DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS SERIES B
Volume **18**, Number **8**, October **2013**

doi:10.3934/dcdsb.2013.18.2051

pp. **2051–2067**

MEAN-SQUARE CONVERGENCE OF NUMERICAL APPROXIMATIONS FOR A CLASS OF BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS

Chuchu Chen and Jialin Hong

State Key Laboratory of Scientific and Engineering Computing
Institute of Computational Mathematics
and Scientific/Engineering Computing
Academy of Mathematics and Systems Science
Chinese Academy of Sciences
Beijing, 100190, China

(Communicated by Zhimin Zhang)

Abstract. This paper is devoted to the fundamental convergence theorem on the mean-square order of numerical approximations for a class of backward stochastic differential equations with terminal condition $\chi = \varphi(W_T + x)$. Our theorem shows that the mean-square order of convergence of a numerical method depends on the order of the one-step approximation for the mean-square deviation only. And some numerical schemes as examples are presented to verify the theorem.

1. **Introduction.** The study of backward stochastic differential equations (BSDEs) is strongly motivated by numerous applications in finance and stochastic control theory [1, 4]. In 1990, E. Pardoux and S. Peng [6] proved the existence and uniqueness of the solution of the general BSDEs. They also revealed the natural connection between the backward stochastic differential equation and the parabolic partial differential equation (PDE) [7].

As we know, few BSDEs can be analyzed exactly, even for the simple linear case. To further investigate this type of equations, numerical methods are necessary. In fact some efforts have been made on this topic. The studies on the important case that the terminal condition is a function of W_T, where W_t is a Brownian motion, are as follows. A family of numerical schemes depending on two parameters θ_1 and θ_2 was proposed in [9] and the L^p-error estimation was considered in [2] when $\theta_1 = \frac{1}{2}$ and $\theta_2 = 1$. [8] developed a scheme of Crank-Nicolson type and proved the second order convergence in the strong L^1- sense for both variables while [2] proved the second order convergence of this scheme in the L^p-sense. Utilizing the variational equation, [11] showed that for $\theta \in [0, 1]$ the strong order of θ-scheme is 1 for both variables and also gave the strong order of the special case $\theta = \frac{1}{2}$. The above papers [2, 8, 9, 11] considered the case that the generator $f = f(t, Y)$, while [10] extended

2010 *Mathematics Subject Classification.* Primary: 60H10, 60H35, 65C20, 65C30.

Key words and phrases. Backward stochastic differential equations, mean-square order of convergence.

Authors are supported by NNSFC (NO. 91130003 and NO. 11021101).

this to $f = f(t, Y, Z)$. Following them, we consider the case where the terminal condition is a function of W_T with $f = f(t, Y, Z)$ being the generator.

The mean-square order is an important convergence index for numerical methods of stochastic system. In the numerical analysis of stochastic ordinary differential equations (SODEs), [5] presents a fundamental convergence theorem which establishes the mean-square order of convergence of a method resting on properties of its one-step approximation only. This fundamental theorem is the most important criterion to evaluate the mean-square order of convergence for numerical approximations of SODEs. However, to the best of our knowledge, there are no such theorem in the numerical analysis of BSDEs, even the terminal condition is reduced to a function of W_T. Hence motivated by Milstein's work, we propose a fundamental convergence theorem on the mean-square orders of numerical approximations for a class of BSDEs with terminal condition $\chi = \varphi(W_T + x)$ in this paper. The theorem shows that the mean-square order of convergence of a numerical method depends on the property of mean-square deviation of one-step approximation only. As we all know, the solution of a BSDE is a pair of stochastic processes. The main difficult problem lies in the estimation of the martingale integrand Z_t. By utilizing the variational equation of the BSDE, this problem is solved in this paper.

The rest of the paper is organized as follows. In section 2 we present some preliminaries including basic assumptions and properties of the BSDE and its variational equation. In section 3 we prove the fundamental convergence theorem on the mean-square order of convergence for general numerical methods. Section 4 presents some specific numerical methods to verify the efficiency of the fundamental convergence theorem and finally conclusions are made in section 5.

2. Preliminaries.

The general form of the BSDE is

$$\begin{aligned} dY(t) &= f(t, Y(t), Z(t)))\mathrm{d}t + Z(t)\mathrm{d}W_t, \quad 0 \le t \le T, \\ Y(T) &= \chi, \end{aligned} \tag{1}$$

where $W = (W^1, \cdots, W^d)^T$ is a standard d-dimensional Brownian motion defined on a complete probability space $(\Omega, \mathcal{F}, P)$, with $\{\mathcal{F}_t, t \in [0, T]\}$ being its natural normal filtration, $\chi \in L^2(\Omega)$ is a $\mathcal{F}_T$-measurable random variable, and $f : [0, T] \times R \times R^d \longmapsto R$ is a Lipschitz function and assumed to be $\mathcal{B}([0, T]) \otimes \mathcal{B} \otimes \mathcal{B}^d/\mathcal{B}$ measurable. The unknowns are a pair of $(\mathcal{F}_t)$-adapted processes $Y(t)$ and $Z(t)$, whose existence and uniqueness are shown in [6]. In this paper, $Y(t)$ and $Z(t)$ are said to be the first and second processes of equation (1), respectively.

Following [2, 8, 9, 10, 11], we focus on the case where there exists a Borel measurable function $\varphi : R^d \longmapsto R$ such that $\chi = \varphi(W_T + x)$ with $x \in R^d$ being deterministic in this paper.

2.1. Basic assumptions and notations.

In this paper, we make the following assumptions:

(i) $\varphi(x) \in C_b^3$, where C_b^k is the set of continuously differentiable functions $\phi(x)$ such that the derivatives $\frac{\partial^l \phi}{\partial x^l}$ exist and are uniformly bounded for $1 \le l \le k$.

(ii) $f(t, y, z) \in C_b^{1,2,2}$, where $C_b^{k/2,k,k}$ is the set of continuously differentiable functions $\phi(t, y, z)$ such that the partial derivatives $\partial_t^{l_0}\partial_y^{l_1}\partial_z^{l_2}\phi(t, y, z)$ exist and are uniformly bounded for $2l_0 + l_1 + l_2 \le k$.

Since it is natural to use some sufficiently broad assumptions which allow us to analyze the numerical approximations conveniently, we do not assert that these assumptions are the optimal ones for our analysis, but they are sufficient. However, investigations of the optimal assumptions are beyond the scope of the paper.

We also make the following notations in sequels.

- $|x|$ is the Euclide norm of the vector or matrix x.
- For $f(t, y, z) : [0, T] \times R \times R^d$, denote the vector $\left(\frac{\partial f}{\partial z_1}, \cdots, \frac{\partial f}{\partial z_d}\right)^T$ of partial derivatives with respect to each component of z by $\frac{\partial f}{\partial z}$; $\frac{\partial^2 f}{\partial y \partial z}$ and $\frac{\partial^2 f}{\partial z^2}$ mean the vector $\left(\frac{\partial^2 f}{\partial y \partial z_1}, \cdots, \frac{\partial f}{\partial y \partial z_d}\right)^T$ and the matrix $\left(\frac{\partial^2 f}{\partial z^2}\right)$, respectively.
- $Y_{t+h,\xi}(t)$ is the process Y(t) under the condition $Y(t+h) = \xi$.
- $\Delta W_{t_{k+1}} = W_{t_{k+1}} - W_{t_k}$.

2.2. **Properties of the BSDE and its variational equation.** If $\chi = \varphi(W_T + x)$, then the solution of BSDE (1) relates to a quasilinear parabolic PDE

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

with $\nabla_x u = (\frac{\partial u}{\partial x_1}, \cdots, \frac{\partial u}{\partial x_d})^T$ and the terminal condition $u(T, x) = \varphi(x)$, see for instance [7] or [11]. By Itô formula, we have the following theorem.

Theorem 2.1. *Let $u(t, x)$ be the solution of the equation (2), $Y_t = u(t, W_t + x)$, $Z_t = \nabla Y_t = \nabla_x u(t, W_t + x)$ and $\nabla Z_t = \nabla_x^2 u(t, W_t + x)$. Under the above assumptions (i) and (ii), we have the following relationships*

$$\begin{aligned} &dY_t = f(t, Y_t, Z_t)dt + Z_t dW_t, \quad t \in [0, T], \\ &Y_T = \varphi(W_T + x), \end{aligned} \tag{3}$$

and

$$\begin{aligned} &d\nabla Y_t = F(t, Y_t, Z_t, \nabla Y_t, \nabla Z_t)dt + \nabla Z_t dW_t, \quad t \in [0, T], \\ &\nabla Y_T = \frac{\partial}{\partial x} \varphi(W_T + x), \end{aligned} \tag{4}$$

where $F(t, Y_t, Z_t, \nabla Y_t, \nabla Z_t) = \nabla Y_t \frac{\partial}{\partial y} f(t, Y_t, Z_t) + \nabla Z_t \frac{\partial}{\partial z} f(t, Y_t, Z_t)$ and the partial derivatives refer to $f = f(t, y, z)$.

Proof. Assume $d = 1$ without loss of generality.

It follows from Itô formula that

$$\begin{aligned} dY_t &= du(t, W_t + x) \\ &= \frac{\partial u}{\partial t}(t, W_t + x)dt + \frac{\partial u}{\partial x}(t, W_t + x)dW_t + \frac{1}{2}\frac{\partial^2 u}{\partial x^2}(t, W_t + x)dt \\ &= f(t, u(t, W_t + x), \frac{\partial}{\partial x}u(t, W_t + x))dt + \frac{\partial}{\partial x}u(t, W_t + x)dW_t \\ &= f(t, Y_t, Z_t)dt + Z_t dW_t. \end{aligned}$$

Similarly,

$$\begin{aligned} dZ_t &= d(\frac{\partial u}{\partial x}(t, W_t + x)) \\ &= \frac{\partial^2 u}{\partial t \partial x}(t, W_t + x)dt + \frac{\partial^2 u}{\partial x^2}(t, W_t + x)dW_t + \frac{1}{2}\frac{\partial^3 u}{\partial x^3}(t, W_t + x)dt \\ &= \frac{\partial}{\partial x}\Big(\frac{\partial u}{\partial t}(t, W_t + x) + \frac{1}{2}\frac{\partial^2 u}{\partial x^2}(t, W_t + x)\Big)dt + \frac{\partial^2 u}{\partial x^2}(t, W_t + x)dW_t \\ &= \Big(\nabla Y_t \frac{\partial}{\partial y} f(t, Y_t, Z_t) + \nabla Z_t \frac{\partial}{\partial z} f(t, Y_t, Z_t)\Big)dt + \nabla Z_t dW_t \\ &= F(t, Y_t, Z_t, \nabla Y_t, \nabla Z_t)dt + \nabla Z_t dW_t. \end{aligned}$$

This completes the proof. □

Remark 1. We denote ∇Y_t (resp. ∇Z_t) as the variation of Y_t (resp. Z_t) with respect to x. Analogously, $\nabla \bar{Y}_k$ (resp. $\nabla \bar{Z}_k$) means variation of the numerical solution Y_k (resp. Z_k) with respect to x in the following.

Remark 2. The equation (4) is called the variational equation of (3). As stated before, when one analyzes the convergence order of a numerical method to the BSDE, the difficulty usually lies in the estimation of the second process Z_t. Here we solve this problem by utilizing the variational equation (4), especially, the relationship $\nabla Y_t = Z_t$.

Next we present the property of the boundedness of solutions for the BSDE (3) and its variational equation (4) under assumptions (i) and (ii), which can be derived from the regularity of the solution $u(t, x)$ for the PDE (2), too. However, the proof here is from the pointview of BSDEs instead of PDEs; see [3] for the boundedness of Y_t in the case of $f = f(t, Y)$.

Throughout this paper, all constants K depend only on T, the coefficients of the equation and its numerical approximation, which may be different from line to line.

Proposition 1. *Under assumptions (i) and (ii) above, the solutions Y_t, Z_t and ∇Z_t of the equations (3) and (4) are uniformly bounded by some positive constant K a.s..*

Proof. We define an approximating sequence by a kind of Picard iteration. Let $Z_t^0 \equiv 0$, and $\{Y_t^{(n)}, Z_t^{(n)}, 0 \le t \le T\}_{n \ge 1}$ be defined recursively by

$$Y_t^{(n)} = \varphi(W_T + x) - \int_t^T f(s, Y_s^{(n)}, Z_s^{(n-1)})\mathrm{d}s - \int_t^T Z_s^{(n)}\mathrm{d}W_s. \quad (5)$$

Note that from equation (5) we have

$$Y_t^{(n)} = E_t\Big[\varphi(W_T + x) - \int_t^T f(s, Y_s^{(n)}, Z_s^{(n-1)})\mathrm{d}s\Big],$$

where $E_t(\cdot) = E(\cdot|\mathcal{F}_t)$ is the conditional expectation with respect to $\mathcal{F}_t$.

Then

$$|Y_t^{(n)}| \le E_t\Big[|\varphi(W_T + x)| + \int_t^T |f(s, Y_s^{(n)}, Z_s^{(n-1)})|\mathrm{d}s\Big].$$

Since the coefficients φ and f are bounded, we see that $Y_t^n, \forall n$ is bounded by a fixed constant depending on T and the bounds of the coefficients.

Next by inductive approach we will prove that for each n there exists a finite constant M_n depending on n such that $|Z_t^{(n)}|$ and $|\nabla Z_t^{(n)}|$ are uniformly bounded by M_n.

Suppose that $|Z_t^{(n-1)}|$ and $|\nabla Z_t^{(n-1)}|$ are uniformly bounded by M_{n-1}. Obviously, this is true for $n = 1$.

The corresponding variational equation of (5) is

$$Z_t^{(n)} = \frac{\partial}{\partial x}\varphi(W_T + x) - \int_t^T F(s)\mathrm{d}s - \int_t^T \nabla Z_s^{(n)}\mathrm{d}W_s, \tag{6}$$

where $F(s) = Z_s^{(n)}\frac{\partial}{\partial y}f(s, Y_s^{(n)}, Z_s^{(n-1)}) + \nabla Z_s^{(n-1)}\frac{\partial}{\partial z}f(s, Y_s^{(n)}, Z_s^{(n-1)})$.

For any nonnegative number A we consider the following BSDE

$$\phi^A(Z_t^{(n)}) = \frac{\partial}{\partial x}\varphi(W_T + x) - \int_t^T F^A(s)ds - \int_t^T \nabla Z_s^{(n)}dW_s, \tag{7}$$

where $F^A(s) = \phi^A(Z_t^{(n)})\frac{\partial}{\partial y}f(s, Y_s^{(n)}, \phi^A(Z_t^{(n-1)})) + \nabla Z_s^{(n-1)}\frac{\partial}{\partial z}f(s, Y_s^{(n)}, \phi^A(Z_t^{(n-1)}))$ and ϕ^A is a truncation function such that $\phi^A(x) = x$ for $|x| \le A$, and in the next we shall determine A.

Note that $\phi^A(Z_t^{(n)})$ also satisfies the following equation

$$\phi^A(Z_t^{(n)}) = E_t\Big[\frac{\partial}{\partial x}\varphi(W_T + x) - \int_t^T F^A(s)ds\Big].$$

Under the bounded condition of $\frac{\partial f}{\partial y}$, $\frac{\partial f}{\partial z}$, $\nabla Z_s^{(n-1)}$ and $\phi^A(Z_t^{(n)})$, we can derive the boundedness of $\phi^A(Z_t^{(n)})$. However, the bound here depends on A and it's not the final bound that we need. So let $h^A(t) = |\phi^A(Z_t^{(n)})|_\infty$ denotes the $L^\infty(\Omega)$-norm of $\phi^A(Z_t^{(n)})$. Then $h^A(t)$ satisfies the inequality

$$h^A(t) \le K_1 + K_2\int_t^T h^A(s)ds,$$

where K_1 depends on the bounds of $\frac{\partial \varphi}{\partial x}$, $\frac{\partial f}{\partial z}$, $\nabla Z_s^{(n-1)}$ and K_2 depends on the bound of $\frac{\partial f}{\partial y}$, thus $K_1 e^{K_2 T} < \infty$. By Gronwall's inequality, we know that there exists a finite constant M_n such that

$$h^A(t) \le M_n.$$

Note that M_n does not depend on A; therefore if we choose $A = M_n$ in (7), then $\phi^A(Z_t^{(n)})$ actually is $Z_t^{(n)}$, which means

$$|Z_t^{(n)}| \le h^A(t) \le M_n \quad \text{a.s.}$$

Similarly, consider the corresponding variational equation of (6)

$$\nabla Z_t^{(n)} = \frac{\partial^2}{\partial x^2}\varphi(W_T + x) - \int_t^T G(s)ds - \int_t^T \nabla^2 Z_s^{(n)}dW_s,$$

where

$$
\begin{aligned}
&G(s)\\
=&\frac{\partial^2}{\partial y^2}f(s,Y_s^{(n)},Z_s^{(n-1)})Z_s^{(n)}(Z_s^{(n)})^T+Z_s^{(n)}\Big(\frac{\partial^2}{\partial y\partial z}f(s,Y_s^{(n)},Z_s^{(n-1)})\Big)^T\nabla Z_s^{(n-1)}\\
&+\nabla Z_s^{(n-1)}\frac{\partial^2}{\partial y\partial z}f(s,Y_s^{(n)},Z_s^{(n-1)})(Z_s^{(n)})^T+\nabla Z_s^{(n-1)}\frac{\partial^2}{\partial z^2}f(s,Y_s^{(n)},Z_s^{(n-1)})\nabla Z_s^{(n-1)}\\
&+\frac{\partial}{\partial y}f(s,Y_s^{(n)},Z_s^{(n-1)})\nabla Z_s^{(n)}+\Big(\frac{\partial}{\partial z}f(s,Y_s^{(n)},Z_s^{(n-1)})\Big)^T\nabla^2 Z_s^{(n-1)}.
\end{aligned}
$$

After the parallel procedure as the demonstration of $Z_t^{(n)}$, we can obtain the boundedness of $\nabla Z_t^{(n)}$.

By induction on n, we get that for all $n\geq 0$, $Y_t^{(n)}$, $Z_t^{(n)}$ and $\nabla Z_t^{(n)}$ are bounded by a constant M_n a.s..

Since $Y_t^{(n)}$, $Z_t^{(n)}$ and $\nabla Z_t^{(n)}$ converge to Y_t, Z_t and ∇Z_t in $L^2(\Omega)$ respectively; see [6] for the proof of the L^2 convergence, there exist subsequences $Y_t^{(n_k)}$, $Z_t^{(n_k)}$ and $\nabla Z_t^{(n_k)}$ converge to Y_t, Z_t and ∇Z_t a.s., respectively.

For the fixed number $\varepsilon=1$, there exists a natural number N such that $|X_t-X_t^{(N)}|\leq 1$, where X stands for Y or Z or ∇Z. Hence by triangle inequality, we have

$$|X_t|\leq|X_t-X_t^{(N)}|+|X_t^{(N)}|\leq M_N+1<\infty.$$

Therefore the proof of this proposition is completed. $\square$

The proofs of the following proposition need that $f(t,x,y)$ grows at most as a linear function of x and y, i.e., there exists some positive constant K, such that

$$\forall(t,x,y),\quad |f(t,x,y)|\leq K(1+|x|+|y|),$$

which follows from assumption (ii).

Proposition 2. *Suppose that $Y(t+h)=\xi\in L^2(\Omega)$ is a $\mathcal{F}_{t+h}$-measurable random variable. Then the solution of the equation (3) satisfies the following inequality ($h<1$):*

$$E|Y(t)|^2+E\int_t^{t+h}|Z(s)|^2\,\mathrm{d}s\leq K(1+E|\xi|^2),\quad 0\leq t\leq T. \tag{8}$$

Proof. By the Itô formula, we have

$$
\begin{aligned}
E|Y(t)|^2+E\int_t^{t+h}|Z(s)|^2\,\mathrm{d}s&=E|\xi|^2-2E\int_t^{t+h}Y(s)f(s,Y(s),Z(s))\,\mathrm{d}s\\
&\leq E|\xi|^2+2E\int_t^{t+h}|Y(s)||f(s,Y(s),Z(s))|\,\mathrm{d}s\\
&\leq E|\xi|^2+2KE\int_t^{t+h}|Y(s)|[1+|Y(s)|+|Z(s)|]\,\mathrm{d}s\\
&\leq E|\xi|^2+h+(2K+3K^2)E\int_t^{t+h}|Y(s)|^2\,\mathrm{d}s+\frac{1}{2}E\int_t^{t+h}|Z(s)|^2\,\mathrm{d}s,
\end{aligned}
$$

where the last inequality follows from the fact $ab\leq\frac{1}{2}(a^2+b^2)$. Immediately, this leads to the following inequality

$$E|Y(t)|^2+\frac{1}{2}E\int_t^{t+h}|Z(s)|^2\,\mathrm{d}s\leq E|\xi|^2+1+KE\int_t^{t+h}|Y(s)|^2\,\mathrm{d}s.$$

Then (8) is nothing other than Gronwall's lemma. This completes the proof. □

Proposition 3. *Under the same assumption of Proposition 2, we have*

$$E|Y_{t+h,\xi}(t)-\xi|^2 \le K(1+E|\xi|^2)h.$$

Proof. From the equation (3), we have

$$\begin{aligned} E|Y_{t+h,\xi}(t)-\xi|^2 &= E|\int_t^{t+h} f(s,Y(s),Z(s))\mathrm{d}s + \int_t^{t+h} Z(s)\mathrm{d}W(s)|^2 \\ &\le 2hE\int_t^{t+h} f(s,Y(s),Z(s))^2\mathrm{d}s + 2E\int_t^{t+h}|Z(s)|^2\mathrm{d}s \\ &\le KhE\int_t^{t+h}(1+|Y(s)|^2+|Z(s)|^2)\mathrm{d}s + 2E\int_t^{t+h}|Z(s)|^2\mathrm{d}s \\ &\le Kh^2 + KhE\int_t^{t+h}|Y(s)|^2\mathrm{d}s + KE\int_t^{t+h}|Z(s)|^2\mathrm{d}s \\ &\le Kh(1+E|\xi|^2), \end{aligned}$$

where the last inequality follows from the boundedness of $|Z(s)|$ and Proposition 2. Therefore the proof is completed. □

Proposition 4. *Under the same assumption as Proposition 2, we have the following estimation*

$$|E(Y_{t+h,\xi}(t))(Y_{t+h,\xi}(t)-\xi)| \le K(1+E|\xi|^2)h.$$

Proof. From the equation (3), we have

$$\begin{aligned} &|E(Y_{t+h,\xi}(t))(Y_{t+h,\xi}(t)-\xi)| \\ &\quad\le E\big(|Y_{t+h,\xi}(t)||\int_t^{t+h} f(s,Y_{t+h,\xi}(s),Z_{t+h,\xi}(s))\,\mathrm{d}s|\big) \\ &\quad\le (E|Y_{t+h,\xi}(t)|^2)^{1/2}(E|\int_t^{t+h} f(s,Y_{t+h,\xi}(s),Z_{t+h,\xi}(s))\,\mathrm{d}s|^2)^{1/2} \\ &\quad\le K(1+E|\xi|^2)h. \end{aligned}$$

This completes the proof. □

The similar properties of the solution of the variational equation (4) are also needed for the analysis in Section 3.

By assumption (ii), we have

$$\begin{aligned} |F(t,Y_t,Z_t,\nabla Y_t,\nabla Z_t)| &= |\nabla Y_t \frac{\partial}{\partial y} f(t,Y_t,Z_t) + \nabla Z_t \frac{\partial}{\partial z} f(t,Y_t,Z_t)| \\ &\le K(|\nabla Y_t| + |\nabla Z_t|), \end{aligned}$$

where K is the upper bounds of $\frac{\partial f}{\partial y}$ and $\frac{\partial f}{\partial z}$, that is to say $F(t,Y_t,Z_t,\nabla Y_t,\nabla Z_t)$ grows at most as a linear function of ∇Y_t and ∇Z_t. Therefore the estimations of $\nabla Y(t)$ and $\nabla Z(t)$ are similar to those of $Y(t)$ and $Z(t)$, here we just give the conclusions without proofs.

Proposition 5. *Suppose that $(\nabla Y(t),\nabla Z(t))$ is the solution of the variational equation (4) with $\nabla Y(t+h)=\eta$, then we have $(h<1)$*

$$E|\nabla Y(t)|^2 + E\int_t^{t+h}|\nabla Z(s)|^2\mathrm{d}s \le K(1+E|\eta|^2).$$

Proposition 6. *Under the same assumption of Proposition 5, we have*

$$E|\nabla Y_{t+h,\eta}(t)-\eta|^2 \le K(1+E|\eta|^2)h,$$
$$|E(\nabla Y_{t+h,\eta}(t))^T(\nabla Y_{t+h,\eta}(t)-\eta)| \le K(1+E|\eta|^2)h.$$

3. Fundamental convergence theorem.

Now we are in the position of the statement and proof of the main result in this paper.

Firstly, we introduce the uniform partition $0=t_0<\cdots<t_N=T$, and let $h=t_{k+1}-t_k, k=0,\cdots,N-1$ for simplicity. Denote the approximation of the solution $(Y(t_k),Z(t_k),\mathcal{F}_{t_k})$ for BSDE (3) at time t_k by $(\bar{Y}_k,\bar{Z}_k,\mathcal{F}_{t_k})$, which means that the numerical approximation $(\bar{Y}_k,\bar{Z}_k)$ is also $\mathcal{F}_{t_k}$-measurable. Define $(\bar{Y}_k,\bar{Z}_k)$ recurrently by

$$\begin{aligned}\bar{Y}_k &= A(\bar{Y}_{k+1},\bar{Z}_{k+1},\bar{Y}_k,\bar{Z}_k,h),\\ \bar{Z}_k &= B(\bar{Y}_{k+1},\bar{Z}_{k+1},\nabla\bar{Y}_{k+1},\bar{Y}_k,\bar{Z}_k,\nabla\bar{Y}_k,h,W(t_{k+1})-W(t_k)),\end{aligned} \tag{9}$$

for some functions A and B such that $\bar{Y}_k$ and $\bar{Z}_k$ is $\mathcal{F}_{t_k}$-measurable. Here the terminal value $\bar{Y}_N$ and $\bar{Z}_N$ are two given random variables satisfying $(E|\bar{Y}_N-Y(T)|^2)^{1/2}\le Kh^{p_1}$ and $(E|\bar{Z}_N-Z(T)|^2)^{1/2}\le Kh^{\min\{p_1,p_2\}}$, respectively, where p_1 and p_2 are positive numbers.

Remark 3. For the case that the terminal condition is $Y(T)=\varphi(W_T+x)$, one may choose $\bar{Y}_N=Y(T)=\varphi(W_T+x)$ and $\bar{Z}_N=\nabla Y(T)=\frac{\partial}{\partial x}\varphi(W_T+x)$.

Taking the generalized θ-method proposed in [10] for example,

$$\begin{aligned}\bar{Y}_k &= E^x_{t_k}[\bar{Y}_{k+1}]+\theta_1 hf(t_k,\bar{Y}_k,\bar{Z}_k)+(1-\theta_1)hE^x_{t_k}[f(t_{k+1},\bar{Y}_{k+1},\bar{Z}_{k+1})],\\ \theta_3 h\bar{Z}_k &= \theta_4 hE^x_{t_k}[\bar{Z}_{k+1}]+(\theta_3-\theta_4)E^x_{t_k}[\bar{Y}_{k+1}\Delta W^T_{t_{k+1}}]\\ &\quad +(1-\theta_2)hE^x_{t_k}[f(t_{k+1},\bar{Y}_{k+1},\bar{Z}_{k+1})\Delta W^T_{t_{k+1}}],\end{aligned}$$

where $\theta_i\in[0,1], i=1,2$, $\theta_3\in(0,1]$ and $\theta_4\in[-1,1]$ constrained by $|\theta_4|\le\theta_3$.

Taking the variation on both sides of the scheme (9), we obtain $(\nabla\bar{Y}_k,\nabla\bar{Z}_k)$, which is the approximation of the solution $(\nabla Y(t_k),\nabla Z(t_k))$ of the variational equation (4). Also in sequels we make the following notations

$$\begin{aligned}Y_k &= Y(t_k)=Y_{t_{k+1},Y(t_{k+1})}(t_k)=Y_{t_N,Y_T}(t_k),\\ \bar{Y}_k &= \bar{Y}_{t_{k+1},\bar{Y}_{k+1}}(t_k)=\bar{Y}_{t_N,\bar{Y}_N}(t_k),\\ \nabla Y_k &= \nabla Y(t_k)=\nabla Y_{t_{k+1},\nabla Y(t_{k+1})}(t_k)=Z_{t_{k+1},Y(t_{k+1})}(t_k)=Z(t_k),\\ \nabla\bar{Y}_k &= \nabla\bar{Y}_{t_{k+1},\nabla\bar{Y}(t_{k+1})}(t_k),\\ \bar{Z}_k &= \bar{Z}_{t_{k+1},\bar{Y}_{k+1}}(t_k).\end{aligned}$$

Next, we introduce a definition "closeness under variation". Since the numerical methods appearing in [2, 8, 9, 10, 11] are all closed under variation, we study the fundamental convergence theorem for this class of numerical methods mainly in this paper, while presenting the result for the rest of numerical methods without proof.

Definition 3.1. If a numerical method is closed under variation, then the numerical result obtained from applying the method to the variational equation (4) is equal to the variation of the numerical solution for equation (3), i.e. the following diagram commutes:

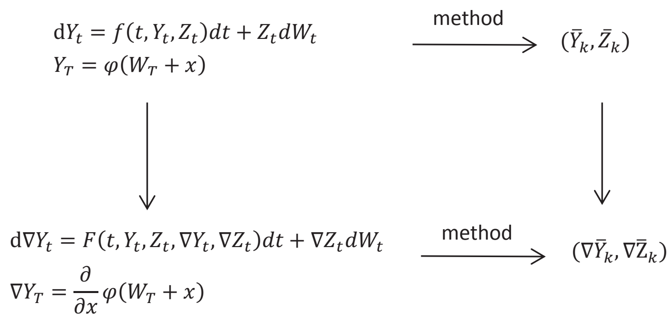

FIGURE 1. *commutative diagram*

3.1. **Statement of the fundamental convergence theorem.** For the numerical approximations of the BSDE (3) which are closed under variation, the theorem on the mean-square order is stated below.

Theorem 3.2. *Suppose that the numerical method (9) for the BSDE (3) is closed under variation and consistent with order p_1+1 for the first process Y and order p_2 for the second one Z in the mean-square sense; more precisely, for arbitrary $0 \le t \le T-h$, the following inequalities hold*

$$(E|Y_{t+h,\xi}(t) - \bar{Y}_{t+h,\xi}(t)|^2)^{1/2} \le KC_1 h^{p_1+1}, \tag{10}$$

$$(E|Z_{t+h,\xi}(t) - \bar{Z}_{t+h,\xi}(t)|^2)^{1/2} \le KC_2 h^{p_2}, \tag{11}$$

where $C_1 = (1+E|\xi|^2)^{1/2}$ and $C_2 = [(1+E|\xi|^2) + (1+E|\nabla\xi|^2)]^{1/2}$ and $\xi \in L^2(\Omega)$ is a $\mathcal{F}_{t+h}$-measurable random variable.

Then for arbitrary N and $k = 0, 1, \cdots, N$, the following inequalities hold

$$[E|Y(t_k) - \bar{Y}_k|^2]^{1/2} \le K\bar{C}_1 h^{p_1}, \tag{12}$$

$$[E|Z(t_k) - \bar{Z}_k|^2]^{1/2} \le K\bar{C}_2 h^{min\{p_1,p_2\}}, \tag{13}$$

where $\bar{C}_1 = (1+E|Y(T)|^2)^{1/2}$ and $\bar{C}_2 = [(1+E|Y(T)|^2) + (1+E|\nabla Y(T)|^2)]^{1/2}$.

Remark 4. The mean-square orders of convergence for the first and second processes are p_1 and $\min\{p_1, p_2\}$, respectively, which will be denoted by $(p_1, \min\{p_1, p_2\})$ in the following.

Remark 5. The conditions (10) and (11) above are parallel to these in Milstein's convergence theorem [5]. Some numerical approximations as examples are presented in Section 4 to verify Theorem 3.2.

Throughout the paper, ξ is the value of Y at time $t+h$, which may be $Y(t+h)$ or the approximation $\bar{Y}_{t+h}$. $\nabla\xi$ is the variation of ξ with respect to x and represents the value of ∇Y at time $t+h$. η and $\nabla\eta$ have the same meaning.

Since the numerical method is closed under variation, after applying the same numerical method to the variational equation (4), we get $(\nabla\bar{Y}_k, \nabla\bar{Z}_k)$. The order of local error of $\nabla\bar{Y}_k$ is the same as that of $\bar{Y}_k$, i.e.,

$$(E|\nabla Y_{t+h,\nabla\xi}(t) - \nabla\bar{Y}_{t+h,\nabla\xi}(t)|^2)^{1/2} \le K(1+E|\nabla\xi|^2)^{1/2} h^{p_1+1}. \tag{14}$$

In another case of the numerical method is not closed under variation, we assume that the order of the local error of $\nabla\bar{Y}_k$ is p_3+1, i.e.,

$$(E|\nabla Y_{t+h,\nabla\xi}(t)-\nabla\bar{Y}_{t+h,\nabla\xi}(t)|^2)^{1/2}\leq K(1+E|\nabla\xi|^2)^{1/2}h^{p_3+1}.$$

Similarly, we have

$$\begin{aligned}&[E|Y(t_k)-\bar{Y}_k|^2]^{1/2}\leq K\bar{C}_1h^{p_1},\\&[E|\nabla Y(t_k)-\nabla\bar{Y}_k|^2]^{1/2}\leq K\bar{C}_2h^{min\{p_1,p_3\}},\\&[E|Z(t_k)-\bar{Z}_k|^2]^{1/2}\leq K\bar{C}_2h^{min\{p_1,p_2,p_3\}}.\end{aligned}$$

which is presented here without proof, since its proof is similar to that of Theorem 3.2.

3.2. **Lemmas.** To prove the above fundamental convergence theorem, we need some lemmas firstly.

From the assumption (ii), we know that $f:\Omega\times[0,T]\times R\times R^d\longmapsto R$ is Lipschitz with respect to x and y for some positive constant K, i.e., $\forall(x_1,y_1),\,(x_2,y_2)\in R\times R^d$,

$$|f(t,x_1,y_1)-f(t,x_2,y_2)|\leq K(|x_1-x_2|+|y_1-y_2|).$$

This property is used in the proof of the following lemma.

Lemma 3.3. *Suppose that $\xi,\ \eta\in L^2(\Omega)$ are the potential values of Y at time $t+h$, which are $\mathcal{F}_{t+h}$-measurable. Then we have*

$$E|Y_{t+h,\xi}(t)-Y_{t+h,\eta}(t)|^2\leq E|\xi-\eta|^2(1+Kh).$$

Proof. As in Proposition 2, following from Itô formula, we obtain

$$\begin{aligned}&E|Y_{t+h,\xi}(t)-Y_{t+h,\eta}(t)|^2+\frac{1}{2}E\int_t^{t+h}|Z_{t+h,\xi}(s)-Z_{t+h,\eta}(s)|^2\,\mathrm{d}s\\&\leq E|\xi-\eta|^2+KE\int_t^{t+h}|Y_{t+h,\xi}(s)-Y_{t+h,\eta}(s)|^2\,\mathrm{d}s.\end{aligned}$$

By Gronwall's lemma, one gets

$$E|Y_{t+h,\xi}(t)-Y_{t+h,\eta}(t)|^2\leq E|\xi-\eta|^2(1+Kh).$$

Therefore the proof of this lemma is completed. □

Lemma 3.4. *For $k=0,1,\cdots,N$, the following inequality holds*

$$E|\bar{Y}_k|^2\leq K(1+E|Y(T)|^2)=K\bar{C}_1.$$

Proof. First of all we need to prove the existence of $E|\bar{Y}_k|^2,\ k=0,\cdots,N$. Suppose that $E|\bar{Y}_{k+1}|^2<\infty$. Then, using the condition (10) in the statement of Theorem 3.2

$$E|Y_{t+h,\bar{Y}_{k+1}}(t)-\bar{Y}_{t+h,\bar{Y}_{k+1}}(t)|^2\leq K(1+E|\bar{Y}_{k+1}|^2)h^{2(p_1+1)},$$

and the conclusion of Proposition 2, the boundedness of $E|\bar{Y}_k|^2$ is obvious. Since $E|\bar{Y}_N|^2=E|Y(T)|^2<\infty$, we have proved the existence of all $E|\bar{Y}_k|^2<\infty,\ k=0,\cdots,N$.

Consider the equation

$$\begin{aligned}E|\bar{Y}_k|^2 = {}& E|\bar{Y}_{k+1}|^2 + E|\bar{Y}_{t_{k+1},\bar{Y}_{k+1}}(t_k) - Y_{t_{k+1},\bar{Y}_{k+1}}(t_k)|^2 \\ &+ E|Y_{t_{k+1},\bar{Y}_{k+1}}(t_k) - \bar{Y}_{k+1}|^2 + 2E\bar{Y}_{k+1}(Y_{t_{k+1},\bar{Y}_{k+1}}(t_k) - \bar{Y}_{k+1}) \\ &+ 2E\bar{Y}_{k+1}(\bar{Y}_{t_{k+1},\bar{Y}_{k+1}}(t_k) - Y_{t_{k+1},\bar{Y}_{k+1}}(t_k)) \\ &+ 2E(\bar{Y}_{t_{k+1},\bar{Y}_{k+1}}(t_k) - Y_{t_{k+1},\bar{Y}_{k+1}}(t_k))(Y_{t_{k+1},\bar{Y}_{k+1}}(t_k) - \bar{Y}_{k+1}),\end{aligned}$$

from Proposition 3, we have

$$E|Y_{t_{k+1},\bar{Y}_{k+1}}(t_k) - \bar{Y}_{k+1}|^2 \leq K(1 + E|\bar{Y}_{k+1}|^2)h.$$

Further, we obtain

$$\begin{aligned}&2|E(\bar{Y}_{t_{k+1},\bar{Y}_{k+1}}(t_k) - Y_{t_{k+1},\bar{Y}_{k+1}}(t_k))(Y_{t_{k+1},\bar{Y}_{k+1}}(t_k) - \bar{Y}_{k+1})| \\ &\quad\leq 2(E|\bar{Y}_{t_{k+1},\bar{Y}_{k+1}}(t_k) - Y_{t_{k+1},\bar{Y}_{k+1}}(t_k)|^2)^{1/2}(E|Y_{t_{k+1},\bar{Y}_{k+1}}(t_k) - \bar{Y}_{k+1}|^2)^{1/2} \\ &\quad\leq K(1 + E|\bar{Y}_{k+1}|^2)h^{p_1+\frac{3}{2}}.\end{aligned}$$

Also from (10), we have

$$\begin{aligned}&2E\bar{Y}_{k+1}^T(\bar{Y}_{t_{k+1},\bar{Y}_{k+1}}(t_k) - Y_{t_{k+1},\bar{Y}_{k+1}}(t_k)) \\ &\quad\leq K(E|\bar{Y}_{k+1}|^2)^{1/2}(E|\bar{Y}_{t_{k+1},\bar{Y}_{k+1}}(t_k) - Y_{t_{k+1},\bar{Y}_{k+1}}(t_k)|^2)^{1/2} \\ &\quad\leq K(1 + E|\bar{Y}_{k+1}|^2)h^{p_1+1},\end{aligned}$$

and following from Proposition 3 and Proposition 4, we have

$$\begin{aligned}&2E\bar{Y}_{k+1}^T(Y_{t_{k+1},\bar{Y}_{k+1}}(t_k) - \bar{Y}_{k+1}) \\ &\quad= -2E|Y_{t_{k+1},\bar{Y}_{k+1}}(t_k) - \bar{Y}_{k+1}|^2 + 2EY_{t_{k+1},\bar{Y}_{k+1}}(t_k)^T(Y_{t_{k+1},\bar{Y}_{k+1}}(t_k) - \bar{Y}_{k+1}) \\ &\quad\leq K(1 + E|\bar{Y}_{k+1}|^2)h.\end{aligned}$$

Therefore we arrive at the inequality

$$E|\bar{Y}_k|^2 \leq E|\bar{Y}_{k+1}|^2 + K(1 + E|\bar{Y}_{k+1}|^2)h \leq (1 + Kh)E|\bar{Y}_{k+1}|^2 + Kh.$$

Hence, we obtain

$$E|\bar{Y}_k|^2 \leq K(1 + E|Y(T)|^2).$$

This completes the proof. □

From the boundedness of Z_t and ∇Z_t, we know

$$\begin{aligned}&|F(t, Y_t^{(1)}, Z_t^{(1)}, \nabla Y_t^{(1)}, \nabla Z_t^{(1)}) - F(t, Y_t^{(2)}, Z_t^{(2)}, \nabla Y_t^{(2)}, \nabla Z_t^{(2)})| \\ &\quad\leq |\nabla Y_t^{(1)} \frac{\partial}{\partial y} f(t, Y_t^{(1)}, Z_t^{(1)}) - \nabla Y_t^{(2)} \frac{\partial}{\partial y} f(t, Y_t^{(1)}, Z_t^{(1)})| \\ &\qquad+ |\nabla Y_t^{(2)} \frac{\partial}{\partial y} f(t, Y_t^{(1)}, Z_t^{(1)}) - \nabla Y_t^{(2)} \frac{\partial}{\partial y} f(t, Y_t^{(2)}, Z_t^{(2)})| \\ &\qquad+ |\nabla Z_t^{(1)} \frac{\partial}{\partial z} f(t, Y_t^{(1)}, Z_t^{(1)}) - \nabla Z_t^{(2)} \frac{\partial}{\partial z} f(t, Y_t^{(1)}, Z_t^{(1)})| \\ &\qquad+ |\nabla Z_t^{(2)} \frac{\partial}{\partial z} f(t, Y_t^{(1)}, Z_t^{(1)}) - \nabla Z_t^{(2)} \frac{\partial}{\partial z} f(t, Y_t^{(2)}, Z_t^{(2)})| \\ &\quad\leq K(|Y_t^{(1)} - Y_t^{(2)}| + |Z_t^{(1)} - Z_t^{(2)}| + |\nabla Y_t^{(1)} - \nabla Y_t^{(2)}| + |\nabla Z_t^{(1)} - \nabla Z_t^{(2)}|)\end{aligned}$$

that is to say $F(t, Y_t, Z_t, \nabla Y_t, \nabla Z_t)$ is Lipschitz with respect to Y_t, Z_t, ∇Y_t and ∇Z_t. So for the variational equation (4) and the variation of numerical method $\nabla \bar{Y}_k$, we have the similar estimates.

Lemma 3.5. *For the variational equation we have*

$$E|\nabla Y_{t+h,\nabla\xi}(t)-\nabla Y_{t+h,\nabla\eta}(t)|^2 \le E|\nabla\xi-\nabla\eta|^2(1+Kh)+KE|\xi-\eta|^2h.$$

Proof. Since

$$\begin{aligned}\nabla Y_{t+h,\nabla\xi}(t)-\nabla Y_{t+h,\nabla\eta}(t) &= \nabla\xi-\nabla\eta \\ &\quad -\int_t^{t+h}[F(s,Y_{t+h,\xi}(s),Z_{t+h,\xi}(s),\nabla Y_{t+h,\nabla\xi}(s),\nabla Z_{t+h,\nabla\xi}(s)) \\ &\quad - F(s,Y_{t+h,\eta}(s),Z_{t+h,\eta}(s),\nabla Y_{t+h,\nabla\eta}(s),\nabla Z_{t+h,\nabla\eta}(s))]\mathrm{d}s \\ &\quad +\int_t^{t+h}[\nabla Z_{t+h,\nabla\xi}(s)-\nabla Z_{t+h,\nabla\eta}(s)]\mathrm{d}W_s\end{aligned}$$

and Itô formula, we have

$$\begin{aligned}&E|\nabla Y_{t+h,\nabla\xi}(t)-\nabla Y_{t+h,\nabla\eta}(t)|^2+E\int_t^{t+h}|\nabla Z_{t+h,\nabla\xi}(s)-\nabla Z_{t+h,\nabla\eta}(s)|^2\mathrm{d}s \\ &= E|\nabla\xi-\nabla\eta|^2-2E\int_t^{t+h}(\nabla Y_{t+h,\nabla\xi}(s)-\nabla Y_{t+h,\nabla\eta}(s))^T(F_1-F_2)\mathrm{d}s \\ &\le E|\nabla\xi-\nabla\eta|^2 \\ &\quad +2KE\int_t^{t+h}|\nabla Y_{t+h,\nabla\xi}(s)-\nabla Y_{t+h,\nabla\eta}(s)|(|Y_{t+h,\xi}(s)-Y_{t+h,\eta}(s)| \\ &\quad +2|\nabla Y_{t+h,\nabla\xi}(s)-\nabla Y_{t+h,\nabla\eta}(s)|+|\nabla Z_{t+h,\nabla\xi}(s)-\nabla Z_{t+h,\nabla\eta}(s)|)\mathrm{d}s \\ &\le E|\nabla\xi-\nabla\eta|^2 \\ &\quad +E\int_t^{t+h}(|Y_{t+h,\xi}(s)-Y_{t+h,\eta}(s)|^2+K|\nabla Y_{t+h,\nabla\xi}(s)-\nabla Y_{t+h,\nabla\eta}(s)|^2 \\ &\quad +\frac{1}{2}|\nabla Z_{t+h,\nabla\xi}(s)-\nabla Z_{t+h,\nabla\eta}(s)|^2)\mathrm{d}s.\end{aligned}$$

Therefore,

$$\begin{aligned}&E|\nabla Y_{t+h,\nabla\xi}(t)-\nabla Y_{t+h,\nabla\eta}(t)|^2+\frac{1}{2}E\int_t^{t+h}|\nabla Z_{t+h,\nabla\xi}(s)-\nabla Z_{t+h,\nabla\eta}(s)|^2\mathrm{d}s \\ &\le E|\nabla\xi-\nabla\eta|^2+E\int_t^{t+h}|Y_{t+h,\xi}(s)-Y_{t+h,\eta}(s)|^2\mathrm{d}s \\ &\quad +KE\int_t^{t+h}|\nabla Y_{t+h,\nabla\xi}(s)-\nabla Y_{t+h,\nabla\eta}(s)|^2\mathrm{d}s \\ &\le E|\nabla\xi-\nabla\eta|^2+KE|\xi-\eta|^2h \\ &\quad +KE\int_t^{t+h}|\nabla Y_{t+h,\nabla\xi}(s)-\nabla Y_{t+h,\nabla\eta}(s)|^2\mathrm{d}s.\end{aligned}$$

Obviously, by the Gronwall's lemma, one gets

$$E|\nabla Y_{t+h,\nabla\xi}(t)-\nabla Y_{t+h,\nabla\eta}(t)|^2 \le E|\nabla\xi-\nabla\eta|^2(1+Kh)+KE|\xi-\eta|^2h.$$

The conclusion of this lemma is proved. □

Lemma 3.6. *For $k=0,1,\cdots,N$, the following inequality holds*

$$E|\nabla\bar{Y}_k|^2 \le K(1+E|\nabla Y(T)|^2).$$

Proof. The proof is the same as that of Lemma 3.4. □

3.3. **The proof of the fundamental convergence theorem.**

$$\begin{aligned} Y(t_k) - \bar{Y}_k &= Y_{t_{k+1},Y_{k+1}}(t_k) - \bar{Y}_{t_{k+1},\bar{Y}_{k+1}}(t_k) \\ &= [Y_{t_{k+1},Y_{k+1}}(t_k) - Y_{t_{k+1},\bar{Y}_{k+1}}(t_k)] \\ &\quad + [Y_{t_{k+1},\bar{Y}_{k+1}}(t_k) - \bar{Y}_{t_{k+1},\bar{Y}_{k+1}}(t_k)] \end{aligned}$$

Taking the mean-square of both sides of the above equation, we obtain

$$\begin{aligned} E|Y(t_k) - \bar{Y}(t_k)|^2 &= E|Y_{t_{k+1},Y_{k+1}}(t_k) - Y_{t_{k+1},\bar{Y}_{k+1}}(t_k)|^2 \\ &\quad + E|Y_{t_{k+1},\bar{Y}_{k+1}}(t_k) - \bar{Y}_{t_{k+1},\bar{Y}_{k+1}}(t_k)|^2 \\ &\quad + 2E[(Y_{t_{k+1},Y_{k+1}}(t_k) - Y_{t_{k+1},\bar{Y}_{k+1}}(t_k)) \\ &\qquad (Y_{t_{k+1},\bar{Y}_{k+1}}(t_k) - \bar{Y}_{t_{k+1},\bar{Y}_{k+1}}(t_k))]. \end{aligned} \tag{15}$$

By the conclusion of Lemma 3.3 we have

$$E|Y_{t_{k+1},Y_{k+1}}(t_k) - Y_{t_{k+1},\bar{Y}_{k+1}}(t_k)|^2 \leq E|Y_{k+1} - \bar{Y}_{k+1}|^2(1 + Kh), \tag{16}$$

and by (10) and Lemma 3.4, we get

$$\begin{aligned} E|Y_{t_{k+1},\bar{Y}_{k+1}}(t_k) - \bar{Y}_{t_{k+1},\bar{Y}_{k+1}}(t_k)|^2 &\leq K(1 + E|\bar{Y}_{k+1}|^2)h^{2p_1+2} \\ &\leq K(1 + E|Y(T)|^2)h^{2p_1+2}. \end{aligned} \tag{17}$$

So the last summand in (15) is

$$\begin{aligned} &2E[(Y_{t_{k+1},Y_{k+1}}(t_k) - Y_{t_{k+1},\bar{Y}_{k+1}}(t_k))(Y_{t_{k+1},\bar{Y}_{k+1}}(t_k) - \bar{Y}_{t_{k+1},\bar{Y}_{k+1}}(t_k))] \\ &\quad \leq K(E|Y_{k+1} - \bar{Y}_{k+1}|^2)^{1/2}(1 + E|Y(T)|^2)^{1/2}h^{p_1+1} \\ &\quad \leq KhE|Y_{k+1} - \bar{Y}_{k+1}|^2 + K(1 + E|Y(T)|^2)h^{2p_1+1}, \end{aligned} \tag{18}$$

where the last inequality follows from the relationship $ab \leq \frac{1}{2}(a^2 + b^2)$.

Let $(\varepsilon_Y^k)^2 := E|Y_k - \bar{Y}_k|^2$. The relations (15)-(18) lead to the inequality $(h < 1)$

$$\begin{aligned} (\varepsilon_Y^k)^2 &\leq (\varepsilon_Y^{k+1})^2(1 + Kh) + K(1 + E|Y(T)|^2)h^{2p_1+2} \\ &\quad + Kh(\varepsilon_Y^{k+1})^2 + K(1 + E|Y(T)|^2)h^{2p_1+1} \\ &\leq (\varepsilon_Y^{k+1})^2(1 + Kh) + K(1 + E|Y(T)|^2)h^{2p_1+1} \end{aligned}$$

Taking account of the fact that $\varepsilon_Y^N \leq Kh^{p_1}$, we have

$$\varepsilon_Y^k \leq K(1 + E|Y(T)|^2)^{1/2}h^{p_1} = K\bar{C}_1 h^{p_1}.$$

We can compute the global error of $\nabla\bar{Y}_k$ in parallel.

$$\begin{aligned} (\varepsilon_{\nabla Y}^k)^2 :&= E|\nabla Y(t_k) - \nabla\bar{Y}_k|^2 \\ &= E|\nabla Y_{t_{k+1},\nabla Y_{k+1}}(t_k) - \nabla Y_{t_{k+1},\nabla\bar{Y}_{k+1}}(t_k)|^2 \\ &\quad + E|\nabla Y_{t_{k+1},\nabla\bar{Y}_{k+1}}(t_k) - \nabla\bar{Y}_{t_{k+1},\nabla\bar{Y}_{k+1}}(t_k)|^2 \\ &\quad + 2E[(\nabla Y_{t_{k+1},\nabla Y_{k+1}}(t_k) - \nabla Y_{t_{k+1},\nabla\bar{Y}_{k+1}}(t_k))^T \\ &\qquad (\nabla Y_{t_{k+1},\nabla\bar{Y}_{k+1}}(t_k) - \nabla\bar{Y}_{t_{k+1},\nabla\bar{Y}_{k+1}}(t_k))]. \end{aligned} \tag{19}$$

By the conclusion of Lemma 3.5, we have

$$\begin{aligned} &E|\nabla Y_{t_{k+1},\nabla Y_{k+1}}(t_k) - \nabla Y_{t_{k+1},\nabla\bar{Y}_{k+1}}(t_k)|^2 \\ &\quad \leq E|\nabla Y_{k+1} - \nabla\bar{Y}_{k+1}|^2(1 + Kh) + KE|Y_{k+1} - \bar{Y}_{k+1}|^2 h \\ &\quad \leq E|\nabla Y_{k+1} - \nabla\bar{Y}_{k+1}|^2(1 + Kh) + K(1 + E|Y(T)|^2)h^{2p_1+1}, \end{aligned}$$

and by (14) and Lemma 3.6, we get

$$E|\nabla Y_{t_{k+1},\nabla\bar{Y}_{k+1}}(t_k)-\nabla\bar{Y}_{t_{k+1},\nabla\bar{Y}_{k+1}}(t_k)|^2 \le K(1+E|\nabla\bar{Y}_{k+1}|^2)h^{2p_1+2}$$
$$\le K(1+E|\nabla Y(T)|^2)h^{2p_1+2}.$$

So the last summand in (19) is

$$\begin{aligned}&2E(\nabla Y_{t_{k+1},\nabla Y_{k+1}}(t_k)-\nabla Y_{t_{k+1},\nabla\bar{Y}_{k+1}}(t_k))^T\\&\quad(\nabla Y_{t_{k+1},\nabla\bar{Y}_{k+1}}(t_k)-\nabla\bar{Y}_{t_{k+1},\nabla\bar{Y}_{k+1}}(t_k))\\&\le K(E|\nabla Y_{k+1}-\nabla\bar{Y}_{k+1}|^2(1+Kh)\\&\quad+K(1+E|Y(T)|^2)h^{2p_1+1})^{1/2}(1+E|\nabla Y(T)|^2)^{1/2}h^{p_1+1}\\&\le KE|\nabla Y_{k+1}-\nabla\bar{Y}_{k+1}|^2h\\&\quad+K(1+E|Y(T)|^2)h^{2p_1+2}+K(1+E|\nabla Y(T)|^2)h^{2p_1+1}.\end{aligned}$$

Therefore we obtain

$$\varepsilon^k_{\nabla Y}\le K\bar{C}_2h^{p_1}.$$

At last,

$$\begin{aligned}(\varepsilon^k_Z)^2:&=E|Z(t_k)-\bar{Z}_k|^2=E|Z_{t_{k+1},Y_{k+1}}(t_k)-\bar{Z}_{t_{k+1},\bar{Y}_{k+1}}(t_k)|^2\\&\le KE|Z_{t_{k+1},Y_{k+1}}(t_k)-Z_{t_{k+1},\bar{Y}_{k+1}}(t_k)|^2\\&\quad+KE|Z_{t_{k+1},\bar{Y}_{k+1}}(t_k)-\bar{Z}_{t_{k+1},\bar{Y}_{k+1}}(t_k)|^2\\&=KE|\nabla Y_{t_{k+1},\nabla Y_{k+1}}(t_k)-\nabla Y_{t_{k+1},\nabla\bar{Y}_{k+1}}(t_k)|^2\\&\quad+KE|Z_{t_{k+1},\bar{Y}_{k+1}}(t_k)-\bar{Z}_{t_{k+1},\bar{Y}_{k+1}}(t_k)|^2\\&\le KE|\nabla Y_{k+1}-\nabla\bar{Y}_{k+1}|^2(1+Kh)+KE|Y_{k+1}-\bar{Y}_{k+1}|^2h+K\bar{C}^2h^{2p_2}\\&\le K\bar{C}_2^2h^{2p_1}+K(1+E|Y(T)|^2)h^{2p_1+1}+K\bar{C}^2h^{2p_2}\\&\le K\bar{C}_2^2h^{\min\{2p_1,2p_2\}}.\end{aligned}$$

So the proof of Theorem 3.2 is completed.

4. **Examples.**

For equation

$$Y_t=\varphi(W_T+x)+\int_t^T f(s,Y_s)\mathrm{d}s-\int_t^T Z_s\mathrm{d}W_s, \tag{20}$$

consider the explicit Euler method of mean-square order $(1,1)$

$$\begin{aligned}\bar{Y}_k&=E_{t_k}(\bar{Y}_{k+1})+hf(t_{k+1},\bar{Y}_{k+1}),\bar{Y}_N=\varphi(W_T),\\h\bar{Z}_k&=E_{t_k}(\bar{Y}_{k+1}\Delta W_{k+1}).\end{aligned}$$

First we check the closeness under variation of the method. The variational equation of (20) is

$$\nabla Y_t=\frac{\partial}{\partial x}\varphi(W_T+x)+\int_t^T\nabla Y_s\frac{\partial}{\partial y}f(s,Y_s)ds-\int_t^T\nabla Z_sdW_s.$$

Applying the method to it, we get

$$\begin{aligned}\overline{\nabla Y}_k&=E_{t_k}(\overline{\nabla Y}_{k+1})+h\overline{\nabla Y}_{k+1}\frac{\partial}{\partial y}f(t_{k+1},\bar{Y}_{k+1}),\\h\overline{\nabla Z}_k&=E_{t_k}(\overline{\nabla Y}_k\Delta W_{k+1}).\end{aligned}$$

In another way, taking variation on both sides of the method, we have

$$\begin{aligned}
\nabla \bar{Y}_k &= E_{t_k}(\nabla \bar{Y}_{k+1}) + h\nabla \bar{Y}_{k+1}\frac{\partial}{\partial y}f(t_{k+1},\bar{Y}_{k+1}),\\
h\nabla \bar{Z}_k &= E_{t_k}(\nabla \bar{Y}_{k+1}\Delta W_{k+1}).
\end{aligned}$$

Obviously, $\overline{\nabla Y}_k = \nabla \bar{Y}_k$ and $\overline{\nabla Z}_k = \nabla \bar{Z}_k$, that is the method is closed under variation. One have

$$\begin{aligned}
|Y_{t_{k+1},\xi}(t_k) - \bar{Y}_{t_{k+1},\xi}(t_k)| &\le \int_{t_k}^{t_{k+1}} E_{t_k}(|f(s,Y_s) - f(t_{k+1},\xi)|)ds\\
&\le Kh^2 + K\int_{t_k}^{t_{k+1}} E_{t_k}(|Y_s - \xi|)ds\\
&\le K(1+E_{t_k}(|\xi|^2))^{1/2}h^2,
\end{aligned}$$

therefore,

$$E(|Y_{t_{k+1},\xi}(t_k) - \bar{Y}_{t_{k+1},\xi}(t_k)|^2) \le K(1+E(|\xi|^2))h^4,$$

that is $p_1 = 1$.

We can rewrite equation (20) as

$$hZ_{t_k} = E_{t_k}(Y_{t_{k+1}}\Delta W_{k+1}) + \int_{t_k}^{t_{k+1}} E_{t_k}(f(s,Y_s)\Delta W_s)ds - \int_{t_k}^{t_{k+1}} E_{t_k}(Z_s - Z_{t_k})ds.$$

Then

$$\begin{aligned}
&h^2|Z_{t_{k+1},\xi}(t_k) - \bar{Z}_{t_{k+1},\xi}(t_k)|^2\\
&= |\int_{t_k}^{t_{k+1}} E_{t_k}(f(s,Y_s)\Delta W_s)ds - \int_{t_k}^{t_{k+1}} E_{t_k}(Z_s - Z_{t_k})ds|^2\\
&\le Kh\int_{t_k}^{t_{k+1}} (E_{t_k}((f(s,Y_s) - f(t_k,Y_{t_k}))\Delta W_s))^2 ds\\
&\quad + Kh\int_{t_k}^{t_{k+1}} (E_{t_k}(Z_s - Z_{t_k}))^2 ds\\
&\le K((1+E_{t_k}(|\xi|^2)) + (1+E_{t_k}(|\nabla\xi|^2)))h^4
\end{aligned}$$

therefore we have

$$E(|Z_{t_{k+1},\xi}(t_k) - \bar{Z}_{t_{k+1},\xi}(t_k)|^2) \le KC^2h^2,$$

that is $p_2 = 1$. The mean-square order (1,1) of the explicit Euler method is known also from the convergence theorem 3.2.

The θ-scheme for equation (20) is considered in [11], which is also closed under variation

$$\begin{aligned}
\bar{Y}_k &= E^x_{t_k}[\bar{Y}_{k+1}] + (1-\theta)hE^x_{t_k}[f(t_{k+1},\bar{Y}_{k+1})] + \theta hf(t_k,\bar{Y}_k),\\
0 &= E^x_{t_k}[\bar{Y}_{k+1}\Delta W^T_{t_{k+1}}] + (1-\theta)hE^x_{t_k}[f(t_{k+1},\bar{Y}_{k+1})\Delta W^T_{t_{k+1}}]\\
&\quad - \{(1-\theta)hE^x_{t_k}[\bar{Z}_{k+1}] + \theta h\bar{Z}_k\},
\end{aligned}$$

They obtain that the mean-square orders of local error of this scheme are $p_1 = 2, p_2 = 3$, if $\theta = \frac{1}{2}$. But the mean-square order of convergence is $(2,1)$. In their Remark 3, they mentioned that for $\bar{Z}_k$, they only prove first-order convergence theoretically, but their numerical experiments show that the convergence rate is higher than one. From our fundamental convergence theorem, we know the mean-square order of convergence for $\bar{Z}_k$ should be 2, which agrees with their numerical experiments.

For the BSDE of the following form

$$Y_t = \varphi(W_T + x) + \int_t^T f(s, Y_s, Z_s)\mathrm{d}s - \int_t^T Z_s \mathrm{d}W_s,$$

a family of numerical methods which is closed under variation is considered in [10],

$$\begin{aligned}
\bar{Y}_k &= E^x_{t_k}[\bar{Y}_{k+1}] + \theta_1 h f(t_k, \bar{Y}_k, \bar{Z}_k) + (1-\theta_1) h E^x_{t_k}[f(t_{k+1}, \bar{Y}_{k+1}, \bar{Z}_{k+1})],\\
\theta_3 h \bar{Z}_k &= \theta_4 h E^x_{t_k}[\bar{Z}_{k+1}] + (\theta_3 - \theta_4) E^x_{t_k}[\bar{Y}_{k+1}\Delta W^T_{t_{k+1}}]\\
&\quad + (1-\theta_2) h E^x_{t_k}[f(t_{k+1}, \bar{Y}_{k+1}, \bar{Z}_{k+1})\Delta W^T_{t_{k+1}}],
\end{aligned}$$

where $\theta_i \in [0,1], i = 1,2$, $\theta_3 \in (0,1]$ and $\theta_4 \in [-1,1]$ constrained by $|\theta_4| \leq \theta_3$.

If $\varphi \in C_b^3$ and $f \in C_b^{1,3,3}$, one has that the mean-square orders of local error are $p_1 = 1, p_2 = 1$; the mean-square order of convergence is $(1,1)$, the results accord with our fundamental convergence theorem. If $\varphi \in C_b^2$ and $f \in C_b^{\frac{1}{2},2,2}$, one has that the mean-square orders of local error are $p_1 = \frac{1}{2}, p_2 = \frac{1}{2}$; the mean-square order of convergence is $(\frac{1}{2}, \frac{1}{2})$, the results match our fundamental convergence theorem. This indicates that the assumptions (i) and (ii) are sufficient.

If we check other schemes mentioned in [2, 8, 9, 10, 11], we will find that they also coincide with the fundamental convergence theorem.

5. **Conclusion.** We consider an important class of BSDEs with final condition $\chi = \varphi(W_T + x)$ and propose the fundamental convergence theorem on the mean-square order of numerical approximations for this class of BSDEs, which shows that the mean-square order of convergence is $(p_1, \min\{p_1, p_2\})$, if the numerical approximation is closed under variation, where $p_1 + 1$ and p_2 are mean-square orders of the one-step approximation for the first and second processes of the BSDE, respectively. The presented examples match our theoretical result. In our following works, We will consider the case of a broader class of BSDEs where the terminal condition is a function of a forward SDE.

Acknowledgments. The authors would like to express their sincere gratitude to Prof. Weidong Zhao for his discussions. The authors also would like to thank the reviewers for their constructive suggestions which helped to improve the presentation of this paper.

REFERENCES

[1] N. El Karoui, S. Peng and M. C. Quenez, *Backward stochastic differential equations in finance*, Math. Finance, **7** (1997), 1–71.

[2] Y. Li and W. Zhao, *L^p-error estimates for numerical schemes for solving certain kinds of backward stochastic differential equations*, Statist. Probab. Lett., **80** (2010), 1612–1617.

[3] J. Ma, P. Protter, J. San Martín and S. Torres, *Numerical method for backward stochastic differential equations*, Ann. Appl. Probab., **12** (2002), 302–316.

[4] J. Ma and J. Yong, "Forward-Backward Stochastic Differential Equaitons and their Applications," Lecture Notes in Mathematics, **1702**, Springer-Verlag, Berlin, 1999.

[5] G. N. Milstein, "Numerical Integration of Stochastic Differential Equations," Mathematics and its Applications, **313**, Kluwer Acdemic Publishers Group, Dordrecht, 1995.

[6] É. Pardoux and S. Peng, *Adapted solution of a backward stochastic differential equation*, Systems Control Lett., **14** (1990), 55–61.

[7] É. Pardoux and S. Peng, *Backward stochastic differential equations and quasilinear parabolic partial differential equations*, in "Stochastic Partial Differential Equations and Their Applications" (Charlotte, NC, 1991), Lecture Notes in Control and Inform. Sci., **176**, Springer, Berlin, (1992), 200–217.

[8] J. Wang, C. Luo and W. Zhao, *Crank-Nicolson scheme and its error estimates for backward stochastic differential equations*, Acta Math. Appl. Sinica English Ser., (2009).

[9] 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.

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

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

Received December 2012; revised April 2013.

E-mail address: chenchuchu@lsec.cc.ac.cn
E-mail address: hjl@lsec.cc.ac.cn