lambda + \alpha_k$ satisfy the root conditions. Let $\{\gamma_l\}_{l=0}^{\infty}$ be defined by*

$$\frac{1}{\alpha_0 + \cdots + \alpha_{k-1}\lambda^{k-1} + \alpha_k\lambda^k} = \gamma_0 + \gamma_1\lambda + \gamma_2\lambda^2 + \cdots\,.$$

Then it holds that

$$\Gamma :\equiv \sup_{l\geq 0}|\gamma_l| < +\infty\,.$$

We also introduce the discrete backward Gronwall inequality — e.g. see [22].

Lemma 2.3. *Suppose that N and k are two nonnegative integers with $k \leq N$ and let Δt be any positive number. Let $\{\eta_n\}$ be a series satisfying*

$$|\eta_n| \leq a + b\Delta t \sum_{j=n+1}^{N} |\eta_j|\,, \quad n = N-k, \cdots, 1, 0\,,$$

where a and b are two positive constants. If $T = N\Delta t$, then

$$|\eta_n| \leq e^{bT}\left(a + b\Delta t \sum_{j=N-k+1}^{N} |\eta_j|\right), \quad n = N-k, \cdots, 1, 0\,.$$

https://doi.org/10.4208/eajam.280515.211015a Published online by Cambridge University Press

3. Convergence Analysis for Scheme I

For the convergence analysis of Scheme I, we adopt the notation (for $1 \le j \le k$):

$$e_y^n := Y_{t_n} - Y^n, \qquad e_z^n := Z_{t_n} - Z^n,$$
$$e_y^{n+j} := Y_{t_{n+j}} - Y^{n+j}, \qquad e_f^n := f(t_n, X_{t_n}, Y_{t_n}) - f(t_n, X^n, Y^n).$$

3.1. Estimates of e_y^n

An estimate of e_y^n in the weak sense is given in the following theorem.

Theorem 3.1. *Let (Y^n, Z^n) be the numerical solution at $t = t_n$ obtained by Scheme I. Suppose that Assumption 2.1 holds with $\alpha_i = \alpha_{k,i}\Delta t$, $i = 0, 1, \cdots, k$, and $f(t, X, Y)$ is Lipschitz continuous with respect to Y with Lipschitz constant L. Then for $0 < \Delta t \le |\alpha_0| L^{-1}$,*

$$\max_{0\le n\le N-k} \mathbb{E}[|e_y^n|] \le C\left(\sum_{i=N-k+1}^{N} \mathbb{E}[|e_y^i|] + \sum_{i=0}^{N-k} \mathbb{E}\left[\left| \mathbb{E}_{t_n}[R_{y,i}^k] \right| \right] \Delta t \right), \tag{3.1}$$

where C is a positive constant depending only on T, L and k.

Proof. For each integer n such that $0 \le n \le N-k$, by subtracting (2.16) from (2.14) one obtains

$$\begin{aligned} \alpha_{k,0} e_y^n &= \alpha_{k,0}(Y_{t_n} - Y^n) \\ &= -\sum_{j=1}^{k} \alpha_{k,j} \mathbb{E}_{t_n}[Y_{t_{n+j}} - Y^{n+j}] - \big(f(t_n, X_{t_n}, Y_{t_n}) - f(t_n, X^n, Y^n) \big) + R_{y,n}^k \\ &= -\sum_{j=1}^{k} \alpha_{k,j} \mathbb{E}_{t_n}[e_y^{n+j}] - e_f^n + R_{y,n}^k . \end{aligned} \tag{3.2}$$

Multiplying both sides of the above equation by Δt and, for each n_0 ($0 < n_0 \le n$), taking the expectation $\mathbb{E}_{t_{n_0}}[\cdot]$ on both sides of the derived equation, one obtains

$$\alpha_{k,0} \Delta t\, \mathbb{E}_{t_{n_0}}[e_y^n] = -\sum_{j=1}^{k} \alpha_{k,j} \Delta t\, \mathbb{E}_{t_{n_0}}[e_y^{n+j}] - \mathbb{E}_{t_{n_0}}[e_f^n]\Delta t - \mathbb{E}_{t_{n_0}}[R_{y,n}^k]\Delta t .$$

Let $\beta_n = -e_f^n / e_y^n$. We then have $|\alpha_i| \le c$ for some positive constant c and $|\beta_n| \le L$, hence

$$\alpha_0 \mathbb{E}_{t_{n_0}}[e_y^n] + \alpha_1 \mathbb{E}_{t_{n_0}}[e_y^{n+1}] + \cdots + \alpha_k \mathbb{E}_{t_{n_0}}[e_y^{n+k}] = \mathbb{E}_{t_{n_0}}[\beta_n e_y^n]\Delta t - \mathbb{E}_{t_{n_0}}[R_{y,n}^k]\Delta t . \tag{3.3}$$

For the fixed integer k, let the numbers $\{\gamma_l\}_{l=0}^{N-n-k}$ be defined as in Lemma 2.2. Multiplying both sides of (3.3) with $n := n+l$ by γ_l ($l = 0, 1, \cdots, N-n-k$) and summing the resulting

https://doi.org/10.4208/eajam.280515.211015a Published online by Cambridge University Press

equations together, we get the summation (denoted by S_l) on the left-hand side

$$\begin{aligned}
S_l &= \Big(\alpha_0\mathbb{E}_{t_{n_0}}[e_y^n] + \alpha_1\mathbb{E}_{t_{n_0}}[e_y^{n+1}] + \cdots + \alpha_k\mathbb{E}_{t_{n_0}}[e_y^{n+k}]\Big)\gamma_0 \\
&\quad + \Big(\alpha_0\mathbb{E}_{t_{n_0}}[e_y^{n+1}] + \alpha_1\mathbb{E}_{t_{n_0}}[e_y^{n+2}] + \cdots + \alpha_k\mathbb{E}_{t_{n_0}}[e_y^{n+1+k}]\Big)\gamma_1 + \cdots \\
&\quad + \Big(\alpha_0\mathbb{E}_{t_{n_0}}[e_y^{N-k}] + \alpha_1\mathbb{E}_{t_{n_0}}[e_y^{N-n+1}] + \cdots + \alpha_k\mathbb{E}_{t_{n_0}}[e_y^{N}]\Big)\gamma_{N-n-k} \\
&= \alpha_0\gamma_0\mathbb{E}_{t_{n_0}}[e_y^n] + \Big(\alpha_0\gamma_1 + \alpha_1\gamma_0\Big)\mathbb{E}_{t_{n_0}}[e_y^{n+1}] + \cdots \\
&\quad + \Big(\alpha_0\gamma_{N-n-k} + \alpha_1\gamma_{N-n-k-1} + \cdots + \alpha_k\gamma_{N-n-2k}\Big)\mathbb{E}_{t_{n_0}}[e_y^{N-k}] \\
&\quad + \Big(\alpha_1\gamma_{N-n-k} + \alpha_2\gamma_{N-n-k-1} + \cdots + \alpha_k\gamma_{N-n-2k+1}\Big)\mathbb{E}_{t_{n_0}}[e_y^{N-k+1}] + \cdots \\
&\quad + \Big(\alpha_{k-1}\gamma_{N-n-k} + \alpha_k\gamma_{N-n-k-1}\Big)\mathbb{E}_{t_{n_0}}[e_y^{N-1}] + \alpha_k\gamma_{N-n-k}\mathbb{E}_{t_{n_0}}[e_y^{N}]\,,
\end{aligned}$$

and the summation (denoted by S_r) on the right-hand side

$$\begin{aligned}
S_r &= \Big(\gamma_0\mathbb{E}_{t_{n_0}}[\beta_n e_y^n] + \gamma_1\mathbb{E}_{t_{n_0}}[\beta_{n+1}e_y^{n+1}] + \cdots + \gamma_{N-n-k}\mathbb{E}_{t_{n_0}}[\beta_{N-k}e_y^{N-k}]\Big)\Delta t \\
&\quad - \Big(\gamma_0\mathbb{E}_{t_{n_0}}[R_{y,n}^k] + \gamma_1\mathbb{E}_{t_{n_0}}[R_{y,n+1}^k] + \cdots + \gamma_{N-n-k}\mathbb{E}_{t_{n_0}}[R_{y,N-k}^k]\Big)\Delta t\,.
\end{aligned}$$

We set $\gamma_l = 0$ for $l < 0$. Then, by noting the fact that

$$\alpha_0\gamma_l + \alpha_1\gamma_{l-1} + \cdots + \alpha_k\gamma_{l-k} = \begin{cases} 1 & \text{for } l = 0\,, \\ 0 & \text{for } l > 0\,, \end{cases} \tag{3.4}$$

we obtain

$$S_l = \mathbb{E}_{t_{n_0}}[e_y^n] + A\,,$$

where

$$\begin{aligned}
A &= \Big(\alpha_1\gamma_{N-n-k} + \alpha_2\gamma_{N-n-k-1} + \cdots + \alpha_k\gamma_{N-n-2k+1}\Big)\mathbb{E}_{t_{n_0}}[e_y^{N-k+1}] + \cdots \\
&\quad + \Big(\alpha_{k-1}\gamma_{N-n-k} + \alpha_k\gamma_{N-n-k-1}\Big)\mathbb{E}_{t_{n_0}}[e_y^{N-1}] + \alpha_k\gamma_{N-n-k}\mathbb{E}_{t_{n_0}}[e_y^{N}] \\
&= \sum_{i=1}^{k}\left(\sum_{j=i}^{k+1-i}\alpha_j\gamma_{N-k-n+i-j}\right)\mathbb{E}_{t_{n_0}}[e_y^{N-k+i}].
\end{aligned}$$

Then from the equality $S_l = S_r$, we derive

$$\mathbb{E}_{t_{n_0}}[e_y^n] - \gamma_0\Delta t\mathbb{E}_{t_{n_0}}[\beta_n e_y^n] = -A + \sum_{i=n+1}^{N-k}\gamma_{i-n}\mathbb{E}_{t_{n_0}}[\beta_i e_y^i]\Delta t + B_n\,,$$

where $B_n = -\sum_{i=n}^{N-k}\gamma_{i-n}\mathbb{E}_{t_{n_0}}[R_{y,i}^k]\Delta t$. Now, by letting $n = n_0$ we obtain

$$(1 - |\gamma_0\beta_0|\Delta t)|e_y^{n_0}| \le A_1 + B_{n_0}^1 + \sum_{i=n_0+1}^{N-k}|\gamma_{i-n_0}|\mathbb{E}_{t_{n_0}}[|\beta_i e_y^i|]\Delta t\,, \tag{3.5}$$

https://doi.org/10.4208/eajam.280515.211015a Published online by Cambridge University Press

where

$$A_1 = \sum_{i=1}^{k}\left(\sum_{j=i}^{k+1-i}\left|\alpha_j\gamma_{N-k-n_0+i-j}\right|\right)\mathbb{E}_{t_{n_0}}[|e_y^{N-k+i}|], \qquad B_{n_0}^1 = \sum_{i=n_0}^{N-k}\left|\gamma_{i-n_0}\right|\left|\mathbb{E}_{t_{n_0}}[R_{y,i}^k]\right|\Delta t\,.$$

Taking the expectation $\mathbb{E}[\cdot]$ on both sides of (3.5) and using Lemma 2.2 and that $|\beta_i| \le L$, one gets

$$(1-\Delta t L|\gamma_0|)\mathbb{E}[|e_y^{n_0}|] \le A_2 + B_n^2 + \Gamma L\sum_{i=n_0+1}^{N-k}\mathbb{E}[|e_y^i|]\Delta t\,,$$

where

$$A_2 = C\sum_{i=N-k+1}^{N}\mathbb{E}[|e_y^i|], \qquad B_{n_0}^2 = C\sum_{i=n_0}^{N-k}\mathbb{E}\Big[\left|\mathbb{E}_{t_{n_0}}[R_{y,i}^k]\right|\Big]\Delta t$$

with C being a positive constant that depends on k. Finally, from the arbitrariness of n_0 we have

$$(1-\Delta t L|\gamma_0|)\mathbb{E}[|e_y^n|] \le A_2 + \tilde{B}_n^2 + \Gamma L\sum_{i=n+1}^{N-k}\mathbb{E}[|e_y^i|]\Delta t\,,$$

where

$$\tilde{B}_n^2 = \Gamma\sum_{i=n}^{N-k}\mathbb{E}\Big[\left|\mathbb{E}_{t_n}[R_{y,i}^k]\right|\Delta t\Big].$$

Thus

$$\mathbb{E}[|e_y^n|] \le K^* + L^*\sum_{i=n+1}^{N-k}\mathbb{E}[|e_y^i|]\Delta t\,, \tag{3.6}$$

where

$$K^* = (A_2 + \tilde{B}_n^2)/(1-\Delta t L|\alpha_0^{-1}|), \qquad L^* = \Gamma L/(1-\Delta t L|\alpha_0^{-1}|)\,.$$

Here we use the fact that $\alpha_0\gamma_0 = 1$. Then from the Gronwall inequality in Lemma 2.3 we get

$$\begin{aligned}\mathbb{E}[|e_y^n|] &\le e^{L^*T}\left(K^* + L^*\Delta t\sum_{i=N-k+1}^{N}\mathbb{E}[|e_y^i|]\right)\\ &\le C\left(\sum_{j=N-k+1}^{N}\mathbb{E}[|e_y^i|] + \sum_{i=n}^{N-k}\mathbb{E}\Big[\left|\mathbb{E}_{t_n}[R_{y,i}^k]\right|\Big]\Delta t\right),\end{aligned} \tag{3.7}$$

so the proof is complete. □

Using a slightly different technique, we can provide a stronger estimate for e_y^n in L_∞ norm in the following Theorem.

https://doi.org/10.4208/eajam.280515.211015a Published online by Cambridge University Press

Theorem 3.2. *Under the conditions in Theorem* 3.1*, the following estimate holds:*

$$\max_{0\le n\le N-k}\|e_y^n\|_\infty \le C\left(\sum_{i=N-k+1}^{N}\|e_y^i\|_\infty + \sum_{i=0}^{N-k}\left\|\mathbb{E}_{t_n}[R_{y,i}^k]\right\|_\infty \Delta t\right), \tag{3.8}$$

where C is a positive constant depending only on T, L and k.

Proof. We use similar arguments to those in the proof of Theorem 3.1 until inequality (3.5) — i.e.

$$(1-\Delta t|\beta_0\gamma_0|)\,|e_y^{n_0}| \le A_1 + B_n^1 + \sum_{i=n_0+1}^{N-k}\left|\gamma_{i-n_0}\right|\mathbb{E}_{t_{n_0}}[|\beta_i e_y^i|]\Delta t\,,$$

with

$$A_1 = \sum_{i=1}^{k}\left(\sum_{j=i}^{k+1-i}\left|\alpha_j\gamma_{N-k-n_0+i-j}\right|\right)\mathbb{E}_{t_{n_0}}\left[|e_y^{N-k+i}|\right], \quad B_{n_0}^1 = \sum_{i=n_0}^{N-k}\left|\gamma_{i-n_0}\right|\left|\mathbb{E}_{t_{n_0}}[R_{y,i}^k]\right|\Delta t\,.$$

Taking the L_∞ norm and using Lemma 2.2, we get

$$(1-\Delta t L|\gamma_0|)\|e_y^{n_0}\|_\infty \le A_2 + B_n^2 + \Gamma L\sum_{i=n_0+1}^{N-k}\|e_y^i\|_\infty\Delta t\,,$$

where

$$A_2 = C\sum_{i=N-k+1}^{N}\|e_y^i\|_\infty\,, \qquad B_{n_0}^2 = \Gamma\sum_{i=n_0}^{N-k}\left\|\mathbb{E}_{t_{n_0}}[R_{y,i}^k]\right\|_\infty\Delta t\,.$$

Again, from the arbitrariness of n_0 we have

$$(1-\Delta t L|\gamma_0|)\|e_y^{n}\|_\infty \le A_2 + \tilde{B}_n^2 + \Gamma L\sum_{i=n+1}^{N-k}\|e_y^i\|_\infty\Delta t\,,$$

where

$$\tilde{B}_n^2 = \Gamma\sum_{i=n}^{N-k}\left\|\mathbb{E}_{t_n}[R_{y,i}^k]\right\|_\infty\Delta t\,.$$

Then we have

$$\|e_y^n\|_\infty \le K^* + L^*\sum_{i=n+1}^{N-k}\|e_y^i\|_\infty\Delta t\,, \tag{3.9}$$

with

$$K^* = (A_2+\tilde{B}_n^2)/(1-\Delta t L|\alpha_0^{-1}|)\,, \qquad L^* = \Gamma L/(1-\Delta t L|\alpha_0^{-1}|)\,.$$

https://doi.org/10.4208/eajam.280515.211015a Published online by Cambridge University Press

Applying the Gronwall inequality, we then obtain

$$\begin{aligned}\|e_y^n\|_\infty &\le e^{L^*T}\left(K^* + L^*\Delta t \sum_{i=N-k+1}^{N} \|e_y^i\|_\infty\right)\\ &\le C\left(\sum_{i=N-k+1}^{N} \|e_y^i\|_\infty + \sum_{i=n}^{N-k} \left\|\mathbb{E}_{t_n}[R_{y,i}^k]\right\|_\infty \Delta t\right),\end{aligned} \tag{3.10}$$

so the proof is complete. □

Combining the error inequality (3.8) and Lemma 2.1, we readily derive the following convergence result.

Theorem 3.3. *If the conditions in Lemma* 2.1 *and Theorem* 3.1 *hold, and the terminal data satisfy*

$$\max_{N-k<j\le N} \|e_y^j\|_\infty = \mathcal{O}(\Delta t)^k, \tag{3.11}$$

then for $0 < \Delta t \le |\alpha_0|L^{-1}$ *we have the following estimate:*

$$\max_{0\le n\le N-k} \|e_y^n\|_\infty \le C(\Delta t)^k, \tag{3.12}$$

where C *is a positive constant depending on* T *and the given data* b, σ, φ *and* f.

Proof. The proof follows by inserting the local estimation (2.18) for $|R_{y,n}^k|$ in Lemma 2.1 and the terminal condition (3.11) into (3.8). □

3.2. Estimates of e_z^n

In subsection 3.1, we have shown that the convergence rate for the Y-solution with the k-step scheme is k, and now turn to estimates for e_z^n. One can easily obtain a *non-optimal* estimate from Theorem 3.3, expressed in the following corollary.

Corollary 3.1. *Under the conditions as in Theorem* 3.3*, the following estimate holds:*

$$\max_{0\le n\le N-k} \|e_z^n\|_\infty \le C(\Delta t)^{k-\frac{1}{2}}, \tag{3.13}$$

where C *is a positive constant depending on* T *and the given data* b, σ, φ *and* f.

Proof. For $n = N-k, \cdots, 1, 0$, subtracting (2.17) from (2.15) we get

$$e_z^n = \sum_{j=1}^{k} \alpha_{k,j}\mathbb{E}_{t_n}[e_y^{n+j}(\Delta W_{n,j})^T] + R_{z,n}^k. \tag{3.14}$$

https://doi.org/10.4208/eajam.280515.211015a Published online by Cambridge University Press

Squaring both sides of this equation and invoking the inequality $(a+b)^2 \le 2(a^2+b^2)$ and the Hölder inequality, we obtain

$$\begin{aligned}|e_z^n|^2 &= \left(\sum_{j=1}^{k} \alpha_{k,j}\mathbb{E}_{t_n}\left[e_y^{n+j}\left(\Delta W_{n,j}\right)^T\right] + R_{z,n}^k\right)^2 \\ &\le 2\left(k\sum_{j=1}^{k}\alpha_{k,j}^2\mathbb{E}_{t_n}\left[\left|e_y^{n+j}(\Delta W_{n,j})^T\right|^2\right] + \left|R_{z,n}^k\right|^2\right) \\ &\le 2k\sum_{j=1}^{k}\alpha_{k,j}^2\mathbb{E}_{t_n}\left[\left|e_y^{n+j}\right|^2\right]\Delta t_{n,j} + 2\left|R_{z,n}^k\right|^2.\end{aligned}$$

Now taking the L_∞ norm on both sides and multiplying by Δt, one has

$$\begin{aligned}\|e_z^n\|_\infty^2\Delta t &\le C\sum_{j=1}^{k}\left|\alpha_{k,j}\Delta t\right|^2\left\|e_y^{n+j}\right\|_\infty^2 + 2\left|R_{z,n}^k\right|^2\Delta t \\ &\le C\left((\Delta t)^{2k} + (\Delta t)^{2k+1}\right). \qquad (3.15)\end{aligned}$$

The desired conclusion then follows from Theorem 3.3 and Lemma 2.1. □

The above corollary is not optimal, as we lose a half-order of the convergence rate, but we can resort to the variational formula and Malliavin derivative arguments.

3.2.1. Optimal estimate of e_z^n

Let us first consider the variational equation of the BSDE

$$\nabla Y_t = \varphi_x'\nabla X_T + \int_t^T \left(f_x'(s,X_s,Y_s)\nabla X_s + f_y'(s,X_s,Y_s)\nabla Y_s\right)\mathrm{d}s - \int_t^T \nabla Z_s \mathrm{d}W_s, \qquad (3.16)$$

where the forward process X_s starts at the time-space point $(t, X_t = x)$, and the variation ∇X_s of X_s with respect to x satisfies the SDE

$$\nabla X_s = 1 + \int_t^s b_x'(r,X_r)\nabla X_r \mathrm{d}r + \int_t^s \sigma_x'(r,X_r)\nabla X_r \mathrm{d}W_r \qquad (3.17)$$

where ∇Y_s and ∇Z_s denote the variations of Y_s and Z_s with respect to the spatial variable x, respectively. From the Feynman-Kac formula, the variation ∇Y_s can be expressed as

$$\nabla Y_s^{t,x} = V(s, X_s^{t,x}, \nabla X_s^{t,x}),$$

where V is the solution of the partial differential equation

$$\bar{L}_{t,x}^0 V(t,x,w) + F(t,x,w,u(t,x),V(t,x,w)) = 0 \qquad (3.18)$$

https://doi.org/10.4208/eajam.280515.211015a Published online by Cambridge University Press

with the terminal condition $V(T,x,w)=\varphi'_x(x)w$, $\bar{L}^0_{t,x}$ is the corresponding second order differential operator, and

$$F(s,x,w,u,v)=f'_x(s,x,u)w+f'_y(s,x,u)v\,.$$

From Eq. (3.16), for $t\in[t_n,T]$ we deduce that

$$\mathbb{E}^x_{t_n}[\nabla Y_t]=\mathbb{E}^x_{t_n}[\nabla Y_T]+\int_t^T \mathbb{E}^x_{t_n}\left[f'_x(s,X_s,Y_s)\nabla X_s+f'_y(s,X_s,Y_s)\nabla Y_s\right]\mathrm{d}s\,, \tag{3.19}$$

and then differentiating Eq. (3.19) with respect to t

$$\frac{\mathrm{d}\mathbb{E}^x_{t_n}[\nabla Y_t]}{\mathrm{d}t}=-\mathbb{E}^x_{t_n}\left[f'_x(t,X_t,Y_t)\nabla X_t+f'_y(t,X_t,Y_t)\nabla Y_t\right],\quad t\in[t_n,T]\,, \tag{3.20}$$

such that

$$\left.\frac{\mathrm{d}\mathbb{E}^x_{t_n}[\nabla Y_t]}{\mathrm{d}t}\right|_{t=t_n}=\sum_{j=0}^k\alpha_{k,j}\mathbb{E}^x_{t_n}[\nabla Y_{t_{n+j}}]+R^k_{\nabla y,n}\,, \tag{3.21}$$

where

$$R^k_{\nabla y,n}=\left.\frac{\mathrm{d}\mathbb{E}^x_{t_n}[\nabla Y_t]}{\mathrm{d}t}\right|_{t=t_n}-\sum_{j=0}^k\alpha_{k,j}\mathbb{E}^x_{t_n}[\nabla Y_{t_{n+j}}].$$

On inserting Eq. (3.21) into Eq. (3.19), we get

$$\begin{aligned}\alpha_{k,0}\nabla Y_{t_n}=&-\sum_{j=1}^k\alpha_{k,j}\mathbb{E}^x_{t_n}[\nabla Y_{t_{n+j}}]-\Big(f'_x(t_n,X_{t_n},Y_{t_n})\nabla X_{t_n}\\&+f'_y(t_n,X_{t_n},Y_{t_n})\nabla Y_{t_n}\Big)-R^k_{\nabla y,n}\,.\end{aligned} \tag{3.22}$$

Under conditions in Lemma 2.1, it is easy to show the boundedness of $\mathbb{E}[|\nabla Y_t|]$ — viz.

$$\sup_{t\in[0,T]}\mathbb{E}[|\nabla Y_t|]\le C\,, \tag{3.23}$$

where $C>0$ is a constant depending only on T and the given data b, σ, φ and f. Furthermore, similar to the local estimate of $|R^k_{y,n}|$ in Lemma 2.1, if $(\bar{L}^0_{t,x})^{k+1}V(t,x,w)$ is bounded for all (t,x,w), we have the local estimate

$$\max_{0\le n\le N-k}|R^k_{\nabla y,n}|\le C(\Delta t)^k\,. \tag{3.24}$$

Moreover, the variation process ∇Y^n of the solution Y^n of Scheme I satisfies

$$\alpha_{k,0}\nabla Y^n=-\sum_{j=1}^k\alpha_{k,j}\mathbb{E}^x_{t_n}[\nabla Y^{n+j}]-f'_x(t_n,X_{t_n},Y^n)\nabla X_{t_n}-f'_y(t_n,X_{t_n},Y^n)\nabla Y^n\,.$$

https://doi.org/10.4208/eajam.280515.211015a Published online by Cambridge University Press

with $\nabla Y^N = \varphi_x' \nabla X_T$. Note that the forward process X_s starts from x at the time $t = t_n$, and ∇Y^n is the variation of Y^n with respect to x.

Let ∇Y^{n+j} denote the values of ∇Y^{n+j} at the space points $X_{t_{n+j}}$, $j = 0, 1, \cdots, k$, and set

$$e_{\nabla y}^{n+j} = \nabla Y_{t_{n+j}} - \nabla Y^{n+j}, \quad n = N-k, \cdots, 1, 0\,.$$

We then reach the following lemma, for which one can resort to the proof of Theorem 3.2.

Lemma 3.1. *Under conditions in Theorem* 3.1*, and furthermore, we assume that* $f_x'(t,x,y)$ *and* $f_y'(t,x,y)$ *are uniformly Lipschitz continuous with a common bounded constant* L*. Then for* $0 < \Delta t \le |\alpha_0| L^{-1}$*, we have*

$$\max_{0\le n\le N-k} \|e_{\nabla y}^n\|_\infty \le C\left(\sum_{i=N-k+1}^{N} \|e_{\nabla y}^i\|_\infty + \sum_{i=0}^{N-k} \left\| \mathbb{E}_{t_n}\left[R_{\nabla y,n}^k \right] \right\|_\infty \Delta t + \sum_{i=0}^{N-k} \|e_y^i\|_\infty \Delta t \right),$$

where C *is a positive constant depending on* T, L *and* k*.*

We now let D denote the Malliavin derivative operator, and $D^{1,2}$ the domain of D in $L^2(\Omega)$ — i.e. $D^{1,2}$ is the closure of the class of smooth random variable F endowed with the norm $\|F\|_{1,2} = \left(\mathbb{E}[|F|^2] + \mathbb{E}[\|DF\|_H^2]\right)^{1/2}$, where $\|\cdot\|_H$ is the norm of a real separable Hilbert space H. The following two lemmas on the Malliavin derivative operator D hold [12].

Lemma 3.2. *For* $F \in D^{1,2}$ *and* $u \in L^2(\Omega; H)$*, the following integration-by-parts formula holds*

$$\mathbb{E}\left[\int_0^T u_t D_t F \mathrm{d}t \right] = \mathbb{E}\left[F \int_0^T u_t \mathrm{d}W_t \right].$$

Lemma 3.3. *Suppose that the solution* Y_t *of the BSDE in* (2.1) *is in* $D^{1,2}$*, then we have*

$$D_t Y_s I_{t\le s} = \nabla Y_s (\nabla X_t)^{-1} \sigma(t, X_t) I_{t\le s}\,.$$

We are now able to give the following estimation of e_z^n in the L_∞ norm.

Theorem 3.4. *Under the conditions of Lemmas* 2.1 *and* 3.1*, for* $0 < \Delta t \le |\alpha_0| L^{-1}$*,*

$$\max_{0\le n\le N-k} \|e_z^n\|_\infty \le C\left(\sum_{j=N-k+1}^{N} \|e_{\nabla y}^i\|_\infty + \sum_{i=0}^{N} \left\| \mathbb{E}_{t_n}\left[R_{\nabla y,i}^k \right] \right\|_\infty \Delta t + \sum_{i=0}^{N-k} \|e_y^i\|_\infty \Delta t + \max_{0\le n\le N-k} \|R_{z,n}^k\|_\infty \right),$$

where C *is positive constant that depends on* T*,* L *and* k*.*

https://doi.org/10.4208/eajam.280515.211015a Published online by Cambridge University Press

Proof. We use the superscript (t_n, x) to emphasise the relation $X_{t_n} = x$ for the Y-solution. The conditions of the theorem imply that $Y \in D^{1,2}$. From Lemmas 3.2 and 3.3,

$$\begin{aligned}\mathbb{E}^x_{t_n}\left[Y^{t_n,x}_{t_{n+j}}(\Delta W_{n,j})^T\right] &= \int_{t_n}^{t_{n+j}} \mathbb{E}^x_{t_n}\left[D_t Y^{t_n,x}_{t_{n+j}}\right]\mathrm{d}t \\ &= \int_{t_n}^{t_{n+j}} \mathbb{E}\Big[\nabla Y^{t_n,x}_{t_{n+j}}(\nabla X^{t_n,x}_t)^{-1}\sigma(t, X^{t_n,x}_t)\Big]\mathrm{d}t\,, \qquad (3.25)\end{aligned}$$

$$\mathbb{E}^x_{t_n}\left[Y^{n+j}(\Delta W_{n,j})^T\right] = \int_{t_n}^{t_{n+j}} \mathbb{E}^x_{t_n}\left[\nabla Y^{n+j}(\nabla X^{t_n,x}_t)^{-1}\sigma(t, X^{t_n,x}_t)\right]\mathrm{d}t\,. \qquad (3.26)$$

On subtracting Eq. (2.17) from Eq. (2.15), we get

$$e^n_z = Z_{t_n} - Z^n = \sum_{j=1}^{k} \alpha_{k,j}\mathbb{E}^x_{t_n}[(Y^{t_n,x}_{t_{n+j}} - Y^{n+j})(\Delta W_{n,j})^T] - R^k_{z,n}\,, \qquad (3.27)$$

and inserting Eqs. (3.25) and (3.26) into Eq. (3.27) yields

$$\begin{aligned}|e^n_z| &\le \sum_{j=1}^{k}\left|\alpha_{k,j}\right|\left|\mathbb{E}^x_{t_n}[(Y_{t_{n+j}} - Y^{n+j})(\Delta W_{n,j})^T]\right| + \left|R^k_{z,n}\right| \\ &\le \sum_{j=1}^{k}\left|\alpha_{k,j}\right|\int_{t_n}^{t_{n+j}}\left|\mathbb{E}^x_{t_n}\left[\left(\nabla Y_{t_{n+j}} - \nabla Y^{n+j}\right)(\nabla X^{t_n,x}_t)^{-1}\sigma(t, X^{t_n,x}_t)\right]\right|\mathrm{d}t + \left|R^k_{z,n}\right| \\ &\le \sum_{j=1}^{k}\left|\alpha_{k,j}\right|\int_{t_n}^{t_{n+j}}\|e^{n+j}_{\nabla y}\|_\infty\left|\mathbb{E}^x_{t_n}[(\nabla X^{t_n,x}_t)^{-1}\sigma(t, X^{t_n,x}_t)]\right|\mathrm{d}t + \left|R^k_{z,n}\right|. \qquad (3.28)\end{aligned}$$

Using Itô's formula, one can show that the process $(\nabla X)^{-1}$ solves the SDE

$$\Phi_s = 1 - \int_{t_n}^{s}\Phi_r\left(b'_x(r, X_r) - (\sigma'_x(r, X_r))^2\right)\mathrm{d}r - \int_{t_n}^{s}\Phi_r\sigma'_x(r, X_r)\mathrm{d}W_r, \qquad (3.29)$$

such that

$$\left|\mathbb{E}^x_{t_n}\left[(\nabla X^{t_n,x}_t)^{-1}\sigma(t, X^{t_n,x}_t)\right]\right| \le C\,.$$

Then taking the L_∞ norm on both sides of Eq. (3.27), we obtain

$$\begin{aligned}\|e^n_z\|_\infty &\le C\sum_{j=1}^{k}|\alpha_{k,j}\Delta t|\|e^{n+j}_{\nabla y}\|_\infty + \|R^k_{z,n}\|_\infty \\ &\le C\left(\sum_{j=N-k+1}^{N}\|e^i_{\nabla y}\|_\infty + \sum_{i=n}^{N}\left\|\mathbb{E}_{t_n}[R^k_{\nabla y,i}]\right\|_\infty\Delta t + \sum_{i=n}^{N-k}\|e^i_y\|_\infty\Delta t + \|R^k_{z,n}\|_\infty\right),\end{aligned}$$

The desired conclusion then follows from Lemma 3.1. □

From Lemma 2.1, Theorems 3.3 and 3.4, and the estimate (3.24), it is easy to obtain the convergence rate for the Z-solution summarised in the following theorem.

https://doi.org/10.4208/eajam.280515.211015a Published online by Cambridge University Press

Theorem 3.5. *Assume that the conditions in Lemmas* 2.1 *and* 3.1 *hold,* $(\bar{L}^0_{t,x})^{k+1}V$ *is bounded with* $\bar{L}^0_{t,x}$ *defined in* (3.18)*, and that the terminal values satisfy*

$$\max_{N-k<j\le N} \|e^j_{\nabla y}\|_\infty = \mathcal{O}(\Delta t)^k.$$

Then, for $0 < \Delta t \le |\alpha_0|L^{-1}$*, there holds*

$$\max_{0\le n\le N-k} \|e^n_z\|_\infty \le C(\Delta t)^k,$$

where C *is a constant depending on* T *and the given data* b*,* σ*,* φ *and* f*.*

4. Conclusion

A sharp convergence analysis for the multistep numerical schemes recently proposed by Zhao *et al.* [20] has been presented. We have shown that the solution Y admits a k-order convergence rate for the k-step scheme, and that a $(k—(1/2))$-order convergence rate for the Z solution follows. However, such an estimation is not optimal, so we introduced variational and Malliavin derivative arguments to obtain the k-order convergence rate. For technical reasons, we considered a special type of decoupled FBSDEs, although the schemes in Ref. [20] are applicable to general coupled FBSDEs. A rigorous convergence analysis of Scheme I for general FBSDEs is a focus for ongoing work. Furthermore, in our analysis we have assumed that the SDE is solved exactly, whereas in real applications it is often not so and a numerical solution is sought — a sharp analysis for associated discretisation schemes is thus also desirable.

Acknowledgments

This work was supported by the National Natural Science Foundations of China under grant numbers 91130003 and 11171189. The authors would like to thank the referees for their helpful comments, which considerably improved the paper.

References

[1] C. Bender and R. Denk, *A forward scheme for backward SDEs*, Stochastic Process. Appl. **117**, 1793-1812 (2007).
[2] B. Bouchard and N. Touzi, *Discrete-time approximation and monte-carlo simulation of backward stochastic differential equations*, Stochastic Process. Appl. **111**, 175-206 (2004).
[3] J. C. Butcher, *Numerical Methods for Ordinary Differential Equations*, John Wiley & Sons, New York (2003).
[4] J. F. Chassagneux and D. Crisan, *Runge-Kutta schemes for backward stochastic differential equaitons*, Ann. Appl. Probab. **24**, 679-720 (2014).
[5] J.P. Lemor E. Gobet and X. Warin, *A regression-based Monte-Carlo method to solve backward stochastic differential equations*, Ann. Appl. Probab. **15**, 2172-2202 (2005).

https://doi.org/10.4208/eajam.280515.211015a Published online by Cambridge University Press

[6] E. Gianin, *Risk measure via g-expectation*, Insurance: Mathematics and Economics **39**, 19-34 (2006).
[7] E. Gobet and C. Labart, *Error expansion for the discretization of backward stochastic differential equations*, Stochastic Process. Appl. **117**, 803-829 (2007).
[8] J. Ma J. Douglas and P. Protter, *Numerical methods for forward-backward stochastic differential equations*, Ann. Appl. Probab. **6**, 940-968 (1996).
[9] N. El Karoui, S. Peng, and M. C. Quenez, *Backward stochastic differential equations in finance*, Math. Finance **7**, 1-71 (1997).
[10] J. Ma and J. Yong, *Forward-backward Stochastic Differential Equations and Their Applications*, Lecture Notes in Math., Springer, Berlin (1999).
[11] G. N. Milstein and M. V. Tretyakov, *Numerical algorithms for forward-backward stochastic differential equations*, SIAM J. Sci. Comput. **28**, 561-582 (2006).
[12] D. Nualart, *The Malliavin Calculus and Related Topics*, Springer Verlag, Berlin (1995).
[13] E. Pardoux and S. Peng, *Adapted solution of a backward stochastic differential equation*, Systems Control Lett. **14**, 55-61 (1990).
[14] S. Peng, *A general stochastic maximum principle for optimal control problems*, SIAM J. Control Optim. **28**, 966-979 (1990).
[15] S. Peng, *Probabilistic interpretation for systems of quasilinear parabolic differential equations*, Stoch. Stoch. Rep. **37**, 61-74 (1991).
[16] S. Peng, *Backward SDE and related g-expectation*, Backward Stochastic Differential Equations, Pitman Research Notes in Math. Series **364**, 141-159 (1997).
[17] J. Zhang, *Some fine properties of backward stochastic differential equations*, Ph.D. thesis, Purdue University, West Lafayette, IN (2001).
[18] J. Zhang, *A numerical scheme for BSDEs*, Ann. Appl. Probab **14**, 459-488 (2004).
[19] W. Zhao, L. Chen, and S. Peng, *A new kind of accurate numerical method for backward stochastic differential equations*, SIAM J. Sci. Comput. **28**, 1563-1581 (2006).
[20] W. Zhao, Y. Fu, and T. Zhou, *New kinds of high-order multi-step schemes for forward backward stochastic differential equations*, SIAM J. Sci. Comput. **36**, A1731-A1751 (2014).
[21] W. Zhao, J. Wang, and S. Peng, *Error estimates of the θ-scheme for backward stochastic differential equations*, Discrete Contin. Dyn. Syst. Ser. B **12**, 905-924 (2009).
[22] W. Zhao, G. Zhang, and L. Ju, *A stable multistep scheme for solving backward stochastic differential equations*, SIAM J. Numer. Anal. **48**, 1369-1394 (2010).
[23] W. Zhao, W. Zhang, and L. Ju, *A numerical method and its error estimates for the decoupled forward-backward stochastic differential equations*, Commun. Comput. Phys. **15**, 618-646 (2014).

https://doi.org/10.4208/eajam.280515.211015a Published online by Cambridge University Press