GENERAL FULL IMPLICIT STRONG TAYLOR APPROXIMATIONS FOR STIFF STOCHASTIC DIFFERENTIAL EQUATIONS

Kai Liu* 1)  
School of Mathematics, Shanghai University of Finance and Economics, Shanghai 200433, China;  
College of Science, Hunan University Of Technology and Business, Changsha 410205, China.  
Email: 732964253@qq.com  
Guiding Gu  
School of Mathematics, Shanghai University of Finance and Economics, Shanghai 200433, China.  
Email: guiding@mail.shufe.edu.cn  

Abstract  
In this paper, we present the backward stochastic Taylor expansions for a Ito process, including backward Ito-Taylor expansions and backward Stratonovich-Taylor expansions. We construct the general full implicit strong Taylor approximations (including Ito-Taylor and Stratonovich-Taylor schemes) with implicitness in both the deterministic and the stochastic terms for the stiff stochastic differential equations (SSDE) by employing truncations of backward stochastic Taylor expansions. We demonstrate that these schemes will converge strongly with corresponding order 1, 2, 3, . . . . Mean-square stability has been investigated for full implicit strong Stratonovich-Taylor scheme with order 2, and it has larger mean-square stability region than the explicit and the semi-implicit strong Stratonovich-Taylor schemes with order 2. We can improve the stability of simulations considerably without too much additional computational effort by using our full implicit schemes. The full implicit strong Taylor schemes allow a larger range of time step sizes than other schemes and are suitable for SSDE with stiffness on both the drift and the diffusion terms. Our numerical experiment show these points.  

Mathematics subject classification: 65C30, 60H35.  
Key words: Stiff stochastic differential equations, Approximations, Backward stochastic Taylor expansions, Full implicit Taylor methods.  

1. Introduction  
We are concerned with numerical methods for the solution of a d-dimensional vector Ito stochastic differential equation  
\[ dX_t = a(t, X_t)dt + b(t, X_t)dW_t, \]  
where \( \{W_t, t \in [0, T]\} \) is an m-dimensional Wiener process with components \( W^1_t, W^2_t, \ldots, W^m_t \), which are independent standard Wiener processes on a filtered probability space \( (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}) \), \( a \) is a d-dimensional vector function from \([0, T] \times \mathbb{R}^d\) to \( \mathbb{R}^d \) and \( b \) is a \( d \times m \)-matrix function from \([0, T] \times \mathbb{R}^d\) to \( \mathbb{R}^{d \times m} \). We interpret (1.1) as a stochastic integral equation  
\[ X_t = X_0 + \int_0^t a(s, X_s) \, ds + \sum_{j=1}^m \int_0^t b^j(s, X_s) \, dW^j_s \]  

* Received July 19, 2019 / Revised version received April 8, 2020 / Accepted November 23, 2020 /  
Published online April 6, 2021 /  
1) Corresponding author
with the $i$-th component being
\[ X_t^i = X_0^i + \int_0^t a^i(s, X_s) \, ds + \sum_{j=1}^m \int_0^t b^{i,j}(s, X_s) \, dW_s^j, \quad i = 1, \ldots, d, \tag{1.3} \]
where the stochastic integrals are Ito stochastic integrals. We call $X = \{X_t, t \in [0,T]\}$ an Ito process.

We also can transform (1.1) into the equivalent Stratonovich form
\[
\frac{dX_t}{dt} = \mathcal{g}(t, X_t) dt + b(t, X_t) \circ dW_t, \tag{1.4}
\]
with
\[
\mathcal{g}(t, X) = a(t, X) - \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^m b^{j,k}(t, X) \frac{\partial b^{i,k}(t, X)}{\partial x_j}, \quad i = 1, \ldots, d, \tag{1.5}
\]
and the form of stochastic integral equation is
\[
X_t = X_0 + \int_0^t \mathcal{g}(s, X_s) \, ds + \sum_{j=1}^m \int_0^t b^j(s, X_s) \circ dW^j_s \tag{1.6}
\]
with the $i$-th component being
\[
X_t^i = X_0^i + \int_0^t \mathcal{g}^i(s, X_s) \, ds + \sum_{j=1}^m \int_0^t b^{i,j}(s, X_s) \circ dW^j_s, \quad i = 1, \ldots, d, \tag{1.7}
\]
where the stochastic integrals are Stratonovich stochastic integrals. For more details, see [3].

The stochastic differential equation can be applied in many different fields. Examples include population dynamics, protein kinetics, genetics, experimental psychology, neuronal activity, investment finance, option pricing, turbulent diffusion, radio-astronomy, helicopter rotor, satellite orbit stability, biological waste treatment, hydrology, air quality, seismology, structural mechanics, fatigue cracking, optical bistability, nematic liquid crystals, blood clotting dynamics, cellular energetics, Josephson tunneling junctions, communications and stochastic annealing. For more examples and details, we refer to [3].

However, explicit solutions of Eqs. (1.1) are rare in practical applications and numerical methods are necessary. The most efficient and widely applicable approach to solving (1.1) is the simulation of sample paths of time discrete approximations, like Euler scheme, Milstein scheme and so on. In this paper, we focus on the time discrete approximations and consider a time discretization $(\tau)_{\Delta}$ with
\[
0 = \tau_0 < \tau_1 < \cdots < \tau_n < \cdots < \tau_N = T \tag{1.8}
\]
of a time interval $[0,T]$, which in the simplest equidistant case has step size
\[
\Delta = T/N. \tag{1.9}
\]
The simplest time discrete approximation is the Euler scheme. For (1.1), it has the form
\[
Y_{n+1} = Y_n + a(\tau_n, Y_n) \, \Delta + b(\tau_n, Y_n) \, \Delta W \tag{1.10}
\]
for $n = 0, 1, \ldots, N-1$ with initial value
\[
Y_0 = X_0, \tag{1.11}
\]
\[
\Delta W = W_{\tau_{n+1}} - W_{\tau_n}. \tag{1.12}
\]
$Y_n$ is a approximation of $X_{\tau_n}$, and we call $Y$ a time discrete approximation of $X$. It had been proved that under right conditions [3]

$$E(|X_T - Y_N|) \leq K\Delta^{1/2},$$

where the constant $K$ does not depend on $\Delta$. We say that $Y$ converges strongly to $X$ with order 0.5. More generally, we introduce a convergence definition which can be found in [3].

**Definition 1.1.** Let $Y$ be a time discrete approximation of an Ito process $X$, then $Y$ is said to converge strongly to $X$ with order $\gamma$ at time $T$ if there exists a positive constant $C > 0$, which does not depend on $\Delta$, and a $\Delta_0 > 0$ such that

$$E(|X_T - Y_N|) \leq C\Delta^{\gamma}$$

for each $\Delta \in (0, \Delta_0)$.

Based on the stochastic Ito-Taylor and Stratonovich-Taylor expansions [3], the general strong Taylor approximations [3] have been designed with order 1.0, 1.5, 2.0, ..., 5.0, 2.0, ..., 0.5, ..., such as the general strong Taylor approximations [3] have been designed with order 1.5. However, if the stochastic part plays an essential role in the dynamics, e.g., as it is with large multiplicative noise, the application of fully implicit methods also involving implicit stochastic terms is unavoidable [8]. Next we focus on this situation which stiffness of (1.1) also appearing in stochastic terms. Thus, we need fully implicit scheme involving implicit stochastic terms. Milstein et al. [8] proposed a balanced implicit method to hand it. For (1.1), it has the form

$$Y_{n+1} = Y_n + a(\tau_{n+1}, Y_{n+1}) \Delta + b(\tau_n, Y_n) \Delta W.$$  

(1.15)

It is semi-implicit scheme and converges strongly with order 0.5, but $Y_{n+1}$ only appears in deterministic terms, so we call it semi-implicit scheme.

Analogously as the general strong Taylor approximations, based on the stochastic Ito-Taylor and Stratonovich-Taylor expansions [3], the general (semi)-implicit strong Taylor approximations [3] have been designed with order 1.0, 1.5, 2.0, ..., 5.0, 2.0, ..., 0.5, ..., 0. Thus, we need fully implicit scheme involving implicit stochastic terms. Milstein et al. [8] proposed a balanced implicit method to hand it. For (1.1), it has the form

$$Y_{n+1} = Y_n + a(\tau_n, Y_n) \Delta + \sum_{j=1}^{m} b^j(\tau_n, Y_n) \Delta W_n^j + C_n(Y_n - Y_{n+1}),$$  

(1.16)

where

$$C_n = c^0(\tau_n, Y_n) \Delta + \sum_{j=1}^{m} c^j(\tau_n, Y_n) |\Delta W_n^j|. $$

(1.17)

It has implicitness in stochastic terms and it has been proved that (1.16) converges strongly with order 0.5. Then, Tian et al. [9] presented implicit Euler-Taylor, implicit Milstein-Taylor and implicit order 1.5 strong Taylor schemes. For (1.1) with 1-dimensional Wiener process, implicit Milstein-Taylor scheme has the form

$$Y_{n+1} = Y_n + a(\tau_{n+1}, Y_{n+1}) \Delta + b(\tau_{n+1}, Y_{n+1}) \Delta W_n$$

$$-\frac{1}{2}((\Delta W_n)^2 + \Delta) b'(\tau_{n+1}, Y_{n+1}) b(\tau_{n+1}, Y_{n+1}).$$

(1.18)
We will see in Section 5 that it is a particular case of our general full implicit strong Taylor approximations with strong convergence order 1.

In recent years, many researches have been done about full implicit method. Wang et al. [10] presented the split-step backward balanced Milstein methods. Ahmad et al. [1] proposed the fully implicit stochastic-\(\alpha\) method. Wang et al. [11] presented a family of fully implicit Milstein methods. Haghighi et al. [2] proposed a class of split-step balanced methods. Mao et al. [5] presented the full-implicit truncated Euler-Maruyama method. These methods are all low strong convergence order which not exceeding 1.5. When we want a high precision and large step methods, we need to construct a high order scheme. In this paper, we shall consider general full implicit strong schemes which can get a higher strong convergence order.

Inspired by the deriving of stochastic Ito-Taylor expansions [3], we deduce the backward stochastic Ito-Taylor expansions. For convenience, we consider 1-dimensional autonomous situation

\[
X_t = X_0 + \int_0^t a(X_s) \, ds + \int_0^t b(X_s) \, dW_s.
\]  

(1.19)

Similar to the deriving of stochastic Ito-Taylor expansions [3], let

\[
L^0 = a \frac{\partial}{\partial x} + \frac{1}{2} b^2 \frac{\partial^2}{\partial x^2}, \quad L^1 = b \frac{\partial}{\partial x}.
\]

(1.20)

If we apply the Ito formula to (1.19), we obtain

\[
X_t = X_0 + \int_0^t \left( a(X_s) - \int_s^t L^0 a(X_v) \, dv - \int_s^t L^1 a(X_v) \, dW_v \right) \, ds
\]

\[
+ \int_0^t \left( b(X_s) - \int_s^t L^0 b(X_v) \, dv - \int_s^t L^1 b(X_v) \, dW_v \right) \, dW_s
\]

\[
= X_0 + a(X_t) \int_0^t ds + b(X_t) \int_0^t dW_s + R
\]

(1.21)

with the remainder

\[
R = - \int_0^t \int_s^t L^0 a(X_v) \, dv \, ds - \int_0^t \int_s^t L^1 a(X_v) \, dW_v \, ds
\]

\[
- \int_0^t \int_s^t L^0 b(X_v) \, dv \, dW_s - \int_0^t \int_s^t L^1 b(X_v) \, dW_v \, dW_s.
\]

(1.22)

Similar to (1.21), we apply the Ito formula to \(L^1 b(X_v)\), in which case we get

\[
X_t = X_0 + a(X_t) \int_0^t ds + b(X_t) \int_0^t dW_s - L^1 b(X_t) \int_0^t \int_s^t dW_v dW_s + \overline{R}
\]

(1.23)

with the remainder

\[
\overline{R} = - \int_0^t \int_s^t L^0 a(X_v) \, dv \, ds - \int_0^t \int_s^t L^1 a(X_v) \, dW_v \, ds
\]

\[
- \int_0^t \int_s^t L^0 b(X_v) \, dv \, dW_s - \int_0^t \int_s^t \int_v^t L^0 L^1 b(X_u) \, du \, dW_v dW_s
\]

\[
+ \int_0^t \int_s^t \int_v^t L^1 L^1 b(X_u) \, dW_u \, dW_v \, dW_s.
\]

(1.24)
If we delete the remainder of (1.23) and apply it on the time discretization \( \tau \), we could obtain a approximation of \( X_t \) that is the implicit Milstein-Taylor scheme (1.18).

Keeping on the above procedure, we could obtain higher order expansions and approximations of \( X_t \). The expansion is the backward stochastic Ito-Taylor expansion which will be formulate in Section 3 and the approximation is the full implicit strong Ito-Taylor schemes which we can see in Section 5. Analogously, we also can deduce the backward stochastic Stratonovich-Taylor expansions and the full implicit strong Stratonovich-Taylor schemes. Different from the stochastic Ito-Taylor expansion, multiple integrals in the expansion, which we will call backward multiple Ito integrals in Section 2.2, are not multiple Ito integrals which have been defined by (5.2.12) in [3]. Multiple Ito integrals have the property of zero expectation (except non-stochastic integrals), see Lemma 4.1, but backward multiple Ito integrals do not have. For instance, the multiple Ito integral in the stochastic Ito-Taylor expansions
\[
\int_0^t \int_0^s dW_v dW_s = \frac{1}{2} (W_t^2 - t),
\]
which expectation is zero, but the backward multiple Ito integral in (1.23)
\[
\int_0^t \int_s^t dW_v dW_s = \int_0^t \left( \int_0^t dW_v - \int_0^s dW_v \right) dW_s \\
= (W_t - W_0)^2 - \int_0^t \int_0^s dW_v dW_s \\
= \frac{1}{2} (W_t^2 + t),
\]
which expectation is \( t \). This makes the construction and proof of our full implicit strong Taylor schemes different from the general strong Taylor schemes [3].

In this paper, the general full implicit strong Taylor schemes are presented. Our paper is organized in the following way. First, in Section 2 we list notations that we may use in this paper. Most of these notations are from [3] and some new notations are introduced in this paper. Next we introduce the backward stochastic Taylor expansions in Section 3. In Section 4, we investigate the moments estimation of truncated error for the backward stochastic Taylor expansions, which is the key point about the convergence’s proof of the general full implicit strong Taylor schemes. Section 5 is devoted to the construction and proof of our full implicit schemes. We construct the general full implicit strong Taylor schemes, and prove them will converge strongly with corresponding order 1, 2, 3, . . . . We discuss the mean-square stability of full implicit order 2 strong Stratonovich-Taylor scheme for linear test equation in Section 6. It has better mean-square stability region compared with the order 2 strong explicit and semi-implicit Stratonovich-Taylor scheme. Numerical experiments on full implicit order 2 strong Stratonovich-Taylor scheme are conducted in Section 7, which show our scheme with order 2 has better stability compared with the order 2 strong explicit and semi-implicit Stratonovich-Taylor scheme.

2. Notations

In this section, we shall list some notations that we will use in the following section. Most of these notations are used by Kloeuden and Platen in [3, Chapter 5]. Some new notations are introduced for this paper.
2.1. Multi-index

Set $\alpha$ is a multi-index with

$$\alpha = (j_1, j_2, \ldots, j_l), \quad j_i \in \{0, 1, \ldots, m\}, \quad i \in \{1, 2, \ldots, l\},$$  \hspace{1cm} (2.1)

where $m$ is the dimension of Wiener process in (1.1).

**Length of Multi-index:** Set $l(\alpha)$ represent the length of multi-index $\alpha$ with

$$l(\alpha) = l$$  \hspace{1cm} (2.2)

for $\alpha$ defined by (2.1) and set $v$ is the multi-index of length zero with

$$l(v) = 0.$$  \hspace{1cm} (2.3)

**Zero in Multi-index:** Set $n(\alpha)$ represent the number of components of a multi-index $\alpha$ which are equal to 0. For instance, $n((1, 0, 1, 1, 0)) = 2$, $n((1, 0, 1, 1, 0, 0)) = 3$.

**Operation of Multi-index:** For $l(\alpha) \geq k$, $k \in \mathbb{N}$, set $k - \alpha$ represent the multi-index which is deleted the first $k$ components of $\alpha$ and $\alpha - k$ represent the multi-index which is deleted the last $k$ components of $\alpha$. We write $k - \alpha$ as $-\alpha$ and $\alpha - k$ as $\alpha - k$ when $k = 1$. For instance, if $\alpha = (1, 2, 3, 4)$, then $2 - \alpha = (3, 4)$ and $\alpha - 2 = (1, 2, 3)$.

$$\alpha * \overline{\alpha} = (j_1, j_2, \ldots, j_k, j_{1l}, j_2, \ldots, j_l).$$  \hspace{1cm} (2.4)

**Set of Multi-index:** We denote the set of all multi-indices by $\mathcal{M}$, so

$$\mathcal{M} = \{ (j_1, j_2, \ldots, j_l) : j_i \in \{0, 1, \ldots, m\}, i \in \{1, \ldots, l\} \} \cup \{v\}.$$  \hspace{1cm} (2.5)

We call a subset $\mathcal{A} \subset \mathcal{M}$ an hierarchical set with

$$\mathcal{A} \neq \emptyset, \quad \sup_{\alpha \in \mathcal{A}} l(\alpha) < \infty, \quad -\alpha \in \mathcal{A} \quad \text{for each} \quad \alpha \in \mathcal{A} \setminus \{v\}.\hspace{1cm} (2.6)$$

For any given hierarchical set $\mathcal{A}$ we define the remainder set $\mathcal{B}(\mathcal{A})$ by

$$\mathcal{B}(\mathcal{A}) = \{ \alpha \in \mathcal{M} \setminus \mathcal{A} : -\alpha \notin \mathcal{A} \}.$$  \hspace{1cm} (2.7)

2.2. Multiple Integral

For $\alpha = (j_1, j_2, \ldots, j_l)$, $\mathcal{H}_\alpha$ and $\overline{\mathcal{H}}_\alpha$ are the functions sets defined in [3, Chapter 5].

**Multiple Ito Integral:**

$$I_{\alpha}[f(\cdot, X_{\cdot})]_{\rho, \tau} := \begin{cases} f(\tau, X_{\tau}), & l = 0, \\ \int_{\rho}^{\tau} I_{\alpha_{-}}[f(\cdot, X_{\cdot})]_{\rho, s} ds, & l \geq 1, \quad j_l = 0, \\ \int_{\rho}^{\tau} I_{\alpha_{-}}[f(\cdot, X_{\cdot})]_{\rho, s} dW^h_s, & l \geq 1, \quad j_l \geq 1, \end{cases}$$  \hspace{1cm} (2.8)

where $f \in \mathcal{H}_\alpha$. 


Multiple Backward Ito Integral:

\[
K_\alpha[f(\cdot, X)]_{\rho,\tau} := \begin{cases} f(\rho, X_{\rho}), & l = 0, \\ \int_\rho^\tau K_\alpha[f(\cdot, X)]_{s,\tau}ds, & l \geq 1, \quad j_l = 0, \end{cases} \tag{2.9}
\]

where \( f \in H_\alpha \).

Multiple Stratonovich Integral:

\[
J_\alpha[f(\cdot, X)]_{\rho,\tau} := \begin{cases} f(\tau, X_{\tau}), & l = 0, \\ \int_\rho^\tau J_\alpha[f(\cdot, X)]_{s,\rho}ds, & l \geq 1, \quad j_l = 0, \end{cases} \tag{2.10}
\]

where \( f \in H_\alpha \).

Multiple Backward Stratonovich Integral:

\[
M_\alpha[f(\cdot, X)]_{\rho,\tau} := \begin{cases} f(\rho, X_{\rho}), & l = 0, \\ \int_\rho^\tau M_\alpha[f(\cdot, X)]_{s,\rho}ds, & l \geq 1, \quad j_l = 0, \end{cases} \tag{2.11}
\]

where \( f \in H_\alpha \).

When \( f = 1 \) and \( \rho, \tau \) are not confused, we write

\[
I_\alpha = I_{\alpha,\rho,\tau} := I_\alpha[f(\cdot, X)]_{\rho,\tau}, \tag{2.12a}
\]

\[
J_\alpha = J_{\alpha,\rho,\tau} := J_\alpha[f(\cdot, X)]_{\rho,\tau}, \tag{2.12b}
\]

\[
K_\alpha = K_{\alpha,\rho,\tau} := K_\alpha[f(\cdot, X)]_{\rho,\tau}, \tag{2.12c}
\]

\[
M_\alpha = M_{\alpha,\rho,\tau} := M_\alpha[f(\cdot, X)]_{\rho,\tau}. \tag{2.12d}
\]

Relations Between Multiple Integral: For \( \alpha = (j_1, j_2, \ldots, j_l), l \geq 1, \) from the definition of multiple integral, by the fact of \( K_{(j)}[f(\cdot, X)]_{\rho,\tau} = I_{(j)}[f(\cdot, X)]_{\rho,\tau} \) we have

\[
K_\alpha[f(\cdot, X)]_{\rho,\tau} = K_{-\alpha}\left[K_{(j_1)}[f(\cdot, X)]_{\cdot,\tau}\right]_{\rho,\tau} = K_{-\alpha}\left[I_{(j_1)}[f(\cdot, X)]_{\cdot,\tau}\right]_{\rho,\tau}
\]

\[
= K_{-\alpha}\left[I_{(j_1)}[f(\cdot, X)]_{\rho,\tau} - I_{(j_1)}[f(\cdot, X)]_{\rho,\tau}\right]_{\rho,\tau}
\]

\[
= I_{(j_1)}[f(\cdot, X)]_{\rho,\tau} K_{-\alpha,\rho,\tau} = K_{-\alpha}\left[I_{(j_1)}[f(\cdot, X)]_{\rho,\tau}\right]_{\rho,\tau}. \tag{2.13}
\]

Repeating iteration of (2.13), we obtain

\[
K_\alpha[f(\cdot, X)]_{\rho,\tau} = I_{(j_1)}[f(\cdot, X)]_{\rho,\tau} K_{-\alpha} - I_{(j_1,j_2)}[f(\cdot, X)]_{\rho,\tau} K_{2-\alpha} + I_{(j_1,j_2,j_3)}[f(\cdot, X)]_{\rho,\tau} K_{3-\alpha} - \cdots - (-1)^k I_{(j_1,j_2,\ldots,j_k)}[f(\cdot, X)]_{\rho,\tau} K_{k-\alpha} - \cdots - (-1)^l I_\alpha[f(\cdot, X)]_{\rho,\tau}. \tag{2.14}
\]
Repeating iteration of (2.14) until eliminating all $K$ in (2.14), we obtain

$$K_\alpha [f(\cdot, X.)]|_{\rho, \tau} = \sum_{\alpha_1 + \cdots + \alpha_k = \alpha} (-1)^{(\alpha)-k} I_{\alpha_1} [f(\cdot, X.)]|_{\rho, \tau} I_{\alpha_2} I_{\alpha_3} \cdots I_{\alpha_k}. \quad (2.15)$$

For $f \equiv 1$,

$$K_\alpha = I_{(j_1)} K_{-\alpha} - I_{(j_1, j_2)} K_{2-\alpha} + I_{(j_1, j_2, j_3)} K_{3-\alpha} - \cdots - (-1)^k I_{(j_1, j_2, \ldots, j_k)} K_{k-\alpha} - \cdots - (-1)^{\alpha} I_\alpha, \quad (2.16)$$

$$K_\alpha = \sum_{\alpha_1 + \cdots + \alpha_k = \alpha} (-1)^{(\alpha)-k} I_{\alpha_1} I_{\alpha_2} I_{\alpha_3} \cdots I_{\alpha_k}. \quad (2.17)$$

Analogous to above, we have

$$M_\alpha [f(\cdot, X.)]|_{\rho, \tau} = J_{(j_1)} [f(\cdot, X.)]|_{\rho, \tau} M_{-\alpha} - J_{(j_1, j_2)} [f(\cdot, X.)]|_{\rho, \tau} M_{2-\alpha} + J_{(j_1, j_2, j_3)} [f(\cdot, X.)]|_{\rho, \tau} M_{3-\alpha} - \cdots - (-1)^k J_{(j_1, j_2, \ldots, j_k)} [f(\cdot, X.)]|_{\rho, \tau} M_{k-\alpha} - \cdots - (-1)^{\alpha} J_\alpha [f(\cdot, X.)]|_{\rho, \tau}, \quad (2.18)$$

$$M_\alpha [f(\cdot, X.)]|_{\rho, \tau} = \sum_{\alpha_1 + \cdots + \alpha_k = \alpha} (-1)^{(\alpha)-k} J_{\alpha_1} [f(\cdot, X.)]|_{\rho, \tau} J_{\alpha_2} J_{\alpha_3} \cdots J_{\alpha_k}, \quad (2.19)$$

$$M_\alpha = J_{(j_1)} M_{-\alpha} - J_{(j_1, j_2)} M_{2-\alpha} + J_{(j_1, j_2, j_3)} M_{3-\alpha} - \cdots - (-1)^k J_{(j_1, j_2, \ldots, j_k)} M_{k-\alpha} - \cdots - (-1)^{\alpha} J_\alpha, \quad (2.20)$$

$$M_\alpha = \sum_{\alpha_1 + \cdots + \alpha_k = \alpha} (-1)^{(\alpha)-k} J_{\alpha_1} J_{\alpha_2} J_{\alpha_3} \cdots J_{\alpha_k}. \quad (2.21)$$

### 2.3. Coefficient Functions

Let $m$ be the dimension of Wiener process and $d$ be the dimension of Ito process $X_1$.

**Ito Coefficient Functions:**

$$L^0 = \frac{\partial}{\partial t} + \sum_{k=1}^{d} a_k \frac{\partial}{\partial x^k} + \frac{1}{2} \sum_{k,l=1}^{d} \sum_{j=1}^{m} b^{k,j} b^{l,j} \frac{\partial^2}{\partial x^k \partial x^l}, \quad (2.22)$$

$$L^j = \sum_{k=1}^{d} b^{k,j} \frac{\partial}{\partial x^k}, \quad j \in \{1, \ldots, m\}, \quad (2.23)$$

where $a_k$ and $b^{k,j}$ are the same as in (1.3).

For $\alpha = (j_1, \ldots, j_l)$,

$$f_\alpha = \begin{cases} f, & l = 0, \\
L^{j_l} f_{\alpha}, & l \geq 1. \end{cases} \quad (2.24)$$
**Stratonovich Coefficient Functions:**

\[ L^0 = \frac{\partial}{\partial t} + \sum_{k=1}^{d} a_k \frac{\partial}{\partial x_k}, \]  

\[ L^j = \sum_{k=1}^{d} b^{k,j} \frac{\partial}{\partial x_k}, \quad j \in \{1, \ldots, m\}, \]  

\[ a = a - \frac{1}{2} \sum_{j=1}^{m} L^j b^j, \]  

where \( a^k \) and \( b^{k,j} \) are the same as in (1.7), and \( b^j \) represent the \( j \)-th column of \( b \).

For \( \alpha = (j_1, \ldots, j_l) \),

\[ f_\alpha = \begin{cases} f, & l = 0, \\ L^{j_1} f_\alpha, & l \geq 1. \end{cases} \]  

Note that \( k - \alpha, \alpha - k, K_\alpha[f(\cdot, X)]_{\rho,\tau}, K_\alpha, M_\alpha[f(\cdot, X)]_{\rho,\tau}, M_\alpha \) are new notations, the others can be found in [3, Chapter 5]. From the relations between multiple integrals, we can easily get that \( K_\alpha[f(\cdot, X)]_{\rho,\tau} \) is meaningful when \( I_\alpha[f(\cdot, X)]_{\rho,\tau} \) is meaningful and \( M_\alpha[f(\cdot, X)]_{\rho,\tau} \) is meaningful when \( J_\alpha[f(\cdot, X)]_{\rho,\tau} \) is meaningful. More details, please find in [3, Chapter 5].

3. **Backward Stochastic Taylor Expansions**

We shall now state and prove the backward Ito-Taylor expansions for the \( d \)-dimensional Ito process

\[ X_t = X_0 + \int_0^t a(s, X_s) \, ds + \sum_{j=1}^{m} \int_0^t b^j(s, X_s) \, dW_s^j \]  

and the backward Stratonovich-Taylor expansions for the \( d \)-dimensional Ito process

\[ X_t = X_0 + \int_0^t a(s, X_s) \, ds + \sum_{j=1}^{m} \int_0^t b^j(s, X_s) \circ dW_s^j, \]

where \( t \in [0, T] \).

**Theorem 3.1.** Let \( \rho \) and \( \tau \) be two stopping times with

\[ 0 \leq \rho(\omega) \leq \tau(\omega) \leq T \]  

w.p.1, let \( A \subset M \) be an hierarchical set, and let \( f : \mathbb{R}^+ \times \mathbb{R}^d \rightarrow \mathbb{R} \). Then the backward Ito-Taylor expansions

\[ f(\rho, X_\rho) = \sum_{\alpha \in A} (-1)^{l(\alpha)} K_\alpha [f_\alpha(\tau, X_\tau)]_{\rho,\tau} + \sum_{\alpha \in B(A)} (-1)^{l(\alpha)} K_\alpha [f_\alpha(\cdot, X_\cdot)]_{\rho,\tau} \]  

holds, provided \( f_\alpha \in \mathcal{H}_\alpha \) for \( \alpha \in A \cup B(A) \) and \( f_\alpha \in C^{1,2} \) for \( \alpha \in A \).

**Theorem 3.2.** Let \( \rho \) and \( \tau \) be two stopping times with

\[ 0 \leq \rho(\omega) \leq \tau(\omega) \leq T \]  

\[ 0 \leq \rho(\omega) \leq \tau(\omega) \leq T \]
w.p.1, let $\mathcal{A} \subset \mathcal{M}$ be an hierarchical set, and let $f : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}$. Then the backward Stratonovich-Taylor expansion

$$f (\rho, X_\tau) = \sum_{\alpha \in \mathcal{A}} (-1)^{l(\alpha)} M_\alpha \left[ L_{\alpha} (\tau, X_\tau) \right]_{\rho, \tau} + \sum_{\alpha \in \mathcal{B}(\mathcal{A})} (-1)^{l(\alpha)} M_\alpha \left[ L_{\alpha} (\cdot, X) \right]_{\rho, \tau}$$

(3.6)

holds, provided $L_{\alpha} \in \mathcal{W}_\alpha$ for $\alpha \in \mathcal{A} \cup \mathcal{B}(\mathcal{A})$ and $L_{\alpha} \in C^{1,2}$ for $\alpha \in \mathcal{A}$.

The proofs of Theorems 3.1 and 3.2 are analogous to the proof of stochastic Ito-Taylor and Stratonovich-Taylor expansions in [3]. We only prove Theorem 3.2 here.

**Proof of Theorem 3.2.** Let $l_1(\mathcal{A}) = \sup_{\alpha \in \mathcal{A}} l(\alpha)$. For $l_1(\mathcal{A}) = 0$, we have $\mathcal{A} = \{v\}$ and $\mathcal{B}(\mathcal{A}) = \{(0), (1), \cdots, (m)\}$. By the Stratonovich-Taylor expansions [3], we have

$$f (\tau, X_\tau) = f (\rho, X_\rho) + \sum_{j=0}^{m} J_{(j)} \left[ L_{j} (\cdot, X) \right]_{\rho, \tau},$$

(3.7)

then

$$f (\rho, X_\rho) = f (\tau, X_\tau) - \sum_{j=0}^{m} M_{(j)} \left[ L_{j}^{(\cdot)} (\cdot, X) \right]_{\rho, \tau},$$

(3.8)

which is (3.6). Suppose that it is true for $l_1(\mathcal{A}) \leq k - 1, k \geq 1$. Let $\mathcal{E} = \{\alpha \in \mathcal{A} : l(\alpha) \leq k - 1\}$ which is an hierarchical set, then

$$f (\rho, X_\rho) = \sum_{\alpha \in \mathcal{E}} (-1)^{l(\alpha)} M_\alpha \left[ L_{\alpha} (\tau, X_\tau) \right]_{\rho, \tau} + \sum_{\alpha \in \mathcal{A} \setminus \mathcal{E}} (-1)^{l(\alpha)} M_\alpha \left[ L_{\alpha} (\cdot, X) \right]_{\rho, \tau}.$$  

(3.9)

For $l_1(\mathcal{A}) = k$, we can get $\mathcal{A} \setminus \mathcal{E} \subseteq \mathcal{B}(\mathcal{E})$, so from (3.9) and (3.8) we have

$$f (\rho, X_\rho) = \sum_{\alpha \in \mathcal{E}} (-1)^{l(\alpha)} M_\alpha \left[ L_{\alpha} (\tau, X_\tau) \right]_{\rho, \tau} + \sum_{\alpha \in \mathcal{A} \setminus \mathcal{E}} (-1)^{l(\alpha)} M_\alpha \left[ L_{\alpha} (\cdot, X) \right]_{\rho, \tau}$$

$$+ \sum_{\alpha \in \mathcal{B}(\mathcal{E}) \setminus (\mathcal{A} \setminus \mathcal{E})} (-1)^{l(\alpha)} M_\alpha \left[ L_{\alpha} (\cdot, X) \right]_{\rho, \tau}$$

$$= \sum_{\alpha \in \mathcal{E}} (-1)^{l(\alpha)} M_\alpha \left[ L_{\alpha} (\tau, X_\tau) \right]_{\rho, \tau}$$

$$+ \sum_{\alpha \in \mathcal{A} \setminus \mathcal{E}} (-1)^{l(\alpha)} M_\alpha \left[ L_{\alpha} (\cdot, X) \right]_{\rho, \tau}$$

$$+ \sum_{\alpha \in \mathcal{B}(\mathcal{E}) \setminus (\mathcal{A} \setminus \mathcal{E})} (-1)^{l(\alpha)} M_\alpha \left[ L_{\alpha} (\cdot, X) \right]_{\rho, \tau}$$

$$= \sum_{\alpha \in \mathcal{A} \setminus \mathcal{E}} (-1)^{l(\alpha)} M_\alpha \left[ L_{\alpha} (\tau, X_\tau) \right]_{\rho, \tau}$$

$$+ \sum_{\alpha \in \mathcal{A} \setminus \mathcal{E}} (-1)^{l(\alpha) + 1} M_\alpha \left[ M_{(j)} \left[ L_{j}^{(\cdot)} (\cdot, X) \right]_{\tau, \tau} \right]_{\rho, \tau}$$

$$+ \sum_{\alpha \in \mathcal{B}(\mathcal{E}) \setminus (\mathcal{A} \setminus \mathcal{E})} (-1)^{l(\alpha)} M_\alpha \left[ L_{\alpha} (\cdot, X) \right]_{\rho, \tau}$$

$$= \sum_{\alpha \in \mathcal{A} \setminus \mathcal{E}} (-1)^{l(\alpha)} M_\alpha \left[ L_{\alpha} (\tau, X_\tau) \right]_{\rho, \tau}$$

$$+ \sum_{\alpha \in \mathcal{A} \setminus \mathcal{E}} (-1)^{l(\alpha) + 1} M_\alpha \left[ M_{(j)} \left[ L_{j}^{(\cdot)} (\cdot, X) \right]_{\tau, \tau} \right]_{\rho, \tau}$$

$$+ \sum_{\alpha \in \mathcal{B}(\mathcal{E}) \setminus (\mathcal{A} \setminus \mathcal{E})} (-1)^{l(\alpha)} M_\alpha \left[ L_{\alpha} (\cdot, X) \right]_{\rho, \tau}.$$
Let \( C \) and the positive constants \( \gamma \), then we have
\[
\sum_{\alpha \in A \setminus E} (-1)^{\ell(\alpha)} M_{\alpha} \left[ f_\alpha (\tau, X_\tau) \right]_{\rho, \tau} + \sum_{\alpha \in B_1} (-1)^{\ell(\alpha)} M_{\alpha} \left[ f_\alpha (\tau, X_\tau) \right]_{\rho, \tau}.
\]
(3.10)

(Here, we can easily get that the element in \( B(\mathcal{E}) \setminus (A \setminus \mathcal{E}) \) has length at most \( k \) and the element in \( \bigcup_{j=0}^m \{ (j) \alpha \in M : \alpha \in A \setminus \mathcal{E} \} \) have length \( k+1 \), so two sets do not have intersection.) with
\[
B_1 = \left[ B(\mathcal{E}) \setminus (A \setminus \mathcal{E}) \right] \bigcup \left[ \bigcup_{j=0}^m \{ (j) \alpha \in M : \alpha \in A \setminus \mathcal{E} \} \right]
\]
\[
= \left[ \{ \alpha \in M \setminus \mathcal{E} : -\alpha \in \mathcal{E} \} \{ \alpha \in M \setminus \mathcal{E} : \alpha \in A \} \right] \bigcup \{ \alpha \in M : -\alpha \in A \setminus \mathcal{E} \}
\]
\[
= \{ \alpha \in M \setminus A : -\alpha \in \mathcal{E} \} \bigcup \{ \alpha \in M \setminus A : -\alpha \in A \setminus \mathcal{E} \}
\]
\[
= \{ \alpha \in M \setminus A : -\alpha \in A \} = B(A).
\]
(3.11)

This completes the proof of the theorem. \( \square \)

4. Moments Estimation of Truncated Error for Backward Expansions

In this section, we investigate the truncated error of backward stochastic Taylor expansions which will be used to prove the convergence of our schemes in next section.

**Theorem 4.1.** Let \( \mathcal{A}_\gamma = \{ \alpha \in M : l(\alpha) + n(\alpha) \leq 2\gamma \} \), \( \Gamma_\gamma = \{ \alpha \in M : l(\alpha) + n(\alpha) \leq 2\gamma + 1 \} \), \( \gamma = 1, 2, 3, \ldots \), then we have
\[
\left| E \left( \sum_{\alpha \in B(A_\gamma)} (-1)^{\ell(\alpha)} K_{\alpha} \left[ f_\alpha (\tau, X_\tau) \right]_{\tau_{n+1}, \tau_n} \right) \right| \leq C_1 R_f \Delta^{\gamma+1},
\]
(4.1)
\[
\left( E \left( \left( \sum_{\alpha \in B(A_\gamma)} (-1)^{\ell(\alpha)} K_{\alpha} \left[ f_\alpha (\tau, X_\tau) \right]_{\tau_{n+1}, \tau_n} \right)^2 \right) \right)^{\frac{1}{2}} \leq C_2 R_f \Delta^{\gamma+\frac{1}{2}},
\]
(4.2)

where
\[
R_f = \sup_{\alpha \in B(A_\gamma)} \left( E \left( \sup_{\tau_n \leq \tau \leq \tau_{n+1}} \left| f_\alpha (\tau, X_\tau) \right|^4 \right) \right)^{\frac{1}{4}}
\]
(4.3)
and the positive constants \( C_1, C_2 \) do not depend on \( \Delta \), provided that \( f_\alpha \in H_\alpha \) for all \( \alpha \in \mathcal{A}_\gamma \cup B(A_\gamma) \), \( f_\alpha \in H_{\alpha -} \) for all \( \alpha \in B(\Gamma_\gamma) \), and \( f_\alpha \in C^{1,2} \) for \( \alpha \in \Gamma_\gamma \).

**Theorem 4.2.** Let \( \mathcal{A}_\gamma = \{ \alpha \in M : l(\alpha) + n(\alpha) \leq 2\gamma \} \), \( \Gamma_\gamma = \{ \alpha \in M : l(\alpha) + n(\alpha) \leq 2\gamma + 1 \} \), \( \gamma = 1, 2, 3, \ldots \), then we have
\[
\left| E \left( \sum_{\alpha \in B(A_\gamma)} (-1)^{\ell(\alpha)} M_{\alpha} \left[ f_\alpha (\tau, X_\tau) \right]_{\tau_{n+1}, \tau_n} \right) \right| \leq C_3 R_f \Delta^{\gamma+1},
\]
(4.4)
\[
\left( E \left( \left( \sum_{\alpha \in B(A_\gamma)} (-1)^{\ell(\alpha)} M_{\alpha} \left[ f_\alpha (\tau, X_\tau) \right]_{\tau_{n+1}, \tau_n} \right)^2 \right) \right)^{\frac{1}{2}} \leq C_4 R_f \Delta^{\gamma+\frac{1}{2}},
\]
(4.5)
where

\[
R_\tau = \sup_{\alpha \in \mathcal{C}} \left( E \left( \sup_{\tau_{n-1} \leq t \leq \tau_{n+1}} |f_\alpha(t, X_t)|^4 |\mathcal{F}_{\tau_n}| \right) \right)^{\frac{1}{4}},
\]

\[C_\gamma = \mathcal{B}(A_\gamma) \cup \mathcal{B}(\Gamma_\gamma) \cup \{ (j_1) \ast \alpha : \alpha = (j_1, j_2, \ldots, j_l) \in \mathcal{B}(A_\gamma) \cup \mathcal{B}(\Gamma_\gamma), j_1 \neq 0 \}, \]

and the positive constants \(C_3, C_4\) do not depend on \(\Delta\), provided that \(f_\alpha \in \mathcal{H}_\gamma\) for all \(\alpha \in A_\gamma \cup \mathcal{B}(A_\gamma)\), \(f_\alpha \in \mathcal{H}_{\gamma-}\) for all \(\alpha \in \mathcal{B}(\Gamma_\gamma)\), and \(f_\alpha \in \mathcal{C}^{1,2}\) for \(\alpha \in \Gamma_\gamma\).

We only prove Theorem 4.2 here, and the proof of Theorem 4.1 is analogous and easier. We list some lemmas before we prove it.

**Lemma 4.1.** Let \(\alpha \in \mathcal{M}\backslash \{v\}\) with \(l(\alpha) \neq n(\alpha)\), let \(f \in \mathcal{H}_\gamma\). Then

\[E(J_\alpha[f(\cdot, X)]_{\tau_n, \tau_{n+1}} |\mathcal{F}_{\tau_n}) = 0, \text{ w.p.1.} \]

See [3, Lemma 5.7.1].

**Lemma 4.2.** Let \(\alpha = (j_1, j_2, \ldots, j_l) \in \mathcal{M}\backslash v\), let \(g \in \mathcal{H}_\gamma\), and let \(q = 1, 2, \ldots\) Then

\[E(J_\alpha[g(\cdot, X)]_{\tau_n, \tau_{n+1}} |\mathcal{F}_{\tau_n}) \leq C_5 R \Delta^{\frac{l(\alpha) + n(\alpha)}{4}} \]

with

\[R = \max \left\{ \left( E\left( \sup_{\tau_{n-1} \leq s \leq \tau_{n+1}} |g(s, X_s)|^{2q} |\mathcal{F}_{\tau_n}| \right) \right)^{\frac{1}{4q}}, \right\} \]

where the positive constant \(C_5\) does not depend on \(\Delta\).

By (5.2.34) and Remark 5.2.8 in [3], the multiple Stratonovich integral \(J_\alpha[g(\cdot, X)]_{\tau_n, \tau_{n+1}}\) can be written as a finite sum of multiple Ito integrals \(I_\alpha^\gamma[g(\cdot, X)]_{\tau_n, \tau_{n+1}}\) and \(I_\beta^\gamma[g(\cdot, X)]_{\tau_n, \tau_{n+1}}\) with \(l(\alpha) + n(\alpha) \leq l(\beta) + n(\beta)\). Thus, by [3, Lemma 5.7.5], we can get Lemma 4.2.

**Lemma 4.3.** Let \(\alpha_k \in \mathcal{M}\backslash v, k = 1, 2, \ldots, l\). Then

\[E(\left| J_{\alpha_1} J_{\alpha_2} \cdots J_{\alpha_l} \right| |\mathcal{F}_{\tau_n}) := E \left( \prod_{k=1}^l J_{\alpha_k, \tau_n, \tau_{n+1}} |\mathcal{F}_{\tau_n} \right) \leq C_6 \Delta^{\frac{l}{4} \sum_{k=1}^l (l(\alpha_k) + n(\alpha_k))}, \]

where the positive constant \(C_6\) does not depend on \(\Delta\), and

\[E(J_{\alpha_1} J_{\alpha_2} \cdots J_{\alpha_l} |\mathcal{F}_{\tau_n}) := E \left( \prod_{k=1}^l J_{\alpha_k, \tau_n, \tau_{n+1}} |\mathcal{F}_{\tau_n} \right) = 0, \]

where \(\sum_{k=1}^l (l(\alpha_k) + n(\alpha_k))\) is odd.

By (5.2.34) and Remark 5.2.8 in [3], the multiple Stratonovich integral \(J_{\alpha}\) can be written as a finite sum of multiple Ito integrals \(I_\alpha\) with \(l(\alpha) + n(\alpha) = l(\beta) + n(\beta)\). By [3, Lemma 5.12.3], the product of multiple Ito integrals \(I_\alpha\) and \(I_\beta\) can be written as a finite sum of multiple Ito integrals \(I_\theta\) with \(l(\theta) + n(\theta) = l(\alpha) + n(\alpha) + l(\beta) + n(\beta)\). So we shall prove (4.10) by [3, Lemma 5.7.5], and prove (4.11) by Lemma 4.1.
Proof of Theorem 4.2. It is easy to show that $\mathcal{A}_1$ is a hierarchical set. By the Stratonovich-Taylor expansion [3], we have

$$f(\tau_{n+1}, X_{\tau_{n+1}}) = \sum_{\alpha \in \mathcal{A}_1} J_\alpha \left[ f_\alpha (\tau_n, X_{\tau_n}) \right]_{\tau_n, \tau_{n+1}} + \sum_{\alpha \in \mathcal{B}(\mathcal{A}_1)} J_\alpha \left[ f_\alpha (\cdot, X_{\cdot}) \right]_{\tau_n, \tau_{n+1}}. \tag{4.12}$$

By (3.6) and (4.12), we have

$$\sum_{\alpha \in \mathcal{A}_1 \setminus v} J_\alpha \left[ f_\alpha (\tau_n, X_{\tau_n}) \right]_{\tau_n, \tau_{n+1}} + \sum_{\alpha \in \mathcal{A}_1 \setminus v} (-1)^{l(\alpha)} M_\alpha \left[ f_\alpha (\tau_{n+1}, X_{\tau_{n+1}}) \right]_{\tau_n, \tau_{n+1}} + \sum_{\alpha \in \mathcal{B}(\mathcal{A}_1)} (-1)^{l(\alpha)} M_\alpha \left[ f_\alpha (\cdot, X_{\cdot}) \right]_{\tau_n, \tau_{n+1}} = 0. \tag{4.13}$$

First, we estimate the third item of (4.13). Analogously, by [3, 5.2.34 and Remark 5.2.8], the multiple Stratonovich integrals $I_\alpha[f_\alpha (\cdot, X)_{\tau_n, \tau_{n+1}}]$ can be written as a finite sum of multiple Itô integrals $I_\beta[I_\alpha[f_\alpha (\cdot, X)_{\tau_n, \tau_{n+1}}]$ and $I_\beta[L^{2n}f_\alpha (\cdot, X)_{\tau_n, \tau_{n+1}}$ with

$$l(\beta) + n(\beta) \geq l(\alpha) + n(\alpha) \geq 2\gamma + 1.$$ 

By Lemma 4.1, we have

$$E \left( I_\beta \left[ f_\alpha (\cdot, X) \right]_{\tau_n, \tau_{n+1}} \mid \mathcal{F}_{\tau_n} \right) = 0, \quad E \left( I_\beta \left[ L^{2n} f_\alpha (\cdot, X) \right]_{\tau_n, \tau_{n+1}} \mid \mathcal{F}_{\tau_n} \right) = 0,$$

when $l(\beta) \neq n(\beta)$. Thus,

$$\left| E \left( \sum_{\alpha \in \mathcal{B}(\mathcal{A}_1)} J_\alpha \left[ f_\alpha (\cdot, X) \right]_{\tau_n, \tau_{n+1}} \mid \mathcal{F}_{\tau_n} \right) \right| \leq C_7 R_7 \Delta^{\gamma+1}, \tag{4.14}$$

where the positive constant $C_7$ does not depend on $\Delta$.

Next, we estimate the sum of first and second items of (4.13). By (2.19) and the stochastic Stratonovich expansions, we have

$$J_\alpha \left[ f_\alpha (\tau_n, X_{\tau_n}) \right]_{\tau_n, \tau_{n+1}} + (-1)^{l(\alpha)} M_\alpha \left[ f_\alpha (\tau_{n+1}, X_{\tau_{n+1}}) \right]_{\tau_n, \tau_{n+1}} = J_\alpha f_\alpha (\tau_n, X_{\tau_n}) + (-1)^{l(\alpha)} \left( \sum_{\alpha \in \mathcal{A}_1 \setminus v, i=1,...,k} (-1)^{l(\alpha)-k} J_{\alpha_1} J_{\alpha_2} \cdots J_{\alpha_k} f_\alpha (\tau_{n+1}, X_{\tau_{n+1}}) \right)$$

$$= J_\alpha f_\alpha (\tau_n, X_{\tau_n}) + \left( \sum_{\alpha \in \mathcal{A}_1 \setminus v, i=1,...,k} (-1)^{k} J_{\alpha_1} J_{\alpha_2} \cdots J_{\alpha_k} \right)$$

$$\times \left( f_\alpha (\tau_n, X_{\tau_n}) + \sum_{\beta \in \mathcal{B}(\mathcal{A}_1)} J_\beta \left[ f_{\beta+\alpha} (\tau_n, X_{\tau_n}) \right]_{\tau_n, \tau_{n+1}} + \sum_{\beta \in \mathcal{B}(\mathcal{A}_1)} J_\beta \left[ f_{\beta+\alpha} (\cdot, X_{\cdot}) \right]_{\tau_n, \tau_{n+1}} \right)$$

$$= \left( \sum_{\alpha \in \mathcal{A}_1 \setminus v, i=1,...,k} (-1)^{k} J_{\alpha_1} J_{\alpha_2} \cdots J_{\alpha_k} \right) f_\alpha (\tau_n, X_{\tau_n})$$

$$+ \left( \sum_{\alpha \in \mathcal{A}_1 \setminus v, i=1,...,k} (-1)^{k} J_{\alpha_1} J_{\alpha_2} \cdots J_{\alpha_k} \right) \sum_{\beta \in \mathcal{B}(\mathcal{A}_1)} J_\beta \left[ f_{\beta+\alpha} (\tau_n, X_{\tau_n}) \right]_{\tau_n, \tau_{n+1}}$$
We consider the coefficient of $J$ consists of some sums of
situation. One is $\alpha = (4.10)$, Lemma 4.2 and
$\eta \neq v$, $l(\eta) + n(\eta) \leq 2\gamma + 1$, $l \geq 2$
and
$\pm J_{\eta_1} \cdots J_{\eta_l} \left[ f_{\eta_l} (\cdot, X) \right]_{\tau_n, \tau_{n+1}}, \ \eta = \eta_l \ast \eta_{l-1} \ast \cdots \ast \eta_1, \ \eta_i \neq v, \ l(\eta) + n(\eta) \geq 2\gamma + 2.$
We consider the coefficient of $J_{\eta_1} \cdots J_{\eta_l} f_{\eta_l} (\tau_n, X_{\tau_n})$. When $l(\eta) + n(\eta) \leq 2\gamma$, there are two
two situation. One is $\alpha = \eta_{l-1} \ast \eta_{l-2} \ast \cdots \ast \eta_1$ whose coefficient is $(-1)^{l-1}$, and the other is $\alpha = \eta_l \ast \eta_{l-1} \ast \cdots \ast \eta_1$ whose coefficient is $(-1)^l$. Thus the coefficient of $J_{\eta_1} \cdots J_{\eta_l} f_{\eta_l} (\tau_n, X_{\tau_n})$
is zero when $l(\eta) + n(\eta) \leq 2\gamma$. When $l(\eta) + n(\eta) = 2\gamma + 1$, from Lemma 4.3, we can get
$E (J_{\eta_1} \cdots J_{\eta_l} f_{\eta_l} (\tau_n, X_{\tau_n})) | F_{\tau_n} = 0.$
From above analysis, and by the Cauchy-Schwarz inequality, Lemmas 4.2 and 4.3, we have
\begin{align*}
| E \left( \sum_{\alpha \in A_1 \setminus v} J_{\alpha} \left[ f_{\alpha} (\tau_n, X_{\tau_n}) \right]_{\tau_n, \tau_{n+1}} + \sum_{\alpha \in A_1 \setminus v} (-1)^{l(\alpha)} M_{\alpha} \left[ f_{\alpha} (\tau_{n+1}, X_{\tau_{n+1}}) \right]_{\tau_{n+1}, \tau_{n+1}} | F_{\tau_n} \right) | \\
= | E \left( \sum_{\alpha \in A_1 \setminus v} \sum_{\alpha_{l+1} \cdots \alpha_k = \alpha} \sum_{\beta \in B(A_1)} (-1)^k J_{\alpha} \cdots J_{\alpha_l} f_{\beta+\alpha_l+\cdots+\alpha_k} (\tau_n, X_{\tau_n}) \right)_{\tau_n, \tau_{n+1}} | F_{\tau_n} | \\
\leq \sum_{\alpha \in A_1 \setminus v} \sum_{\alpha_{l+1} \cdots \alpha_k = \alpha} \sum_{\beta \in B(A_1)} E \left( \left| J_{\alpha} \cdots J_{\alpha_l} f_{\beta+\alpha_l+\cdots+\alpha_k} \right|^2 | F_{\tau_n} \right) E \left( \left| f_{\beta+\alpha_l+\cdots+\alpha_k} (\cdot, X) \right|_{\tau_n, \tau_{n+1}}^2 | F_{\tau_n} \right)^{1/2} \\
\leq C_8 \beta_{\alpha_l+\cdots+\alpha_k} \gamma^{l+1}, \ (4.15)
\end{align*}
where the positive constant $C_8$ does not depend on $\Delta$. The last but one inequality is obtain by
(4.10), Lemma 4.2 and $I(\alpha) + n(\alpha) + l(\beta) + n(\beta) \geq 2\gamma + 2.$
By (4.13)-(4.15), we have
\[
\left| E\left( \sum_{\alpha \in B(A_1)} (-1)^{l(\alpha)} M_{\alpha} \left[ f_{\alpha}(\cdot, X) \right]_{\tau_n, \tau_{n+1}} \right| F_{\tau_n} \right| \leq C_3 R \Delta^{\gamma+1},
\]
where the positive constant $C_3$ does not depend on $\Delta$.

By (2.19), the Minkowski’s inequality, the Cauchy-Schwarz inequality, Lemmas 4.2, and 4.3, we have
\[
\left( E\left( \left( \sum_{\alpha \in B(A_1)} (-1)^{l(\alpha)} \sum_{\alpha_1, \ldots, \alpha_k = \alpha} (-1)^{l(\alpha)-k} J_{\alpha_1} \left[ f_{\alpha_1}(\cdot, X) \right]_{\tau_n, \tau_{n+1}} J_{\alpha_2} J_{\alpha_3} \cdots J_{\alpha_k} \right)^2 | F_{\tau_n} \right) \right)^{\frac{1}{2}}
\leq \sum_{\alpha \in B(A_1)} \sum_{\alpha_1, \ldots, \alpha_k = \alpha} E\left( \left( J_{\alpha_1} \left[ f_{\alpha_1}(\cdot, X) \right]_{\tau_n, \tau_{n+1}} J_{\alpha_2} J_{\alpha_3} \cdots J_{\alpha_k} \right)^2 | F_{\tau_n} \right) \right)^{\frac{1}{2}}
\leq \sum_{\alpha \in B(A_1)} \sum_{\alpha_1, \ldots, \alpha_k = \alpha} \left( E\left( \left( J_{\alpha_1} \left[ f_{\alpha_1}(\cdot, X) \right]_{\tau_n, \tau_{n+1}} J_{\alpha_2} J_{\alpha_3} \cdots J_{\alpha_k} \right)^4 | F_{\tau_n} \right) \right)^{\frac{1}{4}}
\leq C_4 R \Delta^{\gamma + \frac{1}{2}},
\]
where the positive constant $C_4$ does not depend on $\Delta$. The last but one inequation is obtained by (4.10), Lemma 4.2 and $l(\alpha) + n(\alpha) \geq 2\gamma + 1$. This completes the proof of the theorem. □

5. Fully Implicit Schemes

Now, we present our full implicit order $\gamma$ strong Ito-Taylor scheme
\[
Y_{n+1} = Y_n - \sum_{\alpha \in A_1 \setminus \psi} (-1)^{l(\alpha)} K_{\alpha} \left[ f_{\alpha}(\tau_{n+1}, Y_{n+1}) \right]_{\tau_n, \tau_{n+1}}
\]
and the full implicit order $\gamma$ strong Stratonovich-Taylor scheme
\[
Y_{n+1} = Y_n - \sum_{\alpha \in A_1 \setminus \psi} (-1)^{l(\alpha)} M_{\alpha} \left[ f_{\alpha}(\tau_{n+1}, Y_{n+1}) \right]_{\tau_n, \tau_{n+1}},
\]
where $f(t, x) \equiv x$ and $A_\psi$ is the set defined in Section 4. Let
\[
F_{\tau_{n+1}}(x) = x + \sum_{\alpha \in A_1 \setminus \psi} (-1)^{l(\alpha)} K_{\alpha, \tau_n, \tau_{n+1}} f_{\alpha}(\tau_{n+1}, x),
\]
\[
E_{\tau_{n+1}}(x) = x + \sum_{\alpha \in A_1 \setminus \psi} (-1)^{l(\alpha)} M_{\alpha, \tau_n, \tau_{n+1}} f_{\alpha}(\tau_{n+1}, x)
\]
for $n = 0, 1, \ldots, N-1$. Thus we have the following convergence theorems.
Theorem 5.1. Let $Y$ be a discrete time approximation defined by (5.1), $Y_0 = X_0$, $E(|X_0|^2) < \infty$, $f_\alpha \in \mathcal{H}_\alpha$ for all $\alpha \in \mathcal{A}_1 \cup \mathcal{B}(\mathcal{A}_1)$, $f_\alpha \in \mathcal{H}_\alpha$ for all $\alpha \in \mathcal{B}(\Gamma)$, $f_\alpha \in C^{1,2}$ for $\alpha \in \Gamma$, and

$$|f_\alpha(t, x)| \leq K_1 (1 + |x|^2)$$

(5.5)

for all $\alpha \in \{(i) : i = 0, 1, \cdots, m\} \cup \mathcal{B}(\mathcal{A}_1) \cup \mathcal{B}(\Gamma)$, $t \in [0, T]$ and $x \in \mathbb{R}^d$;

$$|f_\alpha(t, x) - f_\alpha(t, y)| \leq K_2 |x - y|$$

(5.6)

for all $\alpha \in \mathcal{A}_1$, $t \in [0, T]$ and $x, y \in \mathbb{R}^d$;

$$|F'_\alpha(x)| \geq K_3, \quad \text{w.p. 1}$$

(5.7)

for all $u = 0, 1, \ldots, N - 1, x \in \mathbb{R}^d$. Then $Y$ converges strongly to $X$ with order $\gamma$, that is

$$E(|Y_N - X_T|) \leq K_4 \Delta^\gamma,$$

(5.8)

where the positive constants $K_1, K_2, K_3$ and $K_4$ do not depend on $\Delta$.

Theorem 5.2. Let $Y$ be a discrete time approximation defined by (5.2), $Y_0 = X_0$, $E(|X_0|^2) < \infty$, $f_\alpha \in \mathcal{H}_\alpha$ for all $\alpha \in \mathcal{A}_1 \cup \mathcal{B}(\mathcal{A}_1)$, $f_\alpha \in \mathcal{H}_\alpha$ for all $\alpha \in \mathcal{B}(\Gamma)$, $f_\alpha \in C^{1,2}$ for $\alpha \in \Gamma$, and

$$|f_\alpha(t, x)| \leq K_5 (1 + |x|^2)$$

(5.9)

for all $\alpha \in \{(i) : i = 0, 1, \ldots, m\} \cup \mathcal{B}(\mathcal{A}_1) \cup \mathcal{B}(\Gamma) \cup \{(j_1) * \alpha : \alpha \in \mathcal{B}(\mathcal{A}_1) \cup \mathcal{B}(\Gamma), j_1 \neq 0\}$, $t \in [0, T]$ and $x \in \mathbb{R}^d$;

$$|f_\alpha(t, x) - f_\alpha(t, y)| \leq K_6 |x - y|$$

(5.10)

for all $\alpha \in \mathcal{A}_1$, $t \in [0, T]$ and $x, y \in \mathbb{R}^d$;

$$|F'_\alpha(x)| \geq K_7, \quad \text{w.p. 1}$$

(5.11)

for all $u = 0, 1, \ldots, N - 1, x \in \mathbb{R}^d$. Then $Y$ converges strongly to $X$ with order $\gamma$, that is

$$E(|Y_N - X_T|) \leq K_8 \Delta^\gamma,$$

(5.12)

where the positive constants $K_5, K_6, K_7$ and $K_8$ do not depend on $\Delta$.

The proofs of Theorem 5.1 and Theorem 5.2 are very similar. We only prove Theorem 5.2 here. To prove this theorem, we need the following convergence theorem given by Milstein [6–8].

Theorem 5.3. Assume that the coefficient functions of (1.1) are continuous for $t \in [0, T]$ and Lipschitz continuous for $x \in \mathbb{R}^d$, $E(X_0^2) < \infty$. Assume that for a one-step discrete time approximation $Y$, the local mean error and mean-square error for all $N = 1, 2, \ldots$, and $n = 0, 1, \ldots, N - 1$ satisfy the estimates

$$E(X_{n+1}^2 - Y_{n+1}^2 | F_n) \leq K \left(1 + |Y_n|^2\right)^{\frac{1}{2}} \Delta^{p_1},$$

(5.13)

$$E \left(|X_{n+1}^0 - Y_{n+1}|^2 | F_n\right) \leq K \left(1 + |Y_n|^2\right)^{\frac{1}{2}} \Delta^{p_2},$$

(5.14)

with $p_2 \geq \frac{1}{2}$ and $p_1 \geq p_2 + \frac{1}{2}$. Then

$$E \left(|X_{k+1}^0 - Y_{k+1}|^2 | F_0\right) \leq K \left(1 + |Y_0|^2\right)^{\frac{1}{2}} \Delta^{p_2 - \frac{1}{2}},$$

(5.15)

holds for each $k = 0, 1, \ldots, N$ and the positive constants $K$ do not depend on $\Delta$, where $X_{k+1}^0$ means $Y_n = X_{\tau_n}$. 
Proof of Theorem 5.2. By (5.2) and (5.4), we have
\[
E_{\tau_{n+1}}(Y_{n+1}) = Y_n. \tag{5.16}
\]
First, we discuss the existence and uniqueness of the solution of the methods. By \( f_\alpha \in C^{1,2} \) for \( \alpha \in \Gamma, \gamma \), it is known that \( E_{\tau_{n+1}}'(x) \) is continuous. Thus, by condition (5.11), we have that \( E_{\tau_{n+1}}'(x) \) is always greater than \( K_7 \) or smaller than \(-K_7\) with probability 1. So, (5.16) exist unique real solution with probability 1. This means the solution of discrete time approximation defined by (5.2) is existence and uniqueness with probability 1.

By the backward Stratonovich-Taylor expansion (3.6) with \( f(t, x) \equiv x \) and (5.4), we have
\[
E_{\tau_{n+1}}(X_{\tau_{n+1}}) = X_{\tau_n} - \sum_{\alpha \in B(A_\gamma)} (-1)^{l(\alpha)} M_\alpha \left[ \mathcal{L}_\alpha \left( \cdot, X \right) \right]_{\tau_n, \tau_{n+1}}. \tag{5.17}
\]
Thus,
\[
E_{\tau_{n+1}}(X_{\tau_{n+1}}) - E_{\tau_{n+1}}(Y_{n+1}) = X_{\tau_n} - Y_n - \sum_{\alpha \in B(A_\gamma)} (-1)^{l(\alpha)} M_\alpha \left[ \mathcal{L}_\alpha \left( \cdot, X \right) \right]_{\tau_n, \tau_{n+1}}, \tag{5.18}
\]
and by mean value theorem we have
\[
E_{\tau_{n+1}}'(Z_{XY})(X_{\tau_{n+1}} - Y_{n+1}) = X_{\tau_n} - Y_n - \sum_{\alpha \in B(A_\gamma)} (-1)^{l(\alpha)} M_\alpha \left[ \mathcal{L}_\alpha \left( \cdot, X \right) \right]_{\tau_n, \tau_{n+1}}, \tag{5.19}
\]
where \( Z_{XY} \) is some vector between \( X_{\tau_{n+1}} \) and \( Y_{n+1} \). Furthermore, by (5.19), (5.11) and (4.5), we have
\[
\left( E \left( (X_{\tau_{n+1}} - Y_{n+1})^2 | \mathcal{F}_{\tau_n} \right) \right)^\frac{1}{2} \leq \frac{1}{K_7} \left( E \left( \left( \sum_{\alpha \in B(A_\gamma)} (-1)^{l(\alpha)} M_\alpha \left[ \mathcal{L}_\alpha \left( \cdot, X \right) \right]_{\tau_n, \tau_{n+1}} \right)^2 | \mathcal{F}_{\tau_n} \right) \right)^\frac{1}{2} \leq \frac{C_\gamma}{K_7} B_\gamma \Delta^{\gamma + \frac{1}{2}}. \tag{5.20}
\]
In addition, by (3.6), (5.2), (4.4), (5.10), the Cauchy-Schwarz inequality, (2.21), (4.10) and (5.20) we have
\[
\left| E \left( (X_{\tau_{n+1}} - Y_{n+1}) | \mathcal{F}_{\tau_n} \right) \right| \leq E \left( \sum_{\alpha \in A_\gamma \setminus v} (-1)^{l(\alpha)} M_\alpha \left[ \mathcal{L}_\alpha \left( \tau_{n+1}, X_{\tau_{n+1}} \right) \right]_{\tau_n, \tau_{n+1}} + \sum_{\alpha \in B(A_\gamma)} (-1)^{l(\alpha)} M_\alpha \left[ \mathcal{L}_\alpha \left( \cdot, X_{\tau_{n+1}} \right) \right]_{\tau_n, \tau_{n+1}} - \sum_{\alpha \in A_\gamma \setminus v} (-1)^{l(\alpha)} M_\alpha \left[ \mathcal{L}_\alpha \left( \tau_{n+1}, Y_{n+1} \right) \right]_{\tau_n, \tau_{n+1}} \right) \left| \mathcal{F}_{\tau_n} \right| \leq E \left( \sum_{\alpha \in A_\gamma \setminus v} (-1)^{l(\alpha)} M_\alpha \left[ \mathcal{L}_\alpha \left( \tau_{n+1}, X_{\tau_{n+1}} \right) - \mathcal{L}_\alpha \left( \tau_{n+1}, Y_{n+1} \right) \right]_{\tau_n, \tau_{n+1}} \right) \left| \mathcal{F}_{\tau_n} \right|.
\]
and \( (5.9) \), we can easily get the expectations, so we do not need to make sure formula (5.22) is true. Thus by (5.22) and (5.9), we can easily get

\[
E \left( \sup_{\tau_n \leq s \leq \tau_{n+1}} |X_s|^4 \middle| \mathcal{F}_{\tau_n} \right) \leq K \left( 1 + X_{\tau_n}^4 \right),
\]

where the constant \( K \) does not depend on \( \Delta \). In (5.22), the expectations are conditional expectations, so we do not need \( E(|X_{\tau_n}|^4) < \infty \) to make sure formula (5.22) is true. Thus by (5.22) and (5.9), we can easily get

\[
\bar{R}_f \leq C_{10} \left( 1 + X_{\tau_n}^2 \right)^{\frac{1}{2}},
\]

where the constant \( C_{10} \) does not depend on \( \Delta \).

Substitute (5.23) into (5.20) and (5.21), we complete the proof of Theorem 5.2 by using Theorem 5.3.

About the boundedness of our numerical solution’s moment, by the mean-square deviation condition (1.6) in \([6]\), or (5.14) in this paper) and the conditional expectation version of [3, Theorem 4.5.4], we have

\[
E[Y_{n+1}^2] = E \left( E(Y_{n+1}^2 \middle| \mathcal{F}_{\tau_n}) \right)
= E \left( E \left( Y_{n+1}^2 - X_{\tau_{n+1}}^2 Y_n + X_{\tau_{n+1}}^2 \middle| \mathcal{F}_{\tau_n} \right) \right)
\leq 2E \left( E \left( Y_{n+1}^2 - X_{\tau_{n+1}}^2 \middle| \mathcal{F}_{\tau_n} \right) \right) + 2E \left( E \left( |X_{\tau_{n+1}}^2| \middle| \mathcal{F}_{\tau_n} \right) \right)
\leq K \left( 1 + E|Y_n|^2 \right),
\]

where the constant \( K \) does not depend on \( \Delta \). Because \( E(|Y_0|^2) < \infty \), then \( E(|Y_k|^2) < \infty \) for \( k = 0, 1, \ldots, N \). Furthermore we have

\[
E(|Y_k|^2) \leq K \left( 1 + E|Y_0|^2 \right)
\]

for \( k = 0, 1, \ldots, N \), where the constant \( K \) does not depend on \( \Delta \). More details, we can find in [6, Lemma 1.2, p. 14]. So, if the condition of fundamental mean-square convergence theorem (Theorem 5.3) is true, the second moment of numerical solution is bounded. And in the proof of
Theorem 5.2 (we prove that the Theorem 5.3 is true for our methods in the proof), we did not use the condition that the moments of numerical solutions are bounded. In the proof, we use the uniform moment estimate (4.5.16) in [3] to estimate $R_x$ in (5.22). But the uniform moment estimate used by us is the conditional expectation version, so the boundedness of numerical solution’s moment is not required. Hence the second moment of our numerical solution is bounded.

Next, we consider (1.1) for the situation of autonomous 1-dimensional case $d = m = 1$ and which $a(t, X_t)$ and $b(t, X_t)$ are scalar functions about $X_t$.

Let $\gamma = 1$ in (5.1), then the full implicit order 1 strong Ito-Taylor scheme has the form

$$Y_{n+1} = Y_n + a(Y_{n+1}) \Delta + b(Y_{n+1}) \Delta W - \frac{1}{2} b(Y_{n+1}) b' (Y_{n+1}) (\Delta W)^2 + \Delta. \quad (5.26)$$

Let $\gamma = 1$ in (5.2), then the full implicit order 1 strong Stratonovich-Taylor scheme has the form

$$Y_{n+1} = Y_n + \underline{g}(Y_{n+1}) \Delta + b(Y_{n+1}) \Delta W - \frac{1}{2} b(Y_{n+1}) b' (Y_{n+1}) (\Delta W)^2. \quad (5.27)$$

By the definition (1.5) of $\underline{g}$, we can see that (5.26) and (5.27) are coincident with the implicit Milstein-Taylor method which had been presented in [9].

Let $\gamma = 2$ in (5.2), then the full implicit order 2 strong Stratonovich-Taylor scheme has the form

$$\begin{align*}
Y_{n+1} &= Y_n + \underline{g}(Y_{n+1}) \Delta + b(Y_{n+1}) \Delta W - \frac{1}{2} b(Y_{n+1}) b' (Y_{n+1}) (\Delta W)^2 \\
&\quad - b(Y_{n+1}) \underline{g}'(Y_{n+1}) M_{(1,0)} - \frac{1}{2} b(Y_{n+1}) \underline{g}'(Y_{n+1}) \Delta^2 \\
&\quad - \underline{g}(Y_{n+1}) b' (Y_{n+1}) M_{(0,1)} + \frac{1}{6} b(Y_{n+1}) b' (Y_{n+1}) b' (Y_{n+1}) (\Delta W)^3 \\
&\quad + \underline{g}(Y_{n+1}) \left(b(Y_{n+1}) b' (Y_{n+1}) \right)^{\prime} M_{(0,1,1)} + b(Y_{n+1}) \left(\underline{g}(Y_{n+1}) b' (Y_{n+1}) \right)^{\prime} M_{(1,0,1)} \\
&\quad + b(Y_{n+1}) \left(\underline{g}(Y_{n+1}) \underline{g}'(Y_{n+1}) \right)^{\prime} M_{(1,1,0)} \\
&\quad - \frac{1}{24} b(Y_{n+1}) \left(b(Y_{n+1}) b' (Y_{n+1}) \right)^{\prime} (\Delta W)^4. \quad (5.28)
\end{align*}$$

Analogously to the order 2 strong Taylor scheme in [3], we need to approximate the multiple backward Stratonovich integrals $M_{(1,0)}, M_{(0,1)}, M_{(1,1,0)}, M_{(1,0,1)}$ and $M_{(0,1,1)},$ which can not be expressed in terms of $\Delta$ and $\Delta W$. By (5.2.45) in [3], we have

$$\begin{align*}
\Delta J_{(j)} &= J_{(j,0)} + J_{(0,j)}, && (5.29a) \\
\Delta J_{(j,j)} &= J_{(j,j,0)} + J_{(j,0,j)} + J_{(0,j,j)}, && (5.29b) \\
J_{(j,j)}(0,j) &= 2J_{(0,j,j)} + J_{(j,0,j)}, && (5.29c) \\
J_{(j,j)}(j,0) &= J_{(j,0,j)} + 2J_{(j,j,0)}. && (5.29d)
\end{align*}$$

Combining these with (2.20), we obtain

$$\begin{align*}
J_{(1,0)} + J_{(0,1)} &= \Delta W \Delta, && (5.30a) \\
M_{(1,0)} &= J_{(0,1)}, && M_{(0,1)} = J_{(1,0)}, \quad M_{(1,1,0)} = J_{(0,1,1)}, \quad M_{(1,0,1)} = J_{(1,0,1)}, \quad M_{(0,1,1)} = J_{(1,1,0)}, \quad (5.30b) \\
M_{(1,0,1)} &= J_{(1,0,1)}, \quad M_{(0,1,1)} = J_{(1,1,0)}. \quad (5.30c)
\end{align*}$$
tends to zero. In fact, this condition will be satisfied as \( \Delta \) can not guarantee the stability or well-posedness of the method. We could easy to confirm that the differential of its solving functions should keep away from the zero, without this condition we could not guarantee the convergence of full implicit strong Taylor schemes besides of the Lipschitz condition and linear growth bound. Actually, it is a necessary condition for all implicit method which the convergence of full implicit strong Taylor schemes besides of the Lipschitz condition and linear growth bound. Additionally, it is a necessary condition for all implicit method which the convergence of full implicit strong Taylor schemes besides of the Lipschitz condition and linear growth bound.

The key point is the property of multiple stochastic integrals which the multiple Ito Integrals have zero expectation (except non-stochastic integrals), see Lemma 4.1, but others not have.

Another point we should point out here, we have an additional condition (5.7) or (5.11) for the convergence of full implicit strong Taylor schemes besides of the Lipschitz condition and linear growth bound. Actually, it is a necessary condition for all implicit method which the differential of its solving functions should keep away from the zero, without this condition we can not guarantee the stability or well-posedness of the method. We could easy to confirm that the condition is satisfied for the linear situation. In fact, this condition will be satisfied as \( \Delta \) tends to zero.

6. Stability Properties

In this section, we investigate our scheme’s mean-square stability for the linear test equation

\[
    dX_t = aX_t dt + bX_t dW_t.
\]
In this case, our full implicit strong Ito-Taylor scheme and the Stratonovich-Taylor scheme has the following

\[ Y_{n+1} = R(\Delta, a, b, \tilde{N})Y_n, \quad (6.2) \]

where \( \tilde{N} \) is a standard normal random variable. The numerical scheme is said to be MS-stable, if

\[ \mathcal{R} := E \left( |R(\Delta, a, b, \tilde{N})|^2 \right) < 1. \quad (6.3) \]

In this section, we consider the full implicit order 2.0 strong Stratonovich-Taylor scheme, and compare it with the order 2.0 strong Taylor scheme \([3]\) and the implicit order 2.0 strong Taylor scheme ((12.2.19) or (12.2.20) with \( \alpha = \beta = \frac{1}{2} \) in \([3]\)). As the order 2.0 strong Taylor scheme is explicit and Stratonovich format, we call it the explicit order 2.0 strong Stratonovich-Taylor scheme in this paper. Similarly, as the implicit order 2.0 strong Taylor scheme is semi-implicit and Stratonovich format, we call it the semi-implicit order 2.0 strong Stratonovich-Taylor scheme in this paper.

Let \( p = a \delta, q = b \delta \). Set \( R_i \) and \( \mathcal{R}_i \) to represent \( R(\delta, a, b, \tilde{N}) \) and \( \mathcal{R} \) for different schemes. For the full implicit order 2.0 strong Stratonovich-Taylor scheme, we have

\[ R_1 = 1/ \left( 1 - \left( p - \frac{1}{2} q^2 \right) + \frac{1}{2} \left( p - \frac{1}{2} q^2 \right)^2 + \left( p - \frac{1}{2} q^2 - 1 \right) q \tilde{N} \right) \nonumber \]

\[ + \frac{1}{2} \left( 1 - p + \frac{1}{2} q^2 \right) q^2 \tilde{N}^2 - \frac{1}{6} (q \tilde{N})^3 + \frac{1}{24} (q \tilde{N})^4 \right), \quad (6.4a) \]

\[ \mathcal{R}_1 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \left( 1 - \left( p - \frac{1}{2} q^2 \right) + \frac{1}{2} \left( p - \frac{1}{2} q^2 \right)^2 + \left( p - \frac{1}{2} q^2 - 1 \right) q x \right) \nonumber \]

\[ + \frac{1}{2} \left( 1 - p + \frac{1}{2} q^2 \right) q^2 x^2 - \frac{1}{6} (q x)^3 + \frac{1}{24} (q x)^4 \right)^2 e^{-x^2/2} dx. \quad (6.4b) \]

For the semi-implicit order 2.0 strong Stratonovich-Taylor Scheme, we have

\[ R_2 = \frac{1 + q \tilde{N} + \frac{1}{2} (q \tilde{N})^2 + \frac{1}{6} (q \tilde{N})^3 + \frac{1}{24} (q \tilde{N})^4}{1 - p + \frac{1}{2} q^2 + \frac{1}{2} (p - \frac{1}{2} q^2)^2}, \quad (6.5a) \]

\[ \mathcal{R}_2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \left( \frac{1 + q x + \frac{1}{2} (q x)^2 + \frac{1}{6} (q x)^3 + \frac{1}{24} (q x)^4}{1 - p + \frac{1}{2} q^2 + \frac{1}{2} (p - \frac{1}{2} q^2)^2} \right)^2 e^{-x^2/2} dx. \quad (6.5b) \]

For the explicit order 2.0 strong Stratonovich-Taylor Scheme, we have

\[ R_3 = 1 + p - \frac{1}{2} q^2 + \frac{1}{2} \left( p - \frac{1}{2} q^2 \right)^2 + \left( 1 + p - \frac{1}{2} q^2 \right) q \tilde{N} \nonumber \]

\[ + \frac{1}{2} \left( 1 + p - \frac{1}{2} q^2 \right) q^2 \tilde{N}^2 + \frac{1}{6} (q \tilde{N})^3 + \frac{1}{24} (q \tilde{N})^4 \right), \quad (6.6a) \]

\[ \mathcal{R}_3 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \left( 1 + p - \frac{1}{2} q^2 + \frac{1}{2} \left( p - \frac{1}{2} q^2 \right)^2 + \left( 1 + p - \frac{1}{2} q^2 \right) q x \right) \nonumber \]

\[ + \frac{1}{2} \left( 1 + p - \frac{1}{2} q^2 \right) q^2 x^2 - \frac{1}{6} (q x)^3 + \frac{1}{24} (q x)^4 \right)^2 e^{-x^2/2} dx. \quad (6.6b) \]
We use the function *integral* in Matlab to compute $R_1, R_2, R_3$ on a $[-2:0.01:2] \times [-2:0.01:2]$ grid. Fig. 6.1 shows that the full implicit scheme have larger MS-stable region when compared with the explicit and semi-implicit schemes. Particularly, the full implicit scheme is MS-stable when $|q|$ is large relative to $|p|$, but the explicit and semi-implicit are not. This means that the full implicit scheme will be suitable for stiff stochastic differential equations with stiffness on stochastic terms and the other schemes are not suitable.

7. Numerical Results

In order to evaluate our full implicit strong Taylor schemes, we test three examples as our numerical experiments. We use the full implicit order 2.0 strong Stratonovich-Taylor scheme to calculate the examples and compare it with the explicit and semi-implicit order 2.0 strong Stratonovich-Taylor scheme. At each step, we need generate $\Delta W$ and the approximations of multiple stochastic integrals which mentioned in Section 5, and solve the implicit equation by Newton-Raphson method. We use the following method which mentioned in [3, Section 9.3] to evaluate our scheme.

We calculate the error of an approximation by the absolute error criterion

$$\epsilon = E(|X_T - Y_N|).$$  \hspace{1cm} (7.1)  

We arrange the simulations into $M$ batches of $N$ simulations each. We denote by $Y_{T,k,j}$ the value of the $k$-th generated trajectory in the $j$-th batch at time $T$ and by $X_{T,k,j}$ the corresponding
value of the Ito process. The average errors
\[ \hat{\epsilon}_j = \frac{1}{N} \sum_{k=1}^{N} |X_{T,k,j} - Y_{T,k,j}| \] (7.2)
of the $M$ batches $j = 1, \ldots, M$ are then independent and approximately Gaussian for large $N$. We estimate the mean of the batch averages
\[ \hat{\epsilon} = \frac{1}{M} \sum_{j=1}^{M} \hat{\epsilon}_j = \frac{1}{NM} \sum_{j=1}^{M} \sum_{k=1}^{N} |X_{T,k,j} - Y_{T,k,j}| \] (7.3)
and then use the formula
\[ \hat{\sigma}^2 = \frac{1}{M-1} \sum_{j=1}^{M} (\hat{\epsilon}_j - \hat{\epsilon})^2 \] (7.4)
to estimate the variance $\sigma^2$ of the batch averages. For the Student t-distribution with $M - 1$ degrees of freedom an $100(1 - \alpha)\%$ confidence interval for $\epsilon$ has the form
\[ (\hat{\epsilon} - \Delta \hat{\epsilon}, \hat{\epsilon} + \Delta \hat{\epsilon}) \] (7.5)
with
\[ \Delta \hat{\epsilon} = t_{1-\alpha,M-1} \sqrt{\frac{\hat{\sigma}^2}{M}}, \] (7.6)
where $t_{1-\alpha,M-1}$ is determined from the Student t-distribution with $M - 1$ degrees of freedom.

In this section, we choose $M = 100$, $N = 100$ as the limit of computing capability and $\alpha = 0.1$ for a 90% confidence interval. And we need to state that $M = 100$, $N = 100$ is not large enough to get a stable confidence interval when the diffusion terms are large, but we still can see the performance from the magnitude of results.

**Example 7.1.** We consider the scalar linear Ito SDE
\[ dX_t = aX_t dt + bX_t dW_t, \] (7.7)
where $W_t$ is a 1-dimensional Wiener process. The explicit solution is
\[ X_t = X_0 \exp \left( \left( a - \frac{1}{2} b^2 \right) t + bW_t \right). \]
This equation has been considered in [2, 11] to test some stiff methods. It could be thought as stiff in deterministic term, if $a$ is large and be thought as stiff in stochastic term, if $b$ is large.

<table>
<thead>
<tr>
<th>$\Delta$</th>
<th>90% confidence interval for the absolute error $\epsilon$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>explicit scheme</td>
</tr>
<tr>
<td>$2^{-2}$</td>
<td>[0.0113, 0.0118]</td>
</tr>
<tr>
<td>$2^{-3}$</td>
<td>[0.0032, 0.0033]</td>
</tr>
<tr>
<td>$2^{-4}$</td>
<td>[0.8361, 0.8679] $\times 10^{-3}$</td>
</tr>
<tr>
<td>$2^{-5}$</td>
<td>[0.2122, 0.2228] $\times 10^{-3}$</td>
</tr>
<tr>
<td>$2^{-6}$</td>
<td>[0.5367, 0.5619] $\times 10^{-4}$</td>
</tr>
</tbody>
</table>
Table 7.2: Numerical results for Example 7.1 with $a = 10, b = 0.1$ and $X_0 = 0.5$.

<table>
<thead>
<tr>
<th>$\Delta$</th>
<th>90% confidence interval for the absolute error $\epsilon$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>explicit scheme</td>
</tr>
<tr>
<td>$2^{-2}$</td>
<td>/</td>
</tr>
<tr>
<td>$2^{-3}$</td>
<td>/</td>
</tr>
<tr>
<td>$2^{-4}$</td>
<td>/</td>
</tr>
<tr>
<td>$2^{-5}$</td>
<td>/</td>
</tr>
<tr>
<td>$2^{-6}$</td>
<td>[390.4453, 392.1787]</td>
</tr>
</tbody>
</table>

Table 7.3: Numerical results for Example 7.1 with $a = -10, b = 0.1$ and $X_0 = 0.5$.

<table>
<thead>
<tr>
<th>$\Delta$</th>
<th>90% confidence interval for the absolute error $\epsilon$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>explicit scheme</td>
</tr>
<tr>
<td>$2^{-2}$</td>
<td>[3.4786, 3.4891]</td>
</tr>
<tr>
<td>$2^{-3}$</td>
<td>[0.0032, 0.0032]</td>
</tr>
<tr>
<td>$2^{-4}$</td>
<td>[0.3990, 0.3995] $\times 10^{-4}$</td>
</tr>
<tr>
<td>$2^{-5}$</td>
<td>[0.5175, 0.5186] $\times 10^{-5}$</td>
</tr>
<tr>
<td>$2^{-6}$</td>
<td>[0.1062, 0.1064] $\times 10^{-5}$</td>
</tr>
</tbody>
</table>

Table 7.4: Numerical results for Example 7.1 with $a = 1, b = 10$ and $X_0 = 0.5$.

<table>
<thead>
<tr>
<th>$\Delta$</th>
<th>90% confidence interval for the absolute error $\epsilon$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>explicit scheme</td>
</tr>
<tr>
<td>$2^{-2}$</td>
<td>/</td>
</tr>
<tr>
<td>$2^{-3}$</td>
<td>/</td>
</tr>
<tr>
<td>$2^{-4}$</td>
<td>/</td>
</tr>
<tr>
<td>$2^{-5}$</td>
<td>/</td>
</tr>
<tr>
<td>$2^{-6}$</td>
<td>[0.1062, 0.1096] $\times 10^{-3}$</td>
</tr>
</tbody>
</table>

Example 7.2. We consider the 2-dimension linear Ito SDE

$$dX_t = AX_t dt + BX_t dW_t,$$  \hspace{1cm} (7.8)

where

$$A = \begin{bmatrix} -a & a \\ a & -a \end{bmatrix}, \quad B = \begin{bmatrix} b & 0 \\ 0 & b \end{bmatrix},$$

and $W_t$ is a 1-dimensional Wiener process. The explicit solution is

$$X_t = P \left[ \begin{array}{c} \exp \left( \rho^+(t) \right) \\ 0 \end{array} \right] P^{-1} X_0, \quad P = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

where

$$\rho^\pm(t) = \left( -a - \frac{1}{2} b^2 \pm a \right) t + b W(t) \quad \text{and} \quad P^{-1} = P.$$

This equation has been considered in [9–11] to test some stiff methods. It could be thought as stiff in deterministic term, if $a$ is large and be thought as stiff in stochastic term, if $b$ is large.
Example 7.3. We consider a 1-dimensional nonlinear Itô SDE
\[ dX_t = -(\alpha + \beta^2 X_t)(1 - X_t^2)dt + \beta(1 - X_t^2)dW_t, \quad (7.9) \]
where \( W_t \) is a 1-dimensional Wiener process. The explicit solution is
\[ X_t = \frac{(1 + X_0) \exp(-2\alpha t + 2\beta W(t)) + X_0 - 1}{(1 + X_0) \exp(-2\alpha t + 2\beta W(t)) - X_0 + 1}, \]
where \(|X_0| < 1\). This equation has been considered in [2] to test some stiff methods. It could be thought as stiff in deterministic term, if \( \alpha \) is large and be thought as stiff in stochastic term, if \( \beta \) is large.
7.11 corresponding to the normal situation which the SDE is nonstiff shows that the three schemes all are unstable. Table 7.2 corresponding to the white area in Fig. 6.1(a) shows that the three schemes all are unstable. Table 7.3 corresponding to the shadow area in Fig. 6.1(b) but not in Fig. 6.1(c) shows that the full implicit has no much difference with the semi-implicit scheme and the implicit schemes have better performance than explicit scheme. Tables 7.4-7.6 corresponding to the shadow area in Fig. 6.1(a) but not in Fig. 6.1(b) shows that the full implicit scheme has better performance than the semi-implicit scheme and the semi-implicit scheme has better performance than the explicit scheme.

Example 7.2 and 7.3 have coincident numerical results with Example 7.1. Tables 7.7 and 7.8 corresponding to the normal situation which the SDE is nonstiff show that the three schemes all are unstable. Table 7.9 corresponding to the shadow area in Fig. 6.1(b) but not in Fig. 6.1(c) shows that the full implicit has no much difference with the semi-implicit scheme and the implicit schemes have better performance than explicit scheme.

<table>
<thead>
<tr>
<th>Δ</th>
<th>90% confidence interval for the absolute error ε</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>explicit scheme</td>
</tr>
<tr>
<td>2^{-2}</td>
<td>/</td>
</tr>
<tr>
<td>2^{-3}</td>
<td>[24.2791 24.3468]</td>
</tr>
<tr>
<td>2^{-4}</td>
<td>[0.2010 0.2013] × 10^{-4}</td>
</tr>
<tr>
<td>2^{-5}</td>
<td>0.6803 0.6823 × 10^{-8}</td>
</tr>
<tr>
<td>2^{-6}</td>
<td>0.5239 0.5274 × 10^{-9}</td>
</tr>
</tbody>
</table>

Table 7.8: Numerical results for Example 7.2 with a = 10, b = 0.1 and X₀ = [1 0]^T.
Table 7.10: Numerical results for Example 7.2 with \( a = 10, b = 10 \) and \( X_0 = [1 \ 0]^T \).

<table>
<thead>
<tr>
<th>( \Delta )</th>
<th>90% confidence interval for the absolute error ( \epsilon )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>explicit scheme</td>
</tr>
<tr>
<td>( 2^{-2} )</td>
<td>/</td>
</tr>
<tr>
<td>( 2^{-3} )</td>
<td>/</td>
</tr>
<tr>
<td>( 2^{-4} )</td>
<td>/</td>
</tr>
<tr>
<td>( 2^{-5} )</td>
<td>/</td>
</tr>
<tr>
<td>( 2^{-6} )</td>
<td>( [0.2797 \times 10^{-5}] )</td>
</tr>
</tbody>
</table>

Table 7.11: Numerical results for Example 7.3 with \( a = 1, b = 0.1 \) and \( X_0 = 0.5 \).

<table>
<thead>
<tr>
<th>( \Delta )</th>
<th>90% confidence interval for the absolute error ( \epsilon )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>explicit scheme</td>
</tr>
<tr>
<td>( 2^{-2} )</td>
<td>( [0.0126, 0.0127] )</td>
</tr>
<tr>
<td>( 2^{-3} )</td>
<td>( [0.0031, 0.0032] )</td>
</tr>
<tr>
<td>( 2^{-4} )</td>
<td>( [0.7847, 0.8033] \times 10^{-3} )</td>
</tr>
<tr>
<td>( 2^{-5} )</td>
<td>( [0.1976, 0.2001] \times 10^{-3} )</td>
</tr>
<tr>
<td>( 2^{-6} )</td>
<td>( [0.4905, 0.5004] \times 10^{-4} )</td>
</tr>
</tbody>
</table>

Table 7.12: Numerical results for Example 7.3 with \( a = 5, b = 0.1 \) and \( X_0 = 0.5 \).

<table>
<thead>
<tr>
<th>( \Delta )</th>
<th>90% confidence interval for the absolute error ( \epsilon )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>explicit scheme</td>
</tr>
<tr>
<td>( 2^{-2} )</td>
<td>( [2.0856, 3.2012] )</td>
</tr>
<tr>
<td>( 2^{-3} )</td>
<td>( [0.0039, 0.0039] )</td>
</tr>
<tr>
<td>( 2^{-4} )</td>
<td>( [0.2075, 0.2082] \times 10^{-3} )</td>
</tr>
<tr>
<td>( 2^{-5} )</td>
<td>( [0.3311, 0.3324] \times 10^{-4} )</td>
</tr>
<tr>
<td>( 2^{-6} )</td>
<td>( [0.7023, 0.7052] \times 10^{-5} )</td>
</tr>
</tbody>
</table>

Schemes all are stable and have no much difference. Tables 7.8 and 7.12 corresponding to the stiff situation which the SSDE is stiff on drift terms show that the full implicit has no much difference with the semi-implicit scheme and the implicit schemes have better performance than the explicit scheme. Tables 7.9, 7.10, 7.13 and 7.14 corresponding to the stiff situation which the SSDE is stiff on diffusion terms show that the full implicit scheme has better performance than other two schemes.

About the computation time of our full implicit schemes, the only difference with semi-implicit is that we need to compute the derivative of additional implicit terms at very step of
Table 7.13: Numerical results for Example 7.3 with $a = 1$, $b = 2$ and $X_0 = 0.5$.

<table>
<thead>
<tr>
<th>$\Delta$</th>
<th>90% confidence interval for the absolute error $\epsilon$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>explicit scheme</td>
</tr>
<tr>
<td>$2^{-2}$</td>
<td>/</td>
</tr>
<tr>
<td>$2^{-3}$</td>
<td>/</td>
</tr>
<tr>
<td>$2^{-4}$</td>
<td>/</td>
</tr>
<tr>
<td>$2^{-5}$</td>
<td>/</td>
</tr>
<tr>
<td>$2^{-6}$</td>
<td>[0.0071, 0.0097]</td>
</tr>
</tbody>
</table>

Table 7.14: Numerical results for Example 7.3 with $a = 5$, $b = 2$ and $X_0 = 0.5$.

<table>
<thead>
<tr>
<th>$\Delta$</th>
<th>90% confidence interval for the absolute error $\epsilon$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>explicit scheme</td>
</tr>
<tr>
<td>$2^{-2}$</td>
<td>/</td>
</tr>
<tr>
<td>$2^{-3}$</td>
<td>/</td>
</tr>
<tr>
<td>$2^{-4}$</td>
<td>/</td>
</tr>
<tr>
<td>$2^{-5}$</td>
<td>/</td>
</tr>
<tr>
<td>$2^{-6}$</td>
<td>/</td>
</tr>
</tbody>
</table>

Newton-Raphson method. So we can improve the stability of simulations considerably without too much additional computational effort by using our full implicit schemes.

Acknowledgments. We would like to thank the anonymous referees (especially the first reviewer) for their valuable comments and suggestions, which greatly improved the original manuscript. This work was supported by the Fundamental Research Funds for the Central Universities of China, and the second author is supported by the National Natural Fund Projects of China (Nos.11771100, 12071332).

References


