THE TRUNCATED MILSTEIN METHOD FOR SUPER-LINEAR
STOCHASTIC DIFFERENTIAL EQUATIONS WITH
MARKOVIAN SWITCHING

WEIJUN ZHAN
Department of Mathematics, Anhui Normal University
Wuhu 241000, China

QIAN GUO∗
Department of Mathematics, Shanghai Normal University
Shanghai 200234, China

YUHAO CONG
Shanghai Customs College, Shanghai 201204, China

(Communicated by Arnulf Jentzen)

Abstract. In this paper, to approximate the super-linear stochastic differential equations modulated by a Markov chain, we investigate a truncated Milstein method with convergence order 1 in the mean-square sense. Under Khasminskii-type conditions, we establish the convergence result by employing a relationship between local and global errors. Finally, we confirm the convergence rate by a numerical example.

1. Introduction. The Markov-modulated stochastic differential equations (SDEs) have been subjected to many research papers. But it is virtually impossible to obtain the exact solution of a highly nonlinear stochastic differential systems. Thus the feasible and efficient numerical methods have an important sense for SDEs with Markovian switching. Up to 2017, only the Euler-Maruyama (EM) scheme has been applied to approximate the SDEs with Markovian switching [6, 7, 12, 13]. As we know, the convergence order of the EM method is 0.5. The Milstein-type scheme is a natural candidate to improve the convergence rate. Recently, Nguyen et al. [9] developed a Milstein-type scheme to deal with Markov-modulated SDEs with globally Lipschitz coefficients, which is a pioneer work in this field. Therefore, it is interesting to extend the idea to more stochastic models, such as the Markov-modulated SDEs under Khasminskii-type conditions.

However, for super-linear SDEs, as we know, Hutzenthaler et al. [4] proved the explicit Euler method is divergent. Then, they presented a modified Euler method, named as tamed EM method, which guarantees to converge to the real solution [3]. Later, Hutzenthaler et al. [2] extended the tamed Euler method to a general class of

2020 Mathematics Subject Classification. Primary: 60H35; Secondary: 65C30.
Key words and phrases. Stochastic differential equation, Milstein method, Khasminskii-type condition, strong convergence, Markovian switching.

The second author is supported by NSFC of China (No:11871343). The third author is supported by NSFC of China (No:11971303).

∗ Corresponding author: Qian Guo.
super-linear SDEs. For the super-linear SDEs without Markov-switching, Wang et al. [11] investigated the tamed Milstein method to improve the convergence order. Recently, Mao [6, 7] proposed another modified Euler method, which called the truncated EM method, for solving SDEs under the Khasminskii condition. Meanwhile, Guo et al. [1] proposed a truncated Milstein-type explicit method for solving the super-linear SDEs. In this paper, we will introduce a truncated Milstein method applied to the SDEs with Markovian switching under the general Khasminskii-type condition.

The rest of this paper is arranged as follows. Section 2 begins with the preliminaries, formulation and some useful lemmas. We will devote to our main results about the convergence of the numerical algorithms in Section 3. Section 4 gives an example to illustrate the performance of the numerical scheme.

2. Mathematical preliminaries.

2.1. Notation. Throughout this paper, unless otherwise specified, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. Suppose $\{\mathcal{F}_t\}_{t \geq 0}$ is a filtration defined on this probability space satisfying the usual conditions (i.e., it is right continuous and $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets) such that $B(t)$ and $\alpha(t)$ are $\{\mathcal{F}_t\}$ adapted. Let $\mathbb{E}$ denote the expectation corresponding to $\mathbb{P}$. If $A$ is a vector or matrix, its transpose is denoted by $A^T$. Let $B(t) = (B^1(t), B^2(t), ..., B^m(t))^T$ be an $m$-dimensional Brownian motion defined on the space. If $x \in \mathbb{R}^d$, then $|x|$ is the Euclidean norm. For two real numbers $a$ and $b$, set $a \vee b = \max(a, b)$ and $a \wedge b = \min(a, b)$. If $G$ is a set, its indicator function is denoted by $I_G$, namely $I_G(x) = 1$ if $x \in G$ and 0 otherwise. Let $\frac{\partial f(x,u)}{\partial x}$ be the partial derivatives of $f$ with respect to $x$. We use $C^{1,2}(\cdot, \cdot)$ to denote the set of real-valued functions that are one time continuously differentiable with respect to the first variable and 2-times continuously differentiable with respect to the second variable.

Consider a $d$-dimensional commutative SDEs with Markov switching

$$dX(t) = b(X(t), \alpha(t))dt + \sigma(X(t), \alpha(t))dB(t) \quad t \geq 0,$$  

with the initial value $X(0) = x_0 \in \mathbb{R}^d$, $b(\cdot, \cdot) : \mathbb{R}^d \times \mathcal{M} \to \mathbb{R}^d$, $\sigma(\cdot, \cdot) : \mathbb{R}^{d \times m} \times \mathcal{M} \to \mathbb{R}^d$, and $X(t) = (X^1(t), X^2(t), ..., X^d(t))^T$, and the $\alpha(t)$ is Markov chain that takes values in the finite set $\mathcal{M} = \{1, 2, ..., m_0\}$. Meanwhile, we assume that $X_0, B(\cdot)$ and $\alpha(\cdot)$ are independent. In some of the proofs, in this paper, we need the more specified notation that $b = (b_1, b_2, ..., b_d)^T$ and $\sigma = (\sigma_{1,j}, \sigma_{2,j}, ..., \sigma_{d,j})^T$. For $j_1, j_2 = 1, 2, ..., m$, we assume that $b$ and $\sigma$ have continuous second-order derivatives and define

$$L^{j_1} \sigma_{j_2}(x, \varphi) = \sum_{i=1}^d \sigma_{j_1}(x, \varphi) \frac{\partial \sigma_{j_2}(x, \varphi)}{\partial x^i}.$$  

And in this paper, the commutative condition is that

$$L^{j_1} \sigma_{j_2} = L^{j_2} \sigma_{j_1}.$$  

For $l = 1, 2, ..., d$, set

$$b'_l(x, \varphi) = \left( \frac{\partial b_l(x, \varphi)}{\partial x^1}, \frac{\partial b_l(x, \varphi)}{\partial x^2}, ..., \frac{\partial b_l(x, \varphi)}{\partial x^d} \right)$$  

and $b''_l(x, \varphi) = \left( \frac{\partial^2 b_l(x, \varphi)}{\partial x^j \partial x^l} \right)_{i,j}$.  

And for $n = 1, 2, ..., m$, $l = 1, 2, ..., d$, set

$$\sigma'_{l,n}(x, \varphi) = \left( \frac{\partial \sigma_{l,n}(x, \varphi)}{\partial x^1}, \frac{\partial \sigma_{l,n}(x, \varphi)}{\partial x^2}, ..., \frac{\partial \sigma_{l,n}(x, \varphi)}{\partial x^d} \right).$$
2.2. Some useful lemmas. We quote the method from Section 2.4 in [8] to construct the continuous-time Markovian chain, with a given generator \( Q = (q_{i_0,j_0}) \in \mathbb{R}^{m_0 \times m_0} \). For the sample paths of \( \alpha(t) \) requires determining its sojourn time in each state and its subsequent moves. The chain sojourns in any given state \( i_0, i_0 \in M \), for a random length of time, \( \eta_{i_0} \), which has an exponential distribution with parameter \(-q_{i_0i_0}\). Subsequently, the process will enter another state. Each state \( j_0 \) (with \( j_0 \in M, j_0 \neq i_0 \)) has a probability \( q_{i_0j_0}/(-q_{i_0i_0}) \) of being the chain’s next residence.

The post jump location is then determined by a discrete random variable \( Z_{i_0} \) which is taking values in \( \{1, 2, \ldots, i_0 - 1, i_0 + 1, \ldots, m_0\} \). Its value is specified by

\[
Z_{i_0} \leq \begin{cases} 
1, & \text{if } U \leq q_{i_01}/(-q_{i_0i_0}), \\
2, & \text{if } q_{i_01}/(-q_{i_0i_0}) \leq U \leq (q_{i_01} + q_{i_02})/(-q_{i_0i_0}), \\
\vdots & \\
 m_0, & \text{if } \sum_{j_0 \neq i_0} q_{i_0j_0}/(-q_{i_0i_0}) \leq U,
\end{cases}
\]

where \( U \) is a random variable uniformly distributed in \((0, 1)\). Thus, the sample path of \( \alpha(t) \) is constructed by sampling from exponential and \( U/(0, 1) \) random variables alternately. With the \( \alpha(t) \) generated above, for \( n = 0, 1, \ldots \), set \( \alpha_n = \alpha_n^h = \alpha(t_n) \) which is the \( h \)-skeleton of the Markovian chain.

For a fixed step size \( \Delta, 0 < \Delta < 1 \), and \( n = 0, 1, 2, \ldots \), denote by \( N_n \) the total number of jumps of the chain in the interval \([t_n, t_{n+1})\) with the sequence of jump times \( t_n = \tau_0^0 < \tau_1^0 < \tau_2^0 < \cdots < \tau_N^0 < t_{n+1} \).

**Lemma 2.1** (cf. Lemma 4.1 in [9]). The following inequality is true

\[
\mathbb{P}(N_n \geq N) \leq q^N \Delta^N, \quad N \geq 1,
\]

where \( q = \max \{-q_{i_0j_0} : j_0 \in M\} \) and \( n = 0, 1, \ldots \). As a consequence, if \( \Delta \leq 1/(2q) \) there is a constant \( C \) independent of \( n \) such that

\[
\mathbb{E}N_n \leq C \Delta
\]

holds.

Thirdly, we will cite the following form of Itô’s formula to find the stochastic expansion of the solution to (1), the proof see [9]. For each pair \((i_0, j_0) \in M \times M, i_0 \neq j_0 \) and \( t \geq 0 \), we define

\[
[M_{i_0j_0}](t) = \sum_{0 \leq s \leq t} I(\alpha(s^-) = i_0)I(\alpha(s) = j_0), \quad \langle M_{i_0j_0} \rangle(t) = \int_0^t q_{i_0j_0}I(\alpha(s^-) = i_0)ds.
\]

Then the process \( M_{i_0j_0}(t), 0 \leq t \leq T \), define by

\[
M_{i_0j_0}(t) = [M_{i_0j_0}](t) - \langle M_{i_0j_0} \rangle(t)
\]

is a purely discontinuous and square integrable martingale with respect to \( F_t \), which is null at the origin. The process \([M_{i_0j_0}](t)\) and \( \langle M_{i_0j_0} \rangle(t) \) are the optional and predictable quadratic variations, respectively. For convenience, we denote \( M_{i_0i_0}(t) = 0 \) for \( i_0 \in M \) and \( 0 \leq t \leq T \).
Lemma 2.2 (cf. Lemma 2.2 in [9]). For a function \( f(\cdot, \cdot) : \mathbb{R}^d \times \mathcal{M} \rightarrow \mathbb{R} \) such that for each \( i_0 \in \mathcal{M} \), \( f(\cdot, \cdot) \in C^{1,2}([0, T] \times \mathbb{R}^d) \) we have

\[
f(X(t), \alpha(t)) = f(x(s), \alpha(s)) + \int_s^t \mathcal{L}f(X(u), \alpha(u))du + \sum_{l=1}^m \int_s^t \langle f'(X(u), \alpha(u)), \sigma_l(X(u), \alpha(u)) \rangle dB_l(u)
\]

\[
+ \sum_{i_0 \neq j_0} \int_s^t (f(X(u), j_0) - f(X(u), i_0))dM_{i_0j_0}(u), \quad 0 \leq s \leq t \leq T.
\]

Where

\[
\mathcal{L}f(x, i_0) = \frac{\partial}{\partial t}f(x, i_0) + \langle b(x, i_0), \frac{\partial}{\partial x}f(x, i_0) \rangle + \frac{1}{2} \text{trace} \left( \sigma^T(x, i_0) \frac{\partial^2 f(x, i_0)}{\partial x_i \partial x_j} \sigma(x, i_0) \right)
\]

\[
+ \sum_{j_0 \in \mathcal{M}} q_{i_0j_0}(f(x, j_0) - f(x, i_0)).
\]

Applying the Itô’s formula (2) to \( f = b \) and \( f = \sigma_l \) for \( l = 1, 2, \ldots, m \), and denote that \( X_n = X(t_n) \), we have following equation

\[
X_{n+1} = X_n + b(X_n, \alpha_n)\Delta + \sum_{l=1}^m \sigma_l(X_n, \alpha_n)\Delta B_l + \sum_{l_1, l_2=1}^m L^{l_2}\sigma_{l_1}(X_n, \alpha_n)I_{l_1, l_2}(n)
\]

\[
+ \sum_{l=1}^m \sum_{i_0 \neq j_0} \int_{t_n}^{t_{n+1}} \int_{t_n}^s (\sigma_l(X(u), j_0) - \sigma_l(X(u), i_0))dM_{i_0j_0}(u)dB_l(s)
\]

\[
+ \sum_{j=1}^6 R_{n,j},
\]

where

\[
R_{n,1} = \int_{t_n}^{t_{n+1}} \int_{t_n}^s \mathcal{L}b(X(u), \alpha(u))duds,
\]

\[
R_{n,2} = \sum_{l=1}^m \int_{t_n}^{t_{n+1}} \int_{t_n}^s L^{l}b(X(u), \alpha(u))dB_l(u)ds,
\]

\[
R_{n,3} = \sum_{i_0 \neq j_0} \int_{t_n}^{t_{n+1}} \int_{t_n}^s (b(X(u), j_0) - b(X(u), i_0))dM_{i_0j_0}(u)ds,
\]

\[
R_{n,4} = \sum_{l=1}^m \int_{t_n}^{t_{n+1}} \int_{t_n}^s \mathcal{L}\sigma(X(u), \alpha(u))dudB_l(s),
\]

\[
R_{n,5} = \sum_{l=1}^m \sum_{i_0 \neq j_0} \int_{t_n}^{t_{n+1}} \int_{t_n}^s (\sigma_l(X(u), i_0) - \sigma_l(X(u), j_0))d(M_{i_0j_0})(u)dB_l(s),
\]

\[
R_{n,6} = \sum_{l_1, l_2=1}^m \int_{t_n}^{t_{n+1}} \int_{t_n}^s (L^{l_2}\sigma_{l_1}(X(u), \alpha(u)) - L^{l_2}\sigma_{l_1}(X_n, \alpha_n))dB_{l_2}(u)dB_{l_1}(s),
\]

and for \( l, l_1, l_2 = 1, 2, \ldots, m, n = 0, 1, \ldots \)

\[
\Delta B_l = B_l(t_{n+1}) - B(t_n), \quad I_{l_1, l_2}(n) = \int_{t_n}^{t_{n+1}} \int_{t_n}^s dB_{l_2}(s_2)dB_{l_1}(s_1).
\]
An explanation for the above choice is that, based on the definition of \([M_{i0,j0}]\), the total contribution of the fifth term after all iterations in the scheme is \(O(\Delta^{1/2})\). The following Lemma gives a more convenient representation for this double integral term.

**Lemma 2.3** (cf. Lemma 4.3 in [9]). If \(N_n \geq 1\), for \(k = 1, 2, \ldots, d\) and \(l = 1, 2, \ldots, m\), we have

\[
\sum_{i_0 \neq j_0} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} (\sigma_l(X(u), j_0) - \sigma_l(X(u), i_0))d[M_{i0,j0}](u)dB_l(s)
\]

\[
= \sum_{i=1}^{N_n} [\sigma_i(X(\tau_i^n), \alpha(\tau_i^n)) - \sigma_i(X(\tau_i^n), \alpha(\tau_i^n))](B_l(t_{n+1}) - B_l(\tau_i^n)).
\]

If \(N_n = 0\) the left-hand side equals 0.

Since we have \(\alpha(\tau_i^n) = \alpha_{n+1}\) and \(\alpha(\tau_i^n) = \alpha_n\) on the set \(\{N_n = 1\}\), it follows from the above Lemma that

\[
\sum_{i_0 \neq j_0} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} (\sigma_l(X(u), j_0) - \sigma_l(X(u), i_0))d[M_{i0,j0}](u)dB_l(s)
\]

\[
= I_{\{N_n = 1\}}[\sigma_l(X(\tau_1^n), \alpha_{n+1}) - \sigma_l(X(\tau_1^n), \alpha_n)](B_l(t_{n+1}) - B_l(\tau_1^n)) + I_{\{N_n \geq 2\}} \sum_{i_0 \neq j_0} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} (\sigma_l(X(u), j_0) - \sigma_l(X(u), i_0))d[M_{i0,j0}](u)dB_l(s).
\]

Denote

\[
R_{n,7} = \sum_{l=1}^{m} I_{\{N_n = 1\}}[(\sigma_l(X(\tau_l^n), \alpha_{n+1}) - \sigma_l(X_n, \alpha_{n+1}))
\]

\[
- (\sigma_l(X(\tau_l^n), \alpha_n) - \sigma_l(X_n, \alpha_n))](B_l(t_{n+1}) - B_l(\tau_l^n))
\]

and

\[
R_{n,8} = \sum_{l=1}^{m} I_{\{N_n \geq 2\}} \sum_{i_0 \neq j_0} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} (\sigma_l(X(u), j_0) - \sigma_l(X(u), i_0))d[M_{i0,j0}](u)dB_l(s).
\]

We can rewrites (3) as

\[
X_{n+1} = X_n + b(X_n, \alpha_n)\Delta + \sum_{l=1}^{m} \sigma_l(X_n, \alpha_n)\Delta B_l + \sum_{l_1, l_2=1}^{m} L_{l_1}^{l_2} \sigma_{l_1}(X_n, \alpha_n)I_{l_1, l_2}(n)
\]

\[
+ \sum_{l=1}^{m} I_{\{N_n = 1\}}[\sigma_l(X_n, \alpha_{n+1}) - \sigma_l(X_n, \alpha_n)](B_l(t_{n+1}) - B_l(\tau_l^n)) + \sum_{j=1}^{8} R_{n,j}.
\]

Next, we write approximation solution \(Y_n\). Since the consideration regarding the terms involving double integer with respect to the optional quadratic variation process \([M_{i0,j0}]\) and the Brownian motions, and the discussion on approximation of the jump times of the Markovian chain \(\alpha\), the component sequences \((Y_n, n \geq 0)\) of the approximation solution should satisfy the following recursive equation,

\[
Y_{n+1} = Y_n + b(Y_n, \alpha_n)\Delta + \sum_{l=1}^{m} \sigma_l(Y_n, \alpha_n)\Delta B_l + \sum_{l_1, l_2=1}^{m} L_{l_1}^{l_2} \sigma_{l_1}(Y_n, \alpha_n)I_{l_1, l_2}(n)
\]
where we set \( x/6 \) WEIJUN ZHAN, QIAN GUO AND YUHAO CONG

\[ \omega \]

for any \( u \) we have \( \omega \) to

we choose a strictly increasing continuous function \( \omega : \mathbb{R}_+ \rightarrow \mathbb{R}_+ \) such that \( \omega(u) \rightarrow \infty \) as \( u \rightarrow \infty \) and

\[
\sup_{|x| \leq u} \frac{b(x,i)}{1+|x|} \sqrt{\sum_{i=1}^{m} |\sigma_l(x,i)|^2 + \sum_{l_1,l_2} |L^{l_2}\sigma_{l_1}(x,i)|} \leq \omega(u),
\]

for any \( u \geq 2, i = 1, ..., m_0 \) and \( l = 1, ..., d \). Denote the inverse function of \( \omega \) by \( \omega^{-1} \). We see that \( \omega^{-1} \) is a strictly increasing continuous function from \([\omega(0), +\infty)\) to \( \mathbb{R}_+ \). We also choose a number \( \Delta^* \in (0,1] \) and a strictly decreasing function \( h : (0, \Delta^*] \rightarrow (0, +\infty) \) such that

\[
\lim_{\Delta \rightarrow 0} h(\Delta) = \infty \quad \text{and} \quad \Delta^{1/2} h(\Delta) \leq K, \forall \Delta \in (0, \Delta^*],
\]

where \( K \) is a positive constant independent of \( \Delta \). For a given step size \( \Delta \in (0,1) \) and any \( x \in \mathbb{R}^d \), let us define a mapping \( \pi_\Delta \) from \( \mathbb{R}^d \) to the closed ball \( \{x \in \mathbb{R}^d : |x| \leq \omega^{-1}(h(\Delta))\} \) by

\[
\pi_\Delta(x) = \left( |x| \land \omega^{-1}(h(\Delta)) \right) \frac{x}{|x|},
\]

where we set \( x/|x| = 0 \) when \( x = 0 \). That is, \( \pi_\Delta \) will map \( x \) to itself when \( |x| \leq \omega^{-1}(h(\Delta)) \) and to \( \omega^{-1}(h(\Delta))x/|x| \) when \( |x| \geq \omega^{-1}(h(\Delta)) \). According to (4) we have

\[
|b(\pi_\Delta(x),i)|^2 \leq h^2(\Delta)(1 + |\pi_\Delta(x)|)^2,
\]

\[
\sum_{l=1}^{m} |\sigma_l(\pi_\Delta(x),i)|^2 \leq h(\Delta)(1 + |\pi_\Delta(x)|)^2,
\]

\[
\sum_{l_1,l_2} |L^{l_2}\sigma_{l_1}(\pi_\Delta(x),i)|^2 \leq h^2(\Delta)(1 + |\pi_\Delta(x)|)^2.
\]

Next we propose our numerical methods to approximation the exact solution of the SDEs (1). For any given step-size \( \Delta \in (0,1) \), the truncated Milstein methods is define by

\[
Y_0 = X_0,
\]

\[
\tilde{Y}_{n+1} = Y_n + b(Y_n, \alpha_n)\Delta + \sum_{l=1}^{m} \sigma_l(Y_n, \alpha_n)\Delta B_n + \sum_{l_1,l_2} L^{l_2}\sigma_{l_1}(Y_n, \alpha_n)I_{l_1,l_2}(n) + \sum_{l=1}^{m} I_{\{N_n=1\}}[\sigma_l(Y_n, \alpha_{n+1}) - \sigma_l(Y_n, \alpha_n)](B_l(t_{n+1}) - B_l(t_n)),
\]

\[
Y_{n+1} = \pi_\Delta(\tilde{Y}_{n+1}),
\]

where \( t_n = n\Delta \), \( \Delta B_n = B(t_{n+1}) - B(t_n) \). We refer to the numerical methods as a truncated Milstein scheme. Consequently, for any nonnegative integer \( n \) we have the following properties

\[
|b(Y_n,i)|^2 \leq h^2(\Delta)(1 + |Y_n|)^2,
\]

\[
\sum_{l=1}^{m} |\sigma_l(Y_n,i)|^2 \leq h(\Delta)(1 + |Y_n|)^2,
\]

\[
|b(Y_n,i)|^2 \leq h^2(\Delta)(1 + |Y_n|)^2,
\]
Proof. For algorithm (6), we have
\[
\sum_{l_1, l_2=1}^{m} |L^{l_2} \sigma_{l_1}(Y_n, i)|^2 \leq h^2(\Delta)(1 + |Y_n|)^2.
\]

3. Main results.

Assumption 3.1. There exist a pair of positive constant \(K\) and \(p > 2\), such that
\[
(x, b(x, i)) + m(p - 1) \sum_{l_1}^{m} |\sigma_l(x, i)|^2 \leq K(1 + |x|^2)
\]
for all \(x \in \mathbb{R}^d, i \in \mathcal{M}\).

We adapt the following lemma from Corollary 3.21 in [8], and the proof of the second result follows with a slight modification of the proof of Theorem 3.23 in [8] while using Itô’s formula first.

Lemma 3.2. Under Assumption 3.1, the SDEs (1) has a unique global solution \(X(t)\). Moreover, for any \(p > 2\), we have
\[
\sup_{0 < t \leq T} \mathbb{E}|X(t)|^p < \infty, \quad \forall T > 0,
\]
and
\[
\mathbb{E}|X(t) - X(s)|^p \leq C|t - s|^\frac{p}{2}, \quad 0 \leq s < t \leq T.
\]

Theorem 3.3. Under Assumption 3.1, there exists a constant \(C\) (stands for generic positive real constants dependent on \(T, p, x_0\), but independent of \(\Delta\)) such that for every \(p > 2\) satisfying Assumption 3.2, we have
\[
\sup_{0 < \Delta < 1 \leq n \Delta \leq T} \mathbb{E}|Y_n|^p \leq C < \infty.
\]

Proof. For algorithm (6), we have
\[
\bar{Y}_{n+1} = Y_n + b(Y_n, \alpha_n)\Delta + \sum_{l_1=1}^{m} \sigma_l(Y_n, \alpha_n)\Delta B_l + \sum_{l_1, l_2=1}^{m} L^{l_2} \sigma_{l_1}(Y_n, \alpha_n)I_{l_1, l_2}(n) \\
+ \sum_{l_1=1}^{m} I_{\{N_n=1\}}[\sigma_l(Y_n, \alpha_{n+1}) - \sigma_l(Y_n, \alpha_n)](B_l(t_{n+1}) - B_l(\tau_1^n))
\]
:= \bar{Y}_n + A_n,
where
\[
A_n = b(Y_n, \alpha_n)\Delta + \sum_{l_1=1}^{m} \sigma_l(Y_n, \alpha_n)\Delta B_l + \sum_{l_1, l_2=1}^{m} L^{l_2} \sigma_{l_1}(Y_n, \alpha_n)I_{l_1, l_2}(n) \\
+ \sum_{l_1=1}^{m} I_{\{N_n=1\}}[\sigma_l(Y_n, \alpha_{n+1}) - \sigma_l(Y_n, \alpha_n)](B_l(t_{n+1}) - B_l(\tau_1^n)).
\]
Squaring on both side, we can get
\[
|\bar{Y}_{n+1}|^2 = |Y_n|^2 + 2\langle Y_n, A_n \rangle + |A_n|^2.
\]
Thus
\[
(1 + |\bar{Y}_{n+1}|^2)^\frac{p}{2} = \left(1 + |Y_n|^2\right)^\frac{p}{2} \left(1 + \frac{2\langle Y_n, A_n \rangle + |A_n|^2}{1 + |Y_n|^2}\right)^\frac{p}{2},
\]
let
\[
\xi_n = \frac{2\langle Y_n, A_n \rangle + |A_n|^2}{1 + |Y_n|^2},
\]
Thanks to the Taylor formula, applying the recursion with \( u > 8 \), we have

\[
(1 + u)^{\frac{p}{2}} \leq 1 + \frac{p}{2} u + \frac{p(p - 2)}{8} u^2 + u^p P_i(u), \quad 2i \leq p \leq 2(i + 1).
\]

where \( P_i(u) \) represents an \( i \)th-order polynomial of \( u \) which coefficients depend only on \( p, i \) is an integer. It follows from (8) that

\[
\mathbb{E}((1 + |\tilde{Y}_{n + 1}|^2)^{\frac{p}{2}} | F_{t_n}) \leq (1 + |Y_n|^2)^{\frac{p}{2}} \mathbb{E}(\xi_n | F_{t_n}) + \frac{p(p - 2)}{8} \mathbb{E}(\xi_n^2 | F_{t_n}) + \mathbb{E}(\xi_n^3 P_i(\xi_n) | F_{t_n})].
\]

Recalling that (7), we have

\[
\mathbb{E}((1 + |\tilde{Y}_{n + 1}|^2)^{\frac{p}{2}} | F_{t_n}) \leq (1 + |Y_n|^2)^{\frac{p}{2}} \mathbb{E}(\xi_n | F_{t_n}) + \frac{p(p - 2)}{8} \mathbb{E}(\xi_n^2 | F_{t_n}) + \mathbb{E}(\xi_n^3 P_i(\xi_n) | F_{t_n})].
\]

We note the fact that \( Y_n \) is independent of \( \Delta B_{n, t_1, t_2} \) and \( B_{t(n + 1)} - B_{t(n)} \) for \( 1 \leq l, l_1, l_2 \leq m \) and according to the [9], we have

\[
\mathbb{E}[\Delta B_{l}^2 = \mathbb{E}[B_{t(n + 1)} - B_{t(n)}] = \mathbb{E}[I_{t_1, t_2}(n)] = 0, \mathbb{E}[\Delta B_{l}^2 = \mathbb{E}[B_{t(n + 1)} - B_{t(n)}] = 0, \mathbb{E}[\Delta B_{l}^2 = \mathbb{E}[B_{t(n + 1)} - B_{t(n)}]| F_{t_n}]^2 \leq C \Delta, \text{ and } \mathbb{E}[I_{t_1, t_2}(n)]^2 \leq C \Delta^2.
\]

Therefore,

\[
\mathbb{E}(\xi_n | F_{t_n}) \leq \mathbb{E}(2(Y_n, b(Y_n, \alpha_n)) \Delta + |A_n|^2 | F_{t_n}) \Delta
\]

Recalling that (7), we have

\[
\mathbb{E} \left[ \sum_{i=1}^{m} I_{\{N_i=n\}} (\sigma_i(Y_n, \alpha_{n+1}) - \sigma_i(Y_n, \alpha_n)) \right]^2 | F_{t_n}
\]

\[
\leq K \mathbb{E} \left[ \sum_{i=1}^{m} I_{\{N_i=n\}} (\sigma_i(Y_n, \alpha_{n+1}))^2 + |\sigma_i(Y_n, \alpha_n)|^2 \right] | F_{t_n}
\]

\[
\leq K h(\Delta) (1 + |Y_n|)^2 \mathbb{E}(I_{\{N_i=n\}})
\]

\[
\leq K h(\Delta) (1 + |Y_n|)^2 \Delta
\]

by using Lemma 2.1. Thus

\[
\mathbb{E}[A_n^2 | F_{t_n}] \leq 4|b(Y_n, \alpha_n)|^2 \Delta^2 + 4 \sum_{l_1, l_2=1}^{m} L_{l_1}^2 \sigma_{l_1}(Y_n, \alpha_n)^2 \Delta^2
\]

\[
+ 4 \sum_{l=1}^{m} \sigma_l(Y_n, \alpha_n)^2 \Delta
\]

\[
+ 4 \mathbb{E} \left[ \sum_{i=1}^{m} I_{\{N_i=n\}} (\sigma_i(Y_n, \alpha_{n+1}) - \sigma_i(Y_n, \alpha_n)) \right] | F_{t_n} \right] \Delta
\]

\[
\leq K(1 + |Y_n|)^2 \Delta^2 + 4m \sum_{i=1}^{m} |\sigma_i(Y_n, \alpha_n)|^2 \Delta.
\]

Therefore, by the (5) and Assumption 3.1, we can get

\[
\mathbb{E}(\xi_n | F_{t_n}) \leq (1 + |Y_n|^2)^{-1} \mathbb{E}\left\{ 2(Y_n, b(Y_n, \alpha_n)) \Delta + 4m \sum_{i=1}^{m} |\sigma_i(Y_n, \alpha_n)|^2 \Delta \right\}
\]

\[
+ (1 + |Y_n|^2)^{-1} \mathbb{E}\left[ K(1 + |Y_n|)^2 \Delta^2 \right]
\]

\[
\leq (1 + |Y_n|^2)^{-1} \mathbb{E}\left\{ 2(Y_n, b(Y_n, \alpha_n)) \Delta + 4m \sum_{i=1}^{m} |\sigma_i(Y_n, \alpha_n)|^2 \Delta \right\} + K \Delta.
\]
For \( \xi_n^2 \), by the Cauchy inequality, we have that

\[
\mathbb{E}[\xi_n^2] \leq \frac{4|Y_n|\mathbb{E}[|A_n|^2]\mathcal{F}_{t_n}}{(1 + |Y_n|^2)^2} + \frac{4|Y_n|\mathbb{E}[|A_n|^3]\mathcal{F}_{t_n}}{(1 + |Y_n|^2)^2} + \frac{\mathbb{E}[|A_n|^4]\mathcal{F}_{t_n}}{(1 + |Y_n|^2)^2}.
\]  

(11)

For the first term of the (11), according to (9), we know

\[
\mathcal{L} \leq |\lambda_{\max}| \|

\]  

(12)

For the third term of the (11), we also have

\[
\mathbb{E}[|A_n|^4] \leq C \Delta^4, \quad \mathbb{E}[\Delta B_k] \leq C \Delta ^2, \quad \text{and} \quad \mathbb{E}|B_{t_{n+1}}(\mathcal{E}) - B_{t_n}(\mathcal{E})|^4 \leq C \Delta ^2. \quad \text{We obtain}
\]

\[
\mathbb{E}[|A_n|^4] \leq 4^3 \mathbb{E}\left\{|b(Y_n, \alpha_n)\Delta| + |\sum_{l=1}^{m} \sigma_l(Y_n, \alpha_n)\Delta B_k|^4
\right.

+ \left| \sum_{l_1, l_2 = 1}^{m} L^{l_2} \sigma_{l_1}(Y_n, \alpha_n) I_{l_1, l_2}(n) \right|^4

+ \left| \sum_{l = 1}^{m} I_{\{N_n = 1\}}[\sigma_l(Y_n, \alpha_{n+1}) - \sigma_l(Y_n, \alpha_n)](B_{t_{n+1}}(\mathcal{E}) - B_{t_n}(\mathcal{E})) \right|^4 \right\}

\]

(13)

Lastly, for the second term of (11), as the same technique, we can get

\[
\frac{\mathbb{E}[4|Y_n||A_n|^3]\mathcal{F}_{t_n}}{(1 + |Y_n|^2)^2} \leq \frac{4|Y_n|}{(1 + |Y_n|^2)^2} \mathbb{E}\left\{|b(Y_n, \alpha_n)|^3 \Delta^3 + \sum_{l=1}^{m} \sigma_l(Y_n, \alpha_n)|^3 |\Delta B_k|^3
\right.

+ \left| \sum_{l_1, l_2 = 1}^{m} L^{l_2} \sigma_{l_1}(Y_n, \alpha_n)|^3 I_{l_1, l_2}(n)|^3

+ \left| \sum_{l = 1}^{m} \sigma_l(Y_n, \alpha_{n+1})| + \left| \sum_{l = 1}^{m} \sigma_l(Y_n, \alpha_n)|^3 \right| (B_{t_{n+1}}(\mathcal{E}) - B_{t_n}(\mathcal{E})) \right|^3 \right\}.
\]
Thanks to the truncated Milstein scheme (6), for any integer $n$ and $T$, we get

$$E|\Delta B_k|^3 \leq (E|\Delta B_k|)^{3/4} \leq C(\Delta^2)^{3/4} \Delta^{3/2},$$

$$E|I_{1,t}I_{2}(n)|^3 \leq C(E|I_{1,t}I_{2}(n)|)^{3/4} \leq \Delta^3,$$ and $E|B(t_{n+1}) - B(t_n)|^3 \leq C\Delta^{3/2}$. Therefore, we have

$$\frac{E[|Y_n||A_{n}^3|F_{n}]}{(1 + |Y_n|^2)^2} \leq C(h^3(\Delta)\Delta^3 + h^{3/2}(\Delta)\Delta^{3/2} + h^{3}(\Delta)\Delta^3 + h^{3/2}(\Delta)\Delta^{3/2}) \leq K\Delta,$$ \hfill (14)

Inserting (12), (13) and (14) into (11), we have

$$E(\xi_n^2 | F_{n}) \leq \frac{16m \sum_{i=1}^{m} |\sigma_i(Y_n, \alpha_n)|^2 \Delta}{(1 + |Y_n|^2)^2} + K\Delta.$$ \hfill (15)

Also we can prove that for any $i \geq 3$,

$$E(\xi_n^i | F_{n}) = O(\Delta).$$ \hfill (16)

Combining (10), (15) and (16), for any $n \geq 0$, and according to Assumption 3.1, we have following inequality

$$\frac{p}{2} \frac{2(Y_n, b(Y_n, \alpha_n)) + 4m \sum_{i=1}^{m} |\sigma_i(Y_n, \alpha_n)|^2}{1 + |Y_n|^2} + \frac{p(p-2)}{8} \frac{16m|Y_n|^2 \sum_{i=1}^{m} |\sigma_i(Y_n, \alpha_n)|^2}{(1 + |Y_n|^2)^2} \leq \frac{2p\left[\langle Y_n, b(Y_n, \alpha_n) \rangle + m(p-1) \sum_{i=1}^{m} |\sigma_i(Y_n, \alpha_n)|^2 \right]}{1 + |Y_n|^2} \leq 2pK.$$ Therefore,

$$E((1 + |\tilde{Y}_{n+1}|^2)^2 \mid F_{n}) \leq E(1 + |Y_n|^2)^2 (1 + K\Delta).$$

Thanks to the truncated Milstein scheme (6), for any integer $n$ satisfying $0 \leq n\Delta \leq T$, we obtain

$$E[(1 + |Y_{n+1}|^2)^2] \leq E[1 + |\pi_{\Delta} \tilde{Y}_{n+1}|^2] \leq E[(1 + |Y_{n+1}|^2)^2] = E\{E[(1 + |\tilde{Y}_{n+1}|^2)^2] \mid F_{n}\} \leq (1 + K\Delta)E[(1 + |Y_{n+1}|^2)^2].$$

Solving the above linear first order difference inequality, we obtain

$$E[(1 + Y_n^2)^2] \leq (1 + K\Delta)^n E[1 + Y_0^2] \leq e^{K\Delta n} (1 + Y_0^2)^2 \leq e^{KT} (1 + Y_0^2)^2.$$ Therefore we get the desired result that

$$\sup_{0 < \Delta < 1} \sup_{0 \leq n \Delta \leq T} E|Y_n|^p \leq \sup_{0 < \Delta < 1} \sup_{0 \leq n \Delta \leq T} E[(1 + Y_n^2)^2] \leq C.$$ The assertion therefore is complete. \hfill \Box

Next, we will establish our main result about convergence rate theorem. We need some additional conditions.

**Assumption 3.4.** Assume that for any $i = 1, 2, \ldots m$ and $l = 1, 2, \ldots, d$, there exists a positive constant $H_3$ and $p \geq 1$ such that

$$|b(x, i)| \vee |\sigma _1(x, i)| \vee |b'(x, i)| \vee |b''(x, i)| \vee |\sigma _l(x, i)| \vee |\sigma _p'(x, i)| \leq H_3(1 + |x|^{p+1}).$$ \hfill (17)
Assumption 3.5. For every $p > 2$, there exists a positive constant $K$ such that

$$\langle x - y, b(x, i) - b(y, i) \rangle + m(p - 1) \sum_{i=1}^{m} |\sigma_i(x, i) - \sigma_i(y, i)|^2 \leq K|x - y|^2$$

for all $x, y \in \mathbb{R}^d$, $i \in M$.

Consider that (6) is a one-step numerical method, we prove the main theorem based on Theorem 2.1 in [10], which presents a relationship between local and global truncation errors. For the reader’s convenience, we introduce the following lemma without its proof to avoid repetition.

Lemma 3.6. (cf. [10]) Suppose

(i) Assumptions 3.1, 3.4 and 3.5 hold;

(ii) The numerical solution $Y_n$ from a one-step approximation has the following orders of accuracy (local error): for some $p_1 \geq 1$ there are $\eta \geq 1$, $\Delta_0 > 0$, and $K > 0$ such that for arbitrary $t_0 \leq t \leq T - \Delta$, $x \in \mathbb{R}^d$, and all $0 < \Delta \leq \Delta_0$:

$$\mathbb{E} \left| X(t_n + \Delta) - Y_{n+1} \middle| X(t_n) = Y_n \right|^2 \leq K(1 + |x|^{2\eta})^{1/2} \Delta \eta,$$

$$\mathbb{E} \left| X(t_n + \Delta) - Y_{n+1} \middle| X(t_n) = Y_n \right|^{2p_1} \leq K(1 + |x|^{2\eta p_1})^{1/(2p_1)} \Delta q_2$$

with

$$q_2 \geq \frac{1}{2}, \quad q_1 \geq q_2 + \frac{1}{2};$$

(iii) The numerical solution $Y_n$ has finite moments, i.e., for some $p_1 \geq 1$ there are $\beta \geq 1$, $\Delta_0 > 0$, and $K > 0$ such that for all $0 < \Delta \leq \Delta_0$ and all $k = 0, \ldots, N$:

$$\mathbb{E}|Y_n|^{2p_1} < K(1 + \mathbb{E}|X_0|^{2p_1\beta}).$$

Then for any $N$ and $k = 0, 1, \ldots, N$ the following inequality holds (global error):

$$\mathbb{E}|X(t_k) - Y_k|^{2p_1} \leq K(1 + \mathbb{E}|X_0|^{2\gamma p_1})^{1/(2\gamma p_1)} \Delta q_2 - \frac{1}{2},$$

where $K > 0$ and $\gamma \geq 1$ do not depend on $h$ and $k$, i.e., the order of accuracy of the one-step numerical method is $q = q_2 - 1/2$.

Theorem 3.7. Let Assumptions 3.1, 3.4 and 3.5 hold. Recalling (17), we may define

$$\omega(u) = H_3 u^{(1+\rho)} \quad u \geq 1,$$

and let

$$h(\Delta) = H_3 \Delta^{-\varepsilon} \quad \text{for some} \quad \varepsilon \in (0, 1/4],$$

and choosing a sufficiently large $p$ such that

$$\varepsilon(p - 2) \geq 3(1 + \rho).$$

Then we can get

$$(\mathbb{E}|X_n - Y_n|^2)^{1/2} \leq C \Delta,$$

where $C$ stands for a generically positive real constant which is independent of $\Delta$ and its value may change between occurrences.
Proof. By applying recursively (3) and (6) we obtain
\[ X_{n+1} - Y_{n+1} = X_n - Y_n + [b(X_n, \alpha_n) - b(Y_n, \alpha_n)]\Delta \]
\[ + \sum_{l=1}^m [\sigma_l(X_n, \alpha_n) - \sigma_l(Y_n, \alpha_n)]\Delta B_l \]
\[ + \sum_{l_1, l_2 = 1}^m [L^{l_2} \sigma_{l_1}(X_n, \alpha_n) - L^{l_2} \sigma_{l_1}(Y_n, \alpha_n)]I_{l_1, l_2}(n) \]
\[ + \sum_{l=1}^m I_{\{N_n=1\}} \left[ (\sigma_l(X_n, \alpha_{n+1}) - \sigma_l(X_n, \alpha_n)) \right. \]
\[ - \left. (\sigma_l(Y_n, \alpha_{n+1}) - \sigma_l(Y_n, \alpha_n)) \right] \times [B_l(t_{n+1}) - B_l(\pi_1^n)] \]
\[ + \sum_{j=1}^8 R_{n,j}. \]
By the Cauchy-Schwarz inequality,
\[ \mathbb{E}[X_{n+1} - Y_{n+1}]^2 \]
\[ \leq 2\mathbb{E}[X_n - Y_n]^2 + \mathbb{E}\left\{ 2(X_n - Y_n, [b(X_n, \alpha_n) - b(Y_n, \alpha_n)]) \Delta \right. \]
\[ + 2[b(X_n, \alpha_n) - b(Y_n, \alpha_n)]^2 \Delta^2 \]
\[ + 2m \sum_{l=1}^m |\sigma(X_n, \alpha_n) - \sigma(Y_n, \alpha_n)|^2 \Delta \]
\[ + 2 \sum_{l_1, l_2 = 1}^m L^{l_2} \sigma_{l_1}(X_n, \alpha_n) - L^{l_2} \sigma_{l_1}(Y_n, \alpha_n)^2 \Delta^2 \]
\[ + 2m \sum_{l=1}^m I_{\{N_n=1\}} |\sigma_l(X_n, \alpha_{n+1}) - \sigma_l(X_n, \alpha_n) - (\sigma_l(Y_n, \alpha_{n+1}) - \sigma_l(Y_n, \alpha_n))|^2 \Delta \]
\[ + 48 \sum_{j=1}^8 |R_{n,j}|^2 \}. \]
Now we prove the local error bound. Therefore, suppose \( X_n = Y_n \), we can get
\[ \mathbb{E}(|X_{n+1} - Y_{n+1}|^2 | X_n = Y_n) \leq C\mathbb{E}\left( \sum_{j=1}^8 |R_{n,j}|^2 | X_n = Y_n \right) \leq C\Delta^3 \]
with Lemma A.1. Thanks to the truncated Milstein scheme (6), for any integer \( n \) satisfying \( 0 \leq n\Delta \leq T \), we obtain
\[ \mathbb{E}|X_{n+1} - Y_{n+1}|^2 = \mathbb{E}|X_{n+1} - \pi_\Delta(\bar{Y}_{n+1})|^2 \]
\[ \leq 2\mathbb{E}|X_{n+1} - \pi_\Delta(X_{n+1})|^2 + 2\mathbb{E}|\pi_\Delta(X_{n+1}) - \pi_\Delta(\bar{Y}_{n+1})|^2. \]
According to (7.21) in [5], we have
\[ |\pi_\Delta(X_{n+1}) - \pi_\Delta(\bar{Y}_{n+1})|^2 \leq |X_{n+1} - \bar{Y}_{n+1}|^2. \]
Meanwhile
\[ \mathbb{E}|X_{n+1} - \pi_\Delta(X_{n+1})|^2 \]
\[ \leq C\mathbb{E}[I_{\{X_{n+1} \geq \omega^{-1}(h(\Delta))\}} | X_{n+1}|^2] \]
Thus
\[ \leq C \left( \frac{\mathbb{E} |X_{n+1}|^p}{(\omega^{-1}(h(\Delta)))^p} \right)^{\frac{p-2}{2}} \leq C(\omega^{-1}(h(\Delta)))^{-(p-2)}. \]

Finally, from (23) we have \( \omega(27) \), and due to the condition (25) we get
\[ q \]
Therefore the local order
\[ \text{In order to obtain the other local order } q_1 \text{ in (18), we only need to prove} \]
\[ \left| \mathbb{E}(\sum_{j=1}^{8} R_{n,j} | X_n = Y_n) \right| \leq C \Delta^2. \]

According to the definition of the operator \( \mathcal{L} \) and Assumption 3.4, we have that
\[ |\mathcal{L}b(x, i_0)| = |b'(x, i_0)|^2 + \frac{1}{2} |b''(x, i_0)| + \left| \sum_{j_0 \in \mathcal{M}} q_{i_0 j_0} (b(x, j_0) - b(x, i_0)) \right| \]
\[ \leq |b'(x, i_0)|^2 + \frac{1}{2} |b''(x, i_0)| + \sum_{j_0 \in \mathcal{M}} q_{i_0 j_0} (|b(x, j_0)| + |b(x, i_0)|) \]
\[ \leq K(1 + |x|^{p+1}). \]

Moreover, according to the Lemma 3.2, we have that
\[ \left| \mathbb{E}(R_{n,1} | X_n = Y_n) \right| = \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} \mathbb{E}|\mathcal{L}b(X(u), \alpha(u))|duds. \]
\[ \leq C \Delta^2. \]

However, for the \( R_{n,2} \) to \( R_{n,8} \), according to [9], we know that them are square integrable martingale, therefore
\[ \left| \mathbb{E}(\sum_{j=2}^{8} R_{n,j} | X_n = Y_n) \right| = 0. \]

So the assertion (18) is also hold with \( q_1 = 2 \).

By Lemma 3.6 with \( q_1 = 2, q_2 = 3/2, \) we obtain
\[ \mathbb{E}|X_n - Y_n|^2 \leq C \Delta^2. \]

The proof therefore is complete. \( \square \)

4. Numerical results.

Example 4.1. Consider a nonlinear scalar hybrid SDEs
\[ dx(t) = b(x(t), \alpha(t))dt + \sigma(x(t), \alpha(t))dB(t) \]
(29)

Here, \( B(t) \) is a scalar Brownian, and \( \alpha(t) \) is a Markovian chain on the state space \( \mathcal{M} = \{1, 2\} \) and they are independent. Let the generator of the Markovian chain that
\[ \Gamma = \begin{pmatrix} -2 & 2 \\ 1 & -1 \end{pmatrix}. \]
Moreover, for $\forall (x, i) \in \mathbb{R} \times M$,

$$b(x, i) = \begin{cases} x - 2x^5 & \text{if } i = 1 \\ 2x - 3x^5 & \text{if } i = 2 \end{cases} \quad \text{and} \quad \sigma(x, i) = \begin{cases} 2x^2 & \text{if } i = 1 \\ x^2 & \text{if } i = 2 \end{cases}$$

**Step 1.** Check the Assumptions.

Due to the example is 1-dimensional, then $m = 1$. For Assumption 3.5, it is straightforward to see that

\[
(x - y)^T (b(x, 1) - b(y, 1)) + |\sigma(x, 1) - \sigma(y, 1)|^2 \\
= (x - y)(x - 2x^5 - (y - 2y^5)) + |2x^2 - 2y^2|^2 \\
= (x - y)^2[1 - 2(y^4 + y^3x + y^2x^2 + yx^3 + x^4) + 4(x + y)^2].
\]

However

\[-2(x^3y + xy^3) = -2xy(x^2 + y^2) \leq (x^2 + y^2)^2 = (x^4 + y^4) + 2x^2y^2.
\]

Hence

\[
(x - y)^T (b(x, 1) - b(y, 1)) + |\sigma(x, 1) - \sigma(y, 1)|^2 \\
\leq (x - y)^2[1 - (y^4 + x^4) + 8(x^2 + y^2)] \\
\leq 31(x - y)^2.
\]

Similarly,

\[
(x - y)^T (b(x, 2) - b(y, 2)) + |\sigma(x, 2) - \sigma(y, 2)|^2 \leq 16(x - y)^2.
\]

So, for any $i \in M$, we have

\[
(x - y)^T (b(x, i) - b(y, i)) + |\sigma(x, i) - \sigma(y, i)|^2 \leq 31|x - y|^2.
\]

In other words, Assumption 3.5 is also fulfilled for any $p$. Moreover,

\[
x^T b(x, 1) + 2(p - 1)|\sigma(x, 1)|^2 \\
= x(x - 2x^5) + 2(p - 1)|2x^2|^2 \\
= x^2 - 2x^6 + 8(p - 1)x^4 \\
= x^2 - 2x^2(x^2 - 2(p - 1))^2 + 2(2(p - 1))^2x^2 \\
\leq 9(p - 1)^2(1 + x^2).
\]

Similarly,

\[
x^T b(x, 2) + 2(p - 1)|\sigma(x, 2)|^2 \leq K(1 + |x|^2),
\]

where $K = \max\{2, 7(p - 1)^2\}$.

Therefore for any $i \in M$, we have

\[
x^T b(x, i) + 2m(p - 1)|\sigma(x, i)|^2 \leq K(1 + |x|^2), \tag{30}
\]

that is, Assumption 3.1 is satisfied for any $p$.

**Step 2.** We need choose $\omega(\cdot)$ and $h(\cdot)$.

According to (4), we can set $\omega(u) = 3u^5$ ($H_3 = 3$) such that

\[
\sup_{|x| \leq u} (|b(x, i)| \vee |\sigma(x, i)|) \leq \sup_{|x| \leq u} 3|x|^5 < 3u^5, \quad u > 3.
\]
and we set $h(\Delta) = \Delta^{-1/10}$, then all the conditions in (5) hold for all $\Delta \in (0,1)$, and obviously we have $\omega^{-1}(h(\Delta)) = (\Delta/3)^{-1/50}$.

**Step 3.** Applying (6), we can obtain the numerical solution. We use the numerical solutions produced by truncated Milstein method with step size $2^{-6}, 2^{-7}, 2^{-8},$ and $2^{-9}$. Meanwhile, it is hardly to find a true solution of (29), therefore the simulations using the step size $2^{-14}$ are regarded as the ‘true solution’. The square roots of the mean square errors at the terminal time $T = 1$ are plotted in Fig. 1 by simulating 1000 paths. We can see that the convergence order is approximately 1.

**Remark 4.2.** Since there are not many multi-dimensional SDEs from the literature that satisfy Assumption 3.6, although this assumption is common in the literature, it is worthy of exploring new criteria for the existence and uniqueness of the global solution of SDEs in the future.

![Figure 1](image_url)  
*Figure 1.* The strong convergence order at the terminal time $T = 1$. The red dashed line is the reference line with the slope of 1.

**Appendix A. The proofs of some error bounds.**

**Lemma A.1.** Suppose Assumptions 3.1, 3.4 and 3.5 hold and for any $\Delta < 1/(2q)$, then we have

$$\sum_{j=1}^{8} \mathbb{E}|R_{n,j}|^2 \leq C \Delta^3.$$  \hfill (31)

**Proof.** **Claim 1.** According to (28), Holder inequality and Theorem 3.3, Therefore,

$$\mathbb{E}|R_{n,1}|^2 = \mathbb{E} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} Lb(X(u), \alpha(u))duds|^2 \leq C \Delta^2 \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} \mathbb{E}|Lb(X(u), \alpha(u))|^2duds \leq C \Delta^2 \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} (1 + \mathbb{E}|X(u)|^{2(\rho+1)})duds \leq C \Delta^4.$$

**Claim 2.** By Cauchy inequality we have

$$L^1b(x,i) = \langle b'(x,i), \sigma_i(x,i) \rangle \leq \frac{1}{2} |b'(x,i)|^2 + \frac{1}{2} |\sigma_i(x,i)|^2 \leq K(1 + |x|^{2(\rho+1)})$$.
Therefore, by the Burkholder-Davis-Gundy (BDG) inequality and Holder inequality we have

\[ \mathbb{E}|R_{n,2}|^2 = \mathbb{E}\left| \sum_{l=1}^{m} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} L^1 b(X(u), \alpha(u)) dB_l(u) ds \right|^2 \]

\[ \leq C \sum_{l=1}^{m} \mathbb{E}\left| \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} L^1 b(X(u), \alpha(u)) dB_l(u) ds \right|^2 \]

\[ \leq C \Delta \sum_{l=1}^{m} \int_{t_n}^{t_{n+1}} \mathbb{E}\left| \int_{t_n}^{s} L^1 b(X(u), \alpha(u)) dB_l(u) ds \right|^2 \]

\[ \leq C \Delta \sum_{l=1}^{m} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} \mathbb{E}|L^1 b(X(u), \alpha(u)) du|^2 ds \]

\[ \leq C \Delta \sum_{l=1}^{m} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} (1 + \mathbb{E}|X(u)|^{4(p+1)}) du ds \]

\[ \leq C \Delta^3. \]

Claim 3. Since \( M_{i_0 j_0}(t) \) is martingale, similarly to \( R_{k,p,3} \) in [9] and Theorem 3.3, we have

\[ \mathbb{E}|R_{n,3}|^2 = \mathbb{E}\left| \sum_{i_0 \neq j_0} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} (b(X(u), j_0) - b(X(u), i_0)) dM_{i_0 j_0}(u) ds \right|^2 \]

\[ \leq C \sum_{i_0 \neq j_0} \mathbb{E}\left| \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} (b(X(u), j_0) - b(X(u), i_0)) dM_{i_0 j_0}(u) ds \right|^2 \]

\[ \leq C \Delta \sum_{i_0 \neq j_0} \int_{t_n}^{t_{n+1}} \mathbb{E}\left| \int_{t_n}^{s} (b(X(u), j_0) - b(X(u), i_0)) dM_{i_0 j_0}(u) \right|^2 ds \]

\[ \leq C \Delta \sum_{i_0 \neq j_0} \int_{t_n}^{t_{n+1}} \mathbb{E}\left| \int_{t_n}^{s} |(b(X(u), j_0) - b(X(u), i_0))| dM_{i_0 j_0}(u) \right|^2 ds \]

\[ \leq C \Delta \sum_{i_0 \neq j_0} \int_{t_n}^{t_{n+1}} \mathbb{E}\left| \int_{t_n}^{s} 2(|b(X(u), j_0)|^2 + |b(X(u), i_0)|^2) dM_{i_0 j_0}(u) \right| ds \]

\[ \leq C \Delta \sum_{i_0 \neq j_0} \int_{t_n}^{t_{n+1}} \mathbb{E}\left| \int_{t_n}^{s} K(1 + |X(u)|^{2(p+1)}) dM_{i_0 j_0}(u) \right| ds \]

\[ \leq C \Delta \sum_{i_0 \neq j_0} \int_{t_n}^{t_{n+1}} K \mathbb{E}(1 + \mathbb{E}(|X(u)|^{2(p+1)}) dM_{i_0 j_0}(u) \right| ds \]

\[ \leq C \Delta \sum_{i_0 \neq j_0} \int_{t_n}^{t_{n+1}} \mathbb{E}([M_{i_0 j_0}](s) - [M_{i_0 j_0}](t_n)) ds \]

\[ \leq C \Delta \sum_{i_0 \neq j_0} \int_{t_n}^{t_{n+1}} \mathbb{E}|N_p| ds \leq C \Delta^3. \]

Claim 4. Similarly to \( R_{n,1} \), by BDG and Holder inequality, we get

\[ \mathbb{E}|R_{n,4}|^2 = \mathbb{E}\left| \sum_{l=1}^{m} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} \mathcal{L}\sigma(X(u), \alpha(u)) dudB_l(s) \right|^2 \leq C \Delta^3. \]
Claim 5. By the definition of \(<M_{i_0,j_0}>\), and Assumption 3.4, we have

\[
\mathbb{E}|R_{n,5}|^2 \\
= \mathbb{E} \left| \sum_{i=1}^{m} \sum_{j \neq j_0} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} (\sigma_1(X(u), i_0) - \sigma_1(X(u), j_0)) d(M_{i_0,j_0})(u) dB_t(s) \right|^2 \\
\leq C \sum_{i=1}^{m} \sum_{j \neq j_0} \int_{t_n}^{t_{n+1}} \mathbb{E} \int_{t_n}^{s} (\sigma_1(X(u), i_0) - \sigma_1(X(u), j_0)) d(M_{i_0,j_0})(u) dB_t(s) \right|^2 ds \\
\leq C \sum_{i=1}^{m} \sum_{j \neq j_0} \int_{t_n}^{t_{n+1}} \mathbb{E} \int_{t_n}^{s} |(\sigma_1(X(u), i_0) - \sigma_1(X(u), j_0)) q_{i_0,j_0} I(\alpha(u) = i_0) du |^2 ds \\
\leq C \Delta \sum_{i=1}^{m} \sum_{j \neq j_0} \int_{t_n}^{t_{n+1}} \mathbb{E} \int_{t_n}^{s} (1 + |X(u)|^{2(\rho+1)}) du ds \\
\leq C \Delta \sum_{i=1}^{m} \sum_{j \neq j_0} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} (1 + \mathbb{E}|X(u)|^{2(\rho+1)}) du ds \\
\leq C \Delta \sum_{i=1}^{m} \sum_{j \neq j_0} \Delta^3 \leq C \Delta^3.
\]

Claim 6. According to (4.30) in [9], we obtain

\[
\mathbb{E}|R_{n,6}|^2 \\
= \mathbb{E} \left| \sum_{i_1,i_2=1}^{m} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} [L^1 \sigma_{i_1}(X(u), \alpha(u)) - L^1 \sigma_{i_1}(X_n, \alpha_n)] dB_{i_2}(u) dB_{i_1}(s) \right|^2 \\
\leq C \sum_{i_1,i_2=1}^{m} \int_{t_n}^{t_{n+1}} \mathbb{E} [L^1 \sigma_{i_1}(X(u), \alpha(u)) - L^1 \sigma_{i_1}(X_n, \alpha_n)]^2 du ds \\
= C \sum_{i_1,i_2=1}^{m} \int_{t_n}^{t_{n+1}} \mathbb{E} [I_{\{N_n=0\}} |L^1 \sigma_{i_1}(X(u), \alpha(u)) - L^1 \sigma_{i_1}(X_n, \alpha_n)|^2] du ds \\
+ C \sum_{i_1,i_2=1}^{m} \int_{t_n}^{t_{n+1}} \mathbb{E} [I_{\{N_n \geq 1\}} |L^1 \sigma_{i_1}(X(u), \alpha(u)) - L^1 \sigma_{i_1}(X_n, \alpha_n)|^2] du ds.
\]

To proceed, note that on \(\{N_n = 0\}\), \(\alpha(u) = \alpha_n\) for \(t_n < u < t_{n+1}\). Thus

\[
\sum_{i_1,i_2=1}^{m} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} \mathbb{E} [I_{\{N_n=0\}} |L^1 \sigma_{i_1}(X(u), \alpha(u)) - L^1 \sigma_{i_1}(X_n, \alpha_n)|^2] du ds \\
= \sum_{i_1,i_2=1}^{m} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} \mathbb{E} [I_{\{N_n=0\}} |L^1 \sigma_{i_1}(X(u), \alpha_n) - L^1 \sigma_{i_1}(X_n, \alpha_n)|^2] du ds
\]
Similarly to Claim 7.

Thus Claim 8. Recalling that we have

\[
\mathbb{E}[\sum_{l_1, l_2=1}^{m} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} \mathbb{E}[I_{\{N_n = 0\}}]\{(1 + |X(u)|^p + |X_n|^p)(X(u) - X_n)\}^2]duds
\]

\[
\leq C \sum_{l_1, l_2=1}^{m} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} \mathbb{E}[X(u) - X_n]duds
\]

\[
\leq C \Delta^3,
\]

where we used the Lemma 3.2 in the first inequality and the last inequality.

For the second of the \( R_{n,6} \), by the Assumption 3.4, Lemma 2.1 and Theorem 3.3, we have

\[
\sum_{l_1, l_2=1}^{m} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} \mathbb{E}[I_{\{N_n \geq 1\}}|L_{l_1}^T \sigma_l(X(u), \alpha(u)) - L_{l_1}^T \sigma_l(X_n, \alpha_n)|^2]duds
\]

\[
\sum_{l_1, l_2=1}^{m} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} \mathbb{E}[I_{\{N_n \geq 1\}}(|L_{l_1}^T \sigma_l(X(u), \alpha(u))|^2 + |L_{l_1}^T \sigma_l(X_n, \alpha_n)|^2)]duds
\]

\[
\leq C \sum_{l_1, l_2=1}^{m} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} \mathbb{E}[I_{\{N_n \geq 1\}}(1 + |X(u)|^{2p} + |X_n|^{2p})]duds
\]

\[
\leq C \sum_{l_1, l_2=1}^{m} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} \mathbb{E}[I_{\{N_n \geq 1\}}]duds
\]

\[
\leq C \sum_{l_1, l_2=1}^{m} \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} \mathbb{E}[I_{\{N_n \geq 1\}}]duds
\]

\[
\leq C \Delta^3.
\]

Thus

\[
\mathbb{E}[R_{n,6}]^2 \leq C \Delta^3.
\]

Claim 7. Similarly to \( R_{k,p,7} \) and \( R_{k,p,8} \) in [9] and Theorem 3.3, we have

\[
\mathbb{E}[R_{n,7}]^2 = \mathbb{E}\left[ \sum_{l=1}^{m} I_{\{N_n = 1\}}\{[\sigma_l(X(\tau_n^l), \alpha_n) - \sigma_l(X_n, \alpha_n) - \sigma_l(X(\tau_n^l), \alpha_n) - \sigma_l(X_n, \alpha_n)]B_l(t_{n+1}) - B_l(\tau_n^l)\}^2 \right]
\]

\[
\leq C \Delta \sum_{l=1}^{m} \mathbb{E}[I_{\{N_n = 1\}}\{[\sigma_l(X(\tau_n^l), \alpha_n) - \sigma_l(X_n, \alpha_n)]B_l(t_{n+1}) - B_l(\tau_n^l)\}^2]
\]

\[
\leq C \Delta \sum_{l=1}^{m} \mathbb{E}[I_{\{N_n = 1\}}\{[1 + |X(\tau_n^l)|^{2p} + |X_n|^{2p}](X(\tau_n^l) - X_n)\}^2]
\]

\[
\leq C \Delta^2 \sum_{l=1}^{m} \mathbb{E}[I_{\{N_n = 1\}}] \leq C \Delta^3.
\]

Claim 8. Recalling that \( \Delta < 1/(2q) \)

\[
\mathbb{E}[R_{n,8}]^2 = \mathbb{E}\left[ \sum_{l=1}^{m} I_{\{N_n \geq 2\}} \right]
\]
Therefore, we obtain
\[
\sum_{i_0 \neq j_0} \int_{t_n}^{t_{n+1}} \int_t^s \left| (\sigma(X(u), j_0) - \sigma(X(u), i_0)) d[M_{t_0j_0}](u) dB_t(s) \right|^2
\]
\[
\leq C \sum_{l=1}^{m} \sum_{N=2}^{\infty} \mathbb{E} \left( |I_{N=1}| \right)
\times \sum_{i_0 \neq j_0} \int_{t_n}^{t_{n+1}} \int_t^s \left| (\sigma(X(u), j_0) - \sigma(X(u), i_0)) d[M_{t_0j_0}](u) dB_t(s) \right|^2 \bigg| I_{N=1} \bigg\}
\leq C \sum_{l=1}^{m} \sum_{N=2}^{\infty} \mathbb{E} \left( |I_{N=1}| \right)
\times \sum_{j=1}^{N} \left| \sigma_l(X(\tau_j^n), \alpha(\tau_j^n)) - \sigma_l(X(\tau_j^n), \alpha(\tau_{j-1}^n)) \right| (B_{t_{n+1}} - B_{t_j^n})^2 \bigg| I_{N=1} \bigg\}
\leq C \Delta \sum_{l=1}^{m} \sum_{N=2}^{\infty} \sum_{j=1}^{N} \mathbb{E} \left( |I_{N=1}| \right) \left| \sigma_l(X(\tau_j^n), \alpha(\tau_j^n)) - \sigma_l(X(\tau_j^n), \alpha(\tau_{j-1}^n)) \right| \bigg| I_{N=1} \bigg\}
\leq C \Delta \sum_{N=2}^{\infty} \mathbb{E}(I_{N=1}) \leq C \Delta \sum_{N=2}^{\infty} (q\Delta)^N \leq C \sum_{N=2}^{\infty} (q\Delta)^{N+1}
\leq C \Delta^3.
\]
Therefore, we obtain
\[
\sum_{j=1}^{8} \mathbb{E} |R_{n,j}|^2 \leq C \Delta^3.
\]

The proof is complete. \qed

Acknowledgments. The authors would like to thank all the referees and the editor for the very useful comments and suggestions, which have helped to improve the paper a lot.

REFERENCES

[1] Q. Guo, W. Liu, X. Mao and R. Yue, The truncated Milstein method for stochastic differential equations with commutative noise, J. Comput. Appl. Math., 338 (2018), 298–310.
[2] M. Hutzenthaler and A. Jentzen, Numerical approximations of stochastic differential equations with non-globally Lipschitz continuous coefficients, Mem. Amer. Math. Soc., 236 (2015).
[3] M. Hutzenthaler, A. Jentzen and P. E. Kloeden, Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients, Ann. Appl. Probab., 22 (2012), 1611–1641.
[4] M. Hutzenthaler, A. Jentzen and P. E. Kloeden, Strong and weak divergence in finite time of Euler’s method for stochastic differential equations with non-globally Lipschitz continuous coefficients, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 467 (2011), 1563–1576.
[5] X. Li, X. Mao and G. Yin, Explicit numerical approximations for stochastic differential equations in finite and infinite horizons: Truncation methods, convergence in p-th moment and stability, IMA J. Numer. Anal., 39 (2019), 847–892.
[6] X. Mao, The truncated Euler–Maruyama method for stochastic differential equations, J. Comput. Appl. Math., 290 (2015), 370–384.
[7] X. Mao, Convergence rates of the truncated Euler-Maruyama method for stochastic differential equations, *J. Comput. Appl. Math.*, 296 (2016), 362–375.

[8] X. Mao and C. Yuan, *Stochastic Differential Equations with Markovian Switching*, Imperial College Press, 2006.

[9] S. L. Nguyen, T. A. Hoang, D. T. Nguyen and G. Yin, Milstein-type procedures for numerical solutions of stochastic differential equations with Markovian switching, *SIAM J. Numer. Anal.*, 55 (2017), 953–979.

[10] M. V. Tretyakov and Z. Zhang, A fundamental mean-square convergence theorem for SDEs with locally Lipschitz coefficients and its applications, *SIAM J. Numer. Anal.*, 51 (2013), 3135–3162.

[11] X. Wang and S. Gan, The tamed Milstein method for commutative stochastic differential equations with non-globally Lipschitz continuous coefficients, *J. Differ. Equations Appl.*, 19 (2013), 466–490.

[12] C. Yuan and X. Mao, Convergence of the Euler–Maruyama method for stochastic differential equations with Markovian switching, *Math. Comput. Simulation*, 64 (2004), 223–235.

[13] S. Zhou, Strong convergence and stability of backward Euler–Maruyama scheme for highly nonlinear hybrid stochastic differential delay equation, *Calcolo.*, 52 (2015), 445–473.

Received September 2019; 1st revision October 2019; final revision October 2020; early access August 2021.

E-mail address: weijunzhan@hotmail.com
E-mail address: qguo@shnu.edu.cn
E-mail address: yhcong@shu.edu.cn