INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 22, Number 6, Pages 777–800

doi: 10.4208/ijnam2025-1034

A CLASS OF RUNGE-KUTTA METHODS FOR BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS

XIAO TANG[1,*] AND JIE XIONG[2]

Abstract. In this paper, we introduce a class of Runge-Kutta (RK) methods for backward stochastic differential equations (BSDEs). The convergence rate is studied and the corresponding order conditions are obtained. For the conditional expectations involved in the methods, we design an approximation algorithm by combining the characteristics of the methods and replacing the increments of Brownian motion with appropriate discrete random variables. An important feature of our approximation algorithm is that interpolation operations can be avoided. The numerical results of four examples are presented to show that our RK methods provide a good approach for solving the BSDEs.

Key words. Backward stochastic differential equations, Runge-Kutta methods, order condition, conditional expectation.

1. Introduction

Consider the backward stochastic differential equation (BSDE) of the integral form

$$y(t) = \varphi(W(T)) + \int_t^T f(s, y(s), z(s))\, ds - \int_t^T z(s)\, dW(s),\ t \in [0, T], \tag{1}$$

where $f : \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^{d\times m} \longrightarrow \mathbb{R}^d$ is a Lipschitz-continuous function, $W(t) = (W^1(t), W^2(t), \cdots, W^m(t))$ is an m-dimensional Wiener process supported by a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}\}_{0\le t\le T})$, and the function $\varphi : \mathbb{R}^m \longrightarrow \mathbb{R}^d$ has continuous and bounded first derivatives.

The existence and uniqueness of the solution of (1) was first proved in [12]. Moreover, by [13] and [14], we know that the solution of (1) can be rewritten as

$$y(t) = u(t, W(t)),\ z(t) = \nabla u(t, W(t)),\ t \in [0, T], \tag{2}$$

where ∇u is the gradient of $u(t, x)$ with respect to the variable x, and $u(t, x)$ is the solution of the terminal value Cauchy problem

$$\begin{cases} \dfrac{\partial u}{\partial t} + \dfrac{1}{2}\displaystyle\sum_{i=1}^{m} \dfrac{\partial^2 u}{\partial x_i^2} + f(t, u, \nabla u) = 0,\ t \in [0, T),\ x \in \mathbb{R}^m, \\ u(T, x) = \varphi(x),\ x \in \mathbb{R}^m. \end{cases} \tag{3}$$

The smoothness of the solution u depends on the smoothness of the functions f and φ (see, e.g., [6, 24]). Specifically, if $f \in C_b^{k,2k,2k},\ \varphi \in C_b^{2k+\epsilon},\ k \in \mathbb{Z}^+,\ \epsilon \in (0, 1)$, then we have $u \in C_b^{k,2k}$, where $C_b^{k,2k,2k}$ denotes the set of continuously differentiable functions $\phi(t, y, z)$ with uniformly bounded partial derivatives $\partial_t^{l_0}\partial_y^{l_1}\partial_z^{l_2}\phi$ for $2l_0 + l_1 + l_2 \le 2k$, $C_b^{k,2k}$ denotes the set of functions $\phi(t, x)$ with uniformly bounded partial derivatives $\partial_t^{l_0}\partial_x^{l_1}\phi$ for $2l_0 + l_1 \le 2k$, and $C_b^{2k+\epsilon}$ denotes the set of functions $\phi(x)$ such that $\partial_x^l\phi,\ l \le 2k$ are uniformly bounded and $\partial_x^{2k}\phi$ is Hölder continuous with index ϵ.

Received by the editors on February 6, 2024 and, accepted on March 4, 2025.
2000 *Mathematics Subject Classification.* 60H10, 65L06, 65L20.
*Corresponding author.

As is well known, it is very difficult to find the analytic solution to most BSDEs. Therefore, developing numerical methods for solving BSDEs is becoming highly desired in practical applications. Up to now, many works on numerical methods of the BSDEs or their extensions forward-backward stochastic differential equations (FBSDEs) have been done. The methods in [5, 8, 9, 10, 11] are developed based on the relation between the BSDEs and the corresponding Cauchy problem. The methods in [3, 15, 16, 19, 20] are developed directly based on the BSDEs.

In recent years, there has been much interest in developing numerical methods based directly on the BSDEs. In particular, the linear multistep methods for solving ordinary differential equations (ODEs) have been successfully extended to solving BSDEs (see, e.g., [1, 17, 18, 21, 23, 25]). However, Runge-Kutta (RK) methods, as another type of important numerical methods for the ODEs, are rarely used to solve the BSDEs. As far as we know, there are currently only two references that have studied RK methods for the BSDEs [2, 4]. The authors of [4] studied a specific second-order RK method. The authors of [2] introduced a class of RK methods and provided rigorous convergence analysis results.

In the present paper, we will introduce a class of RK methods for the BSDEs (1). The order conditions up to third order are obtained for our RK methods. Based on the order conditions, we give two specific explicit RK methods. Combining the characteristics of our RK methods and replacing the increments of Brownian motion with some appropriate discrete random variables, we design an approximation algorithm for the conditional expectations involved in the RK methods. Our RK methods is different from the RK methods proposed in [2]. The main difference lies in the calculation of the internal stages about variable z (see method (7)), which is more conducive to design the approximation algorithm for the conditional expectations (see Remark 1). In addition, no interpolation operations are required for our approximation algorithm of the conditional expectations. What's more, the ideal of our approximate algorithm can be applied to many other methods for solving the BSDEs (see below).

This paper is organized as follows. In section 2, we introduce our RK methods. We study the convergence rate and obtain the corresponding order conditions in section 3. In section 4, the approximation algorithm for the conditional expectations is presented. Finally, we present some numerical results to verify our theoretical results.

2. RK methods for the BSDEs

Under the uniform time stepsize $h = \frac{T}{N},\ t_n = nh,\ n = 0, 1, 2, \cdots, N$ (N is a given positive integer), we have

$$(4)\quad \begin{cases} y(t_n) = y(t_{n+1}) + \int_{t_n}^{t_{n+1}} f(s, y(s), z(s))\ ds - \int_{t_n}^{t_{n+1}} z(s)\ dW(s),\ n < N, \\ y(t_N) = \varphi(W(T)). \end{cases}$$

Inspired by [19, 20], for equation (4), we can establish the following two ordinary differential reference equations

$$(5)\qquad y(t_n) = \mathbb{E}_{t_n}\Big[y(t_{n+1}) + \int_{t_n}^{t_{n+1}} f(s, y(s), z(s))\ ds\Big],$$

(6)

$$0 = \mathbb{E}_{t_n}\Big[y(t_{n+1})\Delta W_{t_n}(h) + \int_{t_n}^{t_{n+1}} f(s, y(s), z(s))\Delta W_{t_n}(s - t_n)\, ds - \int_{t_n}^{t_{n+1}} z(s)\, ds\Big],$$

where $\mathbb{E}_t[\cdot] = \mathbb{E}[\,\cdot \mid \mathcal{F}_t]$ and $\Delta W_t(\theta) = W(t+\theta) - W(t),\ 0 \le \theta \le T - t$.

Learning from the idea of RK methods for the ODEs, we introduce our RK methods for the BSDE (1), which take the form

$$\begin{aligned}
&y_N = \varphi(W(T)),\ z_N = \nabla\varphi(W(T)),\\
&y_n = \mathbb{E}_{t_n}\Big[\ y_{n+1} + h\sum_{i=1}^{q+1} b_i f(t_{n,i},\ Y_{n,i},\ Z_{n,i})\ \Big],\ n = 0, 1, \cdots, N-1,\\
&z_n = \mathbb{E}_{t_n}\Big[\ y_{n+1}H_{n,q+1,1} + h\sum_{i=1}^{q} \beta_i f(t_{n,i},\ Y_{n,i},\ Z_{n,i})H_{n,q+1,i}\ \Big],\\
&Y_{n,1} = y_{n+1},\ Z_{n,1} = z_{n+1},\ Y_{n,q+1} = y_n,\ Z_{n,q+1} = z_n,\\
&Y_{n,i} = \mathbb{E}_{t_{n,i}}\Big[\ y_{n+1} + h\sum_{j=1}^{i} a_{ij} f(t_{n,j},\ Y_{n,j},\ Z_{n,j})\ \Big],\ 2 \le i \le q,\\
&Z_{n,i} = \mathbb{E}_{t_{n,i}}\Big[\ z_{n+1} + h\sum_{j=1}^{i-1} \alpha_{ij} f(t_{n,j},\ Y_{n,j},\ Z_{n,j})H_{n,i,j}\ \Big],
\end{aligned} \tag{7}$$

where q is a given positive integer, and $t_{n,i} = t_{n+1} - c_i h$ with $0 = c_1 < c_2 \le \cdots \le c_{q+1} = 1$. The coefficients $b_i,\ \beta_i,\ a_{ij},\ \alpha_{ij} \in \mathbb{R}$ and satisfy

$$\begin{cases}
\beta_i = 0,\ if\ c_i = 1,\\
\alpha_{ij} = 0,\ if\ c_i \le c_j,\\
\sum_{j=1}^{i} a_{ij} = \sum_{j=1}^{i-1} \alpha_{ij} = c_i,\ 2 \le i \le q.
\end{cases}$$

The random variable $H_{n,i,j}$ satisfies

$$H_{n,i,j} = \begin{cases}
\dfrac{1}{(c_i - c_j)h}(W(t_{n,j}) - W(t_{n,i})),\ if\ c_i > c_j,\\
0,\ else.
\end{cases}$$

3. Convergence analysis of the RK methods for the BSDEs

Although RK method (7) appears to be more complex than the classical RK methods for solving ODE, we can still draw on the analytical approach of the classical RK methods when exploring the convergence of RK method (7). Similar to the classical RK methods, we study the convergence of RK method (7) for the BSDEs by analyzing (mean-square) zero stability and consistency.

3.1. Stability. We first analyze the stability of the method (7). For convenience of description, we rewrite the method (7) as the form

$$\begin{aligned}
&y_n = \mathbb{E}_{t_n}\Big[\ y_{n+1} + h\Phi(t_{n+1}, y_{n+1}, z_{n+1}, h, H)\ \Big],\ n = 0, 1, \cdots, N-1,\\
&z_n = \mathbb{E}_{t_n}\Big[\ y_{n+1}H_{n,q+1,1} + h\Psi(t_{n+1}, y_{n+1}, z_{n+1}, h, H)\ \Big],
\end{aligned} \tag{8}$$

where H can be seen as a vector containing all the random variables $H_{n,i,j},\ 1 \leq j \leq i \leq q+1$, and the functions $\Phi,\ \Psi$ satisfy

$$
\begin{aligned}
&\Phi(t_{n+1}, y_{n+1}, z_{n+1}, h, H) = \sum_{i=1}^{q+1} b_i f(t_{n,i},\ Y_{n,i},\ Z_{n,i}),\\
&\Psi(t_{n+1}, y_{n+1}, z_{n+1}, h, H) = \sum_{i=1}^{q} \beta_i f(t_{n,i},\ Y_{n,i},\ Z_{n,i}) H_{n,q+1,i},\\
&Y_{n,1} = y_{n+1},\ Z_{n,1} = z_{n+1},\ Y_{n,q+1} = y_n,\ Z_{n,q+1} = z_n,\\
&Y_{n,i} = \mathbb{E}_{t_{n,i}}\Big[\ y_{n+1} + h\sum_{j=1}^{i} a_{ij} f(t_{n,j},\ Y_{n,j},\ Z_{n,j})\ \Big],\ 2 \leq i \leq q,\\
&Z_{n,i} = \mathbb{E}_{t_{n,i}}\Big[\ z_{n+1} + h\sum_{j=1}^{i-1} \alpha_{ij} f(t_{n,j},\ Y_{n,j},\ Z_{n,j}) H_{n,i,j}\ \Big].
\end{aligned}
\tag{9}
$$

To investigate the stability, we introduce a perturbed method

$$
\begin{aligned}
&\tilde{y}_n = \mathbb{E}_{t_n}\Big[\ \tilde{y}_{n+1} + h\Phi(t_{n+1}, \tilde{y}_{n+1}, \tilde{z}_{n+1}, h, H)\ \Big] + \xi_n^Y,\ n = 0, 1, \cdots, N-1,\\
&\tilde{z}_n = \mathbb{E}_{t_n}\Big[\ \tilde{y}_{n+1} H_{n,q+1,1} + h\Psi(t_{n+1}, \tilde{y}_{n+1}, \tilde{z}_{n+1}, h, H)\ \Big] + \xi_n^Z,
\end{aligned}
\tag{10}
$$

where the terminal values $\tilde{y}_N,\ \tilde{z}_N$ belong to $L^2(\mathcal{F}_T)$ and the perturbation variables $\xi_n^Y,\ \xi_n^Z$ belong to $L^2(\mathcal{F}_{t_n})$ ($L^2(\mathcal{F}_t)$ denotes the set of all the $\mathcal{F}_t$-measurable and square integrable random variables). Next, we will prove the following result of stability.

Theorem 3.1. *Let $\delta y_n = y_n - \tilde{y}_n,\ \delta z_n = z_n - \tilde{z}_n$. If the function f is Lipschitz-continuous and h is small enough, then we have*

$$
\begin{aligned}
&\max_{0\leq n\leq N-1} \mathbb{E}\big[\ |\ \delta y_n\ |^2\ \big] + h\sum_{n=0}^{N-1} \mathbb{E}\big[\ |\ \delta z_n\ |^2\ \big]\\
&\leq C\Big(\mathbb{E}\big[\ |\ \delta y_N\ |^2\ \big] + h\mathbb{E}\big[\ |\ \delta z_N\ |^2\ \big] + \sum_{n=0}^{N-1} \mathbb{E}\big[\ \frac{1}{h}\ |\ \xi_n^Y\ |^2 + h\ |\ \xi_n^Z\ |^2\ \big]\Big),
\end{aligned}
\tag{11}
$$

where $|\cdot|$ is the Euclidean norm, and C is a positive constant that does not depend on h.

Proof. The proof approach of this theorem is similar to the proof of zero stability for the classical RK methods. But the specific proof process will be more complicated in the mean square sense.

We first prove that

$$
\begin{aligned}
&\Big|\mathbb{E}_{t_n}\big[\Phi(t_{n+1}, y_{n+1}, z_{n+1}, h, H) - \Phi(t_{n+1}, \tilde{y}_{n+1}, \tilde{z}_{n+1}, h, H)\big]\Big|^2\\
&\leq C_1\Big(\frac{1}{h}B_n + \mathbb{E}_{t_n}\big[\ |\ \delta y_{n+1}\ |^2 + |\ \delta z_{n+1}\ |^2\ \big] + |\ \xi_n^Y\ |^2 + |\ \xi_n^Z\ |^2\ \Big),
\end{aligned}
\tag{12}
$$

$$
\begin{aligned}
&\Big|\mathbb{E}_{t_n}\big[\Psi(t_{n+1}, y_{n+1}, z_{n+1}, h, H) - \Psi(t_{n+1}, \tilde{y}_{n+1}, \tilde{z}_{n+1}, h, H)\big]\Big|^2\\
&\leq \frac{C_2}{h}\Big(\frac{1}{h}B_n + \mathbb{E}_{t_n}\big[\ |\ \delta y_{n+1}\ |^2 + |\ \delta z_{n+1}\ |^2\ \big] + |\ \xi_n^Y\ |^2 + |\ \xi_n^Z\ |^2\ \Big)
\end{aligned}
\tag{13}
$$

when h is small enough, where $B_n = \mathbb{E}_{t_n}\big[\ |\ \delta y_{n+1}\ |^2\ \big] - |\ \mathbb{E}_{t_n}\big[\delta y_{n+1}\big]\ |^2$ and C_1, C_2 are positive constants that do not depend on h.

Based on the definition of $H_{n,i,j}$, the Lipshcitz condition and (9), it is not difficult to show that

$$\mathbb{E}_{t_n}\big[\mid H_{n,i,j}\mid^2\big]\le\frac{K}{h},\ K=\max_{c_i>c_j}\ \frac{1}{c_i-c_j},\ 1\le j\le i\le q+1, \tag{14}$$

$$\begin{aligned}&\Big|\mathbb{E}_{t_n}\big[\Phi(t_{n+1},y_{n+1},z_{n+1},h,H)-\Phi(t_{n+1},\tilde{y}_{n+1},\tilde{z}_{n+1},h,H)\big]\Big|^2\\ &\le(q+1)\sum_{i=1}^{q+1}\mid b_i\mid^2\mathbb{E}_{t_n}\big[\mid f(t_{n,i},\ Y_{n,i},\ Z_{n,i})-f(t_{n,i},\ \tilde{Y}_{n,i},\ \tilde{Z}_{n,i})\mid^2\big]\\ &\le 2(q+1)L^2\sum_{i=1}^{q+1}\mid b_i\mid^2\mathbb{E}_{t_n}\big[\mid\delta Y_{n,i}\mid^2+\mid\delta Z_{n,i}\mid^2\big],\end{aligned} \tag{15}$$

(16)

$$\begin{aligned}&\Big|\mathbb{E}_{t_n}\big[\Psi(t_{n+1},y_{n+1},z_{n+1},h,H)-\Psi(t_{n+1},\tilde{y}_{n+1},\tilde{z}_{n+1},h,H)\big]\Big|^2\\ &\le q\sum_{i=1}^{q}\mid\beta_i\mid^2\Big|\mathbb{E}_{t_n}\big[\big(f(t_{n,i},\ Y_{n,i},\ Z_{n,i})-f(t_{n,i},\ \tilde{Y}_{n,i},\ \tilde{Z}_{n,i})\big)H_{n,q+1,i}\big]\Big|^2\\ &\le q\sum_{i=1}^{q}\mid\beta_i\mid^2\mathbb{E}_{t_n}\big[\mid f(t_{n,i},\ Y_{n,i},\ Z_{n,i})-f(t_{n,i},\ \tilde{Y}_{n,i},\ \tilde{Z}_{n,i})\mid^2\big]\mathbb{E}_{t_n}\big[\mid H_{n,q+1,i}\mid^2\big]\\ &\le\frac{2qKL^2}{h}\sum_{i=1}^{q}\mid\beta_i\mid^2\mathbb{E}_{t_n}\big[\mid\delta Y_{n,i}\mid^2+\mid\delta Z_{n,i}\mid^2\big],\end{aligned}$$

where $\tilde{Y}_{n,i},\ \tilde{Z}_{n,i}$ are defined as $Y_{n,i},\ Z_{n,i}$ by replacing $y_{n+1},\ z_{n+1}$ with $\tilde{y}_{n+1},\ \tilde{z}_{n+1}$, $\delta Y_{n,i}=Y_{n,i}-\tilde{Y}_{n,i},\ \delta Z_{n,i}=Z_{n,i}-\tilde{Z}_{n,i}$, and L is the Lipshcitz constant. Note that

$$\begin{aligned}\mathbb{E}_{t_n}\big[\mid\delta Y_{n,1}\mid^2\big]&=\mathbb{E}_{t_n}\big[\mid\delta y_{n+1}\mid^2\big],\\ \mathbb{E}_{t_n}\big[\mid\delta Z_{n,1}\mid^2\big]&=\mathbb{E}_{t_n}\big[\mid\delta z_{n+1}\mid^2\big],\end{aligned}$$

and

$$\begin{aligned}&\mathbb{E}_{t_n}\big[\mid\delta Y_{n,q+1}\mid^2\big]\\ &=\mid\delta y_n\mid^2\\ &\le 3\Big(\mathbb{E}_{t_n}\big[\mid\delta y_{n+1}\mid^2\big]\\ &\quad+h^2\big|\mathbb{E}_{t_n}\big[\Phi(t_{n+1},y_{n+1},z_{n+1},h,H)-\Phi(t_{n+1},\tilde{y}_{n+1},\tilde{z}_{n+1},h,H)\big]\big|^2\\ &\quad+\mid\xi_n^Y\mid^2\Big)\\ &\le 3\Big(\mathbb{E}_{t_n}\big[\mid\delta y_{n+1}\mid^2\big]\\ &\quad+2(q+1)L^2h^2\sum_{i=1}^{q+1}\mid b_i\mid^2\mathbb{E}_{t_n}\big[\mid\delta Y_{n,i}\mid^2+\mid\delta Z_{n,i}\mid^2\big]+\mid\xi_n^Y\mid^2\Big),\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}_{t_n}\big[\mid \delta Z_{n,q+1}\mid^2\big]\\
=&\mid \delta z_n\mid^2\\
\leq&3\Big(\mid \mathbb{E}_{t_n}\big[\delta y_{n+1}H_{n,q+1,1}\big]\mid^2\\
&+h^2\big|\mathbb{E}_{t_n}\big[\Psi(t_{n+1},y_{n+1},z_{n+1},h,H)-\Psi(t_{n+1},\tilde{y}_{n+1},\tilde{z}_{n+1},h,H)\big]\big|^2\\
&+\mid \xi_n^Z\mid^2\Big)\\
=&3\Big(\mid \mathbb{E}_{t_n}\big[\big(\delta y_{n+1}-\mathbb{E}_{t_n}\big[\delta y_{n+1}\big]\big)H_{n,q+1,1}\big]\mid^2\\
&+h^2\big|\mathbb{E}_{t_n}\big[\Psi(t_{n+1},y_{n+1},z_{n+1},h,H)-\Psi(t_{n+1},\tilde{y}_{n+1},\tilde{z}_{n+1},h,H)\big]\big|^2\\
&+\mid \xi_n^Z\mid^2\Big)\\
\leq&3\Big(\frac{K}{h}B_n+2qKL^2h\sum_{i=1}^{q}\mid \beta_i\mid^2\mathbb{E}_{t_n}\big[\mid \delta Y_{n,i}\mid^2+\mid \delta Z_{n,i}\mid^2\big]+\mid \xi_n^Z\mid^2\Big),
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}_{t_n}\big[\mid \delta Y_{n,i}\mid^2\big]\\
\leq&2\Big(\mathbb{E}_{t_n}\big[\mid \delta y_{n+1}\mid^2\big]+2iL^2h^2\sum_{j=1}^{i}\mid a_{ij}\mid^2\mathbb{E}_{t_n}\big[\mid \delta Y_{n,j}\mid^2+\mid \delta Z_{n,j}\mid^2\big]\Big),\\
&\mathbb{E}_{t_n}\big[\mid \delta Z_{n,i}\mid^2\big]\\
\leq&2\Big(\mathbb{E}_{t_n}\big[\mid \delta z_{n+1}\mid^2\big]\\
&+2(i-1)KL^2h\sum_{j=1}^{i-1}\mid \alpha_{ij}\mid^2\mathbb{E}_{t_n}\big[\mid \delta Y_{n,j}\mid^2+\mid \delta Z_{n,j}\mid^2\big]\Big)
\end{aligned}$$

for $2\leq i\leq q$. Take

$$M=\max_{i,j}\{\mid b_i\mid,\mid \beta_i\mid,\mid a_{ij}\mid,\mid \alpha_{ij}\mid\},$$

then it is not difficult to show that

(17)
$$\begin{aligned}
&\sum_{i=1}^{q+1}\mathbb{E}_{t_n}\big[\mid \delta Y_{n,i}\mid^2+\mid \delta Z_{n,i}\mid^2\big]\\
\leq&\frac{3K}{h}B_n+(2q+2)\mathbb{E}_{t_n}\big[\mid \delta y_{n+1}\mid^2+\mid \delta z_{n+1}\mid^2\big]+3\big(\mid \xi_n^Y\mid^2+\mid \xi_n^Z\mid^2\big)\\
&+\big(2q(q+2)K+(2q^2+8q+2)h\big)M^2L^2h\sum_{i=1}^{q+1}\mathbb{E}_{t_n}\big[\mid \delta Y_{n,i}\mid^2+\mid \delta Z_{n,i}\mid^2\big].
\end{aligned}$$

The estimates (12) and (13) follow from the estimates (15)-(17) if h is small enough.

Note that

$$(a+b)^2\leq(1+\lambda_1h)a^2+(1+\frac{1}{\lambda_1h})b^2, \tag{18}$$

where λ_1 could be any given positive constant. Then, by using the estimates (12) and (13), we have

(19)
$$\begin{aligned}
\mid \delta y_n \mid^2 \leq & (1+\lambda_1 h)\big|\mathbb{E}_{t_n}\big[\delta y_{n+1}\big]\big|^2 \\
& + (1+\frac{1}{\lambda_1 h})\Big(h\big|\mathbb{E}_{t_n}\big[\Phi(t_{n+1}, y_{n+1}, z_{n+1}, h, H) - \Phi(t_{n+1}, \tilde{y}_{n+1}, \tilde{z}_{n+1}, h, H)\big]\big| \\
& + \mid \xi_n^Y \mid\Big)^2 \\
\leq & (1+\lambda_1 h)\big|\mathbb{E}_{t_n}\big[\delta y_{n+1}\big]\big|^2 \\
& + 2C_1 h^2 (1+\frac{1}{\lambda_1 h})\Big(\frac{1}{h}B_n + \mathbb{E}_{t_n}\big[\mid \delta y_{n+1} \mid^2 + \mid \delta z_{n+1} \mid^2\big] \\
& + (1+\frac{1}{C_1 h^2}) \mid \xi_n^Y \mid^2 + \mid \xi_n^Z \mid^2\Big),
\end{aligned}$$

(20)
$$\begin{aligned}
\mid \delta z_n \mid^2 \leq & 3\Big(\mid \mathbb{E}_{t_n}\big[\delta y_{n+1} H_{n,q+1,1}\big] \mid^2 \\
& + h^2\big|\mathbb{E}_{t_n}\big[\Psi(t_{n+1}, y_{n+1}, z_{n+1}, h, H) - \Psi(t_{n+1}, \tilde{y}_{n+1}, \tilde{z}_{n+1}, h, H)\big]\big|^2 + \mid \xi_n^Z \mid^2\Big) \\
\leq & 3C_2 h\Big((\frac{1}{h} + \frac{K}{C_2 h^2})B_n + \mathbb{E}_{t_n}\big[\mid \delta y_{n+1} \mid^2 + \mid \delta z_{n+1} \mid^2\big] \\
& + \mid \xi_n^Y \mid^2 + (1+\frac{1}{C_2 h}) \mid \xi_n^Z \mid^2\Big).
\end{aligned}$$

Combining (19), (20) and the definition of B_n leads to

(21)
$$\begin{aligned}
& \mid \delta y_n \mid^2 + \lambda_2 h \mid \delta z_n \mid^2 \\
\leq & \Big((1+\lambda_1 h) - 2C_1(\tfrac{1}{\lambda_1} + h) - 3\lambda_2(K + C_2 h)\Big)\big|\mathbb{E}_{t_n}\big[\delta y_{n+1}\big]\big|^2 \\
& + \Big(2C_1(\tfrac{1}{\lambda_1} + \tfrac{1}{\lambda_1}h + h + h^2) + 3\lambda_2(K + C_2 h + C_2 h^2)\Big)\mathbb{E}_{t_n}\big[\mid \delta y_{n+1} \mid^2\big] \\
& + \Big(2C_1(\tfrac{1}{\lambda_1}h + h^2) + 3\lambda_2 C_2 h^2\Big)\mathbb{E}_{t_n}\big[\mid \delta z_{n+1} \mid^2\big] \\
& + \Big(2(\tfrac{1}{\lambda_1 h} + 1)(1 + C_1 h^2) + 3\lambda_2 C_2 h^2\Big) \mid \xi_n^Y \mid^2 \\
& + \Big(2C_1 h(\tfrac{1}{\lambda_1} + h) + 3\lambda_2 h(1 + C_2 h)\Big) \mid \xi_n^Z \mid^2,
\end{aligned}$$

where λ_2 could be any given positive constant. Let $\lambda_1 = 4C_1(\frac{1}{2} + 3K)$ and $\lambda_2 = \dfrac{1}{\frac{1}{2} + 3K}$, then there is a positive constant C_3 such that

(22)
$$\begin{aligned}
& \mathbb{E}\big[\mid \delta y_n \mid^2 + \frac{\lambda_2}{2}h \mid \delta z_n \mid^2\big] + \frac{\lambda_2}{2}h\mathbb{E}\big[\mid \delta z_n \mid^2\big] \\
\leq & (1 + C_3 h)\big|\mathbb{E}\big[\mid \delta y_{n+1} \mid^2 + \frac{\lambda_2}{2}h \mid \delta z_{n+1} \mid^2\big] + C_3\mathbb{E}\big[\frac{1}{h} \mid \xi_n^Y \mid^2 + h \mid \xi_n^Z \mid^2\big]
\end{aligned}$$

when h is small enough.

On the one hand, the Gronwall inequality leads to

$$
\begin{aligned}
&\max_{0\le n\le N-1}\mathbb{E}\big[\mid \delta y_n\mid^2\big]\\
(23)\qquad &\le \max_{0\le n\le N-1}\mathbb{E}\big[\mid \delta y_n\mid^2+\tfrac{\lambda_2}{2}h\mid \delta z_n\mid^2\big]\\
&\le C_4\Big(\mathbb{E}\big[\mid \delta y_N\mid^2+\tfrac{\lambda_2}{2}h\mid \delta z_N\mid^2\big]+\sum_{n=0}^{N-1}\mathbb{E}\big[\tfrac{1}{h}\mid \xi_n^Y\mid^2+h\mid \xi_n^Z\mid^2\big]\Big),
\end{aligned}
$$

where $C_4>0$ and independent of the stepsize h.

On the other hand, summing the inequality (22) over n, we have

$$
\begin{aligned}
h\sum_{n=0}^{N-1}\mathbb{E}\big[\mid \delta z_n\mid^2\big]\le \frac{2}{\lambda_2}\Big(&\mathbb{E}\big[\mid \delta y_N\mid^2+\frac{\lambda_2}{2}h\mid \delta z_N\mid^2\big]\\
(24)\qquad &+C_3T\max_{0\le n\le N-1}\mathbb{E}\big[\mid \delta y_{n+1}\mid^2+\frac{\lambda_2}{2}h\mid \delta z_{n+1}\mid^2\big]\\
&+C_3\sum_{n=0}^{N-1}\mathbb{E}\big[\frac{1}{h}\mid \xi_n^Y\mid^2+h\mid \xi_n^Z\mid^2\big]\Big).
\end{aligned}
$$

The conclusion (11) follows from (23) and (24). □

3.2. Order conditions. Consistency analysis is actually analyzing local errors. The most commonly used technique for analyzing local errors is Taylor expansion. Therefore, for a clearer description, we first introduce some notations about stochastic Taylor expansions (For details, please refer to the Chapter 5 of [7]).

Let

(25)
$$\mathcal{M}=\Big\{(j_1,j_2,\cdots,j_l)\;\Big|\;j_i\in\{0,1,\cdots,m\},\ i\in\{1,2,\cdots,l\},\ l=1,2,3,\cdots\Big\}\bigcup\{\emptyset\}$$

be the set of all multi-indices. The number of components of $\gamma\in\mathcal{M}$ is denoted by $l(\gamma)$, in particular, $l(\emptyset):=0$. The number of zero in γ is denoted by $n(\gamma)$. The multi-index $(j_1,j_2,\cdots,j_l)$ can abbreviate to $(j)_l$ when $j_1=j_2=\cdots=j_l=j$. For $\gamma=(j_1,j_2,\cdots,j_l)$, $\eta=(k_1,k_2,\cdots,k_p)\in\mathcal{M}$, we define

$$
\begin{aligned}
&\gamma*\eta=(j_1,j_2,\cdots,j_l,k_1,k_2,\cdots,k_p),\\
(26)\qquad &L^\gamma=L^{(j_1)}L^{(j_2)}\cdots L^{(j_l)},
\end{aligned}
$$

where the differential operators $L^{(j)},\ j=0,1,\cdots,m$ satisfy

$$
(27)\qquad L^{(j)}=\begin{cases}\dfrac{\partial}{\partial t}+\dfrac{1}{2}\displaystyle\sum_{i=1}^m\dfrac{\partial^2}{\partial x_i^2},\ j=0,\\[2ex] \dfrac{\partial}{\partial x_j},\ j=1,2,\cdots,m.\end{cases}
$$

For a given positive integer p, take

$$(28)\qquad \mathcal{A}_p=\big\{\gamma\in\mathcal{M}\;\big|\;l(\gamma)+n(\gamma)\le 2p+1\big\}.$$

If the function $\phi(t,x)\in C_b^{1+p,2+2p}$, the multi-indices γ,η belong to $\mathcal{A}_p$ and differ only in the order of the indices, then it is not difficult to show that

$$(29)\qquad L^\gamma\phi=L^\eta\phi.$$

Now, we introduce the following result.

Proposition 3.2. *Let $u_t = u(t, W(t))$, $u_t^\gamma = L^\gamma u(t, W(t))$, if the function $u \in C_b^{1+p,2+2p}$, then for a given positive integer p, we have*

(30)
$$\mathbb{E}_t\big[u(t+h, W(t+h))\big] = u_t + hu_t^{(0)} + \frac{h^2}{2}u_t^{(0)_2} + \cdots + \frac{h^p}{p!}u_t^{(0)_p} + O_t(h^{p+1}),$$

(31)
$$\frac{1}{h}\mathbb{E}_t\big[u(t+h, W(t+h))\Delta W_t^k(h)\big] = u_t^{(k)} + hu_t^{(k,0)} + \cdots + \frac{h^p}{p!}u_t^{(k)*(0)_p} + O_t(h^{p+1}),$$

where $\Delta W_t^k(h) = W^k(t+h) - W^k(t), k \in \{1, 2, \cdots, m\}$, $O_t(h^{p+1}) \in \mathcal{F}_t$ and $\mathbb{E}\big[\mid O_t(h^{p+1}) \mid \big] = O(h^{p+1})$.

Proposition 3.2 can be seen as the special situation of Propositions 2.2 and 2.3 in [2] with $X(t) = W(t)$. Therefore the readers can refer to [2] for its proof.

Before giving the result of order conditions, we first give the definition about the convergence order of the RK method (7).

Definition 3.3. If the estimate

$$\max_{0\le n\le N-1} \mathbb{E}\big[\mid y(t_n) - y_n \mid^2 \big] + h\sum_{n=0}^{N-1}\mathbb{E}\big[\mid z(t_n) - z_n \mid^2 \big] \le Ch^{2p} \tag{32}$$

is fulfilled when h is small enough, then we call the RK method (7) is of order p.

By using Theorem 3.1 and Proposition 3.2, we can obtain a main result, which presents the order conditions of the RK method (7) up to third order.

Theorem 3.4. *Assume $f \in C_b^{4,8,8}$ and $\varphi \in C_b^{8+\epsilon}$, then the solution u of (3) belong to $C_b^{4,8}$. If the condition*

$$\sum_{i=1}^{q+1} b_i = 1 \tag{33}$$

is fulfilled, then the RK method (7) is of order 1. *If (33) and the addition conditions*

$$\sum_{i=1}^{q}\beta_i = 1,\ \sum_{i=2}^{q} b_ic_i + b_{q+1} = \frac{1}{2} \tag{34}$$

are fulfilled, then the RK method (7) is of order 2. *If (33), (34) and the addition conditions*

$$\begin{aligned}&\sum_{i=2}^{q}\beta_ic_i = \frac{1}{2},\ \sum_{i=2}^{q}b_ic_i^2 + b_{q+1} = \frac{1}{3},\\ &\sum_{i=2}^{q}\sum_{j=2}^{i}b_ia_{ij}c_j + \frac{1}{2}b_{q+1} = \frac{1}{6},\ \sum_{i=3}^{q}\sum_{j=2}^{i-1}b_i\alpha_{ij}c_j + \frac{1}{2}b_{q+1} = \frac{1}{6}\end{aligned} \tag{35}$$

are fulfilled, then the RK method (7) is of order 3.

Proof. The proof of this theorem mainly relies on the repeated use of Taylor or stochastic Taylor expansion techniques. There is no essential difference between one-dimensional and multi-dimensional BSDEs in the proof process. Therefore, we prove only the case of $d = m = 1$ for simplicity of description.

Based on (2), (3) and (27), we know the fact that

$$y(t) = u_t,\ z(t) = u_t^{(1)},\ f(t, y(t), z(t)) = -u_t^{(0)}. \tag{36}$$

Take $y_{n+1}=y(t_{n+1}),\ z_{n+1}=z(t_{n+1})$ in (7), then $Y_{n,1}=u_{t_{n+1}},\ Z_{n,1}=u^{(1)}_{t_{n+1}}$. By using Proposition 3.2, the equation (29) and the fact that $\alpha_{21}=c_2$, we can show that

(37)

$$\begin{aligned}Z_{n,2}&=\mathbb{E}_{t_{n,2}}\Big[\ z_{n+1}+h\alpha_{21}f(t_{n+1},\ y_{n+1},\ z_{n+1})H_{n,2,1}\ \Big]\\&=u^{(1)}_{t_{n,2}}+c_2hu^{(1,0)}_{t_{n,2}}+\frac{c_2^2}{2}h^2u^{(1,0,0)}_{t_{n,2}}-\alpha_{21}hu^{(1,0)}_{t_{n,2}}-\alpha_{21}c_2h^2u^{(1,0,0)}_{t_{n,2}}+O_{t_{n,2}}(h^3)\\&=u^{(1)}_{t_{n,2}}-\frac{c_2^2}{2}h^2u^{(1,0,0)}_{t_{n,2}}+O_{t_{n,2}}(h^3).\end{aligned}$$

For $Y_{n,2}$, it is not difficult to show that

(38)

$$\begin{aligned}Y_{n,2}&=\mathbb{E}_{t_{n,2}}\Big[\ y_{n+1}+ha_{21}f(t_{n+1},\ y_{n+1},\ z_{n+1})\ \Big]+ha_{22}f(t_{n,2},\ Y_{n,2},\ Z_{n,2})\\&=\mathbb{E}_{t_{n,2}}\Big[\ y_{n+1}+ha_{21}f(t_{n+1},\ y_{n+1},\ z_{n+1})\ \Big]+ha_{22}f(t_{n,2},\ u_{t_{n,2}},\ u^{(1)}_{t_{n,2}})\\&\quad+ha_{22}(\delta f_1+\delta f_2)\\&=u_{t_{n,2}}+c_2hu^{(0)}_{t_{n,2}}+\frac{c_2^2}{2}h^2u^{(0,0)}_{t_{n,2}}-a_{21}hu^{(0)}_{t_{n,2}}-a_{21}c_2h^2u^{(0,0)}_{t_{n,2}}-a_{22}hu^{(0)}_{t_{n,2}}\\&\quad+ha_{22}(\delta f_1+\delta f_2)+O_{t_{n,2}}(h^3)\\&=u_{t_{n,2}}-(\frac{c_2^2}{2}-a_{22}c_2)h^2u^{(0,0)}_{t_{n,2}}+ha_{22}(\delta f_1+\delta f_2)+O_{t_{n,2}}(h^3),\end{aligned}$$

where we use the equation $a_{21}+a_{22}=c_2$ and

$$\begin{aligned}\delta f_1&=f(t_{n,2},\ Y_{n,2},\ Z_{n,2})-f(t_{n,2},\ Y_{n,2},\ u^{(1)}_{t_{n,2}}),\\\delta f_2&=f(t_{n,2},\ Y_{n,2},\ u^{(1)}_{t_{n,2}})-f(t_{n,2},\ u_{t_{n,2}},\ u^{(1)}_{t_{n,2}}).\end{aligned}$$

Combining (37), (38) and the Lipshcitz condition can lead to

$$\mid\delta f_1\mid=O_{t_{n,2}}(h^2),\ \mid\delta f_2\mid=O_{t_{n,2}}(h^2).$$

Then

$$Y_{n,2}=u_{t_{n,2}}-(\frac{c_2^2}{2}-a_{22}c_2)h^2u^{(0,0)}_{t_{n,2}}+O_{t_{n,2}}(h^3).\tag{39}$$

Let f^y_t and f^z_t denote the partial derivatives of $f(t,y(t),z(t))$ with respect to y and z, respectively. By a first order Taylor expansion, it is easy to obtain

$$\begin{aligned}f(t_{n,2},\ Y_{n,2},\ Z_{n,2})=&\ -u^{(0)}_{t_{n,2}}-(\tfrac{c_2^2}{2}-a_{22}c_2)h^2u^{(0,0)}_{t_{n,2}}f^y_{t_{n,2}}\\&-\tfrac{c_2^2}{2}h^2u^{(1,0,0)}_{t_{n,2}}f^z_{t_{n,2}}+O_{t_{n,2}}(h^3).\end{aligned}\tag{40}$$

Through similar and cumbersome calculations, we can obtain

(41)

$$\begin{aligned}f(t_{n,i},\ Y_{n,i},\ Z_{n,i})=&-u^{(0)}_{t_{n,i}}-\Big(\frac{c_i^2}{2}-\sum_{j=2}^{i}a_{ij}c_j\Big)h^2u^{(0,0)}_{t_{n,i}}f^y_{t_{n,i}}\\&-\Big(\frac{c_i^2}{2}-\sum_{j=2}^{i-1}\alpha_{ij}c_j\Big)h^2u^{(1,0,0)}_{t_{n,i}}f^z_{t_{n,i}}+O_{t_{n,i}}(h^3),\ i=3,4,\cdots,q.\end{aligned}$$

Then we have

(42)

$$
\begin{aligned}
z_n &= \mathbb{E}_{t_n}\Big[\; y_{n+1}H_{n,q+1,1} + h\sum_{i=1}^{q}\beta_i f(t_{n,i},\ Y_{n,i},\ Z_{n,i})H_{n,q+1,i}\;\Big] \\
&= u_{t_n}^{(1)} + hu_{t_n}^{(1,0)} + \frac{1}{2}h^2u_{t_n}^{(1,0,0)} - \beta_1 hu_{t_n}^{(1,0)} - \beta_1 c_2 h^2 u_{t_n}^{(1,0,0)} - \beta_2 hu_{t_n}^{(1,0)} \\
&\quad -\beta_2(1-c_2)h^2u_{t_n}^{(1,0,0)} - \cdots - \beta_q hu_{t_n}^{(1,0)} - \beta_q(1-c_q)h^2u_{t_n}^{(1,0,0)} + O_{t_n}(h^3) \\
&= u_{t_n}^{(1)} + \Big(1-\sum_{i=1}^{q}\beta_i\Big)hu_{t_n}^{(1,0)} + \Big(\frac{1}{2}-\beta_1-\sum_{i=2}^{q}\beta_i(1-c_i)\Big)h^2u_{t_n}^{(1,0,0)} + O_{t_n}(h^3)
\end{aligned}
$$

and

(43)

$$
\begin{aligned}
y_n &= \mathbb{E}_{t_n}\Big[\; y_{n+1} + h\sum_{i=1}^{q} b_i f(t_{n,i},\ Y_{n,i},\ Z_{n,i})\;\Big] + hb_{q+1}f(t_n,\ y_n,\ z_n) \\
&= \mathbb{E}_{t_n}\Big[\; y_{n+1} + h\sum_{i=1}^{q} b_i f(t_{n,i},\ Y_{n,i},\ Z_{n,i})\;\Big] \\
&\quad + hb_{q+1}f(t_n,\ u_{t_n},\ u_{t_n}^{(1)}) + hb_{q+1}(\delta f_3 + \delta f_4) \\
&= u_{t_n} + hu_{t_n}^{(0)} + \frac{1}{2}h^2u_{t_n}^{(0,0)} + \frac{1}{6}h^3u_{t_n}^{(0,0,0)} - b_1hu_{t_n}^{(0)} - b_1h^2u_{t_n}^{(0,0)} - \frac{b_1}{2}h^3u_{t_n}^{(0,0,0)} \\
&\quad - b_2hu_{t_n}^{(0)} - b_2(1-c_2)h^2u_{t_n}^{(0,0)} - \frac{b_2(1-c_2)^2}{2}h^3u_{t_n}^{(0,0,0)} \\
&\quad - \cdots - b_qhu_{t_n}^{(0)} - b_q(1-c_q)h^2u_{t_n}^{(0,0)} - \frac{b_q(1-c_q)^2}{2}h^3u_{t_n}^{(0,0,0)} - b_{q+1}hu_{t_n}^{(0)} \\
&\quad - b_2(\frac{c_2^2}{2} - a_{22}c_2)h^3u_{t_n}^{(0,0)}f_{t_n}^y - b_2\frac{c_2^2}{2}h^3u_{t_n}^{(1,0,0)}f_{t_n}^z \\
&\quad - \cdots - b_q\Big(\frac{c_q^2}{2} - \sum_{j=2}^{q}a_{ij}c_j\Big)h^3u_{t_n}^{(0,0)}f_{t_n}^y - b_q\Big(\frac{c_q^2}{2} - \sum_{j=2}^{q-1}\alpha_{ij}c_j\Big)h^3u_{t_n}^{(1,0,0)}f_{t_n}^z \\
&\quad + b_{q+1}h(\delta f_3 + \delta f_4) + O_{t_n}(h^4) \\
&= u_{t_n} + \Big(1-\sum_{i=1}^{q+1}b_i\Big)hu_{t_n}^{(0)} + \Big(\frac{1}{2} - b_1 - \sum_{i=2}^{q}b_i(1-c_i)\Big)h^2u_{t_n}^{(0,0)} \\
&\quad + \Big(\frac{1}{6} - \frac{b_1}{2} - \sum_{i=2}^{q}\frac{b_i(1-c_i)^2}{2}\Big)h^3u_{t_n}^{(0,0,0)} - \sum_{i=2}^{q}b_i\Big(\frac{c_i^2}{2} - \sum_{j=2}^{i}a_{ij}c_j\Big)h^3u_{t_n}^{(0,0)}f_{t_n}^y \\
&\quad - \Big(b_2\frac{c_2^2}{2} + \sum_{i=3}^{q}b_i\big(\frac{c_i^2}{2} - \sum_{j=2}^{i-1}\alpha_{ij}c_j\big)\Big)h^3u_{t_n}^{(1,0,0)}f_{t_n}^z + b_{q+1}h(\delta f_3 + \delta f_4) + O_{t_n}(h^4),
\end{aligned}
$$

where

$$
\delta f_3 = f(t_n,\ y_n,\ z_n) - f(t_n,\ y_n,\ u_{t_n}^{(1)}),\ \delta f_4 = f(t_n,\ y_n,\ u_{t_n}^{(1)}) - f(t_n,\ u_{t_n},\ u_{t_n}^{(1)}).
$$

If the condition (33) is fulfilled, the equations (42) and (43) mean that

(44)

$$
\begin{aligned}
y(t_n) &= \mathbb{E}_{t_n}\Big[\; y(t_{n+1}) + h\Phi(t_{n+1}, y(t_{n+1}), z(t_{n+1}), h, H)\;\Big] + O_{t_n}(h^2), \\
z(t_n) &= \mathbb{E}_{t_n}\Big[\; y(t_{n+1})H_{n,q+1,1} + h\Psi(t_{n+1}, y(t_{n+1}), z(t_{n+1}), h, H)\;\Big] + O_{t_n}(h).
\end{aligned}
$$

The first conclusion follows from Theorem 3.1 and (44). By some simple calculations, we can easily verify that the second and third conclusions are also correct.

The proof is complete. □

Many specific RK methods can be proposed by using Theorem 3.4. We list two specific RK methods here.

The first one is a second order explicit RK method with $q = 2$, which takes the form

$$
\begin{aligned}
y_n &= \mathbb{E}_{t_n}\Big[\ y_{n+1} + h\big(\frac{1}{2}f(t_{n+1},\ Y_{n,1},\ Z_{n,1})\ \big)\Big] \\
&\quad + \frac{1}{2}hf(t_n,\ Y_{n,2},\ Z_{n,2}), \\
z_n &= \mathbb{E}_{t_n}\Big[\ y_{n+1}H_{n,1} + hf(t_{n+1},\ Y_{n,1},\ Z_{n,1})H_{n,1}\ \Big], \\
Y_{n,1} &= y_{n+1},\ Z_{n,1} = z_{n+1}, \\
Y_{n,2} &= \mathbb{E}_{t_n}\Big[\ y_{n+1} + hf(t_{n+1},\ Y_{n,1},\ Z_{n,1})\ \Big], \\
Z_{n,2} &= \mathbb{E}_{t_n}\Big[\ z_{n+1} + hf(t_{n+1},\ Y_{n,1},\ Z_{n,1})H_{n,1}\ \Big],
\end{aligned}
\tag{45}
$$

where $H_{n,1} = \frac{1}{h}\big(W(t_{n+1}) - W(t_n)\big)$.

The second one is a third order explicit RK method with $q = 3$, which takes the form

$$
\begin{aligned}
y_n &= \mathbb{E}_{t_n}\Big[\ y_{n+1} + h\big(\frac{1}{6}f(t_{n+1},\ Y_{n,1},\ Z_{n,1}) + \frac{2}{3}f(t_{n+\frac{1}{2}},\ Y_{n,2},\ Z_{n,2})\ \big)\Big] \\
&\quad + \frac{1}{6}hf(t_n,\ Y_{n,3},\ Z_{n,3}), \\
z_n &= \mathbb{E}_{t_n}\Big[\ y_{n+1}H_{n,1} + hf(t_{n+\frac{1}{2}},\ Y_{n,2},\ Z_{n,2})H_{n,2}\ \Big], \\
Y_{n,1} &= y_{n+1},\ Z_{n,1} = z_{n+1}, \\
Y_{n,2} &= \mathbb{E}_{t_{n+\frac{1}{2}}}\Big[\ y_{n+1} + \frac{1}{2}hf(t_{n+1},\ Y_{n,1},\ Z_{n,1})\ \Big], \\
Z_{n,2} &= \mathbb{E}_{t_{n+\frac{1}{2}}}\Big[\ z_{n+1} + \frac{1}{2}hf(t_{n+1},\ Y_{n,1},\ Z_{n,1})H_{n,3}\ \Big], \\
Y_{n,3} &= \mathbb{E}_{t_n}\Big[\ y_{n+1} - hf(t_{n+1},\ Y_{n,1},\ Z_{n,1}) + 2hf(t_{n+\frac{1}{2}},\ Y_{n,2},\ Z_{n,2})\ \Big], \\
Z_{n,3} &= \mathbb{E}_{t_n}\Big[\ z_{n+1} - hf(t_{n+1},\ Y_{n,1},\ Z_{n,1})H_{n,1} + 2hf(t_{n+\frac{1}{2}},\ Y_{n,2},\ Z_{n,2})H_{n,2}\ \Big],
\end{aligned}
\tag{46}
$$

where $H_{n,1} = \frac{1}{h}\big(W(t_{n+1}) - W(t_n)\big)$, $H_{n,2} = \frac{2}{h}\big(W(t_{n+\frac{1}{2}}) - W(t_n)\big)$ and $H_{n,3} = \frac{2}{h}\big(W(t_{n+1}) - W(t_{n+\frac{1}{2}})\big)$.

4. Approximation of conditional expectation

To implement the method (7), we first need to compute the conditional expectations involved. In this section, we will introduce our algorithm to approximate the needed conditional expectations. In our approximate algorithm, the increments of Brownian motion will be replaced by some appropriate discrete distributed random variables. An important advantage of our algorithm is that it can avoid the interpolation operations. High-order interpolation polynomials are not only prone to instability, but also costly for high-dimensional problems. Therefore, it is meaningful to avoid interpolation. In addition, the ideal of our approximate algorithm can be applied to many other methods for solving the BSDEs,such as the θ-methods (see, e.g., [20, 22]) and the multistep methods (see, e.g., [17, 18, 21]).

We first introduce the following useful result.

Theorem 4.1. *Assume the function $u \in C_b^{1+p,2+2p}$. If $\xi_1, \xi_2, \cdots, \xi_m$ are the independent and identically distributed random variables and satisfy*

$$\mathbb{E}\big[(\sqrt{h}\xi_i)^q\big] = \begin{cases} 0, \; q = 1, 3, \cdots, 2p+1, \\ (q-1)!! \; h^{\frac{q}{2}}, \; q = 2, 4, \cdots, 2p, \quad i = 1, 2, \cdots, m \\ O(h^{\frac{q}{2}}), \; q \geq (2p+2), \end{cases} \tag{47}$$

for a positive integer p. Then, for the given $t \in \mathbb{R}, x \in \mathbb{R}^m$, we have

$$\begin{aligned} &\mathbb{E}\big[u\big(t+h, x+\Delta W_t(h)\big)\big] = \mathbb{E}\big[u(t+h, x+\sqrt{h}\xi)\big] + O(h^{p+1}), \\ &\mathbb{E}\big[u\big(t+h, x+\Delta W_t(h)\big)\Delta W_t(h)\big] = \mathbb{E}\big[u(t+h, x+\sqrt{h}\xi)\sqrt{h}\xi\big] + O(h^{p+1}), \end{aligned} \tag{48}$$

where $\xi = (\xi_1, \xi_2, \cdots, \xi_m)$.

Proof. Take

$$X(t+h) = x + \Delta W_t(h) = x + \int_t^{t+h} \circ \, dW(s), \tag{49}$$

where the stochastic integral is the Stratonovich stochastic integral. Define

$$\hat{L}^{(j)} = \begin{cases} \dfrac{\partial}{\partial t}, \; j = 0, \\ \dfrac{\partial}{\partial x_j}, \; j = 1, 2, \cdots, m, \end{cases} \tag{50}$$

and let

$$\begin{aligned} &\hat{L}^{\gamma} = \hat{L}^{(j_1)}\hat{L}^{(j_2)}\cdots\hat{L}^{(j_l)}, \; \gamma = (j_1, j_2, \cdots, j_l) \in \mathcal{M}, \\ &J_{\gamma,t,t+h} = \int_t^{t+h}\int_t^{s_l}\cdots\int_t^{s_2} \circ \, dW^{j_1}(s_1) \circ \; dW^{j_2}(s_2) \cdots \circ \; dW^{j_l}(s_l), \end{aligned} \tag{51}$$

where $\circ \; dW^0(s) = ds$. The multiple integral $J_{\gamma,t,t+h}$ will be abbreviated as J_γ hereinafter.

Using Stratonovich-Taylor expansion formula [7], we have

$$u(t+h, X(t+h)) = \sum_{\gamma \in \mathcal{A}_p} \hat{u}_t^{\gamma} J_\gamma + R_p, \tag{52}$$

where $\hat{u}_t^{\gamma} = \hat{L}^{\gamma}u(t,x)$, $\mathbb{E}[R_p] = O(h^{p+1})$ and $\mathcal{A}_p$ is defined by (28).

Let $\Lambda_{j_1j_2\cdots j_l}, \; j_1, j_2, \cdots, j_l \in \{0, 1, \cdots, m\}$ denote the set generated by all the permutations of the indices $j_1, j_2, \cdots, j_l$. For example,

(53)

$$\begin{aligned} &\Lambda_{j_1} = \{(j_1)\}, \\ &\Lambda_{j_1j_2} = \{(j_1, j_2), (j_2, j_1)\}, \\ &\Lambda_{j_1j_2j_3} = \{(j_1, j_2, j_3), (j_1, j_3, j_2), (j_2, j_1, j_3), (j_2, j_3, j_1), (j_3, j_1, j_2), (j_3, j_2, j_1)\}. \end{aligned}$$

From [7], we know that

(54)

$$J_{(j_1,j_2,\cdots,j_l)}J_{(j_{l+1})} = J_{(j_{l+1},j_1,j_2,\cdots,j_l)} + J_{(j_1,j_{l+1},j_2,\cdots,j_l)} + \cdots + J_{(j_1,j_2,\cdots,j_l,j_{l+1})},$$

where $j_1, j_2, \cdots, j_{l+1} \in \{0, 1, 2, \cdots, m\}$. Repeatedly applying the formula (54), we have

(55)
$$\sum_{\gamma\in\Lambda_{j_1j_2\cdots j_l}} J_\gamma = J_{(j_1)} \sum_{\gamma\in\Lambda_{j_2j_3\cdots j_l}} J_\gamma = J_{(j_1)}J_{(j_2)} \sum_{\gamma\in\Lambda_{j_3j_4\cdots j_l}} J_\gamma = \cdots = J_{(j_1)}J_{(j_2)}\cdots J_{(j_l)}.$$

Similar to (29), we can easily verify that

(56) $$\hat{u}_t^\gamma = \hat{u}_t^\eta,\ \gamma,\eta \in \mathcal{A}_p$$

when γ, η differ only in the order of the indices.

Based on (55) and (56), Stratonovich-Taylor expansion (52) can be rewritten as

(57) $$u(t+h, X(t+h)) = \sum_{\gamma\in\mathcal{A}_p} \hat{u}_t^\gamma \hat{J}_\gamma + R_p,$$

where

$$\hat{J}_\gamma = \hat{J}_{(j_1,j_2,\cdots,j_l)} = \frac{1}{l!} J_{(j_1)}J_{(j_2)}\cdots J_{(j_l)}.$$

Using Taylor expansion of the multivariate function, we have

(58) $$u(t+h, x+\sqrt{h}\xi) = \sum_{\gamma\in\mathcal{A}_p} \hat{u}_t^\gamma \tilde{J}_\gamma + \tilde{R}_p,$$

with $\mathbb{E}[\tilde{R}_p] = O(h^{p+1})$ and

$$\tilde{J}_\gamma = \tilde{J}_{(j_1,j_2,\cdots,j_l)} = \frac{h^{\frac{l(\gamma)+n(\gamma)}{2}}}{l!}\xi_{j_1}\xi_{j_2}\cdots\xi_{j_l},$$

where $\xi_j = 1$ if $j = 0$.

If the condition (47) is fulfilled, then it is not difficult to show that

(59)
$$\mathbb{E}\big[\hat{J}_\gamma\big] = \mathbb{E}\big[\tilde{J}_\gamma\big],\ \gamma \in \mathcal{A}_p,$$
$$\mathbb{E}\big[\hat{J}_\gamma\big(W(t+h) - W(t)\big)\big] = \begin{cases} \mathbb{E}\big[\sqrt{h}\tilde{J}_\gamma\xi\big],\ l(\gamma)+n(\gamma) \le 2p, \\ \mathbb{E}\big[\sqrt{h}\tilde{J}_\gamma\xi\big] + O(h^{p+1}),\ l(\gamma)+n(\gamma) = 2p+1. \end{cases}$$

The conclusion (48) follows from (57)-(59).

The proof is complete. □

Remark 1. If the function $u \in C_b^{1+p,3+2p}$, based on Theorem 4.1, we have

(60)
$$\mathbb{E}\big[u^{(1)}\big(t+h, x+\Delta W_t(h)\big)\big] = \mathbb{E}\big[u^{(1)}(t+h, x+\sqrt{h}\xi)\big] + O(h^{p+1}),$$
$$\frac{1}{h}\mathbb{E}\big[u\big(t+h, x+\Delta W_t(h)\big)\Delta W_t(h)\big] = \frac{1}{h}\mathbb{E}\big[u(t+h, x+\sqrt{h}\xi)\sqrt{h}\xi\big] + O(h^p).$$

Note that

(61) $$y(t) = u(t, W(t)),\ z(t) = u^{(1)}(t, W(t)),$$

then combining (60) and (61), it can be seen that the accuracy of the approximation of $\mathbb{E}_{t_n}\big[\ z_{n+1}\big]$ is higher than that of $\frac{1}{h}\mathbb{E}_{t_n}\big[\ y_{n+1}\Delta W_t(h)\big]$ when $\Delta W_t(h)$ is replaced by ξ. This is why our RK method (7) uses $\mathbb{E}_{t_n}\big[\ z_{n+1}\big]$ instead of $\mathbb{E}_{t_n}\big[\ y_{n+1}H_{n,i,1}\big]$ when calculating the internal stages $Z_{n,i},\ 2 \le i \le q$.

Based on the condition (47), we list some discrete random variables in Table 1. It is not difficult to show that the random variable $\hat{\eta}_3$ satisfies the condition (47) with $p = 1$, η_3, η_5 satisfy the condition (47) with $p = 2$ and η_7 satisfies the condition (47) with $p = 3$. For a given positive integer p, the discrete random variable satisfying

TABLE 1. Some discrete random variables.

three-point	$\mathbb{P}(\hat{\eta}_3 = \pm\sqrt{\frac{3}{2}}) = \frac{1}{3},\ \mathbb{P}(\hat{\eta}_3 = 0) = \frac{1}{3}$ $\mathbb{P}(\eta_3 = \pm\sqrt{3}) = \frac{1}{6},\ \mathbb{P}(\eta_3 = 0) = \frac{2}{3}$
five-point	$\mathbb{P}(\eta_5 = \pm\sqrt{2}) = \frac{5}{24},\ \mathbb{P}(\eta_5 = \pm 2\sqrt{2}) = \frac{1}{96},\ \mathbb{P}(\eta_5 = 0) = \frac{9}{16}$
seven-point	$\mathbb{P}(\eta_7 = \pm\sqrt{\frac{3}{2}}) = \frac{25}{108},\ \mathbb{P}(\eta_7 = \pm 2\sqrt{\frac{3}{2}}) = \frac{13}{540},$ $\mathbb{P}(\eta_7 = \pm 3\sqrt{\frac{3}{2}}) = \frac{1}{1620},\ \mathbb{P}(\eta_7 = 0) = \frac{79}{162}$

condition (47) is not unique. The purpose of selecting $\hat{\eta}_3, \eta_3, \eta_5$ and η_7 is to avoid interpolation operations when calculating the conditional expectations (see below).

According to Theorem 3.1, we know that if the method (8) is of order p then the method

$$\begin{aligned} y_n &= \hat{\mathbb{E}}_{t_n}\Big[\ y_{n+1} + h\Phi(t_{n+1}, y_{n+1}, z_{n+1}, h, H)\ \Big],\ n = 0, 1, \cdots, N-1, \\ z_n &= \hat{\mathbb{E}}_{t_n}\Big[\ y_{n+1}H_{n,q+1,1} + h\Psi(t_{n+1}, y_{n+1}, z_{n+1}, h, H)\ \Big] \end{aligned} \tag{62}$$

is also of order p when the conditions

$$\begin{aligned} &\mathbb{E}_{t_n}\big[\ y_{n+1}\big] = \hat{\mathbb{E}}_{t_n}\big[\ y_{n+1}\big] + O(h^{p+1}), \\ &\mathbb{E}_{t_n}\big[\Phi(t_{n+1}, y_{n+1}, z_{n+1}, h, H)\big] = \hat{\mathbb{E}}_{t_n}\big[\Phi(t_{n+1}, y_{n+1}, z_{n+1}, h, H)\big] + O(h^p), \\ &\mathbb{E}_{t_n}\big[\ y_{n+1}H_{n,q+1,1}\big] = \hat{\mathbb{E}}_{t_n}\big[\ y_{n+1}H_{n,q+1,1}\big] + O(h^p), \\ &\mathbb{E}_{t_n}\big[\Psi(t_{n+1}, y_{n+1}, z_{n+1}, h, H)\big] = \hat{\mathbb{E}}_{t_n}\big[\Psi(t_{n+1}, y_{n+1}, z_{n+1}, h, H)\big] + O(h^{p-1}) \end{aligned} \tag{63}$$

are fulfilled, where $\hat{\mathbb{E}}_{t_n}\big[\cdot\big]$ is a approximation of $\mathbb{E}_{t_n}\big[\cdot\big]$. We can observe form (63) that different parts have different requirements for the approximation accuracy of the conditional expectations. It is these differences that make us more flexible in choosing the random variables to approximate the Brownian motion increments.

To remove the conditional expectations, we need spatial discretization. For ease of description, we assume $m = 1$ and let

$$D_n = \big\{x_i \mid x_i = i\sqrt{3h},\ i = 0, \pm 1, \cdots, \pm n\big\},\ n = 0, 1, \cdots, N \tag{64}$$

be a partition of spatial variable x at $t = t_n$. Based on Theorem 4.1 and Table 1, for $x_i \in D_n$, we have

$$\begin{aligned} \mathbb{E}\big[u\big(t_{n+1}, x_i + \Delta W_{t_n}(h)\big)\big] &= \mathbb{E}\big[u(t_{n+1}, x_i + \sqrt{h}\eta_3)\big] + O(h^3), \\ &= \frac{1}{6}u(t_{n+1}, x_{i-1}) + \frac{2}{3}u(t_{n+1}, x_i) \\ &\quad + \frac{1}{6}u(t_{n+1}, x_{i+1}) + O(h^3), \end{aligned} \tag{65}$$

(66)
$$\begin{aligned}\mathbb{E}\big[u\big(t_{n+1},x_i+\Delta W_{t_n}(h)\big)\Delta W_{t_n}(h)\big] &= \mathbb{E}\big[u(t_{n+1},x_i+\sqrt{h}\eta_3)\sqrt{h}\eta_3\big]+O(h^3)\\ &= -\frac{\sqrt{3h}}{6}u(t_{n+1},x_{i-1})+\frac{\sqrt{3h}}{6}u(t_{n+1},x_{i+1})\\ &\quad +O(h^3).\end{aligned}$$

Take $y_N^j=\varphi(x_j),\ z_N^j=\nabla\varphi(x_j)$, and let $y_{n+1}^j,\ z_{n+1}^j,\ f(t_{n+1},\ y_{n+1}^j,\ z_{n+1}^j)$ denote the approximations of $u(t_{n+1},x_j)$, $u^{(1)}(t_{n+1},x_j)$, $-u^{(0)}(t_{n+1},x_j)$, respectively. Combining formulas (63) and (65), we can remove the conditional expectations in the RK method (45) and obtain the following second order explicit method

(67)
$$\begin{aligned}Y_{n,2}^i &= \sum_{j=0}^{2}\Big(y_{n+1}^{i-1+j}+hf(t_{n+1},\ y_{n+1}^{i-1+j},\ z_{n+1}^{i-1+j})\ \Big)P_j,\ n=0,1,\cdots,N-1,\\ Z_{n,2}^i &= \sum_{j=0}^{2}\Big(z_{n+1}^{i-1+j}+f(t_{n+1},\ y_{n+1}^{i-1+j},\ z_{n+1}^{i-1+j})w_j\Big)P_j,\ i\in\{0,\pm1,\cdots,\pm n\},\\ y_n^i &= \sum_{j=0}^{2}\Big(y_{n+1}^{i-1+j}+\frac{1}{2}hf(t_{n+1},\ y_{n+1}^{i-1+j},\ z_{n+1}^{i-1+j})\Big)P_j+\frac{1}{2}hf(t_n,\ Y_{n,2}^i,\ Z_{n,2}^i),\\ z_n^i &= \frac{1}{h}\sum_{j=0}^{2}\Big(y_{n+1}^{i-1+j}+hf(t_{n+1},\ y_{n+1}^{i-1+j},\ z_{n+1}^{i-1+j})\Big)w_jP_j,\end{aligned}$$

where

$$P_0=P_2=\frac{1}{6},\ P_1=\frac{2}{3},\ w_0=-\sqrt{3h},\ w_1=0,\ w_2=\sqrt{3h}. \tag{68}$$

Similarly, for the third order RK method (46), we can replace $W(t_{n+1})-W(t_n)$ with $\sqrt{h}\eta_7$, and replace $W(t_{n+\frac{1}{2}})-W(t_n),\ W(t_{n+1})-W(t_{n+\frac{1}{2}})$ with $\sqrt{\frac{h}{2}}\eta_3$. Let

$$\begin{aligned}\hat{D}_n &= \big\{x_i\mid x_i=i\sqrt{\frac{3}{2}h},\ i=0,\pm1,\cdots,\pm3n\big\},\ n=0,1,\cdots,N,\\ \hat{D}_{n+\frac{1}{2}} &= \hat{D}_n\bigcup\big\{\pm(3n+1)\sqrt{\frac{3}{2}h}\big\},\ n=0,1,\cdots,N-1\end{aligned}\tag{69}$$

be the partitions of spatial variable x at $t=t_n$ and $t=t_n+\frac{h}{2}$, respectively. Then we can remove the conditional expectations in the RK method (46) and obtain the

method
(70)

$$
\begin{aligned}
Y_{n,2}^k &= \sum_{j=0}^{2}\Big(y_{n+1}^{k-1+j} + \frac{1}{2}hf(t_{n+1},\ y_{n+1}^{k-1+j},\ z_{n+1}^{k-1+j})\ \Big)P_j,\ n=0,1,\cdots,N-1,\\
Z_{n,2}^k &= \sum_{j=0}^{2}\Big(z_{n+1}^{k-1+j} + f(t_{n+1},\ y_{n+1}^{k-1+j},\ z_{n+1}^{k-1+j})\hat{w}_{j+2}\Big)P_j,\ k\in\{0,\pm1,\cdots,\pm(3n+1)\},\\
Y_{n,3}^i &= \sum_{j=0}^{6}\Big(y_{n+1}^{i-3+j} - hf(t_{n+1},\ y_{n+1}^{i-3+j},\ z_{n+1}^{i-3+j})\ \Big)\hat{P}_j\\
&\quad +2h\sum_{j=0}^{2}f(t_{n+\frac{1}{2}},\ Y_{n,2}^{i-1+j},\ Z_{n,2}^{i-1+j})P_j,\\
Z_{n,3}^i &= \sum_{j=0}^{6}\Big(z_{n+1}^{i-3+j} - f(t_{n+1},\ y_{n+1}^{i-3+j},\ z_{n+1}^{i-3+j})\hat{w}_j\Big)\hat{P}_j\\
&\quad +4\sum_{j=0}^{2}f(t_{n+\frac{1}{2}},\ Y_{n,2}^{i-1+j},\ Z_{n,2}^{i-1+j})\hat{w}_{j+2}P_j,\\
y_n^i &= \sum_{j=0}^{6}\Big(y_{n+1}^{i-3+j} + \frac{1}{6}hf(t_{n+1},\ y_{n+1}^{i-3+j},\ z_{n+1}^{i-3+j})\Big)\hat{P}_j\\
&\quad +\frac{2}{3}h\sum_{j=0}^{2}f(t_{n+\frac{1}{2}},\ Y_{n,2}^{i-1+j},\ Z_{n,2}^{i-1+j})P_j\\
&\quad +\frac{1}{6}hf(t_n,\ Y_{n,3}^i,\ Z_{n,3}^i),\\
z_n^i &= \frac{1}{h}\sum_{j=0}^{6}y_{n+1}^{i-3+j}\hat{w}_j\hat{P}_j\\
&\quad +2\sum_{j=0}^{2}f(t_{n+\frac{1}{2}},\ Y_{n,2}^{i-1+j},\ Z_{n,2}^{i-1+j})\hat{w}_{j+2}P_j,\ i\in\{0,\pm1,\cdots,\pm3n\},
\end{aligned}
$$

where

$$
\begin{aligned}
&\hat{P}_0 = \hat{P}_6 = \frac{1}{1620},\ \hat{P}_1 = \hat{P}_5 = \frac{13}{540},\ \hat{P}_2 = \hat{P}_4 = \frac{25}{108},\ \hat{P}_3 = \frac{79}{162},\\
&\hat{w}_i = (i-3)\sqrt{\frac{3}{2}h},\ i\in\{0,1,2,\ldots,6\}.
\end{aligned} \tag{71}
$$

To illustrate that the ideal of our approximate algorithm can be applied to other methods of solving the BSDEs, we will take the two-step and three-step explicit Adams methods as examples.

For the BSDE (1), the two-step second order explicit Adams method can be written as

$$
\begin{aligned}
y_n &= \mathbb{E}_{t_n}\Big[\ y_{n+1} + h\big(\frac{3}{2}f(t_{n+1},\ y_{n+1},\ z_{n+1}) - \frac{1}{2}f(t_{n+2},\ y_{n+2},\ z_{n+2})\ \big)\Big],\\
z_n &= \frac{1}{h}\mathbb{E}_{t_n}\Big[\ y_{n+1}\Delta W_{t_n}(h) + hf(t_{n+1},\ y_{n+1},\ z_{n+1})\Delta W_{t_n}(h)\ \Big].
\end{aligned} \tag{72}
$$

If the conditions

$$(73)\quad \begin{aligned} &\mathbb{E}_{t_n}[\, y_{n+1}] = \hat{\mathbb{E}}_{t_n}[\, y_{n+1}] + O(h^3),\\ &\mathbb{E}_{t_n}[\, y_{n+1}\Delta W_{t_n}(h)] = \hat{\mathbb{E}}_{t_n}[\, y_{n+1}\Delta W_{t_n}(h)] + O(h^3),\\ &\mathbb{E}_{t_n}[\, f(t_{n+i},\ y_{n+i},\ z_{n+i})] = \hat{\mathbb{E}}_{t_n}[\, f(t_{n+i},\ y_{n+i},\ z_{n+i})] + O(h^2),\ i=1,2,\\ &\mathbb{E}_{t_n}[\, f(t_{n+1},\ y_{n+1},\ z_{n+1})\Delta W_{t_n}(h)]\\ =\ &\hat{\mathbb{E}}_{t_n}[\, f(t_{n+1},\ y_{n+1},\ z_{n+1})\Delta W_{t_n}(h)] + O(h^2), \end{aligned}$$

are fulfilled, then the method

$$(74)\quad \begin{aligned} y_n &= \hat{\mathbb{E}}_{t_n}\Big[\, y_{n+1} + h\big(\frac{3}{2}f(t_{n+1},\ y_{n+1},\ z_{n+1}) - \frac{1}{2}f(t_{n+2},\ y_{n+2},\ z_{n+2})\ \big)\Big],\\ z_n &= \frac{1}{h}\hat{\mathbb{E}}_{t_n}\Big[\, y_{n+1}\Delta W_{t_n}(h) + hf(t_{n+1},\ y_{n+1},\ z_{n+1})\Delta W_{t_n}(h)\ \Big]. \end{aligned}$$

is also second order. Based on Theorem 4.1 and Table 1, it is not difficult to verify that the conditions in (73) are fulfilled when $\Delta W_{t_n}(h)$ is replaced by $\sqrt{h}\eta_3$, $\Delta W_{t_n}(2h)$ is replaced by $\sqrt{2h}\hat{\eta}_3$. Assume that the values of $y^j_{N-k}, z^j_{N-k},\ x_j \in D_{N-k},\ k=0,1$ are given (We can use the second order RK method above to obtain these values). Then we can remove the conditional expectations in the method (72) and obtain the method

(75)

$$\begin{aligned} y^i_n &= \sum_{j=0}^{2}\Big(y^{i-1+j}_{n+1} + \frac{3}{2}hf(t_{n+1},\ y^{i-1+j}_{n+1},\ z^{i-1+j}_{n+1})\Big)P_j\\ &\quad -\frac{1}{2}h\sum_{j=0}^{2} f(t_{n+2},\ y^{i-1+j}_{n+2},\ z^{i-1+j}_{n+2})\bar{P}_j,\\ z^i_n &= \frac{1}{h}\sum_{j=0}^{2}\Big(y^{i-1+j}_{n+1} + hf(t_{n+1},\ y^{i-1+j}_{n+1},\ z^{i-1+j}_{n+1})\Big)w_jP_j,\ n=0,1,\cdots,N-2, \end{aligned}$$

where

$$(76)\quad \bar{P}_0 = \bar{P}_1 = \bar{P}_2 = \frac{1}{3}.$$

Similarly, for the three-step third order explicit Adams method

$$(77)\quad \begin{aligned} y_n &= \mathbb{E}_{t_n}\Big[\, y_{n+1} + h\big(\frac{23}{12}f(t_{n+1},\ y_{n+1},\ z_{n+1}) - \frac{4}{3}f(t_{n+2},\ y_{n+2},\ z_{n+2})\\ &\qquad + \frac{5}{12}f(t_{n+3},\ y_{n+3},\ z_{n+3})\ \big)\Big],\\ z_n &= \frac{1}{h}\mathbb{E}_{t_n}\Big[\, y_{n+1}\Delta W_{t_n}(h) + h\big(\frac{3}{2}f(t_{n+1},\ y_{n+1},\ z_{n+1})\Delta W_{t_n}(h)\\ &\qquad - \frac{1}{4}f(t_{n+2},\ y_{n+2},\ z_{n+2})\Delta W_{t_n}(2h)\big)\ \Big], \end{aligned}$$

we can replace $\Delta W_{t_n}(h), \Delta W_{t_n}(2h), \Delta W_{t_n}(3h)$ with $\sqrt{h}\eta_7, \sqrt{2h}\eta_3, \sqrt{3h}\eta_5$, respectively.

Assume that the values of $y^j_{N-k}, z^j_{N-k},\ x_j \in \hat{D}_{N-k},\ k=0,1,2$ are given (We can use the third order RK method above to obtain these values). Then we can

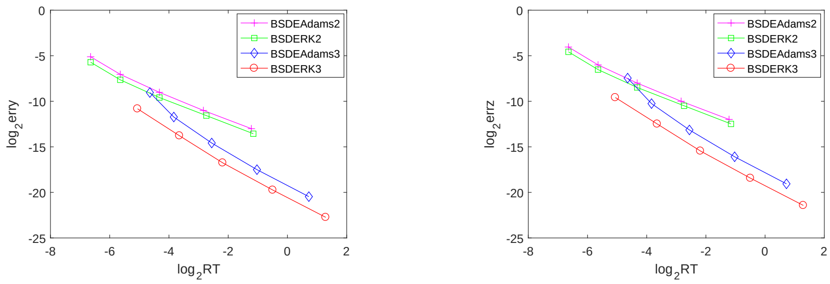

FIGURE 1. RT vs. $erry$ (left), and RT vs. $errz$ (right) for Example 5.1.

remove the conditional expectations in the method (77) and obtain the method

$$
\begin{aligned}
y_n^i =& \sum_{j=0}^{6}\Big(y_{n+1}^{i-3+j}+\frac{23}{12}hf(t_{n+1},\ y_{n+1}^{i-3+j},\ z_{n+1}^{i-3+j})\Big)\hat{P}_j \\
&-\frac{4}{3}h\sum_{j=0}^{2} f(t_{n+2},\ y_{n+2}^{i-2+2j},\ z_{n+2}^{i-2+2j})P_j \\
&+\frac{5}{12}h\sum_{j=0}^{4} f(t_{n+3},\ y_{n+3}^{i-4+2j},\ z_{n+3}^{i-4+2j})\tilde{P}_j, \\
z_n^i =& \frac{1}{h}\sum_{j=0}^{6}\Big(y_{n+1}^{i-3+j}+\frac{3}{2}hf(t_{n+1},\ y_{n+1}^{i-3+j},\ z_{n+1}^{i-3+j})\Big)\hat{w}_j\hat{P}_j \\
&-\frac{1}{4}\sum_{j=0}^{2} f(t_{n+2},\ y_{n+2}^{i-2+2j},\ z_{n+2}^{i-2+2j})\hat{w}_{2j+1}P_j,\ n=0,1,\cdots,N-3,
\end{aligned} \tag{78}
$$

where

$$
\tilde{P}_0=\tilde{P}_4=\frac{1}{96},\ \tilde{P}_1=\tilde{P}_3=\frac{5}{24},\ \tilde{P}_2=\frac{9}{16}. \tag{79}
$$

5. Numerical results

In this section, four numerical examples will be presented to illustrate our theoretical results. For ease of description, we denote the methods (67), (70) (75) and (78) by BSDERK2, BSDERK3, BSDEAdams2 and BSDEAdams3, respectively.

Example 5.1. Inspired by [20], we consider the BSDE with $d=m=1$:

$$
y(t)=sin\big(2W(T)+T\big)+cos\big(2W(T)+T\big)+\int_t^T \frac{4y(s)-z(s)}{y^2(s)+\frac{1}{4}z^2(s)}\,ds-\int_t^T z(s)\,dW(s), \tag{80}
$$

whose analytic solution is given by

$$
y(t)=sin\big(2W(t)+t\big)+cos\big(2W(t)+t\big),\ z(t)=2\Big(cos\big(2W(t)+t\big)-sin\big(2W(t)+t\big)\Big). \tag{81}
$$

We take $T=1$ for all the examples in this section. We observe the error at time $t=0$ and let $erry=\mid y(0)-y_0\mid$, $errz=\mid z(0)-z_0\mid$. Take $h=2^{-4},2^{-5},\cdots,2^{-8}$ in turn, then we can obtain Table 2, where RT denotes the running time and p denotes the numerical convergence order obtained by using the least square fitting.

TABLE 2. The results of Example 5.1.

		$h=2^{-4}$	$h=2^{-5}$	$h=2^{-6}$	$h=2^{-7}$	$h=2^{-8}$	p
BSDEAdams2	$erry$	2.89(−2)	7.59(−3)	1.93(−3)	4.87(−4)	1.22(−4)	1.97
	$errz$	6.07(−2)	1.56(−2)	3.91(−3)	9.80(−4)	2.46(−4)	1.99
	RT	0.01	0.02	0.05	0.14	0.43	
BSDERK2	$erry$	1.91(−2)	5.01(−3)	1.30(−3)	3.29(−4)	8.29(−5)	1.96
	$errz$	4.21(−2)	1.09(−2)	2.80(−3)	6.95(−4)	1.75(−4)	1.98
	RT	0.01	0.02	0.05	0.15	0.45	
BSDEAdams3	$erry$	1.90(−3)	2.95(−4)	4.09(−5)	5.38(−6)	6.91(−7)	2.86
	$errz$	5.70(−3)	8.20(−4)	1.10(−4)	1.44(−5)	1.83(−6)	2.90
	RT	0.04	0.07	0.17	0.49	1.65	
BSDERK3	$erry$	5.52(−4)	7.10(−5)	9.03(−6)	1.14(−6)	1.43(−7)	2.98
	$errz$	1.30(−3)	1.74(−4)	2.23(−5)	2.82(−6)	3.54(−7)	2.96
	RT	0.03	0.08	0.22	0.71	2.45	

We can observe from Table 2 that the methods BSDEAdams2, BSDERK2 can achieve second order and the methods BSDEAdams3, BSDERK3 can achieve third order. This confirms our theoretical results. Using the data in Table 2, we can obtain Figure 1, which can make it easier to compare the computational efficiency of each method.

Example 5.2. We consider the BSDE with $d = 1,\ m = 2$:

$$\begin{aligned} y(t) = {} & \frac{e^{W^1(T)+\frac{1}{4}W^2(T)+T}}{\big(e^{W^1(T)+\frac{1}{4}W^2(T)+T}+1\big)} + \int_t^T \big(z_1(s)+z_2(s)\big)\Big(\frac{17}{20}y(s)-\frac{49}{40}\Big)\, ds \\ & - \int_t^T z_1(s)\, dW^1(s) - \int_t^T z_2(s)\, dW^2(s), \end{aligned} \tag{82}$$

whose analytic solution is given by

$$\begin{aligned} y(t) &= \frac{e^{W^1(t)+\frac{1}{4}W^2(t)+t}}{\big(e^{W^1(t)+\frac{1}{4}W^2(t)+t}+1\big)},\ z_1(t) = \frac{e^{W^1(t)+\frac{1}{4}W^2(t)+t}}{\big(e^{W^1(t)+\frac{1}{4}W^2(t)+t}+1\big)^2}, \\ z_2(t) &= \frac{1}{4}\frac{e^{W^1(t)+\frac{1}{4}W^2(t)+t}}{\big(e^{W^1(t)+\frac{1}{4}W^2(t)+t}+1\big)^2}. \end{aligned} \tag{83}$$

The corresponding numerical results are presented in Table 3 and Figure 2.

Example 5.3. We consider the BSDE with $d = 1,\ m = 3$:

$$\begin{aligned} y(t) = {} & \frac{1}{2}sin\Big(\frac{1}{2}W^1(T)+W^2(T)+W^3(T)+T\Big) \\ & + \int_t^T \frac{\frac{9}{32}y(s)-\frac{1}{10}\big(z_1(s)+z_2(s)+z_3(s)\big)}{y^2(s)+4z_1^2(s)}\, ds \\ & - \int_t^T z_1(s)\, dW^1(s) - \int_t^T z_2(s)\, dW^2(s) - \int_t^T z_3(s)\, dW^3(s), \end{aligned} \tag{84}$$

TABLE 3. The results of Example 5.2.

		$h=2^{-4}$	$h=2^{-5}$	$h=2^{-6}$	$h=2^{-7}$	$h=2^{-8}$	p
BSDEAdams2	$erry$	9.75(−5)	3.19(−5)	8.90(−6)	2.34(−6)	6.00(−7)	1.85
	$errz$	6.32(−4)	1.74(−4)	4.55(−5)	1.16(−5)	2.94(−6)	1.94
	RT	0.13	0.52	3.87	30.96	259.74	
BSDERK2	$erry$	1.29(−4)	3.71(−5)	9.95(−6)	2.57(−6)	6.53(−7)	1.91
	$errz$	4.31(−4)	1.17(−4)	3.03(−5)	7.72(−6)	1.95(−6)	1.95
	RT	0.14	0.58	4.36	34.89	292.99	
BSDEAdams3	$erry$	1.42(−5)	1.98(−6)	2.62(−7)	3.38(−8)	4.29(−9)	2.93
	$errz$	5.19(−5)	7.29(−6)	9.72(−7)	1.26(−7)	1.60(−8)	2.92
	RT	1.08	7.33	57.12	486.65	4394.01	
BSDERK3	$erry$	6.73(−6)	9.03(−7)	1.17(−7)	1.49(−8)	1.87(−9)	2.95
	$errz$	1.43(−5)	1.87(−6)	2.38(−7)	3.00(−8)	3.77(−9)	2.97
	RT	1.34	10.56	92.97	836.14	7083.41	

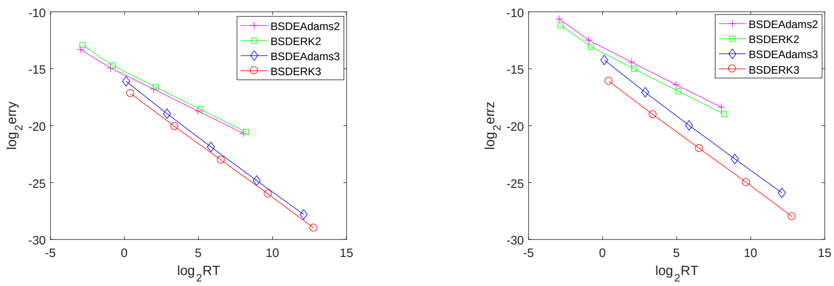

FIGURE 2. RT vs. $erry$ (left), and RT vs. $errz$ (right) for Example 5.2.

whose analytic solution is given by

$$
\begin{aligned}
y(t) &= \frac{1}{2}sin\Big(\frac{1}{2}W^1(t) + W^2(t) + W^3(t) + t\Big), \\
z_1(t) &= \frac{1}{4}cos\Big(\frac{1}{2}W^1(t) + W^2(t) + W^3(t) + t\Big), \qquad (85)\\
z_2(t) = z_3(t) &= \frac{1}{2}cos\Big(\frac{1}{2}W^1(t) + W^2(t) + W^3(t) + t\Big).
\end{aligned}
$$

The corresponding numerical results are presented in Table 4 and Figure 3.

TABLE 4. The results of Example 5.3.

		$h=\frac{1}{8}$	$h=\frac{1}{16}$	$h=\frac{1}{24}$	$h=\frac{1}{32}$	$h=\frac{1}{40}$	p
	$erry$	1.44(−2)	3.99(−3)	1.82(−3)	1.04(−3)	6.70(−4)	1.91
BSDEAdams2	$errz$	2.75(−3)	1.12(−3)	5.50(−4)	3.23(−4)	2.12(−4)	1.60
	RT	0.18	2.11	10.59	33.70	84.89	
	$erry$	7.71(−3)	2.20(−3)	1.01(−3)	5.93(−4)	3.85(−4)	1.86
BSDERK2	$errz$	1.92(−3)	6.23(−4)	2.92(−4)	1.68(−4)	1.09(−4)	1.78
	RT	0.20	2.45	12.40	39.77	101.02	
	$erry$	1.52(−3)	2.20(−4)	6.93(−5)	3.01(−5)	1.57(−5)	2.84
BSDEAdams3	$errz$	1.43(−3)	2.07(−4)	6.46(−5)	2.80(−5)	1.46(−5)	2.85
	RT	8.66	137.57	696.52	2267.53	5633.73	
	$erry$	5.37(−4)	7.05(−5)	2.13(−5)	9.07(−6)	4.67(−6)	2.95
BSDERK3	$errz$	2.49(−4)	3.12(−5)	9.29(−6)	3.93(−6)	2.02(−6)	2.99
	RT	9.75	173.01	917.29	3022.01	7713.01	

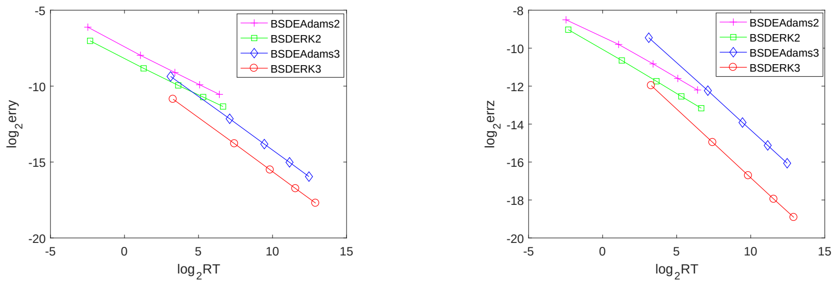

FIGURE 3. RT vs. $erry$ (left), and RT vs. $errz$ (right) for Example 5.3.

Example 5.4. Finally, we consider the BSDE with $d=2,\ m=1$:

$$\begin{aligned} y(t) &= \begin{pmatrix} sin\left(\frac{7}{5}W(T)+T\right) \\ cos\left(\frac{7}{5}W(T)+T\right) \end{pmatrix} \\ &\quad + \int_t^T \left(\frac{25}{49}z_1^2(t)+\frac{25}{49}z_2^2(t)\right) \begin{pmatrix} \frac{49}{50} & -1 \\ 1 & \frac{49}{50} \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} ds \\ &\quad - \int_t^T \begin{pmatrix} z_1(t) \\ z_2(t) \end{pmatrix} dW(s), \end{aligned} \tag{86}$$

whose analytic solution is given by

$$y(t) = \begin{pmatrix} sin\left(\frac{7}{5}W(t)+t\right) \\ cos\left(\frac{7}{5}W(t)+t\right) \end{pmatrix} \quad z(t) = \begin{pmatrix} \frac{7}{5}cos\left(\frac{7}{5}W(t)+t\right) \\ -\frac{7}{5}sin\left(\frac{7}{5}W(t)+t\right) \end{pmatrix}. \tag{87}$$

The corresponding numerical results are presented in Table 5 and Figure 4.

Compared with the Adams multistep methods of the same order, the running time RT of our RK methods will be a little more for the same stepsize h. But when

TABLE 5. The results of Example 5.4.

		$h=2^{-4}$	$h=2^{-5}$	$h=2^{-6}$	$h=2^{-7}$	$h=2^{-8}$	p
BSDEAdams2	$erry$	8.16(−3)	2.71(−3)	7.71(−4)	2.05 − 4)	5.28(−5)	1.83
	$errz$	9.99(−3)	3.23(−3)	9.25(−4)	2.47(−4)	6.38(−5)	1.83
	RT	0.02	0.03	0.06	0.21	0.75	
BSDERK2	$erry$	4.12(−3)	7.75(−4)	1.58(−4)	3.54(−5)	8.43(−6)	2.23
	$errz$	1.09(−2)	2.41(−3)	5.52(−4)	1.31(−4)	3.19(−5)	2.10
	RT	0.02	0.03	0.06	0.22	0.77	
BSDEAdams3	$erry$	2.18(−3)	3.36(−4)	4.66(−5)	6.14(−6)	7.87(−7)	2.86
	$errz$	3.25(−3)	4.98(−4)	6.88(−5)	9.04(−6)	1.16(−6)	2.87
	RT	0.03	0.08	0.23	0.79	3.06	
BSDERK3	$erry$	7.89(−4)	1.16(−4)	1.57(−5)	2.05(−6)	2.61(−7)	2.89
	$errz$	1.08(−3)	1.61(−4)	2.19(−5)	2.85(−6)	3.64(−7)	2.89
	RT	0.04	0.09	0.35	1.24	4.79	

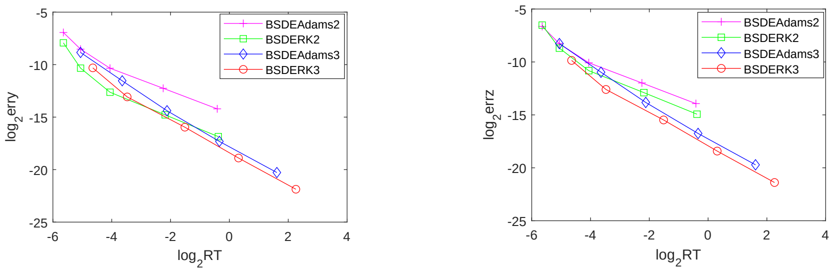

FIGURE 4. RT vs. $erry$ (left), and RT vs. $errz$ (right) for Example 5.4.

we consider both calculation time and accuracy comprehensively, we can observe from Figures 1-4 that our RK methods (especially method BSDERK3) usually have better precision than the Adams multistep methods of the same order under the same running time. This means that our RK method performs better than the Adams multistep methods in the four examples above.

Acknowledgments

This research is partially supported by the National Science Foundation of China under grants 12101525, 61873325, 11831010 and the Southern University of Science and Technology Start up fund Y01286120.

References

[1] J.-F. Chassagneux, Linear multistep schemes for BSDEs, SIAM J. Numer. Anal., 52, 2815-2836 (2014).

[2] J.-F. Chassagneux and D. Crisan, Runge-Kutta schemes for backward stochastic differential equations, Ann. Appl. Prob., 24, 679-720 (2014).

[3] K. W. Chau and C. W. Oosterlee, On the wavelet-based swift method for backward stochastic differential equations, IMA J. Numer. Anal., 38, 1051-1083 (2017).

[4] D. Crisan and K. Manolarakis, Second order discretization of backward SDEs and simulation with the cubature method, Ann. Appl. Prob., 24, 652-678 (2014).

[5] J. Douglas Jr, J. Ma and P. Protter, Numerical methods for forward-backward stochastic differential equations, Ann. Appl. Prob., 6, 940-968 (1996).

[6] E. Gobet and C. Labart, Error expansion for the discretization of backward stochastic differential equations, Stoch. Proc. Appl., 117, 803-829 (2007).

[7] P. E. Kloeden and E. Platen, Numerical Solution of Stochastic Differential Equations, Springer-Verlag, (1992).

[8] J. Ma, P. Protter and J. Yong, Solving forward-backward stochastic differential equations explicitly-a four step scheme, Probab. Theory Related Fields, 98, 339-359 (1994).

[9] J. Ma and J. Zhang, Representation theorems for backward stochastic differential equations, Ann. Appl. Prob., 12, 1390-1418 (2002).

[10] G. N. Milstein and M. V. Tretyakov, Numerical algorithms for forward-backward stochastic differential equations, SIAM J. Sci. Comput., 28, 561-582 (2006).

[11] G. N. Milstein and M. V. Tretyakov, Discretization of forward-backward stochastic differential equations and related quasi-linear parabolic equations, IMA J. Numer. Anal., 27, 24-44 (2007).

[12] E. Pardoux and S. Peng, Adapted solution of a backward stochastic differential equation, Systems Control Lett., 14, 55-61 (1990).

[13] E. Pardoux and S. Peng, Backward stochastic differential equations and quasilinear parabolic partial differential equations, Stochastic Partial Differential Equations and Their Applications, 200-217, Springer (1992).

[14] S. Peng, Probabilistic interpretation for systems of quasilinear parabolic partial differential equations, Stochastics Stochastics Rep., 37, 61-74 (1991).

[15] M.J. Ruijter and C.W. Oosterlee, A fourier cosine method for an efficient computation of solutions to BSDEs, SIAM J. Sci. Comput., 37, A859-A889 (2015).

[16] Y. Sun and W. Zhao, New second-order schemes for forward backward stochastic differential equations, East Asian J. Appl. Math., 8, 399-421 (2018).

[17] X. Tang and J. Xiong, Stability analysis of general multistep methods for Markovian backward stochastic differential equations, IMA J. Numer. Anal., 42, 1789C1805 (2022).

[18] J. Yang, W. Zhao and T. Zhou, A unified probabilistic discretization scheme for FBSDEs: stability, consistency, and convergence analysis, SIAM J. Numer. Anal., 58, 2351-2375 (2020).

[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, 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).

[21] 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).

[22] W. Zhao, Y. Li and G. Zhang, A generalized θ-scheme for solving backward stochastic differential equations, Discrete Contin. Dyn. Syst. Ser. B, 17, 1585-1603 (2012).

[23] W. Zhao, Y. Fu and T. Zhou, New kinds of high-order multistep schemes for coupled forward backward stochastic differential equations, SIAM J. Sci. Comput., 36, A1731-A1751 (2014).

[24] W. Zhao, W. Zhang and L. Ju, A numerical method and its error estimates for the decoupled forward-backward stochastic differential equations, Comnun. Comput. Phys., 15, 618-646 (2014).

[25] W. Zhao, W. Zhang and L. Ju, A multistep scheme for decoupled forward-backward stochastic differential equations, Numer. Math. Theory Methods Appl., 9, 262-288 (2016).

1. School of Mathematics and Computational Science & Hunan Key Laboratory for Computation and Simulation in Science and Engineering, Xiangtan University, Xiangtan, 411105, China
E-mail: `tangx@xtu.edu.cn, 1293060753@qq.com`

2. Department of Mathematics & Shenzhen International Center for Mathematics, Southern University of Science and Technology, Shenzhen, 518055, China
E-mail: `xiongj@sustech.edu.cn`