Weak Convergence Analysis and Improved Error Estimates for Decoupled Forward-Backward Stochastic Differential Equations

Wei Zhang and Hui Min *

Faculty of Science, College of Statistics and Data Science, Beijing University of Technology, Beijing 100124, China; weizhang0313@bjut.edu.cn
* Correspondence: huimin@bjut.edu.cn

Abstract: In this paper, we mainly investigate the weak convergence analysis about the error terms which are determined by the discretization for solving the stochastic differential equation (SDE, for short) in forward-backward stochastic differential equations (FBSDEs, for short), which is on the basis of Itô Taylor expansion, the numerical SDE theory, and numerical FBSDEs theory. Under the weak convergence analysis of FBSDEs, we further establish better error estimates of recent numerical schemes for solving FBSDEs.

Keywords: weak convergence analysis; Itô Taylor expansion; error estimates; forward-backward differential equations.

1. Introduction

This work is concerned with the forward-backward stochastic differential equations (FBSDEs) on \((\Omega, \mathcal{F}, \mathbb{P})\):

\[
X_t = X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)'dW_s, \quad t \in [0, T], \quad (\text{FSDE})
\]

\[
Y_t = \xi + \int_t^T f(s, X_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s, \quad t \in [0, T], \quad (\text{BSDE})
\]

where \(t \in [0, T]\), with \(T > 0\) being the deterministic terminal time; \((\Omega, \mathcal{F}, \mathbb{P})\) is a complete, filtered probability space with filtration \(\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}\) being the natural filtration of the standard \(d\)-dimensional Brownian motion \(W = (W_t)_{0 \leq t \leq T}\); \(X_0 \in \mathcal{F}_0\) is the initial condition for the forward stochastic differential equation (FSDE); \(\xi = \varphi(X_T) \in \mathcal{F}_T\) is the terminal condition for the backward stochastic differential equation (BSDE); the coefficients \(b: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d\) and \(\sigma: [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}\), the generator \(f: [0, T] \times \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \rightarrow \mathbb{R}^m\), and \((X_t, Y_t, Z_t): [0, T] \times \Omega \rightarrow \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^{m \times d}\) are unknown.

We point out that \(b(\cdot, x, y, z), \sigma(\cdot, x, y, z),\) and \(f(\cdot, x, y, z)\) are all \(\mathcal{F}_t\)-adapted for any fixed numbers \(x, y\) and \(z\), and that the two stochastic integrals with respect to \(W_t\) are of the Itô type. A triple \((X_t, Y_t, Z_t)\) is called an \(L^2\)-adapted solution of the decoupled FBSDEs (1) and (2) if, in the probability space \((\Omega, \mathcal{F}, \mathbb{P})\), it is \(\{\mathcal{F}_t\}\)-adapted, square integrable, and satisfies the Equations (1) and (2).

FBSDEs have important applications in many fields, including mathematical finance, partial differential equations, stochastic controls, risk measure, and so on [1–5]. Thus, it is interesting and also important to find solutions of FBSDEs. Usually, it is difficult to get the analytical solutions in an explicit closed form. Thus, numerical methods for solving FBSDEs are desired, especially accurate, effective, and efficient ones. Many numerical schemes for solving BSDE and decoupled FBSDEs have been developed, among which some are numerical methods with low-order convergence rates [6–10] under lower regularity assumptions, while others are high-order numerical methods [11–17]. It is notable
that most of the numerical methods are designed for BSDE, or decoupled FBSDEs. In those methods, the authors place more attention on the truncated error terms, thereby ignoring the errors which are caused by the discretization scheme for solving the SDE (1).

The main purpose of this work is to prove that the weak convergence analysis of the four error terms is determined by the discretization for solving the SDE (1). Our analysis invokes an Ito Taylor expansion, the numerical SDE theory, and the numerical FBSDEs theory. In Section 2, the high-order numerical scheme introduced in [16] is briefly reviewed. Section 3 firstly gives the stability analysis of the proposed Scheme in Section 2; then our main result, weak convergence analysis, together with some useful lemmas is presented; finally, under the weak convergence analysis of FBSDEs, we improve the error estimates about the Schemes [16,17]. Some concluding remarks are made in Section 4.

For a simple representation, let us first introduce the following notations:

- $|·|$: the standard Euclidean norm in the Euclidean space $\mathbb{R}^q$, $\mathbb{R}^m$, and $\mathbb{R}^{m \times d}$.
- $\mathcal{F}_{t}^X(t \leq s \leq T)$: the $\sigma$-field generated by the diffusion process $\{X_t: t \leq r \leq s, X_t = x\}$ starting from the time-space point $(t, x)$. When $s = T$, we use $\mathcal{F}_{T}^X$ to denote $\mathcal{F}_{T}^X$.
- $E_{t}^X[X]$: the conditional expectation of the random variable $X$ under the $\sigma$-field $\mathcal{F}_{t}^X(t \leq s \leq T)$, that is, $E_{t}^X[X] = E[X|\mathcal{F}_{t}^X]$. Let $E_{t}^X[X] = E[X|\mathcal{F}_{t}^X]$.
- $P_{i}[a, b]$: the set of polynomials of degree $i$ defined on $[a, b]$.
- $C_{b}^{k}$: the set of continuously differential functions $\phi: x \in \mathbb{R}^q \to \mathbb{R}$ with uniformly bounded partial derivatives $\partial^{k}_{x} \phi$ for $k_1 \leq k$.
- $C_{b}^{k}x$: the set of functions $\phi: (t, x) \in [0, T] \times \mathbb{R}^q \to \mathbb{R}$ with uniformly bounded partial derivatives $\partial^{l}_{x} \partial^{k}_{t} \phi$ for $l_1 \leq l$ and $k_1 \leq k$.
- $C_{b}^{k,k}$: the set of continuously differential functions $\phi: (t, x, y) \in [0, T] \times \mathbb{R}^q \times \mathbb{R} \to \mathbb{R}$ with uniformly bounded partial derivatives $\partial^{k}_{x} \phi$ and $\partial^{k_1}_{x} \partial^{k_2}_{y} \phi$ for $l_1 \leq l$ and $k_1 + k_2 \leq k$.
- $C_{b}^{k}$: the set of $k$ times continuously differentiable functions $\phi: x \in \mathbb{R}^q \to \mathbb{R}$ for which $\phi$ and all of its partial derivatives of orders up to and including $k$ have polynomial growth.

2. Discretization

For the time interval $[0, T]$, we introduce the following partitions:

$$0 = t_0 < t_1 < \cdots < t_N = T.$$ 

Let $\Delta t_n = t_{n+1} - t_n, n = 0, 1, \ldots, N - 1$ and $\Delta t = \max_{0 \leq n \leq N-1} \Delta t_n$. We also assume that the time partitions have the following regularity:

$$\min_{0 \leq n \leq N-1} \frac{\Delta t_n}{\max_{0 \leq n \leq N-1} \Delta t_n} \leq c_0,$$ 

(3)

where $c_0 \geq 1$ is a constant.

Let $(X^{t, x}, Y^{t, x}, Z^{I, x})$ be the solution of (1) and (2) with date $(t, x)$; that is, $(X^{t, x}, Y^{t, x}, Z^{I, x})$ satisfies

$$\begin{cases} 
X^{t, x}_t = x + \int_t^s b(r, X^{t, x}_r)dr + \int_t^s \sigma(r, X^{t, x}_r)dW_r, & s \in [t, T], \\
Y^{t, x}_s = \phi(X^{t, x}_T) + \int_s^T f(r, X^{t, x}_r, Y^{t, x}_r, Z^{I, x}_r)dr - \int_s^T Z^{I, x}_r dW_r, & s \in [t, T]. 
\end{cases}$$ 

(4)

2.1. Reference Equations

To derive the reference equations, we first define the stochastic process

$$\Delta \tilde{W}_s = \int_{t_n}^{t_{n+1}} (2 - \frac{3(r - t_n)}{\Delta t_n})dW_r, \quad s \in [t_n, t_{n+1}).$$ 

(5)
The process $\Delta \tilde{W}_s = (\Delta \tilde{W}^1_s, \Delta \tilde{W}^2_s, \cdots, \Delta \tilde{W}^d_s)\top$ is a martingale with the properties $E^X_{t_n}[\Delta \tilde{W}_s] = 0$, $E^X_{t_n}[\Delta \tilde{W}^i_s \Delta \tilde{W}^j_s] = 0$ for $i \neq j$, and

$$E^X_{t_n}[(\Delta \tilde{W}^i_s)^2] = 4(s - t_n) - \frac{6(s - t_n)^2}{\Delta t_n} + \frac{3(s - t_n)^3}{\Delta t_n^2}.$$ 

Then when $s = t_{n+1}$, we have $E^X_{t_n}[\Delta \tilde{W}_{t_{n+1}}] = 0$ and $E^X_{t_n}[(\Delta \tilde{W}^i_{t_{n+1}})^2] = \Delta t_n$.

Fixed $n \in N$, let $Y^t_{t_n}X^n = \phi(X^n_{t_n})$ and $(X^n_{t_n}, Y^t_{t_n}X^n, Z^n_{t_n})$ be the solution of (4) with $X^n = x$ for $t \in [t_n, T]$. Denote $f(s, X^n_{t_n}, Y^t_{t_n}X^n, Z^n_{t_n})$ by $f^n_{t_n}X^n$, $s \in [t_n, t_{n+1})$. Then it is easy to get that

$$X^n_{t_{n+1}} = X^n + \int_{t_n}^{t_{n+1}} b(s, X^n_{t_n})ds + \int_{t_n}^{t_{n+1}} \sigma(s, X^n_{t_n})dW_s,$$ (6)

$$Y^t_{t_n}X^n = Y^t_{t_n}X^n + \int_{t_n}^{t_{n+1}} f^n_{t_n}X^n ds - \int_{t_n}^{t_{n+1}} Z^n_{t_n}X^n dW_s,$$ (7)

for $n = 0, 1, \cdots, N - 1$. The forward SDE (6) can be solved independently by existing numerical methods in the above decoupled case, and in this article, we use numerical schemes of the general form

$$X^{n+1} = X^n + \phi(t_n, t_{n+1}, X^n, I_n \in A_t),$$ (8)

where $\beta$ represents the convergence rate in the strong or weak sense and $A_t$ denotes the corresponding hierarchical set; for more details on the notations, readers can be referred to [18]. Our main goal therefore turns to deducing the reference equations for $Y^t_{t_n}X^n$ and $Z^n_{t_n}X^n$ from the BSDE (7).

Since $\int_{t_n}^{t} Z^n_{s}X^n dW_s$ is a martingale for $t > t_n$, by taking the conditional expectation $E^X_{t_n} [\cdot]$ on (7) we get

$$Y^t_{t_n}X^n = E^X_{t_n} [Y^t_{t_n}X^n] + \int_{t_n}^{t_{n+1}} E^X_{t_n} [f^n_{t_n}X^n] ds.$$ (9)

Then, multiplying Equation (7) by $\Delta \tilde{W}^i_{t_{n+1}}$ and taking the conditional mathematical expectation $E^X_{t_n} [\cdot]$ on both sides of the derived equation, we obtain

$$0 = E^X_{t_n} [Y^t_{t_n}X^n \Delta \tilde{W}^i_{t_{n+1}}] + \int_{t_n}^{t_{n+1}} E^X_{t_n} [f^n_{t_n}X^n \Delta \tilde{W}^i_{t_{n+1}}] ds - E^X_{t_n} [\int_{t_n}^{t_{n+1}} Z^n_{t_n}X^n dW_s \cdot \Delta \tilde{W}^i_{t_{n+1}}].$$ (10)

Under the filtration $\mathcal{F}_t$, the integrands in (9) and (10) are deterministic functions of $s \in [t_n, t_{n+1})$, so some numerical integration methods for deterministic integration can be used to approximate these integrals.

First, we apply the trapezoidal rule to obtain

$$\int_{t_n}^{t_{n+1}} E^X_{t_n} [f^n_{t_n}X^n] ds = \frac{1}{2} \Delta t_n f^n_{t_n}X^n + \frac{1}{2} \Delta t_n E^X_{t_n} [f^n_{t_n}X^n] + R^n_y,$$ (11)

with the truncation error

$$R^n_y = \int_{t_n}^{t_{n+1}} \left( E^X_{t_n} [f^n_{t_n}X^n] - \frac{1}{2} f^n_{t_n}X^n - \frac{1}{2} E^X_{t_n} [f^n_{t_n}X^n] \right) ds.$$ (12)

Inserting (11) into (9), we obtain the following reference equation for solving $Y^t_{t_n}X^n$:

$$Y^t_{t_n}X^n = E^X_{t_n} [Y^t_{t_n}X^n] + \frac{1}{2} \Delta t_n f^n_{t_n}X^n + \frac{1}{2} \Delta t_n E^X_{t_n} [f^n_{t_n}X^n] + R^n_y,$$ (13)
Next, for the first integral term on the right-hand side of (10), we easily get the identity
\[ \int_{t_n}^{t_{n+1}} E_{t_n}^{X_n} \left[ f_{t_n,X_n} \Delta \tilde{W}_{t_n+1}^s \right] ds = \Delta t_n E_{t_n}^{X_n} \left[ f_{t_n+1,X_n} \Delta \tilde{W}_{t_n+1}^s \right] + R_n^1, \] (14)
where \( R_n^1 = \int_{t_n}^{t_{n+1}} E_{t_n}^{X_n} \left[ f_{t_n,X_n} \Delta \tilde{W}_{t_n+1}^s \right] ds - \Delta t_n E_{t_n}^{X_n} \left[ f_{t_n+1,X_n} \Delta \tilde{W}_{t_n+1}^s \right] \). For the second integral term on the right-hand side of (10), we have
\begin{align*}
-\frac{1}{2} E_{t_n}^{X_n} & \left[ \int_{t_n}^{t_{n+1}} Z_{t_n,X_n}^n dW_n \cdot \Delta \tilde{W}_{t_n+1}^s \right] = -Z_{t_n,X_n}^n E_{t_n}^{X_n} \left[ \Delta \tilde{W}_{t_n+1}^s \right] + R_n^2 \\
& = -\frac{1}{2} \Delta t_n Z_{t_n,X_n}^n + R_n^2,
\end{align*}
(15)
where \( R_n^2 = \frac{1}{2} \Delta t_n Z_{t_n,X_n}^n - \frac{1}{2} E_{t_n}^{X_n} \left[ \int_{t_n}^{t_{n+1}} Z_{t_n,X_n}^n dW_n \cdot \Delta \tilde{W}_{t_n+1}^s \right] \). From Equations (10), (14) and (15), we obtain the following reference equation for solving \( Z_{t_n,X_n}^n \):
\begin{align*}
\frac{1}{2} \Delta t_n Z_{t_n,X_n}^n & = E_{t_n}^{X_n} \left[ f_{t_{n+1},X_n} \Delta \tilde{W}_{t_n+1}^s \right] + \Delta t_n E_{t_n}^{X_n} \left[ f_{t_{n+1},X_n} \Delta \tilde{W}_{t_n+1}^s \right] + R_n^2 \\
& = \frac{1}{2} \Delta t_n Z_{t_n,X_n}^n + R_n^2.
\end{align*}
(16)
where \( R_n^2 = R_n^1 + R_n^2 \).

2.2. Numerical Scheme

Now we propose a new numerical scheme for solving the decoupled FBSDEs (4). Let \( (X^n, Y^n, Z^n) \) denote an approximation to the analytic solution \( (X_t, Y_t, Z_t) \) of (4) at time \( t = t_n, n = N, N - 1, \ldots, 0 \). To simplify the presentation, we let \( f^n = f(t_n, X^n, Y^n, Z^n) \) for \( n = N, N - 1, \ldots, 0 \). Based on (13) and (16), we propose a numerical scheme, for solving the FBSDEs (4) as follows.

**Scheme 1.** Given random variables \( X_0, Y^N \) and \( Z^N \). Let \( \Delta \tilde{W}_{t_{n+1}} \) \((0 \leq n \leq N - 1)\) be defined by (5) with \( s = t_{n+1} \). For \( n = N - 1, N - 2, \ldots, 0 \), solve random variables \( Y^n \) and \( Z^n \) by
\begin{align*}
Y^n & = E_{t_n}^{X_n} [Y^{n+1} + \frac{1}{2} \Delta t_n f^n + \frac{1}{2} \Delta t_n E_{t_n}^{X_n} [f^{n+1}], \\
Z^n & = E_{t_n}^{X_n} [Y^{n+1} + \Delta t_n E_{t_n}^{X_n} [f^{n+1} + \Delta \tilde{W}_{t_{n+1}}]].
\end{align*}
(17)
(18)

3. Error Estimates

We now provide some theoretical analysis on the numerical stability and convergence of the Scheme 1. Let us denote by \( \bar{Y}_{t_{n+1}}^{n+1} \) and \( \bar{Z}_{t_{n+1}}^{n+1} \) the approximate values of \( Y_{t_{n+1}}^{n+1} \) and \( Z_{t_{n+1}}^{n+1} \) at the time-space \( (t_{n+1}, X^{n+1}) \), respectively, where \( X^{n+1} \) is the approximate solution of \( X_{t_{n+1}}^{n+1} \) calculated by (8), that is, \( Y_{t_{n+1}}^{n+1} = Y_{t_{n+1}}^{n+1} \) and \( Z_{t_{n+1}}^{n+1} = Z_{t_{n+1}}^{n+1} \). To simplify the presentation, we first introduce the following notation:
\[ \bar{f}_t^{n+1} \] and \( \bar{e}_t^{n+1} = Y_t^{n+1} - \bar{Y}_t^{n+1} \), \( e_t^{n+1} = Z_t^{n+1} - \bar{Z}_t^{n+1} \) for \( i = n, \ldots, N \), and
\begin{align*}
R_n^{Y,y} & = E_{t_n}^{X_n} \left[ (Y_{t_{n+1}}^{n+1} - \bar{Y}_{t_{n+1}}^{n+1}) \Delta \tilde{W}_{t_{n+1}}^s \right], \\
R_n^{Z,z} & = E_{t_n}^{X_n} \left[ (Z_{t_{n+1}}^{n+1} - \bar{Z}_{t_{n+1}}^{n+1}) \Delta \tilde{W}_{t_{n+1}}^s \right], \\
R_n^{R} & = E_{t_n}^{X_n} \left[ (f_{t_{n+1},X_n}^{n+1} - \bar{f}_{t_{n+1}}^{n+1}) \Delta \tilde{W}_{t_{n+1}}^s \right].
\end{align*}
(19)
for \( i = n, \ldots, N - 1 \).
3.1. Stability Analysis

The following result on decoupled FBSDEs is well-known by now; for its proof, see Theorem 4.1 in [16] or Theorem 4.1 in [17].

**Theorem 1.** Let \((X_t, Y_t, Z_t), t \in [0, T]\) and \((X^n, Y^n, Z^n), n = 0, 1, \ldots, N,\) be the exact solution of the decoupled FBSDEs (1) and (2) and the approximate solution obtained by Scheme 1, respectively. Assume that the function \(f(t, x, y, z)\) is Lipschitz-continuous with respect to \(x, y,\) and \(z,\) and the Lipschitz constant is \(L.\) Let \(c_0\) be the time partitions regularity parameter defined in (3). Then, for the sufficiently small time-step \(\Delta t_n\) satisfying (3), it holds that

\[
\mathbb{E}[|\epsilon^0|^2] + \Delta t \sum_{i=0}^{N-1} \left( \frac{1 + C\Delta t}{1 - C\Delta t} \right)^{i-n} \mathbb{E}[|\epsilon^i|^2] \\
\leq C' \left( \mathbb{E}[|\epsilon^N|^2] + \Delta t\mathbb{E}[|\epsilon^0|^2] \right) + \sum_{j=1}^{N-1} \left( \frac{1 + C\Delta t}{1 - C\Delta t} \right)^{i-n} \mathbb{E}[|R_{ij,1}^i|^2 + (\Delta t)^2 |R_{ij,2}^i|^2 + |R_{ij,3}^i|^2] \frac{1 - C\Delta t}{C\Delta t} \Delta t (1 - C\Delta t) \tag{20}
\]

for \(n = N - 1, \ldots, 1, 0,\) where \(C\) is a positive constant depending on \(c_0\) and \(L;\) \(C'\) is also a positive constant depending on \(c_0, T,\) and \(L;\) and \(R_{ij,1}^i, R_{ij,2}^i,\) and \(R_{ij,3}^i\) are defined in (12) and (16), respectively.

**Remark 1.** Theorem 1 implies that Scheme 1 is stable, and its solution continuously depends on terminal conditions; that is, for any given positive number \(\epsilon,\) there exists a positive integer \(\delta,\) for different terminal conditions \((Y^N, Z^N)\) and \((Y^n, Z^n),\) if \(\mathbb{E}[||Y^N - Y^n||^2] < \delta\) and \(\mathbb{E}[||Z^N - Z^n||^2] < \delta,\) then for \(0 \leq n \leq N - 1,\) we have

\[
\mathbb{E}[||Y^n - Y^N||^2] + \Delta t \sum_{i=n}^{N} \mathbb{E}[||Z^n - Z^i||^2] < \epsilon.
\]

**Remark 2.** The terms \(R_{ij,1}^i, R_{ij,2}^i,\) and \(R_{ij,3}^i\) in (12) and (16) come from the approximations in their reference equations for \(Y_t\) and \(Z_t.\) The four terms \(R_{ij,1}^i, R_{ij,2}^i, R_{ij,3}^i,\) and \(R_{ij,4}^i\) are determined by the discretizations (8) for solving the SDE in (6), which reflects the local errors (in the weak sense) of the numerical scheme for FDE. Under certain regularity conditions on \(b, c, f,\) and \(\varphi,\) if we get estimates of these terms, a convergence result for the Scheme 1 can be obtained.

In the following lemmas, we will present estimations for \(R_{ij,1}^i, R_{ij,2}^i, R_{ij,3}^i, R_{ij,4}^i, R_{ij,5}^i,\) and \(R_{ij,6}^i\) under certain regularity conditions on \(b, c, f,\) and \(\varphi.\) For the sake of presentation simplicity, we only consider the case \(q = d = 1,\) and the results obtained also hold true for general positive integers \(q\) and \(d.\) In order to get the estimates of \(R_{ij,1}^i, R_{ij,2}^i, R_{ij,3}^i,\) and \(R_{ij,4}^i,\) we need the weak error estimate about the SDE in (16).

3.2. Weak Convergence Analysis

For the convenience, we use the following notions which appeared in [18].

- Denote the set of all multi-indices by \(\mathcal{M},\)

\[
\mathcal{M} = \{ (j_1, j_2, \ldots, j_l) : j_i \in \{0, 1, 2, \ldots, d\}, i \in \{1, 2, \ldots, l\}, \text{ for } l = 1, 2, 3, \ldots \} \cup \{v\}.
\]

- Let \(l(\alpha) := l\) be the length of the multi-index \(\alpha = (j_1, j_2, \ldots, j_l) \in \mathcal{M},\) where \(j_i \in \{0, 1, 2, \ldots, d\}\) for \(i \in \{1, 2, \ldots, l\}, d = 1, 2, 3, \ldots\) Here, \(d\) is the number of components of the Wiener process \(W\) under consideration. For example: \(l((1, 0)) = 2, l((1, 0, 0)) = 3.\)

- Let \(n(\alpha)\) be the number of components of a multi-index \(\alpha\) which are equal to 0. For example: \(n((1, 0)) = 1, n((1, 0, 0)) = 2.\)

- For completeness, denote by \(v\) the multi-index of length zero, that is, with \(l(v) := 0.\)

- If \(l(\alpha) \geq 1,\) write \(-\alpha\) and \(\alpha^-\) for the multi-index in \(\mathcal{M}\) obtained by deleting the first and last component, respectively, of \(\alpha.\) Thus, \(-(1, 0) = (0),\) \((1, 0)^- = (1).\)
• Define $a^+$ to be the multi-index obtained from $a$ by deleting all of the components equal to 0. For $E$, if $a = (1, 2, 0, 1)$, then we have $a^+ = (1, 2, 0, 1)^+ = (1, 2, 1)$.

• Denote by $k_0(a)$ the number of components of $a$ equal to 0 or until the end of $a$ if all its components are zeros.

• Denote by $k_i(a)$, for $i = 1, \ldots, l(a^+)$, the number of components of $a$ between the $i$th nonzero component and the $(i + 1)$th nonzero component or the end of $a$ if $i = l(a^+)$. For example: $a = (0, 1, 2, 0)$ we have $a^+ = (1, 2), l(a^+) = 2$ and $k_0(a) = 1$, $k_1(a) = 0$, $k_2(a) = 1$.

• Define $\omega(a, \beta) := l(a^+) + \sum_{i=0}^{l(a^+)} (k_i(a) + k_i(\beta))$.

• Let $\rho$ and $\tau$ be two stopping times with $0 \leq \rho \leq \tau \leq T$, w.p.1. Then, for a multi-index $a = (j_1, j_2, \ldots, j_l) \in M$, and a process $f \in H_a$ the multi-Itô integral $I_a[f(\cdot)]_{\rho, \tau}$ is recursively defined by

$$
I_a[f(\cdot)]_{\rho, \tau} := \begin{cases} f(\tau) & : l = 0, \\ \int_{\rho}^{\tau} I_a[f(\cdot)]_{\rho, s} ds & : l \geq 1 \text{ and } j_l = 0, \\ \int_{\rho}^{\tau} I_a[f(\cdot)]_{\rho, s} dW_{s}^j & : l \geq 1 \text{ and } j_l \geq 1. 
\end{cases}
$$

For $a = (j_1, j_2, \ldots, j_l)$ with $l \geq 2$ the set $H_a$ is recursively defined to be the totality of adapted right continuous processes $f = \{f(t), t \geq 0\}$ with left-hand limits such that the integral process $\{I_a[f(\cdot)]_{\rho, t}, t \geq 0\}$ considered as a function of $t$ satisfies $I_a[f(\cdot)]_{\rho, t} \in H_{\beta}$.

Based on the upper notations, we give the following lemmas which are valid for both scalar and vector functions. The reader interested in more details is also referred to [18].

**Lemma 1.** Let $\alpha, \beta \in M$ and $f \in H_\alpha$, $g \in H_\beta$. Then for $0 \leq n \leq N - 1$ we have

$$
\mathbb{E}_{t_n}^{X_\alpha} \left[ \left( I_{a_n}^{s_1} \left[ f(s) \right]_{t_n, t_{n+1}} I_{b_j}^{s_1} \left[ g(s) \right]_{t_n, t_{n+1}} \right)^{2q} \right] = 0,
$$

$$
\leq K_{f, g} \cdot \omega(a, b)^{\cdot \cdot \cdot \cdot} \cdot \prod_{i=0}^{l(a^+)} C_{k_i(a) + k_i(\beta)} \cdot \alpha^+ \cdot \beta^+,
$$

with

$$
K_{f, g} = \sup_{s_1, s_2 \in [t_n, t_{n+1}]} \mathbb{E}_{t_n}^{X_\alpha} \left[ \left( f(s_1), g(s_2) \right) \right],
$$

where

$$
C_{k_i(a) + k_i(\beta)} = \frac{(k_i(a) + k_i(\beta))!}{k_i(a) k_i(\beta)!}.
$$

Moreover, (21) holds with equality when $f \equiv g \equiv 1$.

The next lemma provides an estimate for the higher moments of a multiple Itô integral.

**Lemma 2.** Let $\alpha \in M$, $g \in H_\alpha$ and $q = 1, 2, \ldots$. Then for $0 \leq n \leq N - 1$ we have the estimate

$$
\left( \mathbb{E}_{t_n}^{X_\alpha} \left[ \left| I_{a_n}^{s_1} \left[ g(s) \right]_{t_n, t_{n+1}} \right|^{2q} \right] \right)^{1/q} \leq (2(2q - 1) e^T)^{l(a) - n(a)} (\Delta t)^{l(a) + n(a)} R,
$$

where $R = \left( \mathbb{E}_{t_n}^{X_\alpha} \left[ \sup_{t_n \leq s \leq t_{n+1}} \left| g(s)^{2q} \right| \right] \right)^{1/q}$.
Next, we consider the coefficients of stochastic Itô-Taylor expansions. First, we write the definition of diffusion operator $L^0$ as

$$L^0 = \frac{\partial}{\partial t} + \sum_{k=1}^{q} p^k \frac{\partial}{\partial x^k} + \frac{1}{2} \sum_{h,j=1}^{d} \sigma^{kj}(t,x) \frac{\partial^2}{\partial x^h \partial x^j},$$

and for $j \in \{1, \ldots, d\}$ the operator $L^j$ as

$$L^j = \sum_{k=1}^{q} \sigma^{kj} \frac{\partial}{\partial x^k}. \quad (23)$$

For each $\alpha = (j_1, j_2, \ldots, j_l) \in \mathcal{M}$ and function $f \in C^h(\mathbb{R}^+ \times \mathbb{R}^d, \mathbb{R})$ with $h = l(\alpha) + n(\alpha)$ we define recursively $f_\alpha$ as

$$f_\alpha = \begin{cases} f & : l = 0, \\ L^h f_{-\alpha} & : l \geq 0. \end{cases}$$

In order to give weak Taylor expansion schemes, we use the following definitions:

- Call a subset $A \subset \mathcal{M}$ an hierarchical set if it satisfies:
  1. $A$ is nonempty: $A \neq \emptyset$;
  2. The multi-indices in $A$ are uniformly bounded in length: $\sup_{\alpha \in A} l(\alpha) < \infty$;
  3. $-\alpha \in A$ for each $\alpha \in A \setminus \{v\}$.

- For any given hierarchical set $A$ define the remainder set $B(A)$ of $A$ by $B(A) = \{ \alpha \in \mathcal{M} \setminus A : -\alpha \notin A \}$.

In the general multi-dimensional case $q, d = 1, 2, \ldots,$ for $\beta = 1.0, 2.0, 3.0, \ldots,$ define the weak Taylor scheme of order $\beta$ by the vector equation

$$X^{n+1} = X^n + \sum_{\alpha \in \Gamma_{\beta}(0)} f_\alpha(t_n, X^n) I_{\alpha[n,t_{n+1}]} = X^n + \sum_{j=0}^{d} \int_{t_n}^{t_{n+1}} \sigma^j_t dW^j_t, \quad (24)$$

with $dW^0_t = ds$ and

$$\sigma^j_t = \sum_{\alpha \in \Gamma_{\beta}(e)} I_{\alpha - [f_\alpha(t_n, X^n)]_{t_n,t_{n+1}},}$$

for $j = 0, 1, \ldots, d$, where $\alpha = (j_1, j_2, \ldots, j_l(\alpha))$. Here, $X_{t_n} = X^n$ is assumed. The scheme (24) is obtained from

$$X^{n+1}_{t_{n+1}} = X_{t_n} + \sum_{\alpha \in \Gamma_{\beta}(v)} f_\alpha(t_n, X_{t_n}) I_{\alpha[n,t_{n+1}]} + R^\beta_{X_{t_{n+1}}}. \quad (25)$$

Here, $R^\beta_{X_{t_{n+1}}} = \sum_{\alpha \in B(\Gamma_{\beta})} I_{\alpha [f_\alpha(\cdot, X^{n+1}_{t_{n+1}})]_{t_n,t_{n+1}}} \Gamma_{\beta} = \{ \alpha \in \mathcal{M} : l(\alpha) \leq \beta \}$ is an hierarchical set, and $B(\Gamma_{\beta}) = \{ \alpha \in \mathcal{M} : l(\alpha) = \beta + 1 \}$ is the remainder set of $\Gamma_{\beta}$.

We denote $C^p_\beta(\mathbb{R}^d, \mathbb{R})$ by the space of $l$ times continuously differentiable functions $g : \mathbb{R}^d \to \mathbb{R}$ for which $g$ and all its derivatives of orders up to and including $l$ have polynomial growth.

**Lemma 3.** Let $\beta \in \{1, 2, \ldots\}$ be given and suppose that the coefficients $b^k$ belong to the space $C^2_\beta(0, T \times \mathbb{R}^d, \mathbb{R})$ and $\sigma^{kj}$ belong to the space $C^2_\beta([0, T] \times \mathbb{R}^d, \mathbb{R}^{q \times d})$ and satisfy Lipschitz conditions and linear growth bounds for $k = 1, \ldots, q$ and $j = 0, 1, \ldots, d$. Then, for
each $g \in C_P^2([0, T] \times \mathbb{R}^d)$ and for the sufficiently small time-step $\Delta t$ satisfying (3), there exist constants $C \in (0, \infty)$ and $r \in \{1, 2, \ldots\}$ such that

$$\mathbb{E}_{t_n} \left[ \sup_{0 \leq n \leq N-1} \mathbb{E}_{t_{n+1}}^{X_n} \left| X_{t_{n+1}}^{X_n} \right|^2 \right] \leq C(1 + |X|^r),$$

(26)

and

$$\sup_{0 \leq n \leq N-1} \mathbb{E}_{t_n} \mathbb{E}_{t_{n+1}}^{X_n} \left| g(X_{t_{n+1}}^{X_n}) - g(X^{X+1}) \right| \leq C(1 + |X|^r)(\Delta t)^{\beta + 1}.$$  

(27)

Lemma 4. Under the conditions that Lemma 3 holds, then for the sufficiently small time-step $\Delta t_n$ satisfying (3), we have that for any $0 \leq n \leq N - 1$,

$$\mathbb{E}[|R_{y_1}|^2] \leq C(1 + \mathbb{E}[|X_0|^4])(\Delta t)^{2\beta + 2}, \quad \mathbb{E}[|R_{y_2}|^2] \leq C(1 + \mathbb{E}[|X_0|^4])(\Delta t)^{2\beta + 2}.$$  

(28)

Proof. We know that

$$R_{y_1} = \mathbb{E}_t^{X} [Y_{t_{n+1}}^{X_n} - Y_{t_{n+1}}^{q}] \quad \text{and} \quad R_{y_2} = \mathbb{E}_t^{X} [f_{t_{n+1}}^{X_n} - f_{t_{n+1}}^{q}].$$

Using Lemma 3, we can get the results directly. \quad \square

In order to give the estimates of $R_{y_1}$ and $R_{y_2}$, we need the following lemmas. Due to the complexity of the Itô Taylor expansion, we give some useful notations.

We write $P_l = \{1, 2, \ldots, q\}^l$ and

$$F_p(y) = \prod_{k=1}^l y^{p_k},$$

(29)

for all $y = (y_1, y_2, \ldots, y^q)$ and $p = (p_1, \ldots, p_l) \in P_l$, $l = 1, 2, \ldots$.

Then starting from (24) and (29), by a generalization of the Itô formula to semimartingales, we obtain

$$F_p(X^{n+1} - X^n) = \sum_{j=0}^d f_{t_{n+1}}^{X^n} \hat{\sigma}^j_{p}(s) dW_s,$$

(30)

for all $0 \leq n \leq N - 1$ and $p \in P_l$, $l = 1, 2, \ldots$, where

$$\hat{\sigma}^j_{p}(t) = \sum_{k=1}^q \sigma^j_{k} \frac{\partial}{\partial x_k} F_p(X^{n+1} - X^n),$$

for $j = 1, 2, \ldots, d$ and

$$\hat{\sigma}^0_{p}(t) = \frac{\partial}{\partial t} F_p(X^{n+1} - X^n) + \sum_{k=1}^q \sigma^0_{k} \frac{\partial}{\partial x_k} F_p(X^{n+1} - X^n)$$

$$+ \frac{1}{2} \sum_{k,l=1}^q \sum_{j=1}^d \sigma^j_{l} \sigma^k_{l} \frac{\partial^2}{\partial x^k \partial x^l} F_p(X^{n+1} - X^n).$$

Lemma 5. Under the conditions of Lemma 3 for each $p = 1, 2, \ldots$ and the sufficiently small time-step $\Delta t$ satisfying (3), there exist constants $C \in (0, \infty)$ and $r \in \{1, 2, \ldots\}$ such that

$$\mathbb{E}_{t_n}^{X_n} \left[ \sup_{0 \leq n \leq N-1} \mathbb{E}_{t_{n+1}}^{X_n} \left| F_p(X_{t_{n+1}}^{X_n}) \right|^2 \right] \leq C(1 + |X|^r)(\Delta t)^{\beta+1},$$

for all $q = 0, 1, \ldots, p \cdot 2^{(\beta+1) - l}$ and $p \in P_l$ where $l = 1, \ldots, 2(\beta + 1)$.

To give an important proposition, first we need to prove the following lemma.
Lemma 6. Under the conditions of Lemma 3, there exist constants $C \in (0, \infty)$ and $r \in \{1, 2, \ldots \}$ such that

$$F(t) := \|E_{t_n}^X \left[ (F_{t_n}(X_{t_n}^{l_n} - X_{t_n}) - F_{t_n}(X^{n+1} - X^n)) \Delta \tilde{W}_{t_{n+1}} \right]\| \leq C(1 + |X^n|^{2r}) \Delta t^{\beta + 1},$$

for all $\tilde{p} \in P_l, l = 1, 2, \ldots, 2(\beta + 1)$.

Proof. Step 1. The case $l = 1$. In this case, $\tilde{p} = (k)$ with $k \in \{1, 2, \ldots, q\}$. Then by (23) and (29), we have

$$F(t) = \bigg| E_{t_n}^X \left[ (X_{t_n}^{l_n}X^n_k - X_{t_n}^{n+1}X^n) \Delta \tilde{W}_{t_{n+1}} \right]\bigg| \leq \sum_{a \in B(t_{n+1})} E_{t_n}^X \left[ I_a \left[ f^k_i(, X_{t_n}^{l_n}X^n) \right]_{t_n, t_{n+1}} \sum_{i=1}^d I(i) \left[ P(\cdot) \right]_{t_n, t_{n+1}} \right],$$

where $P(\cdot) = 2 - \frac{3(1-l_{n+1})}{2\lambda}$ are the set of backward orthogonal polynomials on $[t_n, t_{n+1}]$ from the definition of $\Delta \tilde{W}_t$ in (5).

By Lemma 1, only the terms with $a^+ = (i)$ in the sum on the right-hand side of the above inequality are not zero. For $a^+ = (i)$, it is easy to check $w(a, b) = l(a)$. By Lemma 1 and the Hölder inequality, we have the estimate

$$F(t) \leq \sum_{a \in B(t_{n+1})} \sum_{i=1}^d C \Delta t^{l(a)} (K_1 K_2^i)^{1/2},$$

with $l(a) = \beta + 1$, where

$$K_1 = E_{t_n}^X \left[ \sup_{s_1 \in [t_n, t_{n+1}]} \left| f^k_i(s_1, X_{s_1}^{l_n}X^n) \right|^2 \right],$$

and

$$K_2^i = E_{t_n}^X \left[ \sup_{s_2 \in [t_n, t_{n+1}]} |P(s_2)|^2 \right].$$

The inequality (26) in Lemma 3 and the polynomial growth bound on $f_a$ give us the estimate

$$K_1 \leq C(1 + |X^n|^{2r}).$$

From the definition of the $P_{X/(s)}$, we also have $K_2^i \leq C$.

Combining the above estimates, we obtain

$$F(t) \leq C(1 + |X^n|^{2r}) \Delta t^{\beta + 1}. \quad (31)$$

Step 2. The case $2 \leq l \leq 2(\beta + 1)$. We take the deterministic Taylor expansion of the function $F_{\tilde{p}}$ at the point $X_{t_n}^{n+1} - X^n$ to obtain

$$F(t) \leq \sum_{r=1}^l \frac{1}{r!} \sum_{k_1=1}^d \cdots \sum_{k_r=1}^d \hat{f}_{k_1, \ldots, k_r}(t), \quad (32)$$
where
\[ F_{k_1,\ldots,k_r}(t) = \left| \mathbb{E}^{X^n}_{t_n} \left( (X^n_{t_{n+1},k_1} - X^{n+1,k_1}) \times \cdots \times (X^n_{t_{n+1},k_r} - X^{n+1,k_r}) \right) \times \frac{\partial^r}{\partial x^{k_1} \cdots \partial x^{k_r}} F_{\bar{\rho}}(X^{n+1} - X^n) \right|. \] (33)

(1) First we estimate \( F_{k_1,\ldots,k_r} \) for \( r = 1 \), that is, \( F_k \) for \( k = 1, \ldots, q \). From the definition of \( F_{\bar{\rho}} \) there exists a \( \bar{\rho} \in P_{l-1} \) and a \( q \in \{1, 2, \ldots, l\} \), such that
\[ \frac{\partial}{\partial x} F_{\bar{\rho}}(X^{n+1} - X^n) = qF_{\bar{\rho}}(X^{n+1} - X^n). \]

By (30), (33), and the definition of \( R^n_X \) in (25), we deduce
\[ F_k(t) = \left| \mathbb{E}^{X^n}_{t_n} \left[ \sum_{a \in B(G_p)} I_a \left[ f^k_a(\cdot, X^n_{t_{n+1}}) \right]_{t_{n+1}} \times q \sum_{j=0}^{d} \int_{t_n}^{t_{n+1}} \sigma^\alpha_{\bar{\rho}}(\cdot) dW^j + \sum_{i=1}^{d} I(i) [P(\cdot)]_{t_{n+1}} \right] \right| \leq \sum_{a \in B(G_p)} \sum_{j=0}^{d} \sum_{i=1}^{d} q \left( \mathbb{E}^{X^n}_{t_n} \left[ I_a \left[ f^k_a(\cdot, X^n_{t_{n+1}}) \right]_{t_{n+1}} \times I(i) [P(\cdot)]_{t_{n+1}} \right] \right) \left( \mathbb{E}^{X^n}_{t_n} \left[ I(i) [P(\cdot)]_{t_{n+1}} \right] \right) \right|^{\frac{1}{2}}. \] (34)

For \( \alpha \in B(G_p) \), using Lemma 1 and the polynomial growth bound on \( f_a \), we have the estimate,
\[ \left( \mathbb{E}^{X^n}_{t_n} \left[ I_a \left[ f^k_a(\cdot, X^n_{t_{n+1}}) \right]_{t_{n+1}} \right] \right)^{\frac{1}{2}} \leq C \sup_{t_n \leq s \leq t_{n+1}} \mathbb{E}^{X^n}_{t_n} \left[ |f^k_a(s, X^n_{t_{n+1}})|^2 \right] (\Delta t)^{\beta+1} \]
\[ \leq C(1 + |X^n|^{2r})(\Delta t)^{\beta+1}. \] (35)

Then, using Lemma 2, we have
\[ \left( \mathbb{E}^{X^n}_{t_n} \left[ I(i) [\sigma^\alpha_{\bar{\rho}}(\cdot)]_{t_{n+1}} \right] \right)^{\frac{1}{2}} \leq C \left( \mathbb{E}^{X^n}_{t_n} \left[ \sup_{t_n \leq s \leq t_{n+1}} |\sigma^\alpha_{\bar{\rho}}(s)|^4 \right] \right)^{\frac{1}{2}} (\Delta t)^{\frac{1}{2}}, \] (36)
and
\[ \left( \mathbb{E}^{X^n}_{t_n} \left[ |I(i) [P(\cdot)]_{t_{n+1}}| \right] \right)^{\frac{1}{2}} \leq C \left( \mathbb{E}^{X^n}_{t_n} \left[ \sup_{t_n \leq s \leq t_{n+1}} |P(s)|^4 \right] \right)^{\frac{1}{2}} (\Delta t)^{\frac{1}{2}}. \] (37)

Combining the estimates (32)–(37), we deduce
\[ F(t) \leq C(1 + |X^n|^{2r})(\Delta t)^{\beta+2}. \] (38)

(2) Now we estimate \( F_{k_1,\ldots,k_r}(t) \) for \( 2 \leq r \leq l \). For each \((k_1, \ldots, k_r) \in P_r \) there is a finite \( q \in \{1, 2, \ldots, l\} \) and \( \bar{\rho} \in P_{l-r} \) such that
\[ \frac{\partial^r}{\partial x^{k_1} \cdots \partial x^{k_r}} F_{\bar{\rho}}(X^{n+1} - X^n) = qF_{\bar{\rho}}(X^{n+1} - X^n). \]
Thus, using the upper equality and Hölder inequality to (33), we obtain
\[
\dot{f}_{k_1, \ldots, k_n}(t) = \left| E^n_{x^n} \left( (X^n_{t_{n+1}} - X^n_{t_n}) \ldots (X^n_{t_{n+1}} - X^n_{t_n}) \right) \times \frac{\partial}{\partial x^n} F_p(X^n - X^n) \times \Delta W_{t_{n+1}} \right|
\leq q \left( E^n_{x^n} \left| F_p(X^n - X^n) \right|^2 \right)^{1/2} \left( E^n_{x^n} \left| \sum_{i=1}^{d} I_i(\hat{P}(\cdot))_{t_{n+1}} \right|^4 \right)^{1/4} \left( \frac{d}{\Delta t} \right)^{1/2} \left( E^n_{x^n} \left| \sum_{i=1}^{d} I_i(\hat{P}(\cdot))_{t_{n+1}} \right|^2 \right)^{1/2} \left( E^n_{x^n} \left| X^n_{t_{n+1}} - X^n_{t_n} \right|^{2\beta} \right)^{1/2}.
\]

By Lemma 3, the polynomial growth bound on \( f_\alpha \) and (26), for each \( k = 1, \ldots, q \) and \( l = 4, \ldots, 2^\beta + 1 \), we deduce
\[
\left( E^n_{x^n} \left| X^n_{t_{n+1}} - X^n_{t_n} \right|^{2\beta} \right)^{1/2} = \left( E^n_{x^n} \left| \sum_{a \in B(T_f)} I_a(\hat{P}(\cdot))_{t_{n+1}} \right|^2 \right)^{1/2} \leq \sum_{a \in B(T_f)} \left( E^n_{x^n} \left| I_{\alpha}(\hat{P}(\cdot))_{t_{n+1}} \right|^2 \right)^{1/2} \leq \sum_{a \in B(T_f)} C \Delta t^{\beta(n)/2} \left( E^n_{x^n} \left| \sup_{t_{n+1} \leq t \leq t_{n+1}} |\tilde{f}_a(s, X^n_{t})|^{2\beta} \right|^2 \right)^{1/2} \leq \sum_{a \in B(T_f)} C \Delta t^{(\beta+1)/2} \left( 1 + E^n_{x^n} \left| \sup_{t_{n+1} \leq t \leq t_{n+1}} |X^n_{t}|^{2\beta} \right|^2 \right)^{1/2} \leq C(1 + |X^n|^{2\beta})(\Delta t)^{\beta+1/2}.
\]

Then by Lemma 5, (37), (39), and (40), we obtain
\[
\dot{f}_{k_1, \ldots, k_n}(t) \leq C(1 + |X^n|^{2\beta})(\Delta t)^{(\beta+1)/2}(\Delta t)^{\beta+1/2}(\Delta t)^{\beta+1/2} \leq C(1 + |X^n|^{2\beta})(\Delta t)^{\beta+3/2}.
\]

Hence, from (32) and (41), we have the estimate
\[
\dot{f}(t) \leq C(1 + |X^n|^{2\beta})(\Delta t)^{\beta+3/2}.
\]

At last, by (31), (38), and (42), we complete the proof. \( \square \)

**Proposition 1.** Under the conditions of Lemma 3, for the sufficiently small time-step \( \Delta t \) satisfying (3), there exist constants \( C \in (0, \infty) \) and \( r \in \{1, 2, \ldots\} \) such that
\[
|E^n_{x^n} \left( g(X^n_{t_{n+1}}) - g(X^n_{t_n}) \right) | \leq C(1 + |X^n|^{2r})(\Delta t)^{\beta+1},
\]
where \( \Delta W_{t_{n+1}}(0 \leq n \leq N - 1) \) can be defined by (5) with \( s = t_{n+1} \).
Proof. For the function $g \in C^2_{p(\beta+1)}$, we use the deterministic Taylor expansion to obtain

$$
\mathbb{E}^{X^\alpha}_{t_n} \left[ (g(X^\alpha_{t_{n+1}}) - g(X^\alpha_{n+1})) \Delta \tilde{W}_{t_{n+1}} \right] = \mathbb{E}^{X^\alpha}_{t_n} \left[ (g(X^\alpha_{t_{n+1}}) - g(X^\alpha_{n+1})) \right] \Delta \tilde{W}_{t_{n+1}} + \mathbb{E}^{X^\alpha}_{t_n} \left[ (g(X^\alpha_{t_{n+1}}) - g(X^\alpha_{n+1})) \right] \Delta \tilde{W}_{t_{n+1}}
$$

Applying Proposition 1, we complete the proof.

Now, by (43), we obtain

$$
R_0(Z) = \frac{1}{2(\beta+1)!} \sum_{\beta \in \mathbb{P}_{2(\beta+1)}} \left[ \partial^\beta g(X^n - \Theta^\beta(Z - X^n)) \right] \times F_p(Z - X^n),
$$

for $Z = X^\alpha_{t_{n+1}}$ and $X^{n+1}$, respectively, where $\Theta^\beta(Z)$ is a $q \times q$ diagonal matrix with diagonal components

$$
\Theta^k_{\beta}(Z) = \hat{\Theta}^k_{\beta}(Z) \in [0, 1],
$$

for $k = 1, 2, \ldots, q$.

Now, by (43), we obtain

$$
\mathbb{E}^{X^\alpha}_{t_n} \left[ (g(X^\alpha_{t_{n+1}}) - g(X^\alpha_{n+1})) \Delta \tilde{W}_{t_{n+1}} \right] \leq \sum_{\beta \in \mathbb{P}_{2(\beta+1)}} \left[ \partial^\beta g(X^n) \times \mathbb{E}^{X^\alpha}_{t_n} \left[ (F_p(X^\alpha_{t_{n+1}} - X^n) - F_p(X^{n+1} - X^n)) \Delta \tilde{W}_{t_{n+1}} \right] \right] + \mathbb{E}^{X^\alpha}_{t_n} \left[ R_0(X^\alpha_{t_{n+1}}) \Delta \tilde{W}_{t_{n+1}} \right] + \mathbb{E}^{X^\alpha}_{t_n} \left[ R_0(X^{n+1}) \Delta \tilde{W}_{t_{n+1}} \right].
$$

Then, by using (44), (45), and applying Lemma 5, and the polynomial growth bound on the derivatives of $g$, we obtain

$$
\mathbb{E}^{X^\alpha}_{t_n} \left[ |R_0(X^\alpha_{t_{n+1}}) \Delta \tilde{W}_{t_{n+1}}| \right] \leq C(1 + |X^n|^{2r})(\Delta t)^{\beta+1}.
$$

Applying Lemma 6 and combining (46) with (47), we complete the proof.

Lemma 7. Under the conditions of Lemma 3, then for the sufficiently small time-step $\Delta t_n$ satisfying (3), we have that for any $0 \leq n \leq N - 1$,

$$
\mathbb{E}[|R_2^n|^2] \leq C(1 + \mathbb{E}[|X|^{2r})](\Delta t)^{2\beta+2}, \quad \mathbb{E}[|R_2^n|^2] \leq C(1 + \mathbb{E}[|X||^{2r}])/(\Delta t)^{2\beta+2}.
$$

Proof. We know that

$$
R_{s_1}^n = \mathbb{E}^{X^\alpha}_{t_n} \left[ (Y^\alpha_{t_{n+1}} - \tilde{Y}^\alpha_{t_{n+1}} \Delta \tilde{W}_{t_{n+1}}) \right], \quad R_{s_2}^n = \mathbb{E}^{X^\alpha}_{t_n} \left[ (f^\alpha_{t_{n+1}} - \tilde{f}^\alpha_{t_{n+1}} \Delta \tilde{W}_{t_{n+1}}) \right].
$$

Applying Proposition 1, we complete the proof.

3.3. Error Estimates

In order to give the error estimates, we also need the convergence order of the truncated error terms $R_2^n$ and $R_2^n$ in (12) and (16) for solving $Y_t$ and $Z_t$ in the BSDE in (2) by the discretizations (17) and (18) in Scheme 1. In the following lemma, we give the convergence order for $R_2^n$ and $R_2^n$. 


Lemma 8. If \( f(t, x, y, z) \in C_b^{2,4,4,4}, b(t, x), \sigma(t, x) \in C_b^{2,4}, \varphi \in C_b^{4+\alpha}, \alpha \in (0, 1) \) and \( |b(t, x)|^2 \leq K(1 + |x|^2), \) \( |\sigma(t, x)|^2 \leq K(1 + |x|^2), \) then for the sufficiently small time-step \( \Delta t_n \) satisfying (3), we have that for any \( 0 \leq n \leq N - 1, \)
\[
E[|R^n_y|^2] \leq C(1 + E[|X^n|^8])(\Delta t)^6, \quad E[|R^n_z|^2] \leq C(1 + E[|X^n|^8])(\Delta t)^6,
\]
where \( C \) is a positive constant depending only on \( T, K, \) and upper bounds of the derivatives of \( b, \sigma, \) \( f \) and \( \varphi. \)

For the proof, the reader is referred to Lemmas 4.2 and 4.5 in [16].

Theorem 2. Assume the conditions of Lemma 3 hold and the initial values satisfy \( E[|e^n_y|^2] = O((\Delta t)^2) \) and \( E[|e^n_z|^2] = O((\Delta t)^2) \). If \( f(t, x, y, z) \in C_b^{2,4,4,4}, b(t, x), \sigma(t, x) \in C_b^{2,4}, \varphi \in C_b^{4+\alpha}, \alpha \in (0, 1) \) and \( |b(t, x)|^2 \leq K(1 + |x|^2), \) \( |\sigma(t, x)|^2 \leq K(1 + |x|^2). \) Then for the sufficiently small time-step \( \Delta t_n \) satisfying (3), we have that for any \( 0 \leq n \leq N - 1, \)
\[
E[|e^n_i|^2] + \Delta t \sum_{i=1}^{N-1} E[|e^n_i|^2] \leq C(\Delta t^{2\beta} + \Delta t^4),
\]
where \( C \) is also a positive constant depending on \( c_0, T, L, K, \) the initial value of \( (X_t)_{t \in [0, T]} \) in (1), and the upper bounds of the derivatives of \( b, \sigma, f \) and \( \varphi. \)

Proof. According to Lemmas 4, 7, and 8, we get
\[
\sum_{i=1}^{N-1} \left( 1 + \frac{C\Delta t}{1 - C\Delta t} \right)^{-i-n} \frac{CE[|R^n_y|^2 + (\Delta t)^2|R^n_{y2}|^2 + |R^n_y|^2]}{\Delta t (1 - C\Delta t)} \leq C(1 + E[|X^n|^4] + E[|X^n|^8])(\Delta t)^{2\beta} + (\Delta t)^4,
\]
and
\[
\sum_{i=1}^{N-1} \left( 1 + \frac{C\Delta t}{1 - C\Delta t} \right)^{-i-n} \frac{C\Delta t E[|\frac{1}{\Delta t}|^2|R^n_{y2}|^2 + |R^n_{y2}|^2 + (\frac{1}{\Delta t})^2|R^n_{y2}|^2]}{(1 - C\Delta t)} \leq C(1 + E[|X^n|^4] + E[|X^n|^8])(\Delta t)^{2\beta} + (\Delta t)^4.
\]
Now by Theorem 1, and the estimates (50) and (51), we complete the proof.

Scheme 2. Let \( K = \max\{K_y, K_z, K_f\}. \) Given random variables \( X^0, Y^{N-1}, \) and \( Z^{N-1}, i = 0, 1, \ldots, K - 1. \) Let \( \{X^n\}_{n=0}^N \) be the numerical solution of the forward SDE in the decoupled FBSDEs by a numerical method for solving the SDE. For \( n = N - K, \ldots, 0, \)

1. solve \( Z^n \) by
\[
Z^n = \frac{1}{\Delta t_n} E^n_{i_n} \left[ Y^{n+1} + \Delta \hat{W}^*_{n,1} \right] + \sum_{i=1}^{K_y} b^n_{i, \epsilon} E^n_{i_n} \left[ f^{n+1} \hat{W}^*_{n,i} \right],
\]
2. solve \( Y^n \) by
\[
Y^n = E^n_{i_n} \left[ Y^{n+1} + \Delta t_n \sum_{i=0}^{K_y} b^n_{i, \epsilon} E^n_{i_n} \left[ f^{n+1} \right] \right].
\]

Regarding the proposed Scheme 2, the reader is referred to Scheme 1 in [17].

Theorem 3. Under the conditions of Lemma 3, furthermore, suppose the initial values satisfy \( \max_{0 \leq i \leq N} E[|e^n_i|^2] = O((\Delta t)^{K_y+1}) \) and \( \max_{0 \leq i \leq N} E[|e^n_i|^2] = O((\Delta t)^{K_y+1-K_z} + (\Delta t)^{K_z+1}). \)
Let $K = \max(K_y, K_z, K_f)$; if the order-$\beta$ weak Taylor Scheme is used to solve the SDE (1) in Scheme 2, then for $0 \leq n \leq N-K$, we have the following error estimate:

$$E[|e^n|^2] + \Delta t \sum_{i=n}^{N-K} E[|e^i|^2] \leq C [(\Delta t)^{2\beta} + (\Delta t)^{2K_y+2} + (\Delta t)^{2K_z+2}]$$

(54)

where $C$ is a constant depending on $c_0, T, L, K$, the initial value of $(X_t)_{t \in [0, T]}$ in (1), and the upper bounds of the derivatives of $b, \sigma, f, \phi$.

Since the proof of the theorem essentially uses the argument developed in the proof of Theorem 2, we omit it.

Now we introduce some classical numerical schemes in the form of (2.2) that can be used in Scheme 1 for solving the forward SDE.

Example 1. The Euler scheme is given by

$$X^{n+1} = X^n + b(t_n, X^n) \Delta t_n + \sigma(t_n, X^n) \Delta W_{n+1}.$$  

We know the Euler scheme with $\beta = 1$, so we can obtain the error estimate as

$$E[|e^n|^2] + \Delta t \sum_{i=n}^{N-1} E[|e^i|^2] \leq C(\Delta t)^2.$$  

Example 2. The order-2.0 weak Taylor scheme is given by

$$X^{n+1} = X^n + b(t_n, X^n) \Delta t_n + \sigma(t_n, X^n) \Delta W_{n+1}$$

$$+ L^0 b(t_n, X^n) \int_{t_n}^{t_{n+1}} \int_{t_n}^{t} 1 dW_r ds + L^1 b(t_n, X^n) \int_{t_n}^{t_{n+1}} \int_{t_n}^{t} 1 dW_r dW_s$$

$$+ L^0 \sigma(t_n, X^n) \int_{t_n}^{t_{n+1}} \int_{t_n}^{t} 1 dW_r + L^1 \sigma(t_n, X^n) \int_{t_n}^{t_{n+1}} \int_{t_n}^{t} 1 dW_r dW_s.$$  

We know the order-2.0 weak Taylor scheme with $\beta = 2$, so we can obtain the error estimate as

$$E[|e^n|^2] + \Delta t \sum_{i=n}^{N-1} E[|e^i|^2] \leq C(\Delta t)^4.$$  

4. Conclusions

We first investigated the weak convergence analysis about the error terms which are determined by the discretization for solving the forward equation in FBSDEs, and we showed that the solution of BSDE $(Y, Z)$ admits a higher-order convergence rate. In most present studies, error estimates can be obtained with an analytic solution of the forward equation in FBSDEs, whereas we could not exactly solve the forward equation in many cases. Therefore, it is very important to study the weak convergence of the numerical solution of the associated discretization schemes. Recently, there has been a lot of research about the numerical schemes and error estimates of FBSDEs with jumps [19–22], corresponding to a rigorous weak convergence analysis of numerical schemes for general FBSDEs where jumps are a focus for ongoing work.

Author Contributions: W.Z. is mainly responsible for providing methods, and H.M. is mainly responsible for scientific research. All authors have read and agreed to the published version of the manuscript.

Funding: The work is supported by the National Natural Science Foundation of China (No. 12001023).

Conflicts of Interest: The authors declare no conflict of interest.
References
1. Antonelli, F. Backwardforward stochastic differential equations. *Ann. Appl. Probab.* **1993**, *3*, 777–793. [CrossRef]
2. Karoui, N.; Peng, S.; Quenez, M. Backward stochastic differential equations in finance. *Math. Financ.* **1997**, *7*, 1–71. [CrossRef]
3. Ma, J.; Yong, J. *Forward-Backward Stochastic Differential Equations and Their Applications*; Lecture Notes in Math.; Springer: Berlin/Heidelberg, Germany, 1999.
4. Pardoux, E.; Peng, S. Adapted solution of a backward stochastic differential equation. *Syst. Control Lett.* **1990**, *14*, 55–61. [CrossRef]
5. Peng, S. Backward SDE and related g-expectation, in backward stochastic differential equations (Paris, 1995–1996). In *Pitman Research Notes in Mathematics Series*; Longman: Harlow, UK, 1997; Volume 364, pp. 141–159.
6. Bender, C.; Denk, R. A forward scheme for backward SDEs. *Stoch. Process. Their Appl.* **2007**, *117*, 1793–1812. [CrossRef]
7. Bouchard, B.; Touzi, N. Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stoch. Process. Their Appl.* **2004**, *111*, 175–206. [CrossRef]
8. Gobet, E.; Labart, C. Error expansion for the discretization of backward stochastic differential equations. *Stoch. Process. Their Appl.* **2007**, *117*, 803–829. [CrossRef]
9. Milstein, G.; Tretyakov, M. Numerical algorithms for forward-backward stochastic differential equations. *SIAM J. Sci. Comput.* **2006**, *28*, 561–582. [CrossRef]
10. Zhao, W.; Zhang, W.; Ju, L. A stable multistep scheme for solving backward stochastic differential equations. *SIAM J. Numer. Anal.* **2010**, *48*, 1369–1394. [CrossRef]
11. Zhao, W.; Zhang, W.; Ju, L. A new kind of accurate numerical method for forward-backward stochastic differential equations. *Commun. Comput. Phys.* **2011**, *9*, 1–34. [CrossRef]
12. Zhao, W.; Zhang, W.; Ju, L. Multistep schemes for decoupled forward-backward stochastic differential equations. *J. Sci. Comput.* **2014**, *60*, 108–128. [CrossRef]
13. Zhao, W.; Zhang, W.; Ju, L. Second order numerical schemes for decoupled forward-backward stochastic differential equations. *J. Sci. Comput.* **2015**, *64*, 1–22. [CrossRef]
14. Zhao, W.; Zhang, W.; Ju, L. Multistep schemes for decoupled forward-backward stochastic differential equations with jumps. *J. Sci. Comput.* **2016**, *68*, 185–217. [CrossRef]
15. Zhao, W.; Zhang, W.; Ju, L. Second order numerical schemes for decoupled forward-backward stochastic differential equations with jumps. *J. Sci. Comput.* **2017**, *72*, 213–244. [CrossRef]