INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 10, Number 4, Pages 876–898

ERROR ESTIMATES OF THE CRANK-NICOLSON SCHEME FOR SOLVING BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS

WEIDONG ZHAO, YANG LI AND LILI JU

Abstract. In this paper, we study error estimates of a special θ-scheme – the Crank-Nicolson scheme proposed in [25] for solving the backward stochastic differential equation with a general generator, $-dy_t = f(t, y_t, z_t)dt - z_t dW_t$. We rigorously prove that under some reasonable regularity conditions on φ and f, this scheme is second-order accurate for solving both y_t and z_t when the errors are measured in the L^p $(p \geq 1)$ norm.

Key words. Backward stochastic differential equations, Crank-Nicolson scheme, θ-scheme, error estimate

1. Introduction

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $T > 0$ a finite time, $\mathbb{F} = (\mathcal{F}_t)_{0\leq t\leq T}$ a filtration satisfying the usual conditions. Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a complete, filtered probability space on which a standard d-dimensional Brownian motion W_t is defined and $\mathcal{F}_0$ contains all the P-null sets of $\mathcal{F}$. Let $L^2 = L^2_{\mathcal{F}}(0,T)$ be the set of all $\mathcal{F}_t$-adapted and mean-square-integrable vector/matrix processes. We consider the backward stochastic differential equation (BSDE)

$$-dy_t = f(t, y_t, z_t)dt - z_t dW_t, \quad \forall\, t \in [0, T), \tag{1.1}$$

with the terminal condition

$$y_T = \xi,$$

where the generator $f = f(t, y_t, z_t)$ is a vector function valued in R^m and is $\mathcal{F}_t$-adapted for each (y, z), and the terminal variable $\xi \in L^2$ is $\mathcal{F}_T$ measurable. Rewriting the BSDE (1.1) in the integral form gives us

$$y_t = \xi + \int_t^T f(s, y_s, z_s)\, ds - \int_t^T z_s\, dW_s, \quad \forall\, t \in [0, T). \tag{1.2}$$

We note that the second integral term on the right-hand side of (1.2) is an Itô-type integral. A process (y_t, z_t): $[0,T] \times \Omega \to R^m \times R^{m\times d}$ is called an L^2-solution of the BSDE (1.2) if, in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, it is $\{\mathcal{F}_t\}$-adapted, square integrable, and satisfies the integral equation (1.2) [16].

In 1990, Pardoux and Peng first proved in [16] the existence and uniqueness of the solution of general nonlinear BSDEs (i.e, f is nonlinear), and later in [17], obtained some relations between BSDEs and stochastic partial differential equations (SPDEs). Since then, the theory of BSDEs has been extensively studied by many researchers and BSDEs have found applications in many fields, such as finance, risk measure, stochastic control, and etc.. Peng obtained the relation between BSDEs and parabolic PDEs in [19], and then the generalized stochastic maximum principle

Received by the editors October 28, 2011 and, in revised form, November 6, 2012.

2000 *Mathematics Subject Classification.* 60H35, 60H10, 65C30.

This work is partially supported by the National Natural Science Foundations of China under grant numbers 91130003 and 11171189, and the Natural Science Foundation of Shandong Province under grant number ZR2011AZ002.

and the dynamic programming principle for stochastic control problems based on BSDEs in [18, 21]. The nonlinear $g-$expectation via a particular nonlinear BSDE was introduced in [20], and in [7] it was found that a dynamic coherent risk measure can be represented by a properly defined g-expectation. Thus, it is very important and useful to study solutions of BSDEs.

In this paper, we consider the case of $\xi = \varphi(W_T)$, and assume that the BSDE (1.2) has a unique solution (y_t, z_t). It was shown in [19] that the solution (y_t, z_t) of (1.2) can be represented as

$$y_t = u(t, W_t), \qquad z_t = \nabla_x u(t, W_t), \qquad \forall t \in [0, T), \tag{1.3}$$

where $u(t, x)$ is the solution of the following parabolic partial differential equation

$$\frac{\partial u}{\partial t} + \frac{1}{2}\sum_{i=1}^{d} \frac{\partial^2 u}{\partial x_i^2} + f(t, u, \nabla_x u) = 0, \tag{1.4}$$

with the terminal condition $u(T, x) = \varphi(x)$, and $\nabla_x u$ is the gradient of u with respect to the spacial variable x. The smoothness of u clearly depends on ϕ and f.

It is well-known that it is often difficult to obtain analytic solutions of BSDEs, so that computing their approximate solutions becomes highly desired. Based on the relation between the BSDEs and the corresponding parabolic PDEs, some numerical algorithms were proposed to solve BSDEs [3, 11, 12, 13, 14, 15, 19, 24], and furthermore, a four step algorithm was proposed in [10] to solve a class of more general equations called forward-backward stochastic differential equations (FBSDEs). In [25], a family of θ-schemes were proposed for solving general BSDEs. In particular, a special case of the θ-scheme – the *Crank-Nicolson* (C-N) scheme was numerically demonstrated to be *second-order accurate.* This accuracy result was theoretically proven in [22, 26] for the simplified case that the generator function f is independent of z_t in (1.2), however, the proof for the cases of general generators remains open till now. A family of multi-step schemes were recently developed in [27] based on the Lagrange interpolation and the Gauss-Hermite quadratures. Accuracies of these multi-step schemes were numerically shown to be of high order for solving the BSDE (1.2), but again the result was only theoretically confirmed for BSDEs with a generator f independent of z_t. There are also some other numerical methods for solving BSDEs (or FBSDEs), which were proposed based on directly discretizing BSDEs or FBSDEs, see [1, 2, 4, 5, 8, 9, 21, 23, 24] and references cited therein.

The aim of this paper is to study error estimates of the special θ scheme – the Crank-Nicolson scheme for solving the general BSDE (1.2) with terminal condition $\xi = \varphi(W_T)$. For the purpose of simple representations, let us first introduce the following notations:

- $\|X\|_{L^p}$ $(p \geq 1)$: the L^p-norm for $X \in L^p$ defined by $\mathbb{E}[|X|^p]^{\frac{1}{p}}$.
- $C_b^{l,k,k}$: the set of continuously differential functions $\psi : [0, T] \times R^d \times R^{m\times d} \to R$ with uniformly bounded partial derivatives $\partial_t^{l_1}\psi$ and $\partial_y^{k_1}\partial_z^{k_2}\psi$ for $l_1 \leq l$ and $k_1 + k_2 \leq k$.
- $C_b^{l,k}$: the set of functions $\psi : (t, x) \in [0, T] \times R^d \to R$ with uniformly bounded partial derivatives $\partial_t^{l_1}\partial_x^{k_1}\psi$ for $l_1 \leq l$ and $k_1 \leq k$.
- C_b^k: the set of functions $\psi : x \in R^d \to R$ with uniformly bounded partial derivatives $\partial_x^{k_1}\psi$ for $k_1 \leq k$.
- $\mathcal{F}_s^{t,x}(t \leq s \leq T)$: the σ-field generated by the Brownian motion $\{x + W_r - W_t, t \leq r \leq s\}$ starting from the time-space point (t, x). Let $\mathcal{F}^{t,x} = \mathcal{F}_T^{t,x}$.
- $\mathbb{E}[X]$: the mathematical expectation of the random variable X.

• $\mathbb{E}_s^{t,x}[X]$: the conditional mathematical expectation of the random variable X under the σ-field $\mathcal{F}_s^{t,x}(t \leq s \leq T)$, that is $\mathbb{E}_s^{t,x}[X] = \mathbb{E}[X|\mathcal{F}_s^{t,x}]$. Let $\mathbb{E}_t^{x}[X] = \mathbb{E}[X|\mathcal{F}_t^{t,x}]$.
• $\partial_x\psi$: the matrix valued function $\partial_x\psi = (\partial_{x^j}\psi^i)_{m\times d}\,(1 \leq i \leq m, 1 \leq j \leq d)$ for vector function $\psi = (\psi^1, \cdots, \psi^m)^\top$.
• C: a generic positive constant, and may be different from line to line.

The rest of the paper is organized as follows. In Section 2, we briefly review the θ-scheme proposed in [25] for solving the BSDE (1.2) and its special case – the Crank-Nicolson scheme. Then we rigorously derive error estimates of the C-N scheme in Section 3. Under some reasonable regularity conditions on φ and f, we prove that the C-N scheme is second-order accurate when the errors are measured in the L^p-norm ($p \geq 1$). Some concluding remarks are finally given in Section 4.

2. Review of the θ-Scheme and the Crank-Nicolson Scheme

In this section, we give a brief review of the θ-scheme proposed in [25]. For the time interval $[0, T]$, let us introduce the following partition

$$0 = t_0 < \cdots < t_N = T$$

with $\Delta t_n = t_{n+1} - t_n$, $n = 0, 1, \cdots, N-1$. Denote by (t_n, x^n) the time-space points, where $x^{n+1} = x^n + W_{t_{n+1}} - W_{t_n}$ for $n = 0, 1, \cdots, N-1$.

2.1. Reference equations and the θ-Scheme. Let (y_t, z_t) be the solution of the BSDE (1.2). Then, for $0 \leq n \leq N-1$, it is easy to obtain

$$y_{t_n} = y_{t_{n+1}} + \int_{t_n}^{t_{n+1}} f(s, y_s, z_s)\,ds - \int_{t_n}^{t_{n+1}} z_s\,dW_s. \tag{2.1}$$

Taking the conditional mathematical expectation $\mathbb{E}_{t_n}^{x^n}[\cdot]$ on both sides of (2.1), we get

$$y_{t_n} = \mathbb{E}_{t_n}^{x^n}[y_{t_{n+1}}] + \int_{t_n}^{t_{n+1}} \mathbb{E}_{t_n}^{x^n}[f(s, y_s, z_s)]\,ds. \tag{2.2}$$

The integrand $\mathbb{E}_{t_n}^{x^n}[f(s, y_s, z_s)]$ on the right-hand side of (2.2) is a deterministic smooth function of time s. We can use the the following rule to approximate the integral in (2.2):

$$\begin{aligned}\int_{t_n}^{t_{n+1}} &\mathbb{E}_{t_n}^{x^n}[f(s, y_s, z_s)]\,ds = \\ &\theta_1\Delta t_n f(t_n, y_{t_n}, z_{t_n}) + (1-\theta_1)\Delta t_n \mathbb{E}_{t_n}^{x^n}[f(t_{n+1}, y_{t_{n+1}}, z_{t_{n+1}})] + R_y^n,\end{aligned} \tag{2.3}$$

where $\theta_1 \in [0, 1]$ and

$$R_y^n = \int_{t_n}^{t_{n+1}} \{\mathbb{E}_{t_n}^{x^n}[f(s, y_s, z_s)] - (1-\theta_1)\mathbb{E}_{t_n}^{x^n}[f(t_{n+1}, y_{t_{n+1}}, z_{t_{n+1}})] - \theta_1 f(t_n, y_{t_n}, z_{t_n})\}\,ds.$$

Inserting (2.3) into (2.2) leads to the first reference equation

$$\begin{aligned}y_{t_n} = \mathbb{E}_{t_n}^{x^n}[y_{t_{n+1}}] &+ \theta_1\Delta t_n f(t_n, y_{t_n}, z_{t_n}) \\ &+ (1-\theta_1)\Delta t_n \mathbb{E}_{t_n}^{x^n}[f(t_{n+1}, y_{t_{n+1}}, z_{t_{n+1}})] + R_y^n.\end{aligned} \tag{2.4}$$

Let $\Delta_{t_n} W_s = W_s - W_{t_n}$ for $t_n \leq s \leq t_{n+1}$. Then $\Delta_{t_n} W_s$ is a standard Brownian motion with mean zero and variance $s - t_n$. Multiplying (2.1) by $\Delta_{t_n} W_{t_{n+1}}^\top$, and then taking the conditional mathematical expectation $\mathbb{E}_{t_n}^{x^n}[\cdot]$ on both sides of the derived equation, we obtain by the Itô isometry formula

$$-\mathbb{E}_{t_n}^{x^n}[y_{t_{n+1}}\Delta_{t_n}W_{t_{n+1}}^\top] = \int_{t_n}^{t_{n+1}} \mathbb{E}_{t_n}^{x^n}[f(s,y_s,z_s)\Delta_{t_n}W_s^\top]\,ds - \int_{t_n}^{t_{n+1}} \mathbb{E}_{t_n}^{x^n}[z_s]\,ds. \tag{2.5}$$

Following similar derivation of the equation (2.4) and using the fact $\Delta_{t_n}W_{t_n}=0$, we can obtain the second reference equation as

$$\begin{aligned} -\mathbb{E}_{t_n}^{x^n}[y_{t_{n+1}}\Delta_{t_n}W_{t_{n+1}}^\top] &= (1-\theta_2)\Delta t_n \mathbb{E}_{t_n}^{x^n}[f(t_{n+1},y_{t_{n+1}},z_{t_{n+1}})\Delta_{t_n}W_{t_{n+1}}^\top] \\ &\quad - \{(1-\theta_3)\Delta t_n \mathbb{E}_{t_n}^{x^n}[z_{t_{n+1}}] + \theta_3\Delta t_n z_{t_n}\} + R_z^n, \end{aligned} \tag{2.6}$$

where

$$\begin{aligned} R_z^n = &\int_{t_n}^{t_{n+1}} \mathbb{E}_{t_n}^{x^n}[f(s,y_s,z_s)\Delta_{t_n}W_{t_{n+1}}^\top]ds \\ &- (1-\theta_2)\Delta t_n \mathbb{E}_{t_n}^{x^n}[f(t_{n+1},y_{t_{n+1}},z_{t_{n+1}})\Delta_{t_n}W_{t_{n+1}}^\top] \\ &- \int_{t_n}^{t_{n+1}} \{\mathbb{E}_{t_n}^{x^n}[z_s]ds - (1-\theta_3)\mathbb{E}_{t_n}^{x^n}[z_{t_{n+1}}] - \theta_3 z_{t_n}\}\,ds. \end{aligned}$$

Define $\Delta W_{n+1} = W_{t_{n+1}} - W_{t_n}$ for $n = N-1,\cdots,1,0$. Clearly $\Delta W_{n+1} = \Delta_{t_n}W_{t_{n+1}}$. Based on the two reference equations (2.4) and (2.6), the following so-called θ-scheme was proposed for solving BSDEs in [25]: for $n = N-1, N-2, \cdots, 0$,

$$\begin{aligned} y^n &= \mathbb{E}_{t_n}^{x^n}[y^{n+1}] + \theta_1\Delta t_n f(t_n,y^n,z^n) \\ &\quad +(1-\theta_1)\Delta t_n \mathbb{E}_{t_n}^{x^n}[f(t_{n+1},y^{n+1},z^{n+1})] \end{aligned} \tag{2.7}$$

$$\begin{aligned} -\mathbb{E}_{t_n}^{x^n}[y^{n+1}\Delta W_{n+1}^\top] &= (1-\theta_2)\Delta t_n \mathbb{E}_{t_n}^{x^n}[f(t_{n+1},y^{n+1},z^{n+1})\Delta W_{n+1}^\top] \\ &\quad -\{(1-\theta_3)\Delta t_n \mathbb{E}_{t_n}^{x^n}[z^{n+1}] + \theta_3\Delta t_n z^n\}. \end{aligned} \tag{2.8}$$

where (y_n, z_n) is an approximation (y_t, z_t) at time t_n. This is a semi-discretization, for further fully discretization in the space and efficient calculation of $\mathbb{E}_{t_n}^{x^n}[\cdot]$, see [25, 26, 27] for details.

2.2. The Crank-Nicolson scheme. Notice that at time $t_N = T$ we only have the information $y_{t_N} = \xi = \phi(W_T)$ and z_{t_N} is out of our knowledge. Thus in order to proceed the θ-scheme we need an *initialization* process by choosing $\theta_1=\theta_2=\theta_3=1$ in (2.7) and (2.8) at $n = N-1$ such that we can solve y^{N-1} and z^{N-1} from Y^N

$$y^{N-1} = \mathbb{E}_{t_{N-1}}^{x^{N-1}}[y^N] + \Delta t_{N-1} f(t_{N-1},y^{N-1},z^{N-1}), \tag{2.9}$$

$$\Delta t_{N-1} z^{N-1} = \mathbb{E}_{t_{N-1}}^{x^{N-1}}[y^N \Delta W_N^\top]. \tag{2.10}$$

In the following steps, by setting $\theta_1=\theta_2=\theta_3=\dfrac{1}{2}$, we obtain a special case of the θ-scheme – the *Crank-Nicolson* scheme for solving general BSDEs:

Given y^N which is an approximation to y_{t_N}.

(i) Solve (y^{N-1}, z^{N-1}) according to (2.9) and (2.10);

(ii) For $n = N-2, N-3, \cdots, 0$, solve (y^n, z^n) by

$$
\begin{aligned}
y^n &= \mathbb{E}_{t_n}^{x^n}[y^{n+1}] + \frac{1}{2}\Delta t_n f(t_n, y^n, z^n) \\
&\quad + \frac{1}{2}\Delta t_n \mathbb{E}_{t_n}^{x^n}[f(t_{n+1}, y^{n+1}, z^{n+1})], \qquad (2.11)
\end{aligned}
$$

$$
\begin{aligned}
\frac{1}{2}\Delta t_n z^n &= \mathbb{E}_{t_n}^{x^n}[y^{n+1}\Delta W_{n+1}^\top] - \frac{1}{2}\Delta t_n \mathbb{E}_{t_n}^{x^n}[z^{n+1}] \\
&\quad + \frac{1}{2}\Delta t_n \mathbb{E}_{t_n}^{x^n}[f(t_{n+1}, y^{n+1}, z^{n+1})\Delta W_{n+1}^\top]. \qquad (2.12)
\end{aligned}
$$

Remark 1. We note that the C-N scheme defined by (2.11) and (2.12) is explicit in solving z^n. Then the Lipchitz condition of f leads to that the C-N scheme has a unique solution (y^n, z^n) for small time partition step Δt.

Note that the truncation errors of (2.11) and (2.12) in the step (ii) are one-order smaller (in time step size) than that of (2.9) and (2.10) in the initialization step (i), thus in order to obtain optimal error estimates of the C-N scheme, we also assume

Assumption 1. $\Delta t_n = \Delta t$ for $0 \le n \le N-2$ and $\Delta t_{N-1} = (\Delta t)^2$ where $\Delta t > 0$ is some positive real number.

In the rest of the paper, we will always assume Assumption 1 holds.

Remark 2. Various experiments presented in [25] showed that the above C-N scheme is numerically second-order accurate. In order to obtain some theoretical results on convergence of the C-N scheme, the authors in [22, 26] confined the discussions to the BSDEs in a simplified form

$$
y_t = \varphi(W_T) + \int_t^T f(s, y_s)\, ds - \int_t^T z_s\, dW_s, \quad t \in (0, T], \qquad (2.13)
$$

i.e., the generator f is independent of z_t. It was rigorously proven based on a variation method that the C-N scheme is second-order accurate for this simplified case.

3. Error estimates of the Crank-Nicolson Scheme

Without loss of generality, we only consider the case of one-dimensional BSDEs (i.e., $m = d = 1$). However we remark that all error estimates obtained in the sequel also hold for multidimensional BSDEs.

3.1. Some important lemmas.

Define $\Delta_{t_i}^{x^i} W_s = x^i + W_s - W_{t_i}$ for $t_i \le s \le t_{i+1}$. Let us first introduce the following lemma.

Lemma 3.1. *If $H \in C_b^{3,5}$, then when Δt is sufficiently small it holds that for $1 \le i \le N-2$,*

(3.1)
$$
\left|\mathbb{E}_{t_{i-1}}^{x^{i-1}}\left[\Delta W_i \int_{t_i}^{t_{i+1}} \left\{H(t, \Delta_{t_i}^{x^i} W_t) - \frac{H(t_i, x^i) + H(t_{i+1}, \Delta_{t_i}^{x^i} W_{t_{i+1}})}{2}\right\} dt\right]\right| \le C(\Delta t)^4,
$$

where $C > 0$ is a generic constant depending only on upper bounds of derivatives of H.

Proof. Since ΔW_i is $\mathcal{F}_{t_i}$-measurable, we have the identity

$$\begin{aligned}
&\mathbb{E}_{t_{i-1}}^{x^{i-1}}\Big[\Delta W_i\int_{t_i}^{t_{i+1}}\Big\{H(t,\Delta_{t_i}^{x^i}W_t)-\frac{H(t_i,x^i)+H(t_{i+1},\Delta_{t_i}^{x^i}W_{t_{i+1}})}{2}\Big\}\,dt\Big]\\
&=\mathbb{E}_{t_{i-1}}^{x^{i-1}}\Big[\Delta W_i\mathbb{E}_{t_i}^{x^i}\Big[\int_{t_i}^{t_{i+1}}\Big\{H(t,\Delta_{t_i}^{x^i}W_t)-\frac{H(t_i,x^i)+H(t_{i+1},\Delta_{t_i}^{x^i}W_{t_{i+1}})}{2}\Big\}\,dt\Big]\Big].
\end{aligned}\tag{3.2}$$

By Itô's formula, we have

$$\begin{aligned}
H(t,\Delta_{t_i}^{x^i}W_t)=H(t_i,x^i)&+\int_{t_i}^{t}\Big(H_t(s,\Delta_{t_i}^{x^i}W_s)+\frac{1}{2}H_{xx}(s,\Delta_{t_i}^{x^i}W_s)\Big)\,ds\\
&+\int_{t_i}^{t}H_x(s,\Delta_{t_i}^{x^i}W_s)\,dW_s.
\end{aligned}\tag{3.3}$$

Notice that

$$\begin{aligned}
&H_t'(s,\Delta_{t_i}^{x^i}W_s)+\frac{1}{2}H_{xx}(s,\Delta_{t_i}^{x^i}W_s)\\
&\quad=H_t(t_i,x_i)+\frac{1}{2}H_{xx}(t_i,x_i)\\
&\quad\quad+\int_{t_i}^{s}\Big(H_{tt}(\tau,\Delta_{t_i}^{x^i}W_\tau)+H_{txx}(\tau,\Delta_{t_i}^{x^i}W_\tau)+\frac{1}{4}H_{xxxx}(\tau,\Delta_{t_i}^{x^i}W_\tau)\Big)\,d\tau\\
&\quad\quad+\int_{t_i}^{s}\Big(H_{tx}(\tau,\Delta_{t_i}^{x^i}W_\tau)+\frac{1}{2}H_{xxx}(\tau,\Delta_{t_i}^{x^i}W_\tau)\Big)\,dW_\tau.
\end{aligned}\tag{3.4}$$

By (3.3) and (3.4) we easily get

$$\begin{aligned}
&\mathbb{E}_{t_i}^{x^i}\Big[\int_{t_i}^{t_{i+1}}H(t,\Delta_{t_i}^{x^i}W_t)\,dt\Big]\\
&\quad=H(t_i,x^i)\Delta t+\frac{1}{2}H_t(t_i,x^i)(\Delta t)^2+\frac{1}{4}H_{xx}(t_i,x^i)(\Delta t)^2\\
&\quad\quad+\int_{t_i}^{t_{i+1}}\int_{t_i}^{t}\int_{t_i}^{s}\Big(\mathbb{E}_{t_i}^{x^i}[H_{tt}(\tau,\Delta_{t_i}^{x^i}W_\tau)+H_{txx}(\tau,\Delta_{t_i}^{x^i}W_\tau)]\\
&\quad\quad\quad+\frac{1}{4}\mathbb{E}_{t_i}^{x^i}[H_{xxxx}(\tau,\Delta_{t_i}^{x^i}W_\tau)]\Big)\,d\tau\,dsdt,
\end{aligned}\tag{3.5}$$

and

$$\begin{aligned}
&\mathbb{E}_{t_i}^{x^i}[\int_{t_i}^{t_{i+1}}H(t_{i+1},\Delta_{t_i}^{x^i}W_{t_{i+1}})\,dt]\\
&\quad=H(t_i,x^i)\Delta t+H_t(t_i,x^i)(\Delta t)^2+\frac{1}{2}H_{xx}(t_i,x^i)(\Delta t)^2\\
&\quad\quad+\int_{t_i}^{t_{i+1}}\int_{t_i}^{t_{i+1}}\int_{t_i}^{s}\Big(\mathbb{E}_{t_i}^{x^i}[H_{tt}(\tau,\Delta_{t_i}^{x^i}W_\tau)]+\mathbb{E}_{t_i}^{x^i}[H_{txx}(\tau,\Delta_{t_i}^{x^i}W_\tau)\\
&\quad\quad\quad+\frac{1}{4}H_{xxxx}(\tau,\Delta_{t_i}^{x^i}W_\tau)]\Big)\,d\tau dsdt.
\end{aligned}\tag{3.6}$$

Then by (3.5) and (3.6) we deduce

$$
\begin{aligned}
\mathbb{E}_{t_i}^{x^i}\Big[\int_{t_i}^{t_{i+1}}\Big(H(t,\Delta_{t_i}^{x^i}W_t)-\frac{H(t_i,x^i)+H(t_{i+1},\Delta_{t_i}^{x^i}W_{t_{i+1}})}{2}\Big)\,dt\Big]&\\
=\int_{t_i}^{t_{i+1}}\int_{t_i}^{t}\int_{t_i}^{s}\Big(\mathbb{E}_{t_i}^{x^i}[H_{tt}(\tau,\Delta_{t_i}^{x^i}W_\tau)]&\\
+\mathbb{E}_{t_i}^{x^i}[H_{txx}(\tau,\Delta_{t_i}^{x^i}W_\tau)+\frac{1}{4}H_{xxxx}(\tau,\Delta_{t_i}^{x^i}W_\tau)]\Big)\,d\tau ds dt&\\
-\frac{1}{2}\int_{t_i}^{t_{i+1}}\int_{t_i}^{t_{i+1}}\int_{t_i}^{s}\Big(\mathbb{E}_{t_i}^{x^i}[H_{tt}(\tau,\Delta_{t_i}^{x^i}W_\tau)]&\\
+\mathbb{E}_{t_i}^{x^i}[H_{txx}(\tau,\Delta_{t_i}^{x^i}W_\tau)+\frac{1}{4}H_{xxxx}(\tau,\Delta_{t_i}^{x^i}W_\tau)]\Big)\,d\tau ds dt.&
\end{aligned}\tag{3.7}
$$

According to the Taylor expansion it holds

$$
\begin{aligned}
H_{tt}(\tau,\Delta_{t_i}^{x^i}W_\tau)&=H_{tt}(\tau,x^i)+H_{ttx}(\tau,x^i+\alpha_1(W_\tau-W_{t_i}))(W_\tau-W_{t_i})\\
&=H_{tt}(\tau,x^{i-1})+H_{ttx}(\tau,x^{i-1}+\alpha_2(W_{t_i}-W_{t_{i-1}}))(W_{t_i}-W_{t_{i-1}})\\
&\quad+H_{ttx}(\tau,x^i+\alpha_1(W_\tau-W_{t_i}))(W_\tau-W_{t_i}),
\end{aligned}\tag{3.8}
$$

where α_1 and α_2 are some positive numbers in $[0,1]$. Then from the fact

$$
\mathbb{E}_{t_{i-1}}^{x^{i-1}}[\Delta W_i H_{tt}(\tau,x^{i-1})]=H_{tt}(\tau,x^{i-1})\mathbb{E}_{t_{i-1}}^{x^{i-1}}[\Delta W_i]=0,
$$

we can get

$$
|\mathbb{E}_{t_{i-1}}^{x^{i-1}}[\Delta W_i H_{tt}(\tau,\Delta_{t_i}^{x^i}W_\tau)]|\le C\Delta t.\tag{3.9}
$$

By using similar analysis and the facts

$$
\mathbb{E}_{t_{i-1}}^{x^{i-1}}[\Delta W_i H_{txx}(\tau,x^{i-1})]=0,\quad \mathbb{E}_{t_{i-1}}^{x^{i-1}}[\Delta W_i H_{xxxx}(\tau,x^{i-1})]=0,
$$

we also can obtain the following estimates:

$$
\begin{aligned}
&|\mathbb{E}_{t_{i-1}}^{x^{i-1}}[\Delta W_i H_{txx}(\tau,\Delta_{t_i}^{x^i}W_\tau)]|\le C\Delta t,\\
&|\mathbb{E}_{t_{i-1}}^{x^{i-1}}[\Delta W_i H_{xxxx}(\tau,\Delta_{t_i}^{x^i}W_\tau)]|\le C\Delta t.
\end{aligned}\tag{3.10}
$$

Combination of (3.7), (3.9) and (3.10) gives us the inequality (3.1). The proof is completed. □

Lemma 3.2. *If $H\in C_b^{3,6}$, then when Δt is sufficiently small it holds that for $0\le i\le N-3$,*

$$
\begin{aligned}
&\Big|\mathbb{E}_{t_i}^{x^i}\Big[\mathbb{E}_{t_{i+1}}^{x^{i+1}}\Big[\int_{t_{i+1}}^{t_{i+2}}\Big\{H(t,\Delta_{t_{i+1}}^{x^{i+1}}W_t)-\frac{H(t_{i+1},x^{i+1})+H(t_{i+2},\Delta_{t_{i+1}}^{x^{i+1}}W_{t_{i+2}})}{2}\Big\}\,dt\Big]\Big]\\
&\quad-\mathbb{E}_{t_i}^{x^i}\Big[\int_{t_i}^{t_{i+1}}\Big\{H(t,\Delta_{t_i}^{x^i}W_t)-\frac{H(t_i,x^i)+H(t_{i+1},\Delta_{t_i}^{x^i}W_{t_{i+1}})}{2}\Big\}\,dt\Big]\Big|\le C(\Delta t)^4,
\end{aligned}\tag{3.11}
$$

where $C>0$ is a generic constant depending only on upper bounds of derivatives of the function H.

Proof. Similar to (3.7), we can obtain the following two inequalities:

$$
\begin{aligned}
&\mathbb{E}_{t_i}^{x^i}\Big[\int_{t_i}^{t_{i+1}}\left(H(t,\Delta_{t_i}^{x^i}W_t)-\frac{H(t_i,x^i)-H(t_{i+1},\Delta_{t_i}^{x^i}W_{t_{i+1}})}{2}\right)dt\Big]\\
&\quad=\int_{t_i}^{t_{i+1}}\int_{t_i}^{t}\int_{t_i}^{s}\Big(\mathbb{E}_{t_i}^{x^i}[H_{tt}(\tau,\Delta_{t_i}^{x^i}W_\tau)]\\
&\qquad\qquad+\mathbb{E}_{t_i}^{x^i}[H_{txx}(\tau,\Delta_{t_i}^{x^i}W_\tau)+\frac{1}{4}H_{xxxx}(\tau,\Delta_{t_i}^{x^i}W_\tau)]\Big)\,d\tau ds dt\\
&\quad-\frac{1}{2}\int_{t_i}^{t_{i+1}}\int_{t_i}^{t_{i+1}}\int_{t_i}^{s}\Big(\mathbb{E}_{t_i}^{x^i}[H_{tt}(\tau,\Delta_{t_i}^{x^i}W_\tau)]\\
&\qquad\qquad+\mathbb{E}_{t_i}^{x^i}[H_{txx}(\tau,\Delta_{t_i}^{x}W_\tau)+\frac{1}{4}H_{xxxx}(\tau,\Delta_{t_i}^{x^i}W_\tau)]\Big)\,d\tau ds dt,
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbb{E}_{t_{i+1}}^{x^{i+1}}\Big[\int_{t_{i+1}}^{t_{i+2}}\left(H(t,\Delta_{t_{i+1}}^{x^{i+1}}W_t)-\frac{H(t_{i+1},\Delta_{t_{i+1}}^{x^{i+1}}W_{t_{i+1}})+H(t_{i+2},\Delta_{t_{i+1}}^{x^{i+1}}W_{t_{i+2}})}{2}\right)dt\Big]\\
&\quad=\int_{t_{i+1}}^{t_{i+2}}\int_{t_{i+1}}^{t}\int_{t_{i+1}}^{s}\Big(\mathbb{E}_{t_{i+1}}^{x^{i+1}}[H_{tt}(\tau,\Delta_{t_{i+1}}^{x^{i+1}}W_\tau)]\\
&\qquad\qquad+\mathbb{E}_{t_{i+1}}^{x^{i+1}}[H_{txx}(\tau,\Delta_{t_{i+1}}^{x^{i+1}}W_\tau)+\frac{1}{4}H_{xxxx}(\tau,\Delta_{t_{i+1}}^{x^{i+1}}W_\tau)]\Big)\,d\tau ds dt\\
&\quad-\frac{1}{2}\int_{t_{i+1}}^{t_{i+2}}\int_{t_{i+1}}^{t_{i+2}}\int_{t_{i+1}}^{s}\Big(\mathbb{E}_{t_{i+1}}^{x^{i+1}}[H''_{tt}(\tau,\Delta_{t_{i+1}}^{x^{i+1}}W_\tau)]\\
&\qquad\qquad+\mathbb{E}_{t_{i+1}}^{x^{i+1}}[H_{txx}(\tau,\Delta_{t_{i+1}}^{x^{i+1}}W_\tau)+\frac{1}{4}H_{xxxx}(\tau,\Delta_{t_{i+1}}^{x^{i+1}}W_\tau)]\Big)\,d\tau ds dt.
\end{aligned}
$$

Since $x^{i+1}=x^i+W_{t_{i+1}}-W_{t_i}$, we have

$$
\begin{aligned}
\Delta_{t_{i+1}}^{x^{i+1}}W_\tau &= x^{i+1}+W_\tau-W_{t_{i+1}}\\
&= x^i+W_{t_{i+1}}-W_{t_i}+W_\tau-W_{t_{i+1}}\\
&= x^i+W_\tau-W_{t_i}\\
&= \Delta_{t_i}^{x^i}W_\tau.
\end{aligned}
$$

Then for $t_i\le\tau\le t_{i+2}$, by the Taylor expansion we have

$$
\begin{aligned}
&H_{tt}(\tau,\Delta_{t_i}^{x^i}W_\tau)\\
&\quad= H_{tt}(t_i,x^i)+H_{ttt}(t_i+\alpha_1(\tau-t_i),x^i+\alpha_1(W_\tau-W_{t_i}))(\tau-t_i)\\
&\qquad+H_{ttx}(t_i+\alpha_1(\tau-t_i),x^i)(W_\tau-W_{t_i})\\
&\qquad+H_{ttxx}(t_i+\alpha_1(\tau-t_i),x^i+\alpha_2(W_\tau-W_{t_i}))\alpha_1(W_\tau-W_{t_i})^2,
\end{aligned}
\tag{3.12}
$$

$$
\begin{aligned}
&H_{txx}(\tau,\Delta_{t_i}^{x^i}W_\tau)\\
&\quad= H_{txx}(t_i,x^i)+H_{ttxx}(t_i+\alpha_3(\tau-t_i),x^i+\alpha_3(W_\tau-W_{t_i}))(\tau-t_i)\\
&\qquad+H_{txxx}(t_i+\alpha_3(\tau-t_i),x^i)(W_\tau-W_{t_i})\\
&\qquad+H_{txxx}(t_i+\alpha_3(\tau-t_i),x^i+\alpha_4(W_\tau-W_{t_i}))\alpha_3(W_\tau-W_{t_i})^2,
\end{aligned}
\tag{3.13}
$$

$$\begin{aligned}
&H_{xxxx}(\tau,\Delta_{t_i}^{x^i}W_\tau)\\
&\quad= H_{xxxx}(t_i,x^i)+H_{txxxx}(t_i+\alpha_5(\tau-t_i),x^i+\alpha_5(W_\tau-W_{t_i}))(\tau-t_i)\\
&\qquad+H_{xxxxx}(t_i+\alpha_5(\tau-t_i),x^i)(W_\tau-W_{t_i})\\
&\qquad+H_{xxxxxx}(t_i+\alpha_5(\tau-t_i),x^i+\alpha_6(W_\tau-W_{t_i}))\alpha_5(W_\tau-W_{t_i})^2,
\end{aligned}\tag{3.14}$$

where $\alpha_i(i=1,2,\cdots,6)$ are some positive numbers in $[0,1]$.

Notice that $H_{ttx}(t_i+\alpha_1(\tau-t_i),x^i)$, $H_{txxx}(t_i+\alpha_3(\tau-t_i),x^i)$, $H_{xxxxx}(t_i+\alpha_5(\tau-t_i),x^i)$ are all $\mathcal{F}_{t_i}$-measurable, thus we have

$$\mathbb{E}_{t_i}^{x^i}[H_{ttx}(t_i+\alpha_1(\tau-t_i),x^i)(W_\tau-W_{t_i})]=0,$$

$$\mathbb{E}_{t_i}^{x^i}[H_{txxx}(t_i+\alpha_3(\tau-t_i),x^i)(W_\tau-W_{t_i})]=0,$$

$$\mathbb{E}_{t_i}^{x^i}[H_{xxxxx}(t_i+\alpha_5(\tau-t_i),x^i)(W_\tau-W_{t_i})]=0.$$

Now under the conditions of the lemma and by the equations (3.12), (3.13) and (3.14), we deduce

$$\begin{aligned}
&\Big|\mathbb{E}_{t_i}^{x^i}\Big[\mathbb{E}_{t_{i+1}}^{x^{i+1}}\Big[\int_{t_{i+1}}^{t_{i+2}}\Big\{H(t,\Delta_{t_i}^{x^i}W_t)-\frac{H(t_{i+1},\Delta_{t_i}^{x^i}W_{t_{i+1}})+H(t_{i+2},\Delta_{t_i}^{x^i}W_{t_{i+2}})}{2}\Big\}\,dt\Big]\Big]\\
&\qquad-\mathbb{E}_{t_i}^{x^i}\Big[\int_{t_i}^{t_{i+1}}\Big\{H(t,\Delta_{t_i}^{x^i}W_t)-\frac{H(t_i,x^i)+H(t_{i+1},\Delta_{t_i}^{x^i}W_{t_{i+1}})}{2}\Big\}\,dt\Big]\Big|\\
&\quad\le\Big|\int_{t_i}^{t_{i+1}}\int_{t_i}^{t}\int_{t_i}^{s}\Big(H_{tt}(t_i,x^i)+H_{txx}(t_i,x^i)+\frac{1}{4}H_{xxxx}(t_i,x^i)\Big)\,d\tau ds dt\\
&\qquad-\frac{1}{2}\int_{t_i}^{t_{i+1}}\int_{t_i}^{t_{i+1}}\int_{t_i}^{s}\Big(H_{tt}(t_i,x^i)+H_{txx}(t_i,x^i)+\frac{1}{4}H_{xxxx}(t_i,x^i)\Big)\,d\tau ds dt\\
&\qquad-\int_{t_{i+1}}^{t_{i+2}}\int_{t_{i+1}}^{t}\int_{t_{i+1}}^{s}\Big(H_{tt}(t_i,x^i)+H_{txx}(t_i,x^i)+\frac{1}{4}H_{xxxx}(t_i,x^i)\Big)\,d\tau ds dt\\
&\qquad+\frac{1}{2}\int_{t_{i+1}}^{t_{i+2}}\int_{t_{i+1}}^{t_{i+2}}\int_{t_{i+1}}^{s}\Big(H_{tt}(t_i,x^i)+H_{txx}(t_i,x^i)+\frac{1}{4}H_{xxxx}(t_i,x^i)\Big)\,d\tau ds dt\Big|\\
&\qquad+C(\Delta t)^4\\
&\quad=(\frac{1}{6}-\frac{1}{4}-\frac{1}{6}+\frac{1}{4})(\Delta t)^3\Big(H_{tt}(t_i,x^i)+H_{txx}(t_i,x^i)+\frac{1}{4}H_{xxxx}(t_i,x^i)\Big)+C(\Delta t)^4\\
&\quad\le C(\Delta t)^4,
\end{aligned}$$

where C is a constant which depends only on upper bound of the derivatives of the function H. □

Upper bounds of R_y^n and R_z^n defined in (2.4) and (2.6) are presented in the following.

Lemma 3.3. *Let (y_t,z_t) be the solution of (1.2), and let R_y^n and R_z^n be the truncation errors defined in (2.4) and (2.6) for the θ-scheme. Suppose Δt is sufficiently small.*

(1) *If the terminal function $\varphi\in C_b^2$ and the generator function $f\in C_b^{1,2,2}$, then it holds that for $0\le n<N-1$*

$$|R_y^n|\le C(\Delta t_n)^2,\quad |R_z^n|\le C(\Delta t_n)^2.\tag{3.15}$$

(2) *In particular, for the Crank-Nicolson scheme ($\theta_i = 1/2,\ i = 1,2,3$), if $\varphi \in C_b^3$ and $f \in C_b^{2,4,4}$, then it holds that for $0 \le n < N-1$*

$$|R_y^n| \le C(\Delta t_n)^3, \quad |R_z^n| \le C(\Delta t_n)^3. \tag{3.16}$$

Note here $C > 0$ is a positive constant depending only on T, the upper bounds of the derivatives of the functions φ and f.

Under the conditions of f and φ in Lemma 3.3, using the Taylor expansion and the properties of the Brownian motion W_t, it is easy to prove Lemma 3.3, see [22, 26] for details.

Lemma 3.4. *Let (y_t, z_t) be the solution of (1.2), and R_y^i and R_z^i be the truncation errors defined in (2.4) and (2.6) for the Crank-Nicolson scheme ($\theta_i = 1/2,\ i = 1,2,3$), i.e.,*

$$R_z^i = \int_{t_i}^{t_{i+1}} \{\mathbb{E}_{t_i}^{x^i}[z_s] - \frac{1}{2}\mathbb{E}_{t_i}^{x^i}[z_{t_{i+1}}] - \frac{1}{2} z_{t_i}\}\, ds, \tag{3.17}$$

(3.18)
$$R_y^i = \int_{t_i}^{t_{i+1}} \{\mathbb{E}_{t_i}^{x^i}[f(s,y_s,z_s)] - \frac{1}{2}\mathbb{E}_{t_i}^{x^i}[f(t_{i+1},y_{t_{i+1}},z_{t_{i+1}})] - \frac{1}{2} f(t_i,y_{t_i},z_{t_i})\}\, ds,$$

for $0 \le i \le N-1$. If $\varphi \in C_b^3$ and $f(t,y,z) \in C_b^{3,6,6}$, then when Δt is sufficiently small it holds that for $0 \le i \le N-3$,

$$|\mathbb{E}_{t_i}^{x^i}[R_y^{i+1}\Delta W_{i+1}]| \le C(\Delta t)^4, \tag{3.19}$$

and

$$|R_z^i - \mathbb{E}_{t_i}^{x^i}[R_z^{i+1}]| \le C(\Delta t)^4, \tag{3.20}$$

where C is a positive constant depending only on T, the upper bounds of derivatives of functions φ and f .

Proof. Under the conditions of the lemma, the solution (y_t, z_t) of the BSDE (1.2) can be represented as

$$y_t = u(t, W_t), \quad z_t = \nabla_x u(t, W_t), \qquad \forall\, t \in [0,T), \tag{3.21}$$

where $u(t,x)$ satisfies the parabolic PDE (1.3).

Let $H(t,W_t) = z_t$, then $H(t,W_t) = \nabla_x u(t,W_t)$ according to (1.3). By the theory of partial differential equations [6], it is easy to check that the functions H satisfy the conditions in Lemma 3.1, thus we can easily get the estimates (3.19) using Lemma 3.1.

Similarly, by letting $H(t,W_t) = f(t,y_t,z_t) = f(t,u(t,W_t),\nabla_x u(t,W_t))$ and using Lemma 3.2, we can get (3.20). The proof is completed. □

3.2. Error estimates. Let (y_t, z_t) be the solution of the BSDE (1.2) with the terminal condition $y_T = \varphi(W_T)$, and (y^n, z^n) be its approximate solution produced by using the Crank-Nicolson scheme. For $0 \le n \le N$, let

$$e_y^n = y_{t_n} - y^n, \quad e_z^n = z_{t_n} - z^n, \quad e_f^n = f(t_n, y_{t_n}, z_{t_n}) - f(t_n, y^n, z^n).$$

By the equations (2.4), (2.6), (2.9) and (2.10), we get that for $n = N-1$,

$$e_y^{N-1} = \mathbb{E}_{t_{N-1}}^{x^{N-1}}[e_y^N] + \Delta t_{N-1} e_f^{N-1} + R_y^{N-1}, \tag{3.22}$$

and

$$\Delta t_{N-1} e_z^{N-1} = \mathbb{E}_{t_{N-1}}^{x^{N-1}}[e_y^N \Delta W_N] + R_z^{N-1}. \tag{3.23}$$

Then using (2.4), (2.6), (2.11) and (2.12), we have that for $0 \le n \le N-2$,

$$e_y^n = \mathbb{E}_{t_n}^{x^n}[e_y^{n+1}] + \frac{1}{2}\Delta t_n \mathbb{E}_{t_n}^{x^n}[e_f^{n+1}] + \frac{1}{2}\Delta t_n e_f^n + R_y^n, \tag{3.24}$$

and

$$\frac{1}{2}\Delta t_n e_z^n = -\frac{1}{2}\Delta t_n \mathbb{E}_{t_n}^{x^n}[e_z^{n+1}] + \mathbb{E}_{t_n}^{x^n}[e_y^{n+1}\Delta W_{n+1}] + \frac{1}{2}\Delta t_n \mathbb{E}_{t_n}^{x^n}[e_f^{n+1}\Delta W_{n+1}] + R_z^n. \tag{3.25}$$

We also note that $\Delta t_{N-1} = (\Delta t)^2$ and $\Delta t_n = \Delta t$ for $0 \le n \le N-2$ according to Assumption 1.

Lemma 3.5. *Suppose that $\varphi \in C_b^3$, $f \in C_b^{3,6,6}$, and the initial error satisfies*

$$\mathbb{E}_{t_{N-1}}^{x^{N-1}}[|y_{t_N} - y^N|^2] \le C(\Delta t_{N-1})^3 \tag{3.26}$$

for some constant $C > 0$. Then when Δt is sufficiently small, it holds that for $N-3 \le n \le N-1$,

$$|e_y^n| \le C(\Delta t)^3, \quad |e_z^n| \le C(\Delta t)^2, \tag{3.27}$$

where $C > 0$ is a generic constant depending only on T, upper bounds of derivatives of φ and f.

Proof. According to Lemma 3.3, we know that

$$|R_y^{N-1}| \le C(\Delta t_{N-1})^2, \qquad |R_z^{N-1}| \le C(\Delta t_{N-1})^2. \tag{3.28}$$

Let L be the Lipschitz constant of f with respect to y and z. Recall that $\Delta t_{N-1} = (\Delta t)^2$. By the inequality (3.28) and Hölder's inequality, we deduce

$$\begin{aligned} |e_z^{N-1}| &\le \frac{1}{\Delta t_{N-1}}\Big(|\mathbb{E}_{t_{N-1}}^{x^{N-1}}[e_y^N \Delta W_N]| + |R_z^{N-1}|\Big) \\ &\le \frac{1}{\Delta t_{N-1}}\mathbb{E}_{t_{N-1}}^{x^{N-1}}[|e_y^N|^2]^{\frac{1}{2}}\mathbb{E}_{t_{N-1}}^{x^{N-1}}[|\Delta W_N|^2]^{\frac{1}{2}} + \frac{|R_z^{N-1}|}{\Delta t_{N-1}} \\ &\le \frac{1}{\Delta t_{N-1}} C^{1/2}(\Delta t_{N-1})^{3/2}(\Delta t_{N-1})^{1/2} + C\Delta t_{N-1} \\ &\le C\Delta t_{N-1} = C(\Delta t)^2, \end{aligned} \tag{3.29}$$

and

$$\begin{aligned} |e_y^{N-1}| &\le \mathbb{E}_{t_{N-1}}^{x^{N-1}}[|e_y^N|] + L\Delta t_{N-1}(|e_y^{N-1}| + |e_z^{N-1}|) + |R_y^{N-1}| \\ &\le \mathbb{E}_{t_{N-1}}^{x^{N-1}}[|e_y^N|^2]^{\frac{1}{2}} + L\Delta t_{N-1}|e_y^{N-1}| + LC(\Delta t_{N-1})^2 + C(\Delta t_{N-1})^2 \\ &\le C^{1/2}(\Delta t_{N-1})^{3/2} + L\Delta t_{N-1}|e_y^{N-1}| + (L+1)C(\Delta t_{N-1})^2 \\ &\le L\Delta t_{N-1}|e_y^{N-1}| + C(\Delta t_{N-1})^{3/2}. \end{aligned}$$

Then we get

$$|e_y^{N-1}| \le \frac{C(\Delta t_{N-1})^{3/2}}{1 - L\Delta t_{N-1}} \le C(\Delta t_{N-1})^{\frac{3}{2}} = C(\Delta t)^3. \tag{3.30}$$

Recall that $\Delta t_n = \Delta t$ for $0 \le n \le N-2$. By Lemma 3.3, we have that for $0 \le n \le N-2$,

$$|R_y^n| \le C(\Delta t)^3, \qquad |R_z^n| \le C(\Delta t)^3. \tag{3.31}$$

Now we estimate e_y^n and e_z^n for $n = N-2, N-3$. From the equation (3.25), the inequalities (3.29), (3.30) and (3.31), we deduce by the Hölder inequality that

$$
\begin{aligned}
|e_z^{N-2}| &\le \mathbb{E}_{t_{N-2}}^{x^{N-2}}[|e_z^{N-1}|] + \frac{2}{\Delta t}\mathbb{E}_{t_{N-2}}^{x^{N-2}}[|e_y^{N-1}|^2]^{\frac{1}{2}}\mathbb{E}_{t_{N-2}}^{x^{N-2}}[|\Delta W_{N-1}|^2]^{\frac{1}{2}} \\
&\quad + \mathbb{E}_{t_{N-2}}^{x^{N-2}}[|e_f^{N-1}|^2]^{\frac{1}{2}}\mathbb{E}_{t_{N-2}}^{x^{N-2}}[|\Delta W_{N-1}|^2]^{\frac{1}{2}} + \frac{2}{\Delta t}|R_z^{N-2}| \\
&\le C(\Delta t)^2 + 2C(\Delta t)^{5/2} \\
&\quad + 2L\mathbb{E}_{t_{N-2}}^{x^{N-2}}[|e_y^{N-1}|^2 + |e_z^{N-1}|^2]^{\frac{1}{2}}\mathbb{E}_{t_{N-2}}^{x^{N-2}}[|\Delta W_{N-1}|^2]^{\frac{1}{2}} + 2C(\Delta t)^2 \\
&\le C(\Delta t)^2 + 2LC(\Delta t)^{5/2} \\
&\le C(\Delta t)^2,
\end{aligned} \tag{3.32}
$$

and

$$
\begin{aligned}
|e_y^{N-2}| &\le \mathbb{E}_{t_{N-2}}^{x^{N-2}}[|e_y^{N-1}|] + \frac{1}{2}\Delta t\mathbb{E}_{t_{N-2}}^{x^{N-2}}[|e_f^{N-1}|] + \frac{1}{2}\Delta t|e_f^{N-2}| + |R_y^{N-2}| \\
&\le C(\Delta t)^3 + \frac{L}{2}\Delta t\mathbb{E}_{t_{N-2}}^{x^{N-2}}[|e_y^{N-1}| + |e_z^{N-1}|] \\
&\quad + \frac{L}{2}\Delta t\mathbb{E}_{t_{N-2}}^{x^{N-2}}[|e_y^{N-2}| + |e_z^{N-2}|] + C(\Delta t)^3 \\
&\le LC(\Delta t)^3 + \frac{L}{2}\Delta t|e_y^{N-2}| + C(\Delta t)^3 \\
&\le \frac{L}{2}\Delta t|e_y^{N-2}| + C(\Delta t)^3,
\end{aligned}
$$

which implies

$$
|e_y^{N-2}| \le \frac{C(\Delta t)^3}{1 - L\Delta t/2} \le C(\Delta t)^3. \tag{3.33}
$$

Similarly we can obtain the estimates

$$
|e_z^{N-3}| \le C(\Delta t)^2, \qquad |e_y^{N-3}| \le C(\Delta t)^3. \tag{3.34}
$$

Thus the proof is completed. □

Lemma 3.6. *Suppose that $\varphi \in C_b^3$, $f \in C_b^{3,6,6}$. Then when Δt is sufficiently small, it holds that for $0 \le n \le N-3$, if $|e_z^{n+1}|^2 \le \Delta t$ and $|e_y^{n+1}|^2 \le \Delta t$, then*

$$
\begin{aligned}
&|e_y^n|^2 + |e_z^n|^2 \\
&\le (1 + C\Delta t)\Big(\frac{1}{2}|\mathbb{E}_{t_n}^{x^n}[e_y^{n+1}]|^2 + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|^2] + |\mathbb{E}_{t_n}^{x^n}[e_z^{n+2}]|^2 + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_z^{n+1}|^2]\Big) \\
&\quad + C\Delta t(|e_y^n|^2 + |e_z^n|^2) \\
&\quad + C\frac{|R_y^n|^2}{\Delta t} + C\frac{|\mathbb{E}_{t_n}^{x^n}[R_y^{n+1}\Delta W_{n+1}]|^2 + |R_z^n - \mathbb{E}_{t_n}^{x^n}[R_z^{n+1}]|^2}{(\Delta t)^3},
\end{aligned} \tag{3.35}
$$

where $C > 0$ is a generic constant depending only on T, upper bounds of derivatives of φ and f.

Proof. For stochastic processes X and Y, let us define

$$
Var^n(X) = \mathbb{E}_{t_n}^{x^n}[|X|^2] - |\mathbb{E}_{t_n}^{x^n}[X]|^2, \quad Cov^n(X,Y) = \mathbb{E}_{t_n}^{x^n}[XY] - \mathbb{E}_{t_n}^{x^n}[X]\mathbb{E}_{t_n}^{x^n}[Y],
$$

for $0 \le n \le N-1$.

By the equations (3.24) and (3.25), for $0 \le n \le N-3$, it holds that

$$
\begin{aligned}
e_z^n = {} & \mathbb{E}_{t_n}^{x^n}[e_z^{n+2}] - \frac{2}{\Delta t}\mathbb{E}_{t_n}^{x^n}[e_y^{n+2}\Delta W_{n+2}] \\
& - \mathbb{E}_{t_n}^{x^n}[e_f^{n+2}\Delta W_{n+2}] + \frac{2}{\Delta t}\mathbb{E}_{t_n}^{x^n}[e_y^{n+2}\Delta W_{n+1}] \\
& + \mathbb{E}_{t_n}^{x^n}[e_f^{n+2}\Delta W_{n+1}] + \mathbb{E}_{t_n}^{x^n}[e_f^{n+1}\Delta W_{n+1}] + \mathbb{E}_{t_n}^{x^n}[e_f^{n+1}\Delta W_{n+1}] \\
& + \frac{2}{\Delta t}\mathbb{E}_{t_n}^{x^n}[R_y^{n+1}\Delta W_{n+1}] + 2\frac{R_z^n - \mathbb{E}_{t_n}^{x^n}[R_z^{n+1}]}{\Delta t} \\
= {} & \mathbb{E}_{t_n}^{x^n}[e_z^{n+2}] - \frac{2}{\Delta t}\mathbb{E}_{t_n}^{x^n}[e_y^{n+2}\Delta W_{n+2}] + \frac{2}{\Delta t}\mathbb{E}_{t_n}^{x^n}[e_y^{n+2}\Delta W_{n+1}] \\
& + \mathbb{E}_{t_n}^{x^n}[e_f^{n+2}\Delta W_{n+1}] - \mathbb{E}_{t_n}^{x^n}[e_f^{n+2}\Delta W_{n+2}] \\
& + 2\mathbb{E}_{t_n}^{x^n}[e_f^{n+1}\Delta W_{n+1}] + \frac{2}{\Delta t}\mathbb{E}_{t_n}^{x^n}[R_y^{n+1}\Delta W_{n+1}] + 2\frac{R_z^n - \mathbb{E}_{t_n}^{x^n}[R_z^{n+1}]}{\Delta t}.
\end{aligned}
\tag{3.36}
$$

From the definition of the conditional mathematical expectation $\mathbb{E}_{t_n}^{x^n}[\cdot]$, we have

$$
\begin{aligned}
& \mathbb{E}_{t_n}^{x^n}[y_{t_{n+2}}\Delta W_{n+1}] \\
& \quad = \mathbb{E}_{t_n}^{x^n}[u(t_{n+2}, x^n + W_{t_{n+2}} - W_{t_n})\Delta W_{n+1}] \\
& \quad = \mathbb{E}_{t_n}^{x^n}[\mathbb{E}_{t_{n+1}}^{x^n + W_{t_{n+1}} - W_{t_n}}[u(t_{n+2}, x^n + W_{t_{n+2}} - W_{t_{n+1}} + W_{t_{n+1}} - W_{t_n})]\Delta W_{n+1}] \\
& \quad = \Delta t\mathbb{E}_{t_n}^{x^n}[\mathbb{E}_{t_{n+1}}^{x^n + W_{t_{n+1}} - W_{t_n}}[\nabla_x u(t_{n+2}, x^n + W_{t_{n+2}} - W_{t_{n+1}} + W_{t_{n+1}} - W_{t_n})]] \\
& \quad = \Delta t\mathbb{E}_{t_n}^{x^n}[\nabla y_{t_n+2}],
\end{aligned}
\tag{3.37}
$$

and

$$
\begin{aligned}
\mathbb{E}_{t_n}^{x^n}[y_{t_{n+2}}(W_{t_{n+2}} - W_{t_n})] &= \frac{1}{\sqrt{4\pi\Delta t}}\int_{-\infty}^{\infty} u(t_{n+2}, x^n + v)v\exp(-\frac{v^2}{4\Delta t})\,dv \\
&= \frac{2\Delta t}{\sqrt{4\pi\Delta t}}\int_{-\infty}^{\infty} \nabla_x u(t_{n+2}, x^n + v)\exp(-\frac{v^2}{4\Delta t})\,dv \\
&= 2\Delta t\mathbb{E}_{t_n}^{x^n}[\nabla y_{t_{n+2}}].
\end{aligned}
\tag{3.38}
$$

Thus we get

$$
\begin{aligned}
\mathbb{E}_{t_n}^{x^n}[y_{t_{n+2}}\Delta W_{n+2}] &= \mathbb{E}_{t_n}^{x^n}[y_{t_{n+2}}(W_{t_{n+2}} - W_{t_n} + W_{t_n} - W_{t_{n+1}})] \\
&= \mathbb{E}_{t_n}^{x^n}[y_{t_{n+2}}(W_{t_{n+2}} - W_{t_n})] - \mathbb{E}_{t_n}^{x^n}[y_{t_{n+2}}(W_{t_{n+1}} - W_{t_n})] \\
&= 2\Delta t\mathbb{E}_{t_n}^{x^n}[\nabla y_{t_{n+2}}] - \Delta t\mathbb{E}_{t_n}^{x^n}[\nabla y_{t_{n+2}}] \\
&= \Delta t\mathbb{E}_{t_n}^{x^n}[\nabla y_{t_{n+2}}].
\end{aligned}
\tag{3.39}
$$

Using (3.37) and (3.39), we obtain the identity

$$
\mathbb{E}_{t_n}^{x^n}[y_{t_{n+2}}\Delta W_{n+1}] = \mathbb{E}_{t_n}^{x^n}[y_{t_{n+2}}\Delta W_{n+2}]. \tag{3.40}
$$

If the Brownian motion starts from the time-space point (t_n, x^n), then y^{n+k} is a function of $x^n + W_{t_{n+k}} - W_{t_n}$, that is, y^{n+k} is a function of $x^n + W_{t_{n+k}} - W_{t_n}$. Then similar to the way to obtain (3.40), it holds

$$
\mathbb{E}_{t_n}^{x^n}[y^{n+2}\Delta W_{n+1}] = \mathbb{E}_{t_n}^{x^n}[y^{n+2}\Delta W_{n+2}]. \tag{3.41}
$$

for the approximate solution of the C-N scheme. Thus (3.40) and (3.41) give us the following identity

$$
\mathbb{E}_{t_n}^{x^n}[e_y^{n+2}\Delta W_{n+1}] = \mathbb{E}_{t_n}^{x^n}[e_y^{n+2}\Delta W_{n+2}]. \tag{3.42}
$$

Similarly we can get

$$\mathbb{E}_{t_n}^{x^n}[e_f^{n+2}\Delta W_{n+1}] = \mathbb{E}_{t_n}^{x^n}[e_f^{n+2}\Delta W_{n+2}]. \tag{3.43}$$

Combining the equations (3.36), (3.42) and (3.43) we obtain

$$e_z^n = \mathbb{E}_{t_n}^{x^n}[e_z^{n+2}] + 2\mathbb{E}_{t_n}^{x^n}[e_f^{n+1}\Delta W_{n+1}] + \frac{2}{\Delta t}\mathbb{E}_{t_n}^{x^n}[R_y^{n+1}\Delta W_{n+1}] + 2\frac{R_z^n - \mathbb{E}_{t_n}^{x^n}[R_z^{n+1}]}{\Delta t}. \tag{3.44}$$

Using the Taylor expansion it is easy to obtain that for $i = N-1, N-2, \cdots, 0$,

$$e_f^i = f(t_i, y_{t_i}, z_{t_i}) - f(t_i, y^i, z^i) = \xi^i e_y^i + \gamma^i e_z^i \tag{3.45}$$

where

$$\xi^i = f_y(t_i, y_{t_i} - \beta^i e_y^i, z_{t_i} - \beta^i e_z^i), \tag{3.46}$$

$$\gamma^i = f_z(t_i, y_{t_i} - \beta^i e_y^i, z_{t_i} - \beta^i e_z^i) \tag{3.47}$$

with some $\beta^i \in [0,1]$. Then by (3.44) and (3.45) we deduce

$$\begin{aligned} e_z^n = {} & \mathbb{E}_{t_n}^{x^n}[e_z^{n+2}] + 2\mathbb{E}_{t_n}^{x^n}[(\xi^{n+1}e_y^{n+1} + \gamma^{n+1}e_z^{n+1})\Delta W_{n+1}] \\ & + \frac{2\mathbb{E}_{t_n}^{x^n}[R_y^{n+1}\Delta W_{n+1}]}{\Delta t} + 2\frac{R_z^n - \mathbb{E}_{t_n}^{x^n}[R_z^{n+1}]}{\Delta t}. \end{aligned} \tag{3.48}$$

From the facts that

$$\begin{aligned} \mathbb{E}_{t_n}^{x^n}[f_y(t_n, y_{t_n}, z_{t_n})\Delta W_{n+1}] = f_y(t_n, y_{t_n}, z_{t_n})\mathbb{E}_{t_n}^{x^n}[\Delta W_{n+1}] = 0, \\ \mathbb{E}_{t_n}^{x^n}[f_z(t_n, y_{t_n}, z_{t_n})\Delta W_{n+1}] = f_z(t_n, y_{t_n}, z_{t_n})\mathbb{E}_{t_n}^{x^n}[\Delta W_{n+1}] = 0, \end{aligned}$$

and the definition of $Cov^n(\cdot)$, we have

$$\begin{aligned} & \mathbb{E}_{t_n}^{x^n}[\xi^{n+1}e_y^{n+1}\Delta W_{n+1}] \\ & \quad = \mathbb{E}_{t_n}^{x^n}[e_y^{n+1}]\mathbb{E}_{t_n}^{x^n}[\xi^{n+1}\Delta W_{n+1}] + Cov^n(e_y^{n+1}, \xi^{n+1}\Delta W_{n+1}) \\ & \quad = \mathbb{E}_{t_n}^{x^n}[e_y^{n+1}]\mathbb{E}_{t_n}^{x^n}[(\xi^{n+1} - f_y(t_n, y_{t_n}, z_{t_n}))\Delta W_{n+1}] \\ & \qquad + Cov^n(e_y^{n+1}, \xi^{n+1}\Delta W_{n+1}), \end{aligned} \tag{3.49}$$

and

$$\begin{aligned} & \mathbb{E}_{t_n}^{x^n}[\gamma^{n+1}e_z^{n+1}\Delta W_{n+1}] \\ & \quad = \mathbb{E}_{t_n}^{x^n}[e_z^{n+1}]\mathbb{E}_{t_n}^{x^n}[(\gamma^{n+1} - f_z(t_n, y_{t_n}, z_{t_n}))\Delta W_{n+1}] \\ & \qquad + Cov^n(e_z^{n+1}, \gamma^{n+1}\Delta W_{n+1}). \end{aligned} \tag{3.50}$$

By (3.46), (3.47) and using the Taylor expansion again, we obtain

$$\begin{aligned} & \xi^{n+1} - f_y(t_n, y_{t_n}, z_{t_n}) \\ & \quad = f_y(t_{n+1}, y_{t_{n+1}}, z_{t_{n+1}}) - f_y(t_n, y_{t_n}, z_{t_n}) \\ & \qquad - f_{yy}(t_{n+1}, y_{t_{n+1}} - \alpha_1^{n+1}e_y^{n+1}, z_{t_{n+1}} - \alpha_1^{n+1}e_z^{n+1})\beta_1^{n+1}e_y^{n+1} \\ & \qquad - f_{yz}(t_{n+1}, y_{t_{n+1}} - \alpha_2^{n+1}e_y^{n+1}, z_{t_{n+1}} - \alpha_2^{n+1}e_z^{n+1})\beta_1^{n+1}e_z^{n+1}, \end{aligned} \tag{3.51}$$

and

$$\begin{aligned} & \gamma^{n+1} - f_z(t_n, y_{t_n}, z_{t_n}) \\ & \quad = f_z(t_{n+1}, y_{t_{n+1}}, z_{t_{n+1}}) - f_z(t_n, y_{t_n}, z_{t_n}) \\ & \qquad - f_{zy}(t_{n+1}, y_{t_{n+1}} - \alpha_3^{n+1}e_y^{n+1}, z_{t_{n+1}} - \alpha_3^{n+1}e_z^{n+1})\beta_2^{n+1}e_y^{n+1} \\ & \qquad - f_{zz}(t_{n+1}, y_{t_{n+1}} - \alpha_4^{n+1}e_y^{n+1}, z_{t_{n+1}} - \alpha_4^{n+1}e_z^{n+1})\beta_2^{n+1}e_z^{n+1}, \end{aligned} \tag{3.52}$$

where $\alpha_i^{n+1} \in [0,1]$ $(i = 1,2,3,4)$ and $\beta_i^{n+1} \in [0,1]$ $(i = 1,2)$. Using the equations (3.51) and (3.52), the Hölder inequality, and the assumption that $|e_z^{n+1}|^2 \le \Delta t$ and $|e_y^{n+1}|^2 \le \Delta t$, we obtain

(3.53)
$$\begin{aligned}
&|\mathbb{E}_{t_n}^{x^n}[(\xi^{n+1} - f_y(t_n, y_{t_n}, z_{t_n}))\Delta W_{n+1}]| \\
&\quad = |\mathbb{E}_{t_n}^{x^n}[(f_y(t_{n+1}, y_{n+1}, z_{t_{n+1}}) - f_y(t_n, y_{t_n}, z_{t_n}))\Delta W_{n+1}] \\
&\qquad - \mathbb{E}_{t_n}^{x^n}[f_{yy}(t_{n+1}, y_{t_{n+1}} - \alpha_1^{n+1} e_y^{n+1}, z_{t_{n+1}} - \alpha_1^{n+1} e_z^{n+1})\beta_1^{n+1} e_y^{n+1}\Delta W_{n+1}] \\
&\qquad - \mathbb{E}_{t_n}^{x^n}[f_{yz}(t_{n+1}, y_{t_{n+1}} - \alpha_2^{n+1} e_y^{n+1}, z_{t_{n+1}} - \alpha_2^{n+1} e_z^{n+1})\beta_1^{n+1} e_z^{n+1}\Delta W_{n+1}]| \\
&\quad \le \mathbb{E}_{t_n}^{x^n}[(f_y(t_{n+1}, y_{t_{n+1}}, z_{t_{n+1}}) - f_y(t_n, y_{t_n}, z_{t_n}))^2]^{\frac{1}{2}} \mathbb{E}_{t_n}^{x^n}[(\Delta W_{n+1})^2]^{\frac{1}{2}} \\
&\qquad + C\mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|^2]^{\frac{1}{2}} \mathbb{E}_{t_n}^{x^n}[|\Delta W_{n+1}|^2]^{\frac{1}{2}} + C\mathbb{E}_{t_n}^{x^n}[|e_z^{n+1}|^2]^{\frac{1}{2}} \mathbb{E}_{t_n}^{x^n}[|\Delta W_{n+1}|^2]^{\frac{1}{2}} \\
&\quad \le C\Delta t + C\sqrt{\Delta t}\mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|^2]^{\frac{1}{2}} + C\sqrt{\Delta t}\mathbb{E}_{t_n}^{x^n}[|e_z^{n+1}|^2]^{\frac{1}{2}} \\
&\quad \le C\Delta t.
\end{aligned}$$

Similarly we can get

$$|\mathbb{E}_{t_n}^{x^n}[(\gamma^{n+1} - f_z(t_n, y_{t_n}, z_{t_n}))\Delta W_{n+1}]| \le C\Delta t. \tag{3.54}$$

Using the Cauchy-Schwarz inequality it gives us that

$$\begin{aligned}
&|Cov^n(e_y^{n+1}, \xi^{n+1}\Delta W_{n+1})| \\
&\quad \le \{Var^n(e_y^{n+1}) Var^n(\xi^{n+1}\Delta W_{n+1})\}^{\frac{1}{2}} \\
&\quad \le \mathbb{E}_{t_n}^{x^n}[(\xi^{n+1}\Delta W_{n+1})^2]^{\frac{1}{2}} \{Var^n(e_y^{n+1})\}^{\frac{1}{2}} \le C\sqrt{\Delta t}\{Var^n(e_y^{n+1})\}^{\frac{1}{2}},
\end{aligned} \tag{3.55}$$

and

$$|Cov^n(e_z^{n+1}, \gamma^{n+1}\Delta W_{n+1})| \le C\sqrt{\Delta t}\{Var^n(e_z^{n+1})\}^{\frac{1}{2}}. \tag{3.56}$$

By the inequalities (3.49), (3.50), (3.53), (3.54), (3.55) and (3.56) we easily obtain

$$\begin{aligned}
&|\mathbb{E}_{t_n}^{x^n}[\xi^{n+1} e_y^{n+1}\Delta W_{n+1}]| \le C\Delta t \mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|] + C\sqrt{\Delta t}\{Var^n(e_y^{n+1})\}^{\frac{1}{2}}, \\
&|\mathbb{E}_{t_n}^{x^n}[\gamma^{n+1} e_z^{n+1}\Delta W_{n+1}]| \le C\Delta t \mathbb{E}_{t_n}^{x^n}[|e_z^{n+1}|] + C\sqrt{\Delta t}\{Var^n(e_z^{n+1})\}^{\frac{1}{2}}.
\end{aligned} \tag{3.57}$$

It follows from (3.48) and (3.57) that

$$\begin{aligned}
|e_z^n| \le {} & |\mathbb{E}_{t_n}^{x^n}[e_z^{n+2}]| + C\Delta t \mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}| + |e_z^{n+1}|] \\
& + C\sqrt{\Delta t}\{(Var^n(e_y^{n+1}))^{\frac{1}{2}} + (Var^n(e_z^{n+1}))^{\frac{1}{2}}\} \\
& + \frac{2|\mathbb{E}_{t_n}^{x^n}[R_y^{n+1}\Delta W_{n+1}]|}{\Delta t} + \frac{2|R_z^n - \mathbb{E}_{t_n}^{x^n}[R_z^{n+1}]|}{\Delta t}.
\end{aligned} \tag{3.58}$$

Then by the inequality $2ab \le \eta\Delta t a^2 + \frac{1}{\eta\Delta t}b^2$ (for any $\eta > 0$) and taking square on both sides of the inequality (3.58), we obtain

$$\begin{aligned}
|e_z^n|^2 \le {} & (1 + \eta\Delta t)|\mathbb{E}_{t_n}^{x^n}[e_z^{n+2}]|^2 + \tilde{C}(1 + \frac{1}{\eta\Delta t})\Big((\Delta t)^2 \mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|^2 + |e_z^{n+1}|^2] \\
& + \Delta t Var^n(e_y^{n+1}) + \Delta t Var^n(e_z^{n+1}) \\
& + \frac{|\mathbb{E}_{t_n}^{x^n}[R_y^{n+1}\Delta W_{n+1}]|^2 + |R_z^n - \mathbb{E}_{t_n}^{x^n}[R_z^{n+1}]|^2}{(\Delta t)^2}\Big),
\end{aligned} \tag{3.59}$$

for some generic constant $\tilde{C} > 0$. By choosing $\eta = 2\tilde{C}$ in the inequality (3.59) it yields that

$$\begin{aligned} |e_z^n|^2 \leq{}& (1 + C\Delta t)|\mathbb{E}_{t_n}^{x^n}[e_z^{n+2}]|^2 \\ &+ \frac{1}{2}(1 + C\Delta t)\left(\Delta t\mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|^2 + |e_z^{n+1}|^2] + Var^n(e_y^{n+1}) + Var^n(e_z^{n+1})\right) \\ &+ \frac{1}{2}(1 + C\Delta t)\frac{|\mathbb{E}_{t_n}^{x^n}[R_y^{n+1}\Delta W_{n+1}]|^2 + |R_z^n - \mathbb{E}_{t_n}^{x^n}[R_z^{n+1}]|^2}{(\Delta t)^3}. \end{aligned} \tag{3.60}$$

By (3.24) we know

$$|e_y^n| \leq |\mathbb{E}_{t_n}^{x^n}[e_y^{n+1}]| + \frac{L}{2}\Delta t\left(|e_y^n| + |e_z^n| + \mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}| + |e_z^{n+1}|]\right) + |R_y^n|. \tag{3.61}$$

Taking square of the inequality (3.61) yields

$$\begin{aligned} |e_y^n|^2 \leq{}& (1 + C\Delta t)|\mathbb{E}_{t_n}^{x^n}[e_y^{n+1}]|^2 + C\Delta t\mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|^2 + |e_z^{n+1}|^2] \\ &+ C\Delta t(|e_y^n|^2 + |e_z^n|^2) + \frac{C|R_y^n|^2}{\Delta t}. \end{aligned} \tag{3.62}$$

From (3.60) and (3.62), we easily deduce

$$\begin{aligned} &|e_y^n|^2 + |e_z^n|^2 \\ &\leq (1 + C\Delta t)\left(|\mathbb{E}_{t_n}^{x^n}[e_y^{n+1}]|^2 + |\mathbb{E}_{t_n}^{x^n}[e_z^{n+2}]|^2 + \frac{1}{2}Var^n(e_y^{n+1}) + \frac{1}{2}Var^n(e_z^{n+1})\right) \\ &\quad + C\Delta t\mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|^2 + |e_z^{n+1}|^2] + C\Delta t(|e_y^n|^2 + |e_z^n|^2) \\ &\quad + C\frac{|R_y^n|^2}{\Delta t} + C\frac{|\mathbb{E}_{t_n}^{x^n}[R_y^{n+1}\Delta W_{n+1}]|^2 + |R_z^n - \mathbb{E}_{t_n}^{x^n}[R_z^{n+1}]|^2}{(\Delta t)^3}. \end{aligned} \tag{3.63}$$

Notice that

$$|\mathbb{E}_{t_n}^{x^n}[e_y^{n+1}]|^2 + \frac{1}{2}Var^n(e_y^{n+1}) = \frac{1}{2}|\mathbb{E}_{t_n}^{x^n}[e_y^{n+1}]|^2 + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|^2], \tag{3.64}$$

$$Var^n(e_z^{n+1}) \leq \mathbb{E}_{t_n}^{x^n}[|e_z^{n+1}|^2], \tag{3.65}$$

then we immediate obtain the result (3.35) by putting (3.64) and (3.65) into (3.63). The proof is completed. □

The following lemma gives us a very useful recursive relation on the growing speed of the errors.

Lemma 3.7. *Suppose that $\varphi \in C_b^3$, $f \in C_b^{3,6,6}$. Then when Δt is sufficiently small, it holds that for $0 \leq n \leq N-5$, if $|e_z^k|^2 \leq \Delta t$ and $|e_y^k|^2 \leq \Delta t$ for $k = n+1, n+2, n+3$, then*

$$\begin{aligned} &|\mathbb{E}_{t_n}^{x^n}[e_z^{n+2}]|^2 + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_z^{n+1}|^2] + \frac{1}{2}|\mathbb{E}_{t_n}^{x^n}[e_y^{n+1}]|^2 + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|^2] \\ &\leq (1 + C\Delta t)\left(\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+2}}^{x^{n+2}}[e_z^{n+4}]|^2] + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_z^{n+3}|^2] \right. \\ &\quad \left. + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+2}}^{x^{n+2}}[e_y^{n+3}]|^2] + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[\mathbb{E}_{t_{n+2}}^{x^{n+2}}[|e_y^{n+3}|^2]]\right) + C(\Delta t)^5, \end{aligned} \tag{3.66}$$

where $C > 0$ is a generic constant depending only on T, upper bounds of derivatives of φ and f.

Proof. Replacing n by $n+1$ in the inequality (3.60) gives us

$$\begin{aligned}\mathbb{E}_{t_n}^{x^n}[|e_z^{n+1}|^2] \leq (1+C\Delta t)&\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+1}}^{x^{n+1}}[e_z^{n+3}]|^2]\\ &+\frac{1+C\Delta t}{2}\Delta t\mathbb{E}_{t_n}^{x^n}[|e_y^{n+2}|^2+|e_z^{n+2}|^2]\\ &+\frac{1+C\Delta t}{2}\mathbb{E}_{t_n}^{x^n}[Var^{n+1}(e_y^{n+2})+Var^{n+1}(e_z^{n+2})]\\ &+C\frac{\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+1}}^{x^{n+1}}[R_y^{n+2}\Delta W_{n+2}]|^2+|R_z^{n+1}-\mathbb{E}_{t_{n+1}}^{x^{n+1}}[R_z^{n+2}]|^2]}{(\Delta t)^3},\end{aligned}\tag{3.67}$$

then we have

$$\begin{aligned}&|\mathbb{E}_{t_n}^{x^n}[e_z^{n+2}]|^2+\frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_z^{n+1}|^2]+\frac{1}{2}|\mathbb{E}_{t_n}^{x^n}[e_y^{n+1}]|^2+\frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|^2]\\ &\leq(1+C\Delta t)\Big(|\mathbb{E}_{t_n}^{x^n}[e_z^{n+2}]|^2+\frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+1}}^{x^{n+1}}[e_z^{n+3}]|^2]+\frac{1}{2}|\mathbb{E}_{t_n}^{x^n}[e_y^{n+1}]|^2\\ &\quad+\frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|^2]+\frac{1}{4}\mathbb{E}_{t_n}^{x^n}[Var^{n+1}(e_y^{n+2})+Var^{n+1}(e_z^{n+2})]\Big)\\ &\quad+\frac{1+C\Delta t}{4}\Delta t\mathbb{E}_{t_n}^{x^n}[|e_y^{n+2}|^2+|e_z^{n+2}|^2]\\ &\quad+C\frac{\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+1}}^{x^{n+1}}[R_y^{n+2}\Delta W_{n+2}]|^2+|R_z^{n+1}-\mathbb{E}_{t_{n+1}}^{x^{n+1}}[R_z^{n+2}]|^2]}{(\Delta t)^3}.\end{aligned}\tag{3.68}$$

By (3.24) and Jensen's inequality, we deduce

$$\begin{aligned}\mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|^2]&\leq\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+1}}^{x^{n+1}}[e_y^{n+2}]|^2]+C\Delta t\mathbb{E}_{t_n}^{x^n}[\mathbb{E}_{t_{n+1}}^{x^{n+1}}[|e_y^{n+2}|^2+|e_z^{n+2}|^2]]\\ &\quad+C\Delta t\mathbb{E}_{t_n}^{x^n}[\mathbb{E}_{t_{n+1}}^{x^{n+1}}[|e_y^{n+1}|^2+|e_z^{n+1}|^2]]+C\frac{\mathbb{E}_{t_n}^{x^n}[\mathbb{E}_{t_{n+1}}^{x^{n+1}}[|R_y^{n+1}|^2]]}{\Delta t}\\ &\leq\mathbb{E}_{t_n}^{x^n}[|e_y^{n+2}|^2]+C\Delta t\mathbb{E}_{t_n}^{x^n}[|e_y^{n+2}|^2+|e_z^{n+2}|^2]\\ &\quad+C\Delta t\mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|^2+|e_z^{n+1}|^2]+C\frac{\mathbb{E}_{t_n}^{x^n}[|R_y^{n+1}|^2]}{\Delta t},\end{aligned}\tag{3.69}$$

which implies

$$\begin{aligned}\mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|^2]&\leq(1+C\Delta t)\mathbb{E}_{t_n}^{x^n}[|e_y^{n+2}|^2]\\ &\quad+C\Delta t\mathbb{E}_{t_n}^{x^n}[|e_z^{n+1}|^2+|e_z^{n+2}|^2]+\frac{C\mathbb{E}_{t_n}^{x^n}[|R_y^{n+1}|^2]}{\Delta t}.\end{aligned}\tag{3.70}$$

By (3.24) we easily get

$$\begin{aligned}|\mathbb{E}_{t_n}^{x^n}[e_y^{n+1}]|^2&\leq(1+C\Delta t)|\mathbb{E}_{t_n}^{x^n}[e_y^{n+2}]|^2+C\Delta t\mathbb{E}_{t_n}^{x^n}[|e_y^{n+2}|^2+|e_z^{n+2}|^2]\\ &\quad+C\Delta t\mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|^2+|e_z^{n+1}|^2]+C\frac{|\mathbb{E}_{t_n}^{x^n}[R_y^{n+1}]|^2}{\Delta t}.\end{aligned}\tag{3.71}$$

Combining the inequalities (3.70) and (3.71), we then obtain

$$\begin{aligned}
&\frac{1}{2}|\mathbb{E}_{t_n}^{x^n}[e_y^{n+1}]|^2+\frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|^2]+\frac{1}{4}\mathbb{E}_{t_n}^{x^n}[Var^{n+1}(e_y^{n+2})]\\
&\quad\le(1+C\Delta t)\left(\frac{1}{2}|\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+1}}^{x^{n+1}}[e_y^{n+2}]|^2]+\frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_y^{n+2}|^2]\right)\\
&\qquad+C\Delta t\mathbb{E}_{t_n}^{x^n}[|e_z^{n+1}|^2+|e_z^{n+2}|^2]+C\frac{\mathbb{E}_{t_n}^{x^n}[|R_y^{n+1}|^2]}{\Delta t}\\
&\qquad+\frac{1}{4}\mathbb{E}_{t_n}^{x^n}[\mathbb{E}_{t_{n+1}}^{x^{n+1}}[|e_y^{n+2}|^2]-|\mathbb{E}_{t_{n+1}}^{x^{n+1}}[e_y^{n+2}]|^2]\\
&\quad\le(1+C\Delta t)\mathbb{E}_{t_n}^{x^n}[|e_y^{n+2}|^2]+C\Delta t\mathbb{E}_{t_n}^{x^n}[|e_z^{n+1}|^2+|e_z^{n+2}|^2]+C\frac{\mathbb{E}_{t_n}^{x^n}[|R_y^{n+1}|^2]}{\Delta t}.
\end{aligned}\tag{3.72}$$

Notice that

$$\begin{aligned}
&|\mathbb{E}_{t_n}^{x^n}[e_z^{n+2}]|^2+\frac{1}{4}\mathbb{E}_{t_n}^{x^n}[Var^{n+1}(e_z^{n+2})]\\
&\quad=|\mathbb{E}_{t_n}^{x^n}[e_z^{n+2}]|^2+\frac{1}{4}\mathbb{E}_{t_n}^{x^n}[\mathbb{E}_{t_{n+1}}^{x^{n+1}}[|e_z^{n+2}|^2]-|\mathbb{E}_{t_{n+1}}^{x^{n+1}}[e_z^{n+2}]|^2]\\
&\quad\le\mathbb{E}_{t_n}^{x^n}[|e_z^{n+2}|^2].
\end{aligned}\tag{3.73}$$

Using (3.68), (3.72) and (3.73) we then deduce

$$\begin{aligned}
&|\mathbb{E}_{t_n}^{x^n}[e_z^{n+2}]|^2+\frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_z^{n+1}|^2]+\frac{1}{2}|\mathbb{E}_{t_n}^{x^n}[e_y^{n+1}]|^2+\frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|^2]\\
&\quad\le(1+C\Delta t)\left(\mathbb{E}_{t_n}^{x^n}[|e_y^{n+2}|^2+|e_z^{n+2}|^2]+\frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+1}}^{x^{n+1}}[e_z^{n+3}]|^2]\right)\\
&\qquad+C\frac{\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+1}}^{x^{n+1}}[R_y^{n+2}\Delta W_{n+2}]|^2+|R_z^{n+1}-\mathbb{E}_{t_{n+1}}^{x^{n+1}}[R_z^{n+2}]|^2]}{(\Delta t)^3}\\
&\qquad+C\frac{\mathbb{E}_{t_n}^{x^n}[|R_y^{n+1}|^2]}{\Delta t}.
\end{aligned}\tag{3.74}$$

Replacing n by $n+2$ in (3.59) and (3.62) gives us

$$\begin{aligned}
\mathbb{E}_{t_n}^{x^n}[|e_y^{n+2}|^2]&\le(1+C\Delta t)\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+2}}^{x^{n+2}}[e_y^{n+3}]|^2]+C\Delta t\mathbb{E}_{t_n}^{x^n}[|e_y^{n+3}|^2+|e_z^{n+3}|^2]\\
&\qquad+C\Delta t\mathbb{E}_{t_n}^{x^n}[|e_z^{n+2}|^2]+\frac{C\mathbb{E}_{t_n}^{x^n}[|R_y^{n+2}|^2]}{\Delta t},
\end{aligned}\tag{3.75}$$

and

$$\begin{aligned}
\mathbb{E}_{t_n}^{x^n}[|e_z^{n+2}|^2]&\le(1+C\Delta t)\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+2}}^{x^{n+2}}[e_z^{n+4}]|^2]+\frac{1+C\Delta t}{2}\Delta t\mathbb{E}_{t_n}^{x^n}[|e_y^{n+3}|^2+|e_z^{n+3}|^2]\\
&\qquad+\frac{1+C\Delta t}{2}\mathbb{E}_{t_n}^{x^n}[Var^{n+2}(e_y^{n+3})+Var^{n+2}(e_z^{n+3})]\\
&\qquad+C\frac{\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+2}}^{x^{n+2}}[R_y^{n+3}\Delta W_{n+3}]|^2+|R_z^{n+2}-\mathbb{E}_{t_{n+2}}^{x^{n+2}}[R_z^{n+3}]|^2]}{(\Delta t)^3}.
\end{aligned}\tag{3.76}$$

Lemma 3.4 tells us that

$$|\mathbb{E}_{t_i}^{x^i}[R_y^{i+1}\Delta W_{i+1}]|\le C(\Delta t)^4,\qquad|R_z^i-\mathbb{E}_{t_i}^{x^i}[R_z^{i+1}]|\le C(\Delta t)^4\tag{3.77}$$

for $0\le i\le N-1$. Also notice that

$$\frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+1}}^{x^{n+1}}[e_z^{n+3}]|^2]+\frac{1}{2}\mathbb{E}_{t_n}^{x^n}[Var^{n+2}(e_z^{n+3})]\le\frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_z^{n+3}|^2]$$

and

$$\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+2}}^{x^{n+2}}[e_y^{n+3}]|^2] + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[Var^{n+2}(e_y^{n+3})] = \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+2}}^{x^{n+2}}[e_y^{n+3}]|^2] + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[\mathbb{E}_{t_{n+2}}^{x^{n+2}}[|e_y^{n+3}|^2]].$$

Then by using (3.31), (3.74), (3.75), (3.76), Lemma 3.4 and above identities, we have

(3.78)

$$\begin{aligned}
&|\mathbb{E}_{t_n}^{x^n}[e_z^{n+2}]|^2 + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_z^{n+1}|^2] + \frac{1}{2}|\mathbb{E}_{t_n}^{x^n}[e_y^{n+1}]|^2 + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|^2] \\
&\quad \le (1+C\Delta t)\Big(\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+2}}^{x^{n+2}}[e_y^{n+3}]|^2] + \mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+2}}^{x^{n+2}}[e_z^{n+4}]|^2] \\
&\qquad + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[Var^{n+2}(e_y^{n+3}) + Var^{n+2}(e_z^{n+3})] + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+1}}^{x^{n+1}}[e_z^{n+3}]|^2]\Big) \\
&\qquad + C\Delta t\mathbb{E}_{t_n}^{x^n}[|e_y^{n+3}|^2 + |e_z^{n+3}|^2] \\
&\qquad + C\frac{\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+1}}^{x^{n+1}}[R_y^{n+2}\Delta W_{n+2}]|^2 + |R_z^{n+1} - \mathbb{E}_{t_{n+1}}^{x^{n+1}}[R_z^{n+2}]|^2]}{(\Delta t)^3} \\
&\qquad + C\frac{\mathbb{E}_{t_n}^{x^n}[|R_y^{n+1}|^2 + |R_y^{n+2}|^2]}{\Delta t} \\
&\qquad + C\frac{\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+2}}^{x^{n+2}}[R_y^{n+3}\Delta W_{n+3}]|^2 + |R_z^{n+2} - \mathbb{E}_{t_{n+2}}^{x^{n+2}}[R_z^{n+3}]|^2]}{(\Delta t)^3} \\
&\quad \le (1+C\Delta t)\Big(\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+2}}^{x^{n+2}}[e_z^{n+4}]|^2] + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_z^{n+3}|^2] \\
&\qquad + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+2}}^{x^{n+2}}[e_y^{n+3}]|^2] + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[\mathbb{E}_{t_{n+2}}^{x^{n+2}}[|e_y^{n+3}|^2]]\Big) + C(\Delta t)^5.
\end{aligned}$$

The proof is then completed. □

Now let us present the main result on convergence of the Crank-Nicolson scheme for solving general BSDEs.

Theorem 3.8. *Let (y_t, z_t) be the solution of the BSDE (1.2) with the terminal condition $y_T = \varphi(W_T)$, and (y^n, z^n) be its approximate solution produced by using the Crank-Nicolson scheme. Suppose that $\varphi \in C_b^3$, $f \in C_b^{3,6,6}$, and the initial error satisfies*

$$\mathbb{E}_{t_{N-1}}^{x^{N-1}}[|y_{t_N} - y^N|^2] \le C(\Delta t_{N-1})^3$$

for some constant $C > 0$. Then we have the following L^2 error estimate: when Δt is sufficiently small, it holds that for $0 \le n \le N-1$,

$$\max_{0\le l\le n\le N-1} \mathbb{E}_{t_l}^{x_l}[|y_{t_n} - y^n|^2 + |z_{t_n} - z^n|^2] \le C_{T,\phi,f}(\Delta t)^4, \tag{3.79}$$

where $C_{T,\phi,f} > 0$ is a generic constant depending only on T, upper bounds of derivatives of φ and f.

Proof. Let us first choose constants $C_* > 0$ and $1 > \delta_* > 0$ such that Lemmas 3.4-3.7 hold for this C_* when the time step size $\Delta t \le \delta_*$. Set $C_s = C_* e^{C_* T}(3+T)$. Then we define some constants in the below:

$$C_{T,\phi,f} = 3C_s + 2C_*^2 + 4C_*^3,$$

$$\delta = min\{\delta_*, \frac{1}{2C_*}, \frac{1}{C_{T,\phi,f}}\}.$$

In the following, we will show by *induction* that that for any k with $0 \le k \le N-1$, it holds that

$$|e_y^k|^2 + |e_z^k|^2 \le C_{T,\phi,f}(\Delta t)^4, \tag{3.80}$$

when the time step size $\Delta t < \delta$. We would like to point out that (3.80) also implies

$$|e_y^k|^2 \le \Delta t, \qquad |e_z^k|^2 \le \Delta t, \tag{3.81}$$

since $C_{T,\phi,f}(\Delta t)^4 \le C_{T,\phi,f}(1/C_{T,\phi,f})(\Delta t)^3 \le (\Delta t)^3 \le \Delta t$.

First by Lemma 3.5, we know (3.80) clearly holds for $k = N-1, N-2, N-3$ since $C_{T,\phi,f} \ge C_*$ and $\delta \le \delta_*$.

Now we assume that (3.80) holds for all $n+1 \le k \le N-1$. Note that it also means (3.81) holds for all $n+1 \le k \le N-1$. We will show that (3.80) also hold for $k = n$ under such assumption.

Using Lemma 3.7 we get

$$\begin{aligned}
&|\mathbb{E}_{t_n}^{x^n}[e_z^{n+2}]|^2 + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_z^{n+1}|^2] + \frac{1}{2}|\mathbb{E}_{t_n}^{x^n}[e_y^{n+1}]|^2 + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|^2] \\
&\quad \le (1 + C_*\Delta t)\big(\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+2}}^{x^{n+2}}[e_z^{n+4}]|^2] + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_z^{n+3}|^2] \\
&\qquad + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{n+2}}^{x^{n+2}}[e_y^{n+3}]|^2] + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[\mathbb{E}_{t_{n+2}}^{x^{n+2}}[|e_y^{n+3}|^2]]\Big) + C_*(\Delta t)^5 \\
&\quad \le (1 + C_*\Delta t)^2\big(\mathbb{E}_{t_{n+2}}^{x^{n+2}}[|\mathbb{E}_{t_{n+4}}^{x^{n+4}}[e_z^{n+6}]|^2] + \frac{1}{2}\mathbb{E}_{t_{n+2}}^{x^{n+2}}[|e_z^{n+5}|^2] \\
&\qquad + \frac{1}{2}\mathbb{E}_{t_{n+2}}^{x^{n+2}}[|\mathbb{E}_{t_{n+4}}^{x^{n+4}}[e_y^{n+5}]|^2] + \frac{1}{2}\mathbb{E}_{t_{n+2}}^{x^{n+2}}[\mathbb{E}_{t_{n+4}}^{x^{n+4}}[|e_y^{n+5}|^2]]\Big) \\
&\qquad + [C_* + (1 + C_*\Delta t)C_*](\Delta t)^5.
\end{aligned} \tag{3.82}$$

If $N-n-3$ is even, by repeating the above process we get

$$\begin{aligned}
&|\mathbb{E}_{t_n}^{x^n}[e_z^{n+2}]|^2 + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_z^{n+1}|^2] + \frac{1}{2}|\mathbb{E}_{t_n}^{x^n}[e_y^{n+1}]|^2 + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|^2] \\
&\quad \le (1 + C_*\Delta t)^{\frac{N-n-3}{2}}\Big(\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{N-3}}^{x^{N-3}}[e_z^{N-1}]|^2] + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_z^{N-2}|^2] \\
&\qquad + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|\mathbb{E}_{t_{N-3}}^{x^{N-3}}[e_y^{N-2}]|^2] + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[\mathbb{E}_{t_{N-3}}^{x^{N-3}}[|e_y^{N-2}|^2]]\Big) \\
&\qquad + C_*(\Delta t)^5 \sum_{i=1}^{\frac{N-n-3}{2}} (1 + C_*\Delta t)^{i-1} \\
&\quad \le e^{C_*(N-3)\Delta t/2}\Big(\mathbb{E}_{t_n}^{x^n}[|e_z^{N-1}|^2] + \frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_z^{N-2}|^2] + \mathbb{E}_{t_n}^{x^n}[|e_y^{N-2}|^2]\Big) \\
&\qquad + C_*(\Delta t)^4((N-3)\Delta t)e^{C_*(N-3)\Delta t/2} \\
&\quad \le 3C_* e^{C_*T}(\Delta t)^4 + C_* T e^{C_*T}(\Delta t)^4 \\
&\quad = C_s(\Delta t)^4.
\end{aligned} \tag{3.83}$$

Similarly, if $N-n-3$ is odd, then we obtain

$$\begin{aligned}
&|\mathbb{E}_{t_n}^{x^n}[e_z^{n+2}]|^2+\frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_z^{n+1}|^2]+\frac{1}{2}|\mathbb{E}_{t_n}^{x^n}[e_y^{n+1}]|^2+\frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|^2]\\
&\quad\leq e^{C_*(N-4)\Delta t/2}\Big(\mathbb{E}_{t_n}^{x^n}[|e_z^{N-2}|^2]+\frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_z^{N-3}|^2]+\mathbb{E}_{t_n}^{x^n}[|e_y^{N-3}|^2]\Big)\\
&\qquad\quad+C_*(\Delta t)^4((N-4)\Delta t/2)e^{C_*(N-3)\Delta t/2}\\
&\quad\leq 3C_*e^{C_*T}(\Delta t)^4+C_*Te^{C_*T}(\Delta t)^4\\
&\quad=C_s(\Delta t)^4.
\end{aligned}\tag{3.84}$$

Now by the above results and Lemmas 3.6 and 3.3, we have

$$\begin{aligned}
|e_y^n|^2+|e_z^n|^2\leq{}&\frac{1+C_*\Delta t}{1-C_*\Delta t}\Big(\frac{1}{2}|\mathbb{E}_{t_n}^{x^n}[e_y^{n+1}]|^2+\frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_y^{n+1}|^2]+|\mathbb{E}_{t_n}^{x^n}[e_z^{n+2}]|^2\\
&+\frac{1}{2}\mathbb{E}_{t_n}^{x^n}[|e_z^{n+1}|^2]\Big)+\frac{C_*}{1-C_*\Delta t}\frac{|R_y^n|^2}{\Delta t}\\
&+\frac{C_*}{1-C_*\Delta t}\frac{|\mathbb{E}_{t_n}^{x^n}[R_y^{n+1}\Delta W_{n+1}]|^2+|R_z^n-\mathbb{E}_{t_n}^{x^n}[R_z^{n+1}]|^2}{(\Delta t)^3}\\
\leq{}&3C_s(\Delta t)^4+2C_*^2(\Delta t)^5+4C_*^3(\Delta t)^5\\
\leq{}&C_{T,\phi,f}(\Delta t)^4.
\end{aligned}\tag{3.85}$$

Thus we prove (3.80) by induction.

For any integer l with $0\leq l\leq n$, taking the conditional mathematical expectation $\mathbb{E}_{t_l}^{x_l}[\cdot]$ on both sides of (3.80), we immediately obtain

$$\max_{0\leq l\leq n\leq N-1}\mathbb{E}_{t_l}^{x_l}[|e_y^n|^2+|e_z^n|^2]\leq C_{T,\phi,f}(\Delta t)^4,\tag{3.86}$$

for $n=N-1,\cdots,0$. The proof is completed. □

Theorem 3.9. *Under the conditions of Theorem 3.8, we have the following L^p ($p\geq 1$) error estimate: when Δt is sufficiently small, it holds that for $0\leq n\leq N-1$,*

$$\max_{0\leq l\leq n\leq N-1}\mathbb{E}_{t_l}^{x_l}[|y_{t_n}-y^n|^p+|z_{t_n}-z^n|^p]\leq C_{p,T,\phi,f}(\Delta t)^{2p},\tag{3.87}$$

where $C_{p,T,\phi,f}>0$ is a generic constant depending only on p, T, upper bounds of derivatives of φ and f.

Proof. There are two cases to discuss.

Case I: $p\geq 2$. For any $0\leq n\leq N-1$, taking the power of $\frac{p}{2}$ on both sides of the inequality (3.80) of Theorem 3.8 gives us

$$|e_y^n|^p+|e_z^n|^p\leq(|e_y^n|^2+|e_z^n|^2)^{\frac{p}{2}}\leq C_{p,T,\phi,f}(\Delta t)^{2p},\tag{3.88}$$

where $C_{p,T,\phi,f}=(C_{T,\phi,f})^{\frac{p}{2}}$. Thus we obtain

$$\max_{0\leq l\leq n\leq N-1}\mathbb{E}_{t_l}^{x_l}[|e_y^n|^p+|e_z^n|^p]\leq C_{p,T,\phi,f}(\Delta t)^{2p}.\tag{3.89}$$

Case II: $1\leq p<2$. By the Jensen inequality and the inequality (3.86) of Theorem 3.8 we have

$$\max_{0\leq l\leq n\leq N-1}\mathbb{E}_{t_l}^{x_l}[|e_y^n|^p]^{\frac{2}{p}}\leq\max_{0\leq l\leq n\leq N-1}\mathbb{E}_{t_l}^{x_l}[|e_y^n|^{p\frac{2}{p}}]=\max_{0\leq l\leq n\leq N-1}\mathbb{E}_{t_l}^{x_l}[|e_y^n|^2]\leq C_{T,\phi,f}(\Delta t)^4,$$

which implies

$$\max_{0\leq l\leq n\leq N-1}\mathbb{E}_{t_l}^{x_l}[|e_y^n|^p]\leq(C_{T,\phi,f})^{\frac{p}{2}}(\Delta t)^{2p}.\tag{3.90}$$

Similarly we get

$$\max_{0\le l\le n\le N-1} \mathbb{E}_{t_l}^{x_l}[|e_z^n|^p] \le (C_{T,\phi,f})^{\frac{p}{2}}(\Delta t)^{2p}. \tag{3.91}$$

Combination of the inequalities (3.90) and (3.91) leads to

$$\max_{0\le l\le n\le N-1} \mathbb{E}_{t_l}^{x_l}[|e_y^n|^p + |e_z^n|^p] \le C_{p,T,\phi,f}(\Delta t)^{2p}, \tag{3.92}$$

where $C_{p,T,\phi,f} = 2(C_{T,\phi,f})^{\frac{p}{2}}$. □

Remark 3. We would like to remark that Assumption 1 can be slightly relaxed and Theorems 3.8 and 3.9 still hold. For example, one just need assume that $\Delta t_{N-1} = O((\Delta t)^2)$, $\Delta t_n = O(\Delta t)$ for $0 \le n \le N-2$, and $|\Delta t_n - \Delta t_{n+1}| = O((\Delta t)^2)$ for $0 \le n \le N-3$. It basically means that the time steps must be asymptotically uniform in local in order for the C-N scheme to obtain second-order accuracy in solving BSDEs, but they could be non-uniform in global.

4. Conclusions

In this paper, we study error estimates of a special θ-scheme – the Crank-Nicolson scheme for solving backward stochastic differential equations. We rigorously prove that under some reasonable regularity conditions on φ and f, this scheme is second-order accurate for solving both y_t and z_t if the errors are measured in the L^p $(p \ge 1)$ norm. A key idea in the proof is to use cancelation of local errors in neighbor time steps. Some of future work includes application of the techniques developed in this paper to error analysis of some other accurate schemes for solving BSDEs such as the multi-step scheme proposed in [27] and design of high-order schemes for solving FBSDEs.

References

[1] C. Bender and R. Denk, A forward scheme for backward SDEs, Stochastic Process. Appl., 117 (2007) 1793-1812.

[2] B. Bouchard and N. Touzi, Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations, Stochastic Process. Appl., 111 (2004) 175-206.

[3] D. Chevance, *Numerical methods for backward stochastic differential equations*, in Numerical Methods in Finance, p. 232-244, Publ. Newton Inst., Cambridge Univ. Press, Cambridge, 1997.

[4] J. Cvitanic and J. Zhang, The steepest descent method for forward-backward SDEs, Electron. J. Probab., 10 (2005) 1468-1495.

[5] F. Delarue and S. Menozzi, A forward-backward stochastic algorithm for quasi-linear PDEs, Ann. Appl. Probab., 16 (2006) 140-184.

[6] L. C. Evans, Partial Differential Equations, Providence, RI, American Mathematical Society, 1998.

[7] E. R. Gianin, Risk measures via g-expectations, Insurance Math. Econom., 39 (2006) 19–34.

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

[9] E. Gobet, J.–P. Lemmor and X. Warin, A regression-based Monte Carlo method to solve backward stochastic differential equations, Ann. Appl. Probab., 15 (2005) 2172-2202.

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

[11] J. Ma and J. Yong, Forward-backward stochastic differential equations and their applications, Lecture Notes in Mathematics, 1702, Springer-Verlag, Berlin, 1999.

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

[13] J. Ma, J. Shen and Y. Zhao, On Numerical approximations of forward-backward stochastic differential equations, SIAM J.Numer. Anal., 46 (2008) 2636-2661.

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

[15] 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 (2007) 24-44.
[16] E. Pardoux and S. Peng, Adapted solution of a backward stochastic differential equation, Systems Control Lett., 14 (1990) 55-61.
[17] E. Pardoux and S. Peng, Backward doubly stochastic differential equations and systems of quasilinear SPDEs, Probab. Theory Relat. Fields, 98, 4 (1994) 209-227.
[18] S. Peng, A general stochastic maximum principle for optimal control problems, SIAM J. Control Optim., 28 (1990) 966-979.
[19] S. Peng, Probabilistic interpretation for systems of quasilinear parabolic partial differential equations, Stochastics Stochastics Rep., 37 (1991) 61-74.
[20] S. Peng, *Backward SDE and Related g-Expectation*, in Backward stochastic differential equations (Paris, 1995-1996), 141-159, Pitman Res. Notes Math. Ser., 364, Longman, Harlow, 1997.
[21] S. Peng, A linear approximation algorithm using BSDE, Pacific Economic Review, 4 (1999) 285-291.
[22] J. Wang, C. Luo and W. Zhao, Crank-Nicolson scheme and its error estimates for Backward Stochastic Differential Equations, Acta Mathematicae Applicatae Sinica(English Series), 2010. DOI: 10.1007/s10255-009-9051-z.
[23] J. Zhang, A numerical scheme for BSDEs, Ann. Appl. Probab., 14 (2004) 459-488.
[24] Y. Zhang and W. Zheng, Discretizing a backward stochastic differential equation, Int. J. Math. Math. Sci., 32 (2002) 103-116.
[25] W. Zhao, L. Chen and S. Peng, A new kind of accurate numerical method for backward stochastic differential equations, SIAM J. Sci. Comput., 28 (2006) 1563-1581.
[26] W. Zhao, J. Wang and S. Peng, Error Estimates of the θ-Scheme for Backward Stochastic Differential Equations, Dis. Cont. Dyn. Sys. B, 12 (2009) 905-924.
[27] W. Zhao, G. Zhang and L. Ju, A stable multistep scheme for solving backward stochastic differential equations, SIAM J. Numer. Anal., 48 (2010) 1369-1394.

School of Mathematics, Shandong University, Jinan, Shandong, 250100, China
E-mail: `wdzhao@sdu.edu.cn, liyang19820816@163.com`

Department of Mathematics, University of South Carolina, Columbia, SC 29208, USA
E-mail: `ju@math.sc.edu`