

## Accepted Manuscript

Convergence and stability of two classes of theta-Milstein schemes for stochastic differential equations

Xiaofeng Zong, Fuke Wu, Guiping Xu

PII: S0377-0427(17)30653-2  
DOI: <https://doi.org/10.1016/j.cam.2017.12.025>  
Reference: CAM 11444

To appear in: *Journal of Computational and Applied Mathematics*

Received date : 23 March 2017  
Revised date : 8 November 2017

Please cite this article as: X. Zong, F. Wu, G. Xu, Convergence and stability of two classes of theta-Milstein schemes for stochastic differential equations, *Journal of Computational and Applied Mathematics* (2018), <https://doi.org/10.1016/j.cam.2017.12.025>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



# Convergence and stability of two classes of theta-Milstein schemes for stochastic differential equations

Xiaofeng Zong,<sup>\*</sup>      Fuke Wu,<sup>†</sup>      Guiping Xu,<sup>‡</sup>

November 8, 2017

## Abstract

This paper examines convergence and stability of the two classes of theta-Milstein schemes for stochastic differential equations (SDEs) with non-global Lipschitz continuous coefficients: the split-step theta-Milstein (SSTM) scheme and the stochastic theta-Milstein (STM) scheme. For  $\theta \in [1/2, 1]$ , this paper concludes that the two classes of theta-Milstein schemes converge strongly to the exact solution with the order 1. For  $\theta \in [0, 1/2]$ , under the additional linear growth condition for the drift coefficient, these two classes of the theta-Milstein schemes are also strongly convergent with the standard order. This paper also investigates exponential mean-square stability of these two classes of the theta-Milstein schemes. For  $\theta \in (1/2, 1]$ , these two theta-Milstein schemes can share the exponential mean-square stability of the exact solution. For  $\theta \in [0, 1/2]$ , similar to the convergence, under the additional linear growth condition, these two theta-Milstein schemes can also reproduce the exponential mean-square stability of the exact solution.

**Keywords:** SDEs; Strong convergence rate; Exponential mean-square stability; Stochastic theta-Milstein scheme; Split-step theta-Milstein scheme.

---

<sup>\*</sup>School of Automation, China University of Geosciences, Wuhan 430074, P.R. China, and Hubei Key Laboratory of Advanced Control and Intelligent Automation for Complex Systems, Wuhan 430074, Peoples Republic of China. xfzong87816@gmail.com. The research of this author was supported in part by the Fundamental Research Funds for the Central Universities, China University of Geosciences(Wuhan) under Grant No. CUG170610, and the National Natural Science Foundations of China (No. 61703378 and No. 61473125).

<sup>†</sup>School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan, Hubei 430074, P.R. China, wufuke@mail.hust.edu.cn. The research of this author was supported in part by the National Natural Science Foundations of China (No. 61703378 and No. 61473125).

<sup>‡</sup>School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan, Hubei 430074, P.R. China, gpXu@mail.hust.edu.cn.

## 1 Introduction

During the last decades, stochastic differential equations (SDEs) have become increasingly important tools to describe the real world since stochastic models have wide applications in biological systems, neural network, financial engineering and wireless communications. Most SDEs cannot be solved explicitly, so numerical approximations become important tools to study stochastic models. When a numerical scheme is put forward, it is crucial that this numerical scheme can converge to the exact solution. Moreover, it also need to describe the asymptotic properties of the numerical solution such as boundedness and stability.

Most of the existing convergence theory for numerical methods of SDEs requires the global Lipschitz condition (see [1–4]). However, many well-known stochastic systems do not satisfy the global Lipschitz condition, for example, the stochastic Duffing-van der Pol oscillator [5,6], stochastic Lorenz equation [6, 7], experimental psychology model [6], stochastic Ginzburg-Landau equation [1, 6], stochastic Lotka-Volterra equations [1, 6, 8] and volatility processes [1, 6] and so on. For some SDEs without the global Lipschitz condition, the classical explicit numerical schemes may not converge to the exact solution in the strong mean-square sense (for example Euler-Maruyama (EM) scheme and Milstein scheme, please see [9, 10]). Hence, numerical approximations for SDEs without the global Lipschitz condition have received more and more attention in recent years.

For the SDEs with the one-sided Lipschitz condition (which is weaker than the global Lipschitz condition) on the drift term and the global Lipschitz condition on the diffusion term, Hu [11] examined convergence of the backward Euler-Maruyama (BEM) scheme and obtained that the optimal rate of convergence is 0.5. Under an additional polynomial condition, Higham et al. [12, 13] and Bastani et al. [14] proved that BEM and split-step BEM schemes converge strongly to the exact solution with the optimal rate 0.5. Mao and Szpruch [15] showed that under a dissipative condition on the drift coefficient and the super linear growth condition on the diffusion coefficient, the BEM scheme is convergent with strong order of a half. Under a monotone condition, strong convergence of BEM scheme and theta-EM scheme were investigated by [16, 17]. Recently, there are various explicit Euler schemes with one half order were proposed to approximate the SDEs with non-global Lipschitz coefficients, see [18–21] and the references therein. However, it is still an important and difficult work to look for more appropriate conditions to obtain the strong convergence rate of the numerical schemes for SDEs without the global Lipschitz condition.

The explicit Milstein scheme for SDEs developed by Milstein [2] can obtain a strong order of convergence higher than Euler-type schemes. Hu et al. [22] extended this scheme to solve SDEs with time delay. Implicit Milstein schemes under the global Lipschitz condition were studied by Tian and Burrage [23], Wang et al. [24], Alcock & Burrage [25]. The tamed Milstein scheme was investigated in [26] for SDEs with non-global Lipschitz coefficients. Recently, Higham et al. [27] proposed a new Milstein scheme, named as  $(\theta, \sigma)$ -Milstein scheme, and investigated its strong convergence when SDEs hold polynomial growth for the diffusion term. However, the strong convergence rates of stochastic theta Milstein (STM) schemes ( $\theta \in (0, 1)$ ) for SDEs with non-global Lipschitz coefficients has not yet been established.

Stability of numerical solutions is another central problem for numerical analysis. The mean-square stability of numerical methods for linear stochastic differential equations have been studied by [28–30]. For nonlinear SDEs, Higham et al. [31] showed that the BEM and the split-step BEM can reproduce the exponential mean-square stability of the exact solution. Recently, without the global Lipschitz, Huang [32] and authors [33–35] presented conditions under which the stochastic theta method and split-step theta method can not only reproduce the exponential mean-square stability of the exact solution, but also preserve the bound of the Lyapunov exponent for sufficient small stepsize, which measures the decay rate of the numerical solutions. However, there is little work on the mean-square stability of the theta-Milstein schemes for SDEs with non-Lipschitz continuous coefficients.

The main aim is to examine the boundedness and the convergence rate as well as the exponential mean-square stability of theta Milstein schemes for SDEs with non-Lipschitz continuous coefficients. Besides the classical STM method, we propose a split-step theta Milstein (SSTM) method to approximate SDEs. Under the one-sided Lipschitz condition, we show that for  $\theta \in [1/2, 1]$ , the two classes of theta-Milstein schemes are bounded, strongly convergent, and exponentially mean-square stable, but for  $\theta \in [0, 1/2]$ , the linear growth condition on drift is added to obtain the moment boundedness, strong convergence, and stability. The simulations show that SSTM method may work better than STM method since SSTM method has smaller least squares residual and faster decay.

This paper is organized as follows. The next section presents some necessary notations and preliminaries, and then introduces STM scheme and SSTM scheme. Section 3 establishes the uniform boundedness of the  $p$ th moments of the theta-Milstein schemes and shows that for  $\theta \in [1/2, 1]$ , the two classes of theta-Milstein schemes are bounded in the sense of moment, but for  $\theta \in [0, 1/2]$ , the linear growth condition on drift is added to obtain the moment boundedness of the theta-Milstein schemes. Section 4 proves that the theta-Milstein schemes strongly converge to the exact solution with the order 1. Section 5 investigates the exponential mean-square stability of these two classes of theta-Milstein schemes. Section 6 introduces the numerical experiments to confirm the theoretical results. Section 7 concludes the paper.

## 2 Notations and preliminaries

Throughout this paper, unless otherwise specified, we use the following notations. Let  $|\cdot|$  denote both the Euclidean norm in  $\mathbb{R}^n$ . If  $x$  is a vector, its transpose is denoted by  $x^T$  and the inner product is denoted by  $\langle x, y \rangle = x^T y$  for  $x, y \in \mathbb{R}^n$ . If  $A$  is a matrix, its trace norm is denoted by  $|A| = \sqrt{\text{trace}(A^T A)}$ .  $a \vee b$  represents  $\max\{a, b\}$  and  $a \wedge b$  denotes  $\min\{a, b\}$ .  $\mathbb{N}_+$  represents the positive integer set, namely,  $\mathbb{N}_+ = \{1, 2, 3, \dots\}$ . Let  $(\Omega, \mathfrak{F}, \mathbb{P})$  be a complete probability space with a filtration  $\{\mathfrak{F}_t\}_{t \geq 0}$  satisfying the usual conditions, that is, it is right continuous and increasing while  $\mathfrak{F}_0$  contains all  $\mathbb{P}$ -null sets. Let  $(w(t) = (w^1(t), \dots, w^d(t))^T)_{t \geq 0}$  be a  $d$ -dimensional Brownian motion defined on this probability space.

Let  $f : \mathbb{R}^n \mapsto \mathbb{R}^n$  and  $g : \mathbb{R}^n \mapsto \mathbb{R}^{n \times d}$  be Borel measurable functions. Let us consider the

$n$ -dimensional SDE of the form

$$dx(t) = f(x(t))dt + g(x(t))dw(t), \quad t > 0 \quad (2.1)$$

with initial data  $x(0) = x_0 \in \mathbb{R}^n$ . Let  $g(x) = [g_1(x), \dots, g_d(x)]$  with  $g_i(x) = [g_{1i}(x), \dots, g_{ni}(x)]^T \in \mathbb{R}^n$ . Then the last term in (2.1) has the form  $g(x(t))dw(t) = \sum_{j=1}^d g_j(x(t))dw^j(t)$ .

Assume that  $f$  and  $\{g_i\}_{i=1}^d$  satisfy the following assumption:

**Assumption 2.1.** Assume that the functions  $f, \{g_j\}_{j=1}^d \in C^1$  and there exist constants  $\mu \in \mathbb{R}$  and  $c > 0$  such that for any  $x, y \in \mathbb{R}^n$

$$\langle x - y, f(x) - f(y) \rangle \leq \mu |x - y|^2, \quad (2.2)$$

$$\sum_{j=1}^d |g_j(x) - g_j(y)|^2 \leq c |x - y|^2. \quad (2.3)$$

**Remark 2.1.** Condition (2.2) on the drift  $f$  is known as the onesided Lipschitz condition and (2.3) on the diffusions  $\{g_i\}_{i=1}^d$  is the global Lipschitz condition. Note that  $f \in C^1$  implies that the drift  $f$  is local Lipschitz continuous. Hence, SDE (2.1) admits a unique local solution (see [36, Theorem 3.15]). Moreover, from conditions (2.2) and (2.3), we have the following one-sided linear growth condition,

$$\langle x, f(x) \rangle \vee \sum_{j=1}^d |g_j(x)|^2 \leq \alpha + \beta |x|^2, \quad x \in \mathbb{R}^n, \quad (2.4)$$

where

$$\alpha := \frac{1}{2} |f(0)|^2 \vee 2 \sum_{j=1}^d |g_j(0)|^2 \text{ and } \beta := \left(\mu + \frac{1}{2}\right) \vee 2c.$$

Therefore, from [37, Theorem 3.6], we know that the unique local solution to SDE (2.1) is actually global. Moreover, we have the following  $p$ th moment boundedness.

Under Assumption 2.1, we can present the following lemma (see Lemma 3.2 in [12]).

**Lemma 2.2.** Under Assumption 2.1, for each  $p \geq 2$  there is  $C = C(p, T) > 0$  such that

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |x(t)|^p \right] \leq C(1 + \mathbb{E}|x_0|^p).$$

In what follows, for the purpose of simplicity, let  $C$  represent a generic positive constant independent of the stepsize  $\Delta$ , whose value may change with each appearance.

Fixed any time  $T > 0$  and given a stepsize  $\Delta = T/N$  for certain integer  $N$ , we introduce the split-step theta-Milstein (SSTM) approximation  $\{z_k\}_{k \geq 0}$

$$\begin{cases} y_k = z_k + \theta f(y_k)\Delta, \\ z_{k+1} = z_k + f(y_k)\Delta + g(y_k)\Delta w_k + \sum_{i,j=1}^d L^i g_j(y_k) I_{i,j}^{t_k, t_{k+1}}, \quad k = 0, 1, 2, \dots, N, \end{cases} \quad (2.5)$$

where  $y_0 = x(0)$ ,  $z_0 = y_0 - \theta f(y_0)\Delta$ ,  $\theta \in [0, 1]$ ,  $\Delta w_k = w((k+1)\Delta) - w(k\Delta)$  is the Brownian increment,

$$L^i = \sum_{j=1}^n g_{j,i}(x) \frac{\partial}{\partial x_j} \text{ and } I_{i,j}^{t_k, t_{k+1}} = \int_{t_k}^{t_{k+1}} \int_{t_k}^s dw^i(u) dw^j(s).$$

We follow the idea in [26] and assume the noise term to be commutative, which will avoid additional computational effort when approximating the iterated Itô integrals  $I_{i,j}^{t_k, t_{k+1}}$ . Here, we call the noise term be commutative if

$$L^{i_1} g_{j, i_2} = L^{i_2} g_{j, i_1}, \quad i_1, i_2 = 1, \dots, d, j = 1, \dots, n.$$

The commutative noise has been discussed in [1, 26]. Note that  $I_{i,j}^{t_k, t_{k+1}} + I_{j,i}^{t_k, t_{k+1}} = \Delta w_k^i \Delta w_k^j$  for  $i \neq j$ , where  $\Delta w_k^i = w^i((k+1)\Delta) - w^i(k\Delta)$ . Hence, (2.5) can be rewritten as follows,

$$\begin{cases} y_k = z_k + \theta f(y_k) \Delta, \\ z_{k+1} = z_k + f(y_k) \Delta + g(y_k) \Delta w_k + \frac{1}{2} \sum_{i,j=1}^d L^i g_j(y_k) (\Delta w_k^i \Delta w_k^j - \delta_{ji} \Delta), \end{cases} \quad (2.6)$$

where  $\delta_{ji} = 1$  for  $i = j$  and  $\delta_{ji} = 0$  for  $i \neq j$ .

This scheme  $\{z_k\}_{k \geq 0}$  can be considered as an extension from the split-step theta method developed by the previous works [32, 34, 35]. It is interesting that the approximation  $\{y_k\}_{k \geq 0}$  in (2.6) is in fact the stochastic theta-Milstein (STM) approximation

$$y_{k+1} = y_k + \theta f(y_{k+1}) \Delta + (1 - \theta) f(y_k) \Delta + g(y_k) \Delta w_k + \frac{1}{2} \sum_{i,j=1}^d L^i g_j(y_k) (\Delta w_k^i \Delta w_k^j - \delta_{ji} \Delta), \quad (2.7)$$

which can be proved by substituting  $z_k = y_k - \theta f(y_k) \Delta$  into the second equation in (2.6). The STM (2.7), investigated in [38] for linear scalar SDEs, can be consider as the special case  $\sigma = 0$  of the  $(\theta, \sigma)$ -Milstein scheme developed by Higham et al. [27]. When  $\theta = 0$ ,  $\{z_k\}_{k \geq 0}$  and  $\{y_k\}_{k \geq 0}$  are the classical Milstein approximation proposed in Milstein [2]. When  $\theta = 1$ ,  $\{y_k\}_{k \geq 0}$  becomes the drift-implicit Milstein scheme, which was investigated in [39].

Since theta-Milstein schemes (2.6) and (2.7) are semi-implicit when  $\theta \in (0, 1]$ , to guarantee that they are well defined, we restrict the stepsize  $\Delta$  satisfying  $\theta \mu \Delta < 1$ . This, together with the one-sided Lipschitz condition (2.2), implies the equation

$$y = z + \theta \Delta f(y)$$

has unique solution  $y = F_{\Delta, \theta}(z)$  for any  $z \in \mathbb{R}^n$ .

### 3 Uniform boundedness of $p$ th moments

In order to investigate the boundedness of  $p$ th moments and convergence of the two classes theta-Milstein approximations, we need the following assumption, which is standard for the classical Milstein scheme.

**Assumption 3.1.** Assume that the functions  $f, \{g_j\}_{j=1}^d \in C^2$  and there exists a constant  $\sigma$  such that for any  $x, y \in \mathbb{R}^n$  and  $i = 1, \dots, d, j = 1, \dots, d$ ,

$$|L^i g_j(x) - L^i g_j(y)|^2 \leq \sigma |x - y|^2. \quad (3.1)$$

Then let us investigate the  $p$ th moment boundedness of the theta-Milstein approximations for  $\theta \in [1/2, 1]$  and  $\theta \in [0, 1/2)$ , respectively.

**Theorem 3.1.** *Let Assumptions 2.1 and 3.1 hold,  $\theta \in [1/2, 1]$  and let  $\Delta < \Delta^* = 1/(2\theta\beta)$ . Then for each  $p \geq 2$ ,*

$$\mathbb{E} \left[ \sup_{k\Delta \in [0, T]} |z_k|^p \right] \leq C \quad (3.2)$$

and

$$\mathbb{E} \left[ \sup_{k\Delta \in [0, T]} |y_k|^p \right] \leq C. \quad (3.3)$$

*Proof.* Let  $Q_k = \sum_{j_1, j_2=1}^d L^{j_1} g_{j_2}(y_k) (\Delta w_k^{j_1} \Delta w_k^{j_2} - \delta_{j_1 j_2} \Delta)$ . By the second equation in (2.6),

$$\begin{aligned} |z_{k+1}|^2 &= |z_k|^2 + |f(y_k)|^2 \Delta^2 + \frac{1}{4} |Q_k|^2 + 2\Delta z_k^T f(y_k) + \langle g(y_k) \Delta w_k, Q_k \rangle \\ &\quad + 2\langle z_k + f(y_k) \Delta, g(y_k) \Delta w_k + \frac{1}{2} Q_k \rangle + |g(y_k) \Delta w_k|^2 \\ &= |z_k|^2 + 2\Delta y_k^T f(y_k) + 2|g(y_k) \Delta w_k|^2 + (1 - 2\theta) |f(y_k)|^2 \Delta^2 \\ &\quad + \frac{1}{2} |Q_k|^2 + \frac{2}{\theta} \langle y_k - (1 - \theta) z_k, g(y_k) \Delta w_k \rangle + \frac{1}{\theta} \langle y_k - (1 - \theta) z_k, Q_k \rangle, \end{aligned}$$

where we used the equation  $z_k = y_k - \theta f(y_k) \Delta$ . Note that  $\theta \in [1/2, 1]$ . By (2.4), we have

$$\begin{aligned} |z_{k+1}|^2 &\leq |z_k|^2 + 2(\alpha + \beta |y_k|^2) \Delta + 2|g(y_k) \Delta w_k|^2 + \frac{1}{2} |Q_k|^2 \\ &\quad + \frac{2}{\theta} \langle y_k, g(y_k) \Delta w_k \rangle - 2 \frac{1 - \theta}{\theta} \langle z_k, g(y_k) \Delta w_k \rangle + \frac{1}{\theta} \langle y_k, Q_k \rangle - \frac{1 - \theta}{\theta} \langle z_k, Q_k \rangle, \end{aligned}$$

which implies

$$\begin{aligned} |z_{k+1}|^2 &\leq |z_0|^2 + 2\alpha T + \beta \Delta \sum_{i=0}^k |y_i|^2 + 2 \sum_{i=0}^k |g(y_i) \Delta w_i|^2 \\ &\quad + \frac{1}{\theta} \sum_{i=0}^k \langle y_i, Q_i \rangle + \frac{2}{\theta} \sum_{i=0}^k \langle y_i, g(y_i) \Delta w_i \rangle \\ &\quad - 2 \frac{1 - \theta}{\theta} \sum_{i=0}^k \langle z_i, g(y_i) \Delta w_i \rangle + \frac{1}{2} \sum_{i=0}^k |Q_i|^2 - \frac{1 - \theta}{\theta} \sum_{i=0}^k \langle z_i, Q_i \rangle. \end{aligned}$$

Recall the elementary inequality: for  $x_1, \dots, x_l \geq 0$ ,  $p \geq 1$ ,  $l = 1, 2, \dots, N$ ,

$$\left( \sum_{i=1}^l x_i \right)^p \leq l^{p-1} \sum_{i=1}^l x_i^p. \quad (3.4)$$

We therefore have

$$\begin{aligned} \frac{1}{8^{p-1}} |z_{k+1}|^{2p} &\leq (|z_0|^2 + 2\alpha T)^p + \beta^p \Delta^p \left( \sum_{i=0}^k |y_i|^2 \right)^p + 2^p \left( \sum_{i=0}^k |g(y_i) \Delta w_i|^2 \right)^p \\ &\quad + 2^p \left| \sum_{i=0}^k \langle y_i, Q_i \rangle \right|^p + 4^p \left| \sum_{i=0}^k \langle y_i, g(y_i) \Delta w_i \rangle \right|^p \\ &\quad + 2^p \left| \sum_{i=0}^k \langle z_i, g(y_i) \Delta w_i \rangle \right|^p + 2^{-p} \left( \sum_{i=0}^k |Q_i|^2 \right)^p + \left| \sum_{i=0}^k \langle z_i, Q_i \rangle \right|^p. \end{aligned}$$

Note that  $y_i$  is  $\mathfrak{F}_i$ -measurable while  $\Delta w_i$  is independent of  $\mathfrak{F}_i$ . Hence, for any integer  $m < N$ ,

$$\begin{aligned}
 \mathbb{E} \left( \sup_{0 \leq k \leq m} \sum_{i=0}^k |g(y_i) \Delta w_i|^2 \right)^p &\leq N^{p-1} \mathbb{E} \left[ \sup_{0 \leq k \leq m} \sum_{i=0}^k |g(y_i) \Delta w_i|^{2p} \right] \\
 &= N^{p-1} \mathbb{E} \left[ \sum_{i=0}^m |g(y_i) \Delta w_i|^{2p} \right] \\
 &\leq N^{p-1} d^{2p-1} \sum_{i=0}^m \sum_{j=1}^d \mathbb{E} |g_j(y_i)|^{2p} \mathbb{E} |\Delta w_i^j|^{2p} \\
 &\leq C \Delta \sum_{i=0}^m \mathbb{E} [\alpha + \beta |y_i|^2]^p \\
 &\leq C \Delta \sum_{i=0}^m \mathbb{E} [\alpha^p + \beta^p |y_i|^{2p}] \\
 &\leq C + C \Delta \sum_{i=0}^m \mathbb{E} [|y_i|^{2p}], \tag{3.5}
 \end{aligned}$$

where we also used (2.4). Using the Burkholder-Davis-Gundy inequality, (2.4), and Assumption 3.1, we have

$$\begin{aligned}
 \mathbb{E} \left[ \sup_{0 \leq k \leq m} \left| \sum_{i=0}^k y_i^T Q_i \right|^p \right] &= \mathbb{E} \left[ \sup_{0 \leq k \leq m} \left| \sum_{i=0}^k \sum_{j_1, j_2=1}^d y_i^T L^{j_1} g_{j_2}(y_i) (\Delta w_i^{j_1} \Delta w_i^{j_2} - \delta_{j_1 j_2} \Delta) \right|^p \right] \\
 &\leq d^{2p-2} \sum_{j_1, j_2=1}^d \mathbb{E} \left[ \sup_{0 \leq k \leq m} \left| \sum_{i=0}^k y_i^T L^{j_1} g_{j_2}(y_i) (\Delta w_i^{j_1} \Delta w_i^{j_2} - \delta_{j_1 j_2} \Delta) \right|^p \right] \\
 &\leq C \sum_{j_1, j_2=1}^d \mathbb{E} \left[ \sum_{i=0}^m |y_i|^2 |L^{j_1} g_{j_2}(y_i)|^2 \Delta^2 \right]^{p/2} \\
 &\leq C \Delta^p (m+1)^{p/2-1} \sum_{i=0}^m \mathbb{E} \left[ |y_i|^2 \sum_{j_1, j_2=1}^d |L^{j_1} g_{j_2}(0)|^2 + \sigma |y_i|^2 \right]^{p/2} \\
 &\leq C \Delta^{p/2} + C \Delta^{p/2+1} \sum_{i=0}^m \mathbb{E} [|y_i|^{2p}], \tag{3.6}
 \end{aligned}$$

$$\begin{aligned}
 \mathbb{E} \left[ \sup_{0 \leq k \leq m} \left| \sum_{i=0}^k y_i^T g(y_i) \Delta w_i \right|^p \right] &= \mathbb{E} \left[ \sup_{0 \leq k \leq m} \left| \sum_{i=0}^k \sum_{j=1}^d y_i^T g_j(y_i) \Delta w_i^j \right|^p \right] \\
 &\leq d^{p-1} \sum_{j=1}^d \mathbb{E} \left[ \sup_{0 \leq k \leq m} \left| \sum_{i=0}^k y_i^T g_j(y_i) \Delta w_i^j \right|^p \right]
 \end{aligned}$$

$$\begin{aligned}
 &\leq C \sum_{j=1}^d \mathbb{E} \left[ \sum_{i=0}^m |y_i|^2 |g_j(y_i)|^2 \Delta \right]^{p/2} \\
 &\leq C \Delta^{p/2} (m+1)^{p/2-1} \mathbb{E} \sum_{i=0}^m |y_i|^p [\alpha + \beta |y_i|^2]^{p/2} \\
 &\leq C + C \Delta \sum_{i=0}^m [1 + \mathbb{E}[|y_i|^{2p}], \tag{3.7}
 \end{aligned}$$

and

$$\begin{aligned}
 \mathbb{E} \left[ \sup_{0 \leq k \leq m} \left| \sum_{i=0}^k z_i^T g(y_i) \Delta w_i \right|^p \right] &\leq C \sum_{j=1}^d \mathbb{E} \left[ \sum_{i=0}^m |z_i|^2 |g_j(y_i)|^2 \Delta \right]^{p/2} \\
 &\leq C \Delta^{p/2} (m+1)^{p/2-1} \mathbb{E} \sum_{i=0}^m |z_i|^p [\alpha + \beta |y_i|^2]^{p/2} \\
 &\leq C \Delta \sum_{i=0}^m [1 + \mathbb{E}[|y_i|^{2p}] + C \Delta \sum_{i=0}^m \mathbb{E}[|z_i|^{2p}]. \tag{3.8}
 \end{aligned}$$

Similarly to (3.6), we can obtain

$$\begin{aligned}
 \mathbb{E} \left[ \sup_{0 \leq k \leq m} \left| \sum_{i=0}^k z_i^T Q_i \right|^p \right] &\leq C \Delta^p (m+1)^{p/2-1} \sum_{i=0}^m \mathbb{E} \left[ |z_i|^2 \left( \sum_{j_1, j_2=1}^d |L^{j_1} g_{j_2}(0)|^2 + \sigma |y_i|^2 \right) \right]^{p/2} \\
 &\leq C + C \Delta \sum_{i=0}^m \mathbb{E}[|y_i|^{2p}] + C \Delta \sum_{i=0}^m \mathbb{E}[|z_i|^{2p}]. \tag{3.9}
 \end{aligned}$$

By the inequality (3.4) and Assumption 3.1, we have

$$\begin{aligned}
 \mathbb{E} \left( \sup_{0 \leq k \leq m} \sum_{i=0}^k |Q_i|^2 \right)^p &\leq N^{p-1} \mathbb{E} \left[ \sup_{0 \leq k \leq m} \sum_{i=0}^k |Q_i|^{2p} \right] \\
 &\leq N^{p-1} d^{4p-2} \sum_{j_1, j_2=1}^d \mathbb{E} \left[ \sum_{i=0}^m |L^{j_1} g_{j_2}(y_i)|^2 (\Delta w_i^{j_1} \Delta w_i^{j_2} - \delta_{j_1 j_2} \Delta)^{2p} \right] \\
 &= N^{p-1} d^{4p-2} \sum_{j_1, j_2=1}^d \sum_{i=0}^m \mathbb{E} |L^{j_1} g_{j_2}(y_i)|^{2p} \mathbb{E} |\Delta w_i^{j_1} \Delta w_i^{j_2} - \delta_{j_1 j_2} \Delta|^{2p} \\
 &= C \sum_{j_1, j_2=1}^d \Delta \sum_{i=0}^m \mathbb{E} [|L^{j_1} g_{j_2}(0)|^2 + \sigma |y_i|^2]^p \\
 &\leq C + C \Delta \sum_{i=0}^m \mathbb{E}[|y_i|^{2p}]. \tag{3.10}
 \end{aligned}$$

Combining (3.5)-(3.10) with (3.5) yields

$$\begin{aligned}
 \mathbb{E} \left[ \sup_{0 \leq k \leq m+1} |z_k|^{2p} \right] &\leq C + C \Delta \sum_{i=0}^m \mathbb{E}[|y_i|^{2p}] + C \Delta \sum_{i=0}^m \mathbb{E}[|y_i|^{2p}] \\
 &\leq C + C \Delta \sum_{i=0}^m \mathbb{E} \left[ \sup_{0 \leq k \leq i} |y_k|^{2p} \right] + C \Delta \sum_{i=0}^m \mathbb{E} \left[ \sup_{0 \leq k \leq i} |z_k|^{2p} \right]. \tag{3.11}
 \end{aligned}$$

Using  $z_k = y_k - \theta f(y_k)\Delta$  and (2.4), we have

$$\begin{aligned} |z_k|^2 &= |y_k|^2 - 2\theta y_k^T f(y_k)\Delta + \theta^2 \Delta^2 |f(y_k)|^2 \\ &\geq (1 - 2\theta\Delta\beta)|y_k|^2 - 2\theta\Delta\alpha, \end{aligned}$$

that is,

$$(1 - 2\theta\Delta\beta)|y_k|^2 \leq |z_k|^2 + 2\theta\Delta\alpha, \quad (3.12)$$

which together with (3.11) implies that for  $\Delta < \Delta^*$

$$\mathbb{E}\left[\sup_{0 \leq k \leq m+1} |z_k|^{2p}\right] \leq C + C\Delta \sum_{i=0}^m \mathbb{E}\left[\sup_{0 \leq k \leq i} |z_k|^{2p}\right].$$

Using the discrete-type Gronwall inequality and noting that  $(m+1)\Delta \leq T$  give

$$\mathbb{E}\left[\sup_{0 \leq k \leq m+1} |z_k|^{2p}\right] \leq C. \quad (3.13)$$

This together with (3.12) gives the desired assertion (3.3).  $\square$

The following theorem gives the moment boundedness for  $\theta \in [0, 1/2)$ .

**Theorem 3.2.** *Let Assumptions 2.1 and 3.1 hold,  $\theta \in [0, 1/2)$  and let  $\Delta < \Delta_1 = 1/(2\theta\beta)$  ( $\Delta_1 = \infty$  if  $\theta = 0$ ). If function  $f$  satisfies the linear growth condition*

$$|f(x)|^2 \leq K(1 + |x|^2), \quad (3.14)$$

then for each  $p \geq 2$ ,

$$\mathbb{E}\left[\sup_{k\Delta \in [0, T]} |y_k|^p\right] \leq C \quad (3.15)$$

and

$$\mathbb{E}\left[\sup_{k\Delta \in [0, T]} |z_k|^p\right] \leq C. \quad (3.16)$$

*Proof.* By (2.6),

$$\begin{aligned} |z_{k+1}|^2 &= |z_k|^2 + 2\langle z_k, f(y_k)\Delta + g(y_k)\Delta w_k + \frac{1}{2}Q_k \rangle \\ &\quad + |f(y_k)\Delta + g(y_k)\Delta w_k + \frac{1}{2}Q_k|^2. \end{aligned} \quad (3.17)$$

Note that

$$z_k = y_k - \theta f(y_k)\Delta. \quad (3.18)$$

Substituting this into (3.17) produces

$$\begin{aligned} |z_{k+1}|^2 &\leq |z_k|^2 + 2\langle y_k, f(y_k)\Delta + g(y_k)\Delta w_k + \frac{1}{2}Q_k \rangle \\ &\quad + \theta^2 \Delta^2 |f(y_k)|^2 + 2|f(y_k)\Delta + g(y_k)\Delta w_k + \frac{1}{2}Q_k|^2 \\ &\leq |z_k|^2 + 2y_k^T f(y_k)\Delta + 2y_k^T g(y_k)\Delta w_k + y_k^T Q_k \\ &\quad + 7\Delta^2 |f(y_k)|^2 + 6|g(y_k)\Delta w_k|^2 + 3|Q_k|^2 \\ &\leq |z_k|^2 + (2\alpha + 7|f(0)|^2\Delta)\Delta + (2\beta + 7K\Delta)|y_k|^2\Delta \\ &\quad + 2y_k^T g(y_k)\Delta w_k + y_k^T Q_k + 6|g(y_k)\Delta w_k|^2 + 3|Q_k|^2, \end{aligned}$$

which implies

$$\begin{aligned} |z_{k+1}|^2 &\leq |z_0|^2 + (2\alpha + 7|f(0)|^2\Delta)T + (2\beta + 7K\Delta)\Delta \sum_{i=0}^k |y_i|^2 \\ &\quad + 2 \sum_{i=0}^k y_i^T g(y_i) \Delta w_i + \sum_{i=0}^k y_i^T Q_i + 6 \sum_{i=0}^k |g(y_i) \Delta w_i|^2 + 3 \sum_{i=0}^k |Q_i|^2. \end{aligned}$$

Similar to (3.5), we have

$$\begin{aligned} \frac{1}{6^{p-1}} |z_{k+1}|^{2p} &\leq (|z_0|^2 + (2\alpha + 7|f(0)|^2\Delta)T)^p + (2\beta + 7K\Delta)^p \Delta^p \left( \sum_{i=0}^k |y_i|^2 \right)^p \\ &\quad + 2^p \left| \sum_{i=0}^k y_i^T g(y_i) \Delta w_i \right|^p + \left| \sum_{i=0}^k y_i^T Q_i \right|^p \\ &\quad + 6^p \left( \sum_{i=0}^k |g(y_i) \Delta w_i|^2 \right)^p + 3^p \left( \sum_{i=0}^k |Q_i|^2 \right)^p. \end{aligned} \quad (3.19)$$

Combining (3.5)-(3.7), (3.10) with (3.19) yields

$$\mathbb{E} \left[ \sup_{0 \leq k \leq m+1} |z_k|^{2p} \right] \leq C + C\Delta \sum_{i=0}^m \mathbb{E}[|y_i|^{2p}],$$

which together with (3.12) implies

$$\mathbb{E} \left[ \sup_{0 \leq k \leq m+1} |z_k|^{2p} \right] \leq C + C\Delta \sum_{i=0}^m \mathbb{E}[|y_i|^{2p}] \leq C + C\Delta \sum_{i=0}^m \mathbb{E} \left[ \sup_{0 \leq k \leq i} |z_k|^{2p} \right].$$

Note that  $(m+1)\Delta \leq T$ . Using the discrete-type Gronwall inequality gives the desired assertion (3.16). Similarly, using (3.12) and (3.16) gives another desired assertion (3.15).  $\square$

**Remark 3.1.** For  $\theta \in [0, 1/2)$ , the linear growth condition (3.14) may be necessary for guaranteeing the moment boundedness of the theta-Milstein schemes. This fact can be found in [33, Lemma 3.2] and [9, 10].

## 4 Convergence of the theta-Milstein approximations

Let us now introduce appropriate continuous-time interpolations corresponding to the discrete numerical approximations. More accurately, let us define the continuous solutions  $z(t)$  and  $y(t)$  for  $t \in [t_k, t_{k+1})$  as follows

$$\begin{cases} z(t) = z(t_k) + (t - t_k)f(y_k) + g(y_k)(w(t) - w(t_k)) + \sum_{i,j=1}^d L^i g_j(y_k) I_{i,j}^{t_k,t}, \\ y(t) = F_{\Delta,\theta}(z(t)), \end{cases} \quad (4.1)$$

with  $z(0) = z_0 = y_0 - \theta \Delta f(y_0)$ , where  $t_k = k\Delta$ . Note that  $z(t) = y(t) - \theta f(y(t))\Delta$ . Then the continuous-time approximations  $z(t)$  and  $y(t)$  are  $\mathfrak{F}_t$ -measurable. In our analysis it will be more natural to work with the equivalent definition of  $z(t)$

$$z(t) = z(0) + \int_0^t f(y(\check{s}))ds + \int_0^t g(y(\check{s}))dw(s) + \sum_{i,j=1}^d \int_0^t \int_{\check{s}}^s L^i g_j(y(\check{u}))dw^i(u)dw^j(s), \quad (4.2)$$

where  $\check{s} = t_k$  for  $s \in [t_k, t_{k+1})$ . Note that  $z(t_k) = z_k$  and  $y(t_k) = y_k$ . We refer to  $z(t)$  and  $y(t)$  as the continuous-time extensions of the discrete approximations  $z_k$  and  $y_k$ , respectively.

In what follows, we also use Taylor's formula frequently. If a function  $h : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is twice differentiable, the following Taylor's formula holds

$$h(z(s)) - h(z(\check{s})) = h'(z(\check{s}))(z(s) - z(\check{s})) + \bar{R}_s(h), \quad (4.3)$$

where  $\bar{R}_s(h)$  is the remainder term defined by

$$\bar{R}_s(h) = \int_0^1 (1-r)h''(z(\check{s}) + r(z(s) - z(\check{s}))) (z(s) - z(\check{s}), z(s) - z(\check{s}))dr. \quad (4.4)$$

Here for any  $u, v \in \mathbb{R}^n$  the derivatives have the following expression

$$h'(\cdot)(u) = \sum_{i=1}^n \frac{\partial h}{\partial x^i} u^i, \quad h''(\cdot)(u, v) = \sum_{i,j=1}^n \frac{\partial^2 h}{\partial x^i \partial x^j} u^i v^j.$$

Replacing  $z(s) - z(\check{s})$  in (4.3) with (4.1) and rearranging lead to

$$h(z(s)) - h(z(\check{s})) = h'(z(\check{s})) \left( \int_{\check{s}}^s g(y(\check{u}))dw(u) \right) + R_s(h), \quad (4.5)$$

where

$$R_s(h) = h'(z(\check{s})) \left( (s - \check{s})f(y(\check{s})) + \sum_{i,j=1}^d L^i g_j(y(\check{s}))I_{i,j}^{\check{s},s} \right) + \bar{R}_s(h).$$

In order to obtain the strong convergent rate, we also need the following assumption.

**Assumption 4.1.** *There exist positive constants  $D$  and  $q$  such that for all  $x \in \mathbb{R}^n$  and  $j = 1, \dots, d$ ,*

$$|f'(x)| \vee |f''(x)| \leq D(1 + |x|^q) \quad (4.6)$$

and

$$|g_j''(x)| \leq D. \quad (4.7)$$

Let us firstly give the convergence theorem of the theta-Milstein schemes for  $\theta \in [1/2, 1]$  as below.

**Theorem 4.1.** *Let Assumptions 2.1, 3.1 and 4.1 hold,  $\theta \in (1/2, 1]$  and let  $\Delta < \Delta^* = 1/(2\theta\beta)$ . Then for any  $p \geq 2$ ,*

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |x(t) - z(t)|^p \right] \leq C\Delta^p \quad (4.8)$$

and

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |x(t) - y(t)|^p \right] \leq C\Delta^p. \quad (4.9)$$

To prove Theorem 4.1, we need the following lemmas.

**Lemma 4.2.** *Let the conditions in Theorem 4.1 hold. Then for any  $p \geq 2$ ,*

$$\sup_{0 \leq k \leq N} [\mathbb{E}|f(y_k)|^p] \vee \sup_{0 \leq k \leq N} [\mathbb{E}|f'(y_k)|^p] \vee \sup_{0 \leq k \leq N} [\mathbb{E}|f''(y_k)|^p] \leq C \quad (4.10)$$

and

$$\sup_{0 \leq k \leq N} [\mathbb{E}|g(y_k)|^p] \vee \sup_{0 \leq k \leq N} [\mathbb{E}|L^1 g(y_k)|^p] \leq C. \quad (4.11)$$

Moreover, replacing  $y_k$  by  $z_k$ , (4.10) and (4.11) are also true.

*Proof.* By the polynomial growth condition (4.6), the global Lipschitz condition for  $g$  and  $L^1 g$  and Theorem 3.1 give the desired assertions immediately.  $\square$

Similarly, we can use the polynomial growth condition (4.6), the global Lipschitz condition for  $g$  and  $L^1 g$  and Lemma 2.2 to obtain the following lemma, whose proof is omitted.

**Lemma 4.3.** *Let conditions in Theorem 4.1 hold. Then for any  $p \geq 2$ ,  $i, j = 1, \dots, d$ ,*

$$\sup_{0 \leq t \leq T} [\mathbb{E}|f(x(t))|^p] \vee \sup_{0 \leq t \leq T} [\mathbb{E}|f'(x(t))|^p] \vee \sup_{0 \leq t \leq T} [\mathbb{E}|f''(x(t))|^p] \leq C \quad (4.12)$$

and

$$\sup_{0 \leq t \leq T} [\mathbb{E}|g_j(y(t))|^p] \vee \sup_{0 \leq t \leq T} [\mathbb{E}|L^i g_j(y(t))|^p] \leq C. \quad (4.13)$$

**Lemma 4.4.** *Let conditions in Theorem 4.1 hold. For any  $p \geq 2$ ,*

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |z(t)|^p \right] \vee \mathbb{E} \left[ \sup_{t \in [0, T]} |y(t)|^p \right] \leq C \quad (4.14)$$

and

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |z(t) - y(t)|^p \right] \leq C \Delta^p. \quad (4.15)$$

*Proof.* For any  $p \geq 2$  and  $t \in [0, T]$ , by (4.2),

$$\begin{aligned} \mathbb{E} \left( \sup_{s \in [0, t]} |z(s)|^p \right) &\leq 4^{p-1} \mathbb{E}|z(0)|^p + 4^{p-1} \mathbb{E} \left| \int_0^t |f(y(\check{s}))| ds \right|^p + 4^{p-1} \mathbb{E} \left( \sup_{s \in [0, t]} \left| \int_0^s g(y(\check{s})) dw(s) \right|^p \right) \\ &\quad + 4^{p-1} \mathbb{E} \left( \sup_{s \in [0, t]} \left| \sum_{i,j=1}^d \int_0^s \int_{\check{v}}^v L^i g_j(y(\check{u})) dw^i(u) dw^j(v) \right|^p \right). \end{aligned} \quad (4.16)$$

Using Lemma 4.2, we have

$$\mathbb{E} \left| \int_0^t |f(y(\check{s}))| ds \right|^p \leq C \mathbb{E} \int_0^t |f(y(\check{s}))|^p ds \leq C. \quad (4.17)$$

Applying the Burkholder-Davis-Gundy inequality and Lemma 4.2 yield

$$\begin{aligned} \mathbb{E} \left( \sup_{s \in [0, t]} \left| \int_0^s g(y(\check{s})) dw(s) \right|^p \right) &\leq d^{p-1} \sum_{j=1}^d c_p \mathbb{E} \left( \int_0^t |g_j(y(\check{s}))|^2 ds \right)^{p/2} \\ &\leq C \sum_{j=1}^d \int_0^t \mathbb{E}|g_j(y(\check{s}))|^p ds \leq C \end{aligned} \quad (4.18)$$

and

$$\begin{aligned}
 & \mathbb{E} \left( \sup_{s \in [0, t]} \left| \sum_{i,j=1}^d \int_0^s \int_{\check{v}}^v L^i g_j(y(\check{u})) dw^i(u) dw^j(v) \right|^p \right) \\
 & \leq c_p d^{2p-2} \sum_{i,j=1}^d \mathbb{E} \left( \int_0^t \left| \int_{\check{v}}^v L^i g_j(y(\check{u})) dw(u) \right|^2 dv \right)^{p/2} \\
 & \leq C \Delta^{p/2} d^{2p-2} \sum_{i,j=1}^d \int_0^t \mathbb{E} |L^i g_j(y(\check{v}))|^p dv \\
 & \leq C,
 \end{aligned} \tag{4.19}$$

where  $c_p$  is a constant dependent on  $p$ . Therefore, (4.17)-(4.19) and (4.16) produce

$$\mathbb{E} \left[ \sup_{s \in [0, T]} |z(s)|^p \right] \leq C,$$

which together with (3.12) gives

$$\mathbb{E} \left[ \sup_{s \in [0, T]} |y(s)|^p \right] \leq C.$$

Note that  $z(t) = y(t) - \theta f(y(t))\Delta$ . Hence, (4.15) follows from (4.14).  $\square$

**Lemma 4.5.** *Let conditions in Theorem 4.1 hold. For any  $p \geq 2$  and  $t \in [0, T]$ ,*

$$\mathbb{E}|x(t) - x(\check{t})|^p \leq C \Delta^{p/2} \tag{4.20}$$

and

$$\mathbb{E}|z(t) - z(\check{t})|^p \leq C \Delta^{p/2}. \tag{4.21}$$

*Proof.* By (2.1) and the Burkholder-Davis-Gundy inequality, we have

$$\begin{aligned}
 \mathbb{E}|x(t) - x(\check{t})|^p &= C \mathbb{E} \left| \int_{\check{t}}^t f(x(s)) ds \right|^p + C \mathbb{E} \left| \int_{\check{t}}^t g(x(s)) dw(s) \right|^p \\
 &\leq C \Delta^{p-1} \int_{\check{t}}^t \mathbb{E} |f(x(s))|^p ds + C \mathbb{E} \left| \int_{\check{t}}^t |g(x(s))|^2 ds \right|^{p/2} \leq C \Delta^{p/2},
 \end{aligned}$$

where we also used Lemma 4.3. By (4.1), we have

$$\begin{aligned}
 \mathbb{E}|z(t) - z(\check{t})|^p &= C \Delta^p \mathbb{E} |f(y(\check{t}))|^p + C \mathbb{E} |g(y(\check{t}))(w(t) - w(\check{t}))|^p \\
 &\quad + C \mathbb{E} \left| \sum_{i,j=1}^d L^i g_j(y(\check{t})) I_{i,j}^{t_k, t} \right|^p \\
 &\leq C \Delta^p \mathbb{E} |f(y(\check{t}))|^p + C \Delta^{p/2} \mathbb{E} |g(y(\check{t}))|^p + C \Delta^p \mathbb{E} \left| \sum_{i,j=1}^d L^i g_j(y(\check{t})) \right|^p.
 \end{aligned}$$

Then we obtain (4.21) from Lemma 4.2.  $\square$

**Lemma 4.6.** *For any  $p \geq 2$  and  $t \in [0, T]$ ,*

$$\mathbb{E}|R_t(f)|^p \leq C \Delta^p. \tag{4.22}$$

*Proof.* By the definition of  $R_t(f)$ , we have

$$\begin{aligned} \mathbb{E}|R_t(f)|^p &\leq \Delta^p 3^{p-1} \mathbb{E}|f'(z(\check{t}))(f(y(\check{t})))|^p + 3^{p-1} \mathbb{E}|\bar{R}_t(f)|^p \\ &\quad + \Delta^p 3^{p-1} \frac{1}{2} \mathbb{E}|f'(z(\check{t}))(\sum_{i,j=1}^d L^i g_j(y(\check{t})) I_{i,j}^{\check{t},t})|^p. \end{aligned}$$

By Lemmas 4.4 and 4.5 as well as Assumption 4.1, we can obtain  $\mathbb{E}|\bar{R}_t(f)|^p \leq C\Delta^p$ . By Lemma 4.2, we can obtain the desired assertion (4.22).  $\square$

*Proof of Theorem 4.1.* Let  $e(t) = x(t) - z(t)$ . For the purpose of simplification, define  $G_j(s) = g_j(x(s)) - g_j(y(\check{s})) - \sum_{i=1}^d \int_{\check{s}}^s L^i g_j(y(\check{u})) dw^i(u)$ . By (4.2) and (2.1), we have

$$e(t) = e(0) + \int_0^t [f(x(s)) - f(y(\check{s}))] ds + \sum_{j=1}^d \int_0^t G_j(s) dw^j(s). \quad (4.23)$$

Applying Itô's formula to  $|e(t)|^2$  gives

$$\begin{aligned} |e(t)|^2 &= |e(0)|^2 + 2 \int_0^t \langle e(s), f(x(s)) - f(y(\check{s})) \rangle ds \\ &\quad + \sum_{j=1}^d \int_0^t |G_j(s)|^2 ds + 2 \sum_{j=1}^d \int_0^t \langle e(s), G_j(s) \rangle dw^j(s) \\ &\leq \theta^2 \Delta^2 |f(x(0))|^2 + \sum_{j=1}^d \int_0^t |G_j(s)|^2 ds + 2 \sum_{j=1}^d \int_0^t \langle e(s), G_j(s) \rangle dw^j(s) + 2\mu \int_0^t |e(s)|^2 ds \\ &\quad + 2 \int_0^t \langle e(s), f(z(\check{s})) - f(y(\check{s})) \rangle ds + 2 \int_0^t \langle e(s), f(z(s)) - f(z(\check{s})) \rangle ds \\ &=: \theta^2 \Delta^2 |f(x(0))|^2 + J_1(t) + J_2(t) + 2\mu \int_0^t |e(s)|^2 ds + J_3(t) + J_4(t). \end{aligned} \quad (4.24)$$

Applying Itô's formula for  $g_j(x(s))$  gives

$$g_j(x(s)) = g_j(x(\check{s})) + \sum_{i=1}^d \int_{\check{s}}^s L^i g_j(x(u)) dw^i(u) + \int_{\check{s}}^s (g_j \diamond (f, g))(x(u)) du, \quad (4.25)$$

where  $(g_j \diamond (f, g))(x(u)) = g_j'(x(u))f(x(u)) + \frac{1}{2} \text{trace}\{g(x(u))^T g_j''(x(u))g(x(u))\}$ . Substituting (4.25) into  $G_j(s)$  and using the Burkholder-Davis-Gundy inequality and Hölder's inequality yield that for

any  $p \geq 2$ ,

$$\begin{aligned}
 \mathbb{E}|G_j(s)|^p &\leq 4^{p-1}\mathbb{E}|g_j(x(\check{s})) - g_j(y(\check{s}))|^p + 4^{p-1}\mathbb{E}\left|\sum_{i=1}^d \int_{\check{s}}^s [L^i g_j(x(u)) - L^i g_j(x(\check{u}))]dw^i(u)\right|^p \\
 &\quad + 4^{p-1}\mathbb{E}\left|\sum_{i=1}^d \int_{\check{s}}^s [L^i g_j(x(\check{u})) - L^i g_j(y(\check{u}))]dw^i(u)\right|^p + 4^{p-1}\mathbb{E}\left|\int_{\check{s}}^s (g_j \diamond (f, g))(x(u))du\right|^p \\
 &\leq C\mathbb{E}|x(\check{s}) - y(\check{s})|^p + C\mathbb{E}\left|\int_{\check{s}}^s |x(u) - x(\check{u})|^2 ds\right|^{p/2} \\
 &\quad + C\mathbb{E}\left|\int_{\check{s}}^s |x(\check{u}) - y(\check{u})|^2 ds\right|^{p/2} + C\Delta^p \\
 &\leq C\mathbb{E}|x(\check{s}) - z(\check{s})|^p + C\mathbb{E}|z(\check{s}) - y(\check{s})|^p + C\Delta^{p/2-1}\mathbb{E}\int_{\check{s}}^s |x(u) - x(\check{u})|^p ds \\
 &\quad + C\Delta^{p/2-1}\mathbb{E}\int_{\check{s}}^s |x(\check{u}) - z(\check{u})|^p ds + C\Delta^{p/2-1}\mathbb{E}\int_{\check{s}}^s |z(\check{u}) - y(\check{u})|^p ds + C\Delta^p \\
 &\leq C\mathbb{E}|e(\check{s})|^p + C\Delta^p,
 \end{aligned}$$

where we also used Assumptions 2.1 and 3.1 as well as Lemmas 2.2, 4.2, 4.5. Hence,

$$\mathbb{E}\left[\sup_{t \in [0, r]} |J_1(t)|^{p/2}\right] = C\mathbb{E}\int_0^r |G(s)|^p ds \leq C\int_0^r \mathbb{E}|e(\check{s})|^p ds + C\Delta^p. \quad (4.26)$$

Using the Burkholder-Davis-Gundy inequality and the Hölder inequality yields that for any  $p \geq 2$  and  $r \in [0, T]$

$$\begin{aligned}
 \mathbb{E}\left[\sup_{t \in [0, r]} |J_2(t)|^{p/2}\right] &= 2^{p/2}\mathbb{E}\sup_{t \in [0, r]} \left|\sum_{j=1}^d \int_0^t \langle e(s), G_j(s) \rangle dw^j(s)\right|^{p/2} \\
 &\leq 2^{p/2}d^{p/2-1}c_p \sum_{j=1}^d \mathbb{E}\left(\int_0^r |\langle e(s), G_j(s) \rangle|^2 ds\right)^{p/4} \\
 &\leq 2^{p/2}d^{p/2-1}c_p \sum_{j=1}^d \mathbb{E}\left(\sup_{s \in [0, r]} |e(s)|^2 \int_0^r |G_j(s)|^2 ds\right)^{p/4},
 \end{aligned}$$

where  $c_p$  is a constant dependent on  $p$ . Recalling the fundamental inequality:  $2ab \leq \frac{\kappa_1}{d}a^2 + \frac{d}{\kappa_1}b^2$  for any  $a, b, \kappa_1 > 0$  yields

$$\begin{aligned}
 \mathbb{E}\left[\sup_{t \in [0, r]} |J_2(t)|^{p/2}\right] &\leq \kappa_1 \mathbb{E}\left[\sup_{s \in [0, r]} |e(s)|^p\right] + \frac{2^p c_p^2}{\kappa_1} d^{p-1} \sum_{j=1}^d \mathbb{E}\left(\int_0^r |G_j(s)|^2 ds\right)^{p/2} \\
 &\leq \kappa_1 \mathbb{E}\left[\sup_{s \in [0, r]} |e(s)|^p\right] + C\int_0^r \mathbb{E}|e(\check{s})|^p ds + C\Delta^p.
 \end{aligned} \quad (4.27)$$

Similarly, using Hölder's inequality, Assumption 4.1 and Lemma 4.4, we have

$$\begin{aligned}
 \mathbb{E}\left[\sup_{t \in [0, r]} |J_3(t)|^{p/2}\right] &\leq C\mathbb{E}\int_0^t |e(s)|^p ds + C\mathbb{E}\int_0^r |f(z(\check{s})) - f(y(\check{s}))|^p ds \\
 &\leq C\int_0^r \mathbb{E}|e(s)|^p ds + C\int_0^r \sqrt{\mathbb{E}(1 + |z(\check{s})|^q + |y(\check{s})|^q)^2 \mathbb{E}|z(\check{s}) - y(\check{s})|^{2p}} ds \\
 &\leq C\int_0^r \mathbb{E}|e(s)|^p ds + C\Delta^p.
 \end{aligned} \quad (4.28)$$

Now, let us estimate  $J_4$ . Using (4.5) produces

$$\begin{aligned}
 J_4(t) &= 2 \int_0^t \left\langle e(s), f'(z(\check{s})) \left( \int_{\check{s}}^s g(y(\check{s})) dw(u) \right) + R_s(f) \right\rangle ds \\
 &= 2 \int_0^t \left\langle e(s), f'(z(\check{s})) \left( \int_{\check{s}}^s g(y(\check{s})) dw(u) \right) \right\rangle ds + 2 \int_0^t \langle e(s), R_s(f) \rangle ds \\
 &\leq \sum_{j_1=1}^d J_{j_1 41}(t) + \int_0^t |e(s)|^2 ds + \int_0^t |R_s(f)|^2 ds,
 \end{aligned} \tag{4.29}$$

where  $J_{j_1 41}(t) = 2 \int_0^t \langle e(s), f'(z(\check{s})) (\int_{\check{s}}^s g_{j_1}(y(\check{s})) dw^{j_1}(u)) \rangle ds$ . It is easy to deduce from Lemma 4.6 that

$$\mathbb{E} \left[ \sup_{t \in [0, r]} \left| \int_0^t |R_s(f)|^2 ds \right|^{p/2} \right] \leq C \Delta^p. \tag{4.30}$$

Noting that  $e(s) = e(\check{s}) + \int_{\check{s}}^s [f(x(u)) - f(y(\check{s}))] du + \sum_{j_2=1}^d \int_{\check{s}}^s G_{j_2}(u) dw^{j_2}(u)$ , we have

$$\begin{aligned}
 J_{j_1 41}(t) &= 2 \int_0^t \int_{\check{s}}^s \langle e(\check{s}), f'(z(\check{s})) (g_{j_1}(y(\check{s}))) \rangle dw^{j_1}(u) ds \\
 &\quad + 2 \sum_{j_2=1}^d \int_0^t \left\langle \int_{\check{s}}^s G_{j_2}(u) dw^{j_2}(u), f'(z(\check{s})) \left( \int_{\check{s}}^s g_{j_1}(y(\check{s})) dw^{j_1}(u) \right) \right\rangle ds \\
 &\quad + 2 \int_0^t \left\langle \int_{\check{s}}^s [f(x(u)) - f(y(\check{s}))] du, f'(z(\check{s})) \left( \int_{\check{s}}^s g_{j_1}(y(\check{s})) dw^{j_1}(u) \right) \right\rangle ds \\
 &=: J_{j_1 411}(t) + J_{j_1 412}(t) + J_{j_1 413}(t),
 \end{aligned}$$

where  $J_{j_1 412}(t) = 2 \sum_{j_2=1}^d J_{j_1 j_2 412}(t)$ ,  $J_{j_1 j_2 412}(t) = \int_0^t \left\langle \int_{\check{s}}^s G_{j_2}(u) dw^{j_2}(u), f'(z(\check{s})) \left( \int_{\check{s}}^s g_{j_1}(y(\check{s})) dw^{j_1}(u) \right) \right\rangle ds$ . For any  $t \in [0, T]$ , let  $k_t = \max\{k > 0, t_k < t\}$ . Define  $\hat{s} := t_i$  for  $t_{i-1} < s \leq t_i$  and  $\hat{s} := t$  for  $t_{k_t} < s \leq t$ . Applying Fubini's theorem, we can obtain

$$\begin{aligned}
 J_{j_1 411}(t)/2 &= \sum_{i=1}^{k_t} \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^s \langle e(t_{i-1}), f'(z(t_{i-1})) (g_{j_1}(y(t_{i-1}))) \rangle dw^{j_1}(u) ds \\
 &\quad + \int_{t_{k_t}}^t \int_{t_{k_t}}^s \langle e(t_{k_t}), f'(z(t_{k_t})) (g_{j_1}(y(t_{k_t}))) \rangle dw^{j_1}(u) ds \\
 &= \sum_{i=1}^{k_t} \int_{t_{i-1}}^{t_i} \int_u^{t_i} \langle e(t_{i-1}), f'(z(t_{i-1})) (g_{j_1}(y(t_{i-1}))) \rangle ds dw^{j_1}(u) \\
 &\quad + \int_{t_{k_t}}^t \int_u^t \langle e(t_{k_t}), f'(z(t_{k_t})) (g_{j_1}(y(t_{k_t}))) \rangle ds dw^{j_1}(u) \\
 &= \sum_{i=1}^{k_t} \int_{t_{i-1}}^{t_i} (t_i - u) \langle e(t_{i-1}), f'(z(t_{i-1})) (g_{j_1}(y(t_{i-1}))) \rangle ds dw^{j_1}(u) \\
 &\quad + \int_{t_{k_t}}^t (t - u) \langle e(t_{k_t}), f'(z(t_{k_t})) (g_{j_1}(y(t_{k_t}))) \rangle ds dw^{j_1}(u) \\
 &= \int_0^t (\hat{u} - u) \langle e(\check{u}), f'(z(\check{u})) (g_{j_1}(y(\check{u}))) \rangle dw^{j_1}(u).
 \end{aligned}$$

For any  $\kappa_2 > 0$ , using the Burkholder-Davis-Gundy inequality, Hölder's inequality and Lemma 4.2 gives

$$\begin{aligned}
 \mathbb{E} \left[ \sup_{t \in [0, r]} |J_{j_1 411}(t)|^{p/2} \right] &\leq c_p \mathbb{E} \left( \int_0^r 2 |(\hat{u} - u) \langle e(\tilde{u}), f'(z(\tilde{u}))(g_{j_1}(y(\tilde{u}))) \rangle|^2 du \right)^{p/4} \\
 &\leq c_p \mathbb{E} \left( \int_0^r 2 |e(\tilde{u})|^2 \Delta^2 |f'(z(\tilde{u}))(g_{j_1}(y(\tilde{u})))|^2 du \right)^{p/4} \\
 &\leq c_p \mathbb{E} \left( 2 \sup_{s \in [0, r]} |e(s)|^2 \Delta^2 \int_0^r |f'(z(\tilde{u}))(g_{j_1}(y(\tilde{u})))|^2 du \right)^{p/4} \\
 &\leq \kappa_2 \mathbb{E} \left[ \sup_{s \in [0, r]} |e(s)|^p \right] + \frac{c_p^2 2^{p/2}}{\kappa_2} \Delta^p \mathbb{E} \left( \int_0^r |f'(z(\tilde{u}))(g_{j_1}(y(\tilde{u})))|^2 du \right)^{p/2} \\
 &\leq \kappa_2 \mathbb{E} \left[ \sup_{s \in [0, r]} |e(s)|^p \right] + C \Delta^p.
 \end{aligned} \tag{4.31}$$

Similarly,

$$\begin{aligned}
 \mathbb{E} \left[ \sup_{t \in [0, r]} |J_{j_1 j_2 412}(t)|^{p/2} \right] &\leq C \mathbb{E} \int_0^r \left| \left\langle \int_{\tilde{s}}^s G_{j_2}(u) dw^{j_2}(u), f'(z(\tilde{s})) \left( \int_{\tilde{s}}^s g_{j_1}(y(\tilde{s})) dw^{j_1}(u) \right) \right\rangle \right|^{p/2} ds \\
 &\leq C \int_0^r \sqrt{\mathbb{E} \left| \int_{\tilde{s}}^s G_{j_2}(u) dw^{j_2}(u) \right|^p \mathbb{E} \left| \int_{\tilde{s}}^s f'(z(\tilde{s}))(g_{j_1}(y(\tilde{s}))) dw^{j_1}(u) \right|^p} ds \\
 &\leq C \int_0^r \sqrt{\mathbb{E} \left| \int_{\tilde{s}}^s |G_{j_2}(u)|^2 du \right|^{p/2} \mathbb{E} \left| \int_{\tilde{s}}^s |f'(z(\tilde{s}))(g_{j_1}(y(\tilde{s})))|^2 du \right|^{p/2}} ds \\
 &\leq C \int_0^r \sqrt{\Delta^{p-2} \int_{\tilde{s}}^s \mathbb{E} |G_{j_2}(u)|^p du \int_{\tilde{s}}^s \mathbb{E} |f'(z(\tilde{s}))(g_{j_1}(y(\tilde{s})))|^p du} ds \\
 &\leq C \int_0^r \sqrt{\Delta^p [\mathbb{E} |e(\tilde{s})|^p + \Delta^p]} ds \\
 &\leq C \int_0^r \mathbb{E} |e(\tilde{s})|^p ds + C \Delta^p.
 \end{aligned}$$

Then

$$\mathbb{E} \left[ \sup_{t \in [0, r]} |J_{j_1 412}(t)|^{p/2} \right] \leq C \int_0^r \mathbb{E} |e(\tilde{s})|^p ds + C \Delta^p. \tag{4.32}$$

In order to estimate  $J_{413}$ , we divide it into the following three parts

$$\begin{aligned}
 J_{j_1 413}(t) &= 2 \int_0^t \left\langle \int_{\tilde{s}}^s [f(x(u)) - f(x(\tilde{u}))] du, f'(z(\tilde{s})) \left( \int_{\tilde{s}}^s g_{j_1}(y(\tilde{s})) dw^{j_1}(v) \right) \right\rangle ds \\
 &\quad + 2 \int_0^t \left\langle \int_{\tilde{s}}^s [f(z(\tilde{u})) - f(y(\tilde{s}))] du, f'(z(\tilde{s})) \left( \int_{\tilde{s}}^s g_{j_1}(y(\tilde{s})) dw^{j_1}(v) \right) \right\rangle ds \\
 &\quad + 2 \int_0^t \left\langle \int_{\tilde{s}}^s [f(x(\tilde{u})) - f(z(\tilde{u}))] du, f'(z(\tilde{s})) \left( \int_{\tilde{s}}^s g_{j_1}(y(\tilde{s})) dw^{j_1}(v) \right) \right\rangle ds \\
 &=: I_{j_1 1}(t) + I_{j_1 2}(t) + I_{j_1 3}(t).
 \end{aligned} \tag{4.33}$$

Using the Burkholder-Davis-Gundy inequality and Hölder's inequality give that for any  $p \geq 2$ ,

$$\begin{aligned}
 & \mathbb{E} \left[ \sup_{t \in [0, r]} |I_{j11}(t)|^{p/2} \right] \\
 & \leq C \mathbb{E} \int_0^r \left| \left\langle \int_{\check{s}}^s [f(x(u)) - f(x(\check{u}))] du, f'(z(\check{s})) \left( \int_{\check{s}}^s g_{j1}(y(\check{s})) dw^{j1}(u) \right) \right\rangle \right|^{p/2} ds \\
 & \leq C \int_0^r \sqrt{\mathbb{E} \left| \int_{\check{s}}^s [f(x(u)) - f(x(\check{u}))] du \right|^p \mathbb{E} \left| \int_{\check{s}}^s f'(z(\check{s})) (g_{j1}(y(\check{s}))) dw^{j1}(u) \right|^p} ds \\
 & \leq C \int_0^r \sqrt{\Delta^{p-1} \int_{\check{s}}^s \mathbb{E} |f(x(u)) - f(x(\check{u}))|^p du \Delta^{p/2-1} \int_{\check{s}}^s \mathbb{E} |f'(z(\check{s})) (g_{j1}(y(\check{s})))|^p duds} \\
 & \leq C \int_0^r \sqrt{\Delta^{p-1} \Delta \Delta^{p/2} \Delta^{p/2-1} \Delta} ds \leq C \Delta^p,
 \end{aligned} \tag{4.34}$$

where we used Lemma 4.2 and 4.5. Similarly,

$$\begin{aligned}
 \mathbb{E} \left[ \sup_{t \in [0, r]} |I_{j12}(t)|^{p/2} \right] & \leq (2T)^{p/2} \mathbb{E} \int_0^r \left| \left\langle \int_{\check{s}}^s [f(z(\check{u})) - f(y(\check{u}))] du, f'(z(\check{s})) \left( \int_{\check{s}}^s g_{j1}(y(\check{s})) dw^{j1}(u) \right) \right\rangle \right|^{p/2} ds \\
 & \leq C \int_0^r \sqrt{\mathbb{E} \left| \int_{\check{s}}^s [f(z(\check{u})) - f(y(\check{u}))] du \right|^p \mathbb{E} \left| \int_{\check{s}}^s f'(z(\check{s})) (g_{j1}(y(\check{s}))) dw^{j1}(u) \right|^p} ds \\
 & \leq C \int_0^r \sqrt{\Delta^p \mathbb{E} |f(z(\check{s})) - f(y(\check{s}))|^p \Delta^{p/2-1} \int_{\check{s}}^s \mathbb{E} |f'(z(\check{s})) (g_{j1}(y(\check{s})))|^p duds} \\
 & \leq C \int_0^r \sqrt{\Delta^p \Delta^p \Delta^{p/2}} ds \leq C \Delta^p.
 \end{aligned} \tag{4.35}$$

Applying Fubini's theorem obtains

$$\begin{aligned}
 I_{j13}(t) & = 2 \int_0^t \int_{\check{s}}^s (s - \check{s}) \langle f(x(\check{s})) - f(z(\check{s})), f'(z(\check{s})) (g_{j1}(y(\check{s}))) \rangle dw^{j1}(v) ds \\
 & = 2 \int_0^t \int_v^{\check{v}} (s - \check{s}) \langle f(x(\check{s})) - f(z(\check{s})), f'(z(\check{s})) (g_{j1}(y(\check{s}))) \rangle ds dw^{j1}(v).
 \end{aligned}$$

Therefore,

$$\begin{aligned}
 \mathbb{E} \left[ \sup_{t \in [0, r]} |I_{j13}(t)|^{p/2} \right] & \leq C \mathbb{E} \left( \int_0^r \left| \int_v^{\check{v}} (s - \check{v}) \langle f(x(\check{v})) - f(z(\check{v})), f'(z(\check{v})) (g_{j1}(y(\check{v}))) \rangle ds \right|^2 dv \right)^{p/4} \\
 & \leq C \mathbb{E} \left( \int_0^r \left| \int_v^{\check{v}} (s - \check{v}) ds \langle f(x(\check{v})) - f(z(\check{v})), f'(z(\check{v})) (g_{j1}(y(\check{v}))) \rangle \right|^2 dv \right)^{p/4} \\
 & \leq C \Delta^p \int_0^r \mathbb{E} \left| \langle f(x(\check{v})) - f(z(\check{v})), f'(z(\check{v})) (g_{j1}(y(\check{v}))) \rangle \right|^{p/2} dv \\
 & \leq C \Delta^p.
 \end{aligned} \tag{4.36}$$

By (4.34)-(4.36), we obtain from (4.33)

$$\mathbb{E} \left[ \sup_{t \in [0, r]} |J_{j1413}(t)|^{p/2} \right] \leq C \Delta^p,$$

which together with (4.31) and (4.32) implies

$$\begin{aligned} \mathbb{E} \left[ \sup_{t \in [0, r]} |J_{j_1 41}(t)|^{p/2} \right] &\leq 3^{p/2-1} \left( \mathbb{E} \left[ \sup_{t \in [0, r]} |J_{j_1 411}(t)|^{p/2} \right] + \mathbb{E} \left[ \sup_{t \in [0, r]} |J_{j_1 412}(t)|^{p/2} \right] \right. \\ &\quad \left. + \mathbb{E} \left[ \sup_{t \in [0, r]} |J_{j_1 413}(t)|^{p/2} \right] \right) \\ &\leq 3^{p/2-1} \kappa_2 \mathbb{E} \left[ \sup_{s \in [0, r]} |e(s)|^p \right] + C \int_0^r \mathbb{E} |e(\tilde{s})|^p ds + C \Delta^p. \end{aligned} \quad (4.37)$$

Combining (4.37), (4.30) with (4.29) gives

$$\mathbb{E} \left[ \sup_{t \in [0, r]} |J_4(t)|^{p/2} \right] \leq 3^{p-2} \kappa_2 d \mathbb{E} \left[ \sup_{s \in [0, r]} |e(s)|^p \right] + C \int_0^r \mathbb{E} |e(\tilde{s})|^p ds + C \Delta^p + C \int_0^r \mathbb{E} |e(s)|^p ds.$$

Let  $\kappa = (\kappa_1 + 3^{p-2} \kappa_2 d) 6^{p/2-1}$ . Combining (4.26), (4.27), (4.28), (4.38) and (4.24) produces

$$\begin{aligned} \mathbb{E} \left[ \sup_{t \in [0, r]} |e(t)|^p \right] &\leq 6^{p/2-1} \sum_{l=1}^4 \mathbb{E} \left[ \sup_{t \in [0, r]} |J_l(t)|^{p/2} \right] + C \Delta^p + C \int_0^r |e(s)|^p ds \\ &\leq \kappa \mathbb{E} \left[ \sup_{s \in [0, r]} |e(s)|^p \right] + C \Delta^p + C \int_0^r \mathbb{E} \left[ \sup_{u \in [0, s]} |e(u)|^p \right] ds. \end{aligned}$$

Therefore, letting  $\kappa_1$  and  $\kappa_2$  be sufficiently small such that  $\kappa < 1$ , we have

$$\mathbb{E} \left[ \sup_{t \in [0, r]} |e(t)|^p \right] \leq C \int_0^r \mathbb{E} \left[ \sup_{u \in [0, s]} |e(s)|^p \right] ds + C \Delta^p. \quad (4.38)$$

Thus, the result (4.8) follows from the Gronwall inequality. The desired assertion (4.9) follows from (4.8) and Lemma 4.4.  $\square$

For  $\theta \in [0, 1/2]$ , we can use the similar techniques to obtain the following convergence theorem. Its proof is omitted.

**Theorem 4.7.** *Let Assumptions 2.1, 3.1 and 4.1 hold,  $\theta \in [0, 1/2)$  and let  $\Delta < \Delta_1 = 1/(2\theta\beta)$  ( $\Delta_1 = \infty$ , if  $\theta = 0$ ). If the function  $f$  satisfies the linear growth condition (3.14), then for any  $p \geq 2$  we have*

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |x(t) - z(t)|^p \right] \leq C \Delta^p$$

and

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |x(t) - y(t)|^p \right] \leq C \Delta^p.$$

**Remark 4.1.** Convergence for  $\theta \in (1/2, 1]$  in Theorem 4.1 are partially based on the Assumption 2.1. This does not imply that condition (2.3) in Assumption 2.1 is necessary. For example, Kloeden and Neuenkirch [40] showed that the semi-implicit Milstein scheme (STM scheme with  $\theta = 1$ ) converges to the exact solution for the Cox-Ingersoll-Ross process, whose diffusion term is non-Lipschitz continuous.

**Remark 4.2.** Higham and his coauthors [12, 13] showed that the split step backward Euler and backward Euler schemes strongly converge to the exact solution with the standard order 0.5. In fact, by the similar skills used in the proof of Theorem 4.1, we can show that the convergence rate 0.5 holds for theta-Euler schemes (split-step theta and stochastic theta schemes) with  $\theta \in (1/2, 1]$ . Moreover, for the SDEs with additive noise, that is,  $g(x)$  is a constant matrix, then  $L^i g_j(x) = 0$ , the theta-Milstein schemes are degenerated to the theta-Euler schemes. Hence, our convergence results show that the theta-Euler schemes ( $\theta \in (1/2, 1]$ ) for such SDEs is convergent with order 1, which is a new result for the SDEs with one-sided Lipschitz and polynomial Lipschitz continuous drift.

## 5 Stability analysis of the theta-Milstein schemes for SDEs

For the purpose of stability, assume that  $f(0) = g_j(0) = 0$ . This shows that (2.1) admits a trivial solution. Then inequality (3.1) in Assumption 3.1 becomes

$$|L^i g_j(x)|^2 \leq \sigma |x|^2, \quad i, j = 1, \dots, d. \quad (5.1)$$

In this section, we assume that  $f$  and  $g$  satisfy the following local Lipschitz condition, which is classical for the nonlinear SDEs.

**Assumption 5.1.** (*Local Lipschitz condition*)  $f$  and  $g$  satisfy the local Lipschitz condition, that is, for each  $j > 0$  there exists a positive constant  $K_j$  such that for any  $x, y \in \mathbb{R}^n$  with  $|x| \vee |y| \leq j$ ,

$$|f(x) - f(y)| \vee |g_i(x) - g_i(y)| \leq K_j |x - y|, \quad i = 1, \dots, d. \quad (5.2)$$

To investigate stability of numerical approximations, let us firstly give the stability criterion of SDE (2.1) (also see [37]):

**Theorem 5.1.** *Let Assumption 5.1 hold. If there exists a positive constants  $\gamma$  such that for all  $x, y \in \mathbb{R}^n$ ,*

$$2x^T f(x) + \sum_{j=1}^d |g_j(x)|^2 \leq -\gamma |x|^2, \quad (5.3)$$

*then the solution of (2.1) has the property*

$$\mathbb{E}|x(t)|^2 \leq C(x_0)e^{-\gamma t}, \quad (5.4)$$

*where  $C(x_0)$  is a positive constant dependent on the initial data  $x_0$ .*

Under the one-sided Lipschitz condition (2.2), we define

$$\Delta^* = \begin{cases} \frac{1}{\mu\theta}, & \theta \in (0, 1), \text{ and } \mu > 0, \\ \infty, & \theta = 0 \text{ or } \mu < 0. \end{cases}$$

When  $\Delta < \Delta^*$ , the two classes of theta-Milstein schemes are well defined. Moreover, we have the following stability result.

**Theorem 5.2.** *Let the conditions in Theorem 5.1 hold. Under condition (5.1) and the one-sided Lipschitz condition (2.2), the following two stability results hold.*

(i) *Let  $\theta \in [0, 1/2]$ . If there exists a positive constant  $K$  such that for any  $x \in \mathbb{R}^n$ ,*

$$|f(x)|^2 \leq K|x|^2, \quad (5.5)$$

*then for any  $\Delta < \Delta_1^* = \frac{\gamma}{(1-2\theta)K + 0.5d^2\sigma} \wedge \Delta^*$ ,  $k \in \mathbb{N}_+$ ,*

$$\mathbb{E}|z_k|^2 \leq C(x_0)e^{-\gamma_\Delta k\Delta} \quad \text{and} \quad \mathbb{E}|y_k|^2 \leq C(x_0)e^{-\gamma_\Delta k\Delta},$$

*where  $\gamma_\Delta = -\frac{1}{\Delta} \log \left( 1 - \frac{\gamma - (1-2\theta)K\Delta - 0.5d^2\sigma\Delta}{(1+\theta\Delta\sqrt{K})^2} \Delta \right)$ .*

(ii) *Let  $\theta \in (1/2, 1]$ . Then for any  $\Delta < \Delta_2^* = \frac{2\gamma}{d^2\sigma} \wedge \Delta^*$ ,  $k \in \mathbb{N}_+$ ,*

$$\mathbb{E}|z_k|^2 \leq C(x_0)e^{-\gamma_\Delta k\Delta} \quad \text{and} \quad \mathbb{E}|y_k|^2 \leq C(x_0)e^{-\gamma_\Delta k\Delta},$$

*where  $\gamma_\Delta = -\frac{1}{\Delta} \log \left( 1 - \frac{(2\theta-1)(\gamma - 0.5d^2\sigma\Delta)\Delta}{2\theta-1 + (\gamma - 0.5d^2\sigma\Delta)\Delta\theta^2} \right)$ .*

*Proof.* For any large  $l > 0$ , define the stopping time

$$\lambda_l = \inf\{i > 0 : |y_i| > l \text{ or } |z_i| > l\}. \quad (5.6)$$

It is observed that  $|y_{k-1}| \leq l$ ,  $|z_{k-1}| \leq l$  for  $k \in [0, \lambda_l]$ . By (2.6),

$$|z_k|^2 \leq 4 \left( |z_{k-1}|^2 + |f(y_{k-1})|^2 \Delta^2 + |g(y_{k-1})\Delta w_{k-1}|^2 + \frac{1}{4}|Q_k|^2 \right).$$

Then we can obtain that

$$\begin{aligned} \mathbb{E}[|z_k|^2 \mathbf{1}_{[0, \lambda_l]}(k)] &\leq 4l^2 + 4\mathbb{E}[|f(y_{k-1})|^2 \mathbf{1}_{[0, \lambda_l]}(k)]\Delta^2 + 4d \sum_{j=1}^d \mathbb{E}[|g_j(y_{k-1})|^2 |\Delta w_{k-1}^j|^2 \mathbf{1}_{[0, \lambda_l]}(k)] \\ &\quad + 2d \sum_{j_1, j_2=1}^d \mathbb{E}[|L^{j_1} g_{j_2}(y_k)|^2 \mathbf{1}_{[0, \lambda_l]}(k)] |(\Delta w_k^{j_1} \Delta w_k^{j_2} - \delta_{j_1 j_2} \Delta)|^2 \\ &\leq 4\mathbb{E}[|f(y_{k-1})|^2 \mathbf{1}_{[0, \lambda_l]}(k)]\Delta^2 + 4d \sum_{j=1}^d \left( \mathbb{E}[|g_j(y_{k-1})|^4 \mathbf{1}_{[0, \lambda_l]}(k)] \mathbb{E}[|\Delta w_{k-1}^j|^4] \right)^{1/2} \\ &\quad + 4l^2 + 2d \sum_{j_1, j_2=1}^d \left( \mathbb{E}[|L^{j_1} g_{j_2}(y_k)|^4 \mathbf{1}_{[0, \lambda_l]}(k)] \mathbb{E}[|\Delta w_k^{j_1} \Delta w_k^{j_2} - \delta_{j_1 j_2} \Delta|^4] \right)^{1/2}, \end{aligned}$$

where we used Hölder's inequality. Making use of the property  $\mathbb{E}[|\Delta w_k^j|^{2p}] = (2p-1)!!\Delta^p$ , we have  $\mathbb{E}[|\Delta w_{k-1}^j|^4] = 3\Delta^2$  and  $\mathbb{E}[|\Delta w_k^{j_1} \Delta w_k^{j_2} - \delta_{j_1 j_2} \Delta|^4] \leq 2^3(\mathbb{E}[|\Delta w_{k-1}^{j_1}|^8] + \Delta^4) + \mathbb{E}[|\Delta w_k^{j_1}|^4]\mathbb{E}[|\Delta w_k^{j_2}|^4] = (2^3(7!! + 1) + 9)\Delta^4$ , where  $(2i-1)!! = (2i-1)(2i-3)\cdots 3 \cdot 1$  for  $i = 1, 2, \dots$ . By Assumption

5.1 and (5.1), it is easy to deduce that  $|f(y_{k-1})|^2 \mathbf{1}_{[0, \lambda_l]}(k) \leq K_l^2 l^2$ ,  $|g_j(y_{k-1})|^4 \mathbf{1}_{[0, \lambda_l]}(k) \leq K_l^4 l^4$  and  $|L^i g_j(y_{k-1})|^4 \mathbf{1}_{[0, \lambda_l]}(k) \leq \sigma^2 l^4$ . Therefore,

$$\mathbb{E}[|z_k|^2 \mathbf{1}_{[0, \lambda_l]}(k)] \leq 4l^2 + 4d^2 K_l^2 l^2 \Delta^2 + 8K_l^2 l^2 \Delta + 4\sqrt{7!!} + 2\sigma l^2 \Delta^2 =: K(l), \quad (5.7)$$

which together with  $z_k = y_k - \theta \Delta f(y_k, y_{k-m})$  and  $2x^T f(x) \leq -\gamma|x|^2$  yields

$$|z_k|^2 = |y_k|^2 - 2\theta \Delta y_k^T f(y_k) + |f(y_k)|^2 \Delta^2 \theta^2 \geq (1 + \gamma \theta \Delta) |y_k|^2. \quad (5.8)$$

Hence,

$$\mathbb{E}[|y_k|^2 \mathbf{1}_{[0, \lambda_l]}(k)] \leq K(l). \quad (5.9)$$

Note that  $Q_k = \sum_{j_1, j_2=1}^d L^{j_1} g_{j_2}(y_k) (\Delta w_k^{j_1} \Delta w_k^{j_2} - \delta_{j_1 j_2} \Delta)$ . By (2.6), we have

$$\begin{aligned} |z_{k+1}|^2 &= |z_k|^2 + |f(y_k)|^2 \Delta^2 + \frac{1}{4} |Q_k|^2 + \langle g(y_k) \Delta w_k, Q_k \rangle \\ &\quad + 2 \langle z_k + f(y_k) \Delta, g(y_k) \Delta w_k + \frac{1}{2} Q_k \rangle + |g(y_k) \Delta w_k|^2 + 2 \Delta z_k^T f(y_k) \\ &= |z_k|^2 + 2 \Delta y_k^T f(y_k) + \Delta \sum_{j=1}^d |g_j(y_k)|^2 + (1 - 2\theta) |f(y_k)|^2 \Delta^2 \\ &\quad + \frac{1}{2} \Delta^2 \sum_{i_1, i_2=1}^d |L^{j_1} g_{j_2}(y_k)|^2 + m_k, \end{aligned} \quad (5.10)$$

where  $m_k = 2 \langle y_k + (1 - \theta) f(y_k) \Delta, g(y_k) \Delta w_k + \frac{1}{2} Q_k \rangle + \sum_{i,j=1}^d g_j(y_k)^T g_i(y_k) (\Delta w_k^i \Delta w_k^j - \delta_{ij} \Delta) + \langle g(y_k) \Delta w_k, Q_k \rangle + \frac{1}{4} \sum_{j_1, j_2, j_3, j_4=1}^d L^{j_1} g_{j_2}(y_k)^T L^{j_3} g_{j_4}(y_k) (\Delta W_k^{j_1, j_2} \Delta W_k^{j_3, j_4} - 2 \delta_{j_1, j_2, j_3, j_4} \Delta)$ , where  $\Delta W_k^{j_1, j_2} = \Delta w_k^{j_1} \Delta w_k^{j_2} - \delta_{j_1 j_2} \Delta$ ,  $\delta_{j_1, j_2, j_3, j_4} = 1$  if  $j_1 = j_2 = j_3 = j_4$ , otherwise,  $\delta_{j_1, j_2, j_3, j_4} = 0$ . Using condition (5.1) and the coupled condition (5.3), we get

$$|z_{k+1}|^2 \leq |z_k|^2 - \gamma \Delta |y_k|^2 + (1 - 2\theta) \Delta^2 |f(y_k)|^2 + 0.5 \Delta^2 d^2 \sigma |y_k|^2 + m_k. \quad (5.11)$$

**Case (i)**  $\theta \in [0, 1/2]$ : Note that the linear growth condition of  $f$  and (5.3) imply

$$\sum_{j=1}^d |g_j(x)|^2 \leq -2x^T f(x) \leq |x|^2 + |f(x)|^2 \leq (1 + K) |x|^2. \quad (5.12)$$

From conditions (5.1), (5.5) and (5.12), there is a positive constant  $\bar{C}$  such that

$$|m_k| \leq [\bar{C} + \bar{C} |\Delta w_k|^2 + \bar{C} |\Delta w_k|^4] |y_k|^2. \quad (5.13)$$

Noting that  $y_k$  and  $\mathbf{1}_{[0, \lambda_l]}(k)$  are  $\mathfrak{F}_{k\Delta}$ -measurable while  $\Delta w_k$  is independent of  $\mathfrak{F}_{k\Delta}$ , then we obtain from (5.9) that  $\mathbb{E}[m_k \mathbf{1}_{[0, \lambda_l]}(k)] = 0$ . The linear growth condition (5.5) gives

$$|z_{k+1}|^2 \leq |z_k|^2 + [(1 - 2\theta) K \Delta + 0.5 d^2 \sigma \Delta - \gamma] \Delta |y_k|^2 + m_k. \quad (5.14)$$

Then for  $\Delta < \Delta_1^*$  and sufficiently large  $l$ , by (5.11) and using linear growth condition (5.5), we have

$$\begin{aligned} |z_{k \wedge \lambda_l}|^2 &\leq |z_0|^2 + [(1-2\theta)K\Delta + 0.5d^2\sigma\Delta - \gamma]\Delta \sum_{i=0}^{(k \wedge \lambda_l)-1} |y_i|^2 + \sum_{i=0}^{(k \wedge \lambda_l)-1} m_i \\ &\leq |z_0|^2 + \sum_{i=0}^{k-1} \mathbf{1}_{[0, \lambda_l]}(i) m_i. \end{aligned} \quad (5.15)$$

Hence, for  $\Delta < \Delta_1^*$ ,  $\mathbb{E}[|z_{k \wedge \lambda_l}|^2] \leq \mathbb{E}|z_0|^2 = \mathbb{E}|y_0 - \theta\Delta f(y_0)|^2 =: \bar{K}$ . Note that

$$\mathbb{E}[|z_{\lambda_l}|^2 \mathbf{1}_{\{\lambda_l < k\}}] \leq \mathbb{E}[|z_{k \wedge \lambda_l}|^2] \leq \bar{K}. \quad (5.16)$$

We now claim  $|z_{\lambda_l}| > l$ . Otherwise,  $|z_{\lambda_l}| \leq l$ . By the definition of  $\lambda_l$ ,  $|y_{\lambda_l}| > l$  and  $|y_{\lambda_l-i}| \leq l$  for  $i > 0$ . Then by  $z_k = y_k - \theta\Delta f(y_k)$ ,

$$\begin{aligned} |z_{\lambda_l}|^2 &= |y_{\lambda_l}|^2 - 2\theta\Delta y_{\lambda_l}^T f(y_{\lambda_l}) + |f(y_{\lambda_l})|^2 \Delta^2 \theta^2 \\ &\geq (1 + \gamma\theta\Delta) |y_{\lambda_l}|^2 \\ &> (1 + \gamma\theta\Delta) l^2 \geq l^2, \end{aligned}$$

which is a contradiction. By (5.16), for any  $k > 0$

$$\mathbb{P}\{\lambda_l < k\} \leq \frac{\bar{K}}{l^2} \rightarrow 0, \quad l \rightarrow \infty, \quad (5.17)$$

that is, as  $l \rightarrow \infty$ ,  $\lambda_l \uparrow \infty$  a.s. Define  $\mu_\Delta = \gamma - (1-2\theta)K\Delta - 0.5d^2\sigma\Delta$ . For  $\Delta < \Delta_1^*$ ,  $\mu_\Delta > 0$ . Using the linear growth condition (5.5) gives

$$|z_k|^2 = |y_k - \theta\Delta f(y_k)|^2 \leq (1 + \theta\Delta\sqrt{K})^2 |y_k|^2.$$

Substituting this into (5.14) yields that for  $\Delta < \Delta_1^*$

$$|z_{k+1}|^2 \leq \left(1 - \frac{\mu_\Delta \Delta}{(1 + \theta\Delta\sqrt{K})^2}\right) |z_k|^2 + m_k,$$

which implies

$$e^{\gamma_\Delta k \Delta} |z_k|^2 \leq |z_0|^2 + \sum_{j=0}^{k-1} e^{\gamma_\Delta j \Delta} m_j,$$

where  $\gamma_\Delta = -\frac{1}{\Delta} \log \left(1 - \frac{\mu_\Delta \Delta}{(1 + \theta\Delta\sqrt{K})^2}\right)$ . Replacing  $k$  by  $k \wedge \lambda_l$  yields

$$e^{\gamma_\Delta (k \wedge \lambda_l) \Delta} |z_{k \wedge \lambda_l}|^2 \leq |z_0|^2 + \sum_{j=0}^{k-1} e^{\gamma_\Delta j \Delta} \mathbf{1}_{[0, \lambda_l]}(j) m_j.$$

Taking expectation gives

$$\mathbb{E}[e^{\gamma_\Delta (k \wedge \lambda_l) \Delta} |z_{k \wedge \lambda_l}|^2] \leq \mathbb{E}|z_0|^2.$$

By Fatou's Lemma, letting  $l \rightarrow \infty$ , we have

$$e^{\gamma_{\Delta} k \Delta} \mathbb{E}[|z_k|^2] \leq \mathbb{E}|z_0|^2, \quad (5.18)$$

which together with (5.8) implies the desired assertion.

**Case (ii)**  $\theta \in (1/2, 1]$ : For  $\theta \in (1/2, 1]$ , by  $\Delta f(y_k) = (y_k - z_k)/\theta$ , we know that  $m_k$  defined in (5.10) has the following form

$$\begin{aligned} m_k &= 2\langle y_k + \frac{1-\theta}{\theta}(y_k - z_k), g(y_k)\Delta w_k + \frac{1}{2}Q_k \rangle + \sum_{i,j=1}^d g_j(y_k)^T g_i(y_k)(\Delta w_k^i \Delta w_k^j - \delta_{ij}\Delta) \\ &\quad + \langle g(y_k)\Delta w_k, Q_k \rangle + \frac{1}{4} \sum_{j_1, j_2, j_3, j_4=1}^d L^{j_1} g_{j_2}(y_k)^T L^{j_3} g_{j_4}(y_k)(\Delta W_k^{j_1, j_2} \Delta W_k^{j_3, j_4} - 2\delta_{j_1, j_2, j_3, j_4}\Delta), \end{aligned}$$

and from (5.3), we get

$$\sum_{j=1}^d |g_j(y_k)|^2 \leq -2y_k^T f(y_k) = -\frac{2}{\Delta\theta} [y_k^T (y_k - z_k)] \leq \frac{1}{\theta\Delta} (3|y_k|^2 + |z_k|^2). \quad (5.19)$$

This, together with (5.1), produces that there is a positive constants  $\underline{C}$  such that

$$|m_k| \leq \underline{C}[1 + |\Delta w_k|^2 + |\Delta w_k|^4]|y_k|^2 + \underline{C}[1 + |\Delta w_k|^2]|z_k|^2. \quad (5.20)$$

We therefore have  $\mathbb{E}[m_k \mathbf{1}_{[0, \lambda_l]}(k)] = 0$ . By (5.11) and  $\Delta f(y_k) = (y_k - z_k)/\theta$ ,

$$|z_{k+1}|^2 \leq \frac{(1-\theta)^2}{\theta^2} |z_k|^2 + \left[ \frac{1-2\theta}{\theta^2} - (\gamma - 0.5d^2\sigma\Delta)\Delta \right] |y_k|^2 + \frac{4\theta-2}{\theta^2} z_k^T y_k + m_k. \quad (5.21)$$

Note that  $\Delta < \frac{2\gamma}{\sigma}$  implies  $2\theta - 1 + (\gamma - 0.5d^2\sigma\Delta)\Delta\theta^2 > 0$ . Since

$$2z_k^T y_k \leq \frac{2\theta - 1 + (\gamma - 0.5d^2\sigma\Delta)\Delta\theta^2}{2\theta - 1} |y_k|^2 + \frac{2\theta - 1}{2\theta - 1 + (\gamma - 0.5d^2\sigma\Delta)\Delta\theta^2} |z_k|^2,$$

we get from (5.21)

$$\begin{aligned} |z_{k+1}|^2 &\leq \left[ \frac{(1-\theta)^2}{\theta^2} + \frac{(2\theta-1)^2}{\theta^2(2\theta-1+(\gamma-0.5d^2\sigma\Delta)\Delta\theta^2)} \right] |z_k|^2 + m_k \\ &= |z_k|^2 + \frac{(1-2\theta)(\gamma-0.5d^2\sigma\Delta)\Delta}{2\theta-1+(\gamma-0.5d^2\sigma\Delta)\Delta\theta^2} |z_k|^2 + m_k. \end{aligned} \quad (5.22)$$

By the similar techniques used in case of  $\theta \in [0, 1/2]$ , we can prove that  $\lambda_l \uparrow \infty$  a.s., as  $l \rightarrow \infty$ .

Note that (5.22) implies

$$e^{\gamma_{\Delta} k \Delta} |z_k|^2 \leq |z_0|^2 + \sum_{j=0}^{k-1} e^{\gamma_{\Delta} j \Delta} m_j,$$

where  $\gamma_{\Delta} = -\frac{1}{\Delta} \log \left( 1 - \frac{(2\theta-1)(\gamma-0.5d^2\sigma\Delta)\Delta}{2\theta-1+(\gamma-0.5d^2\sigma\Delta)\Delta\theta^2} \right)$ . Then we obtain the desired assertion by repeating the proof process of the case (i).  $\square$

**Remark 5.1.** Theorem 5.2 shows that the two theta-Milstein schemes can share the exponential mean-square stability of the exact solution. In fact, the upper bound  $\gamma$  of the decay rate can also be reproduced arbitrarily accurately for sufficiently small stepsize  $\Delta$  since  $\lim_{\Delta \rightarrow 0} \gamma_\Delta = \gamma$ .

Theorem 5.2 shows the explicit stepsize bounds  $\Delta_1^*$  and  $\Delta_2^*$  for the two classes methods to be exponentially stable with different choice of  $\theta$ . Now, it is difficult to obtain the optimal stepsize bound for nonlinear SDEs. However, for linear scalar case, the optimal stepsize bound can be deduced, which is revealed in the following.

Let us now examine the linear scalar system

$$dx(t) = \mu x(t)dt + cx(t)dw^0(t), \quad (5.23)$$

where  $\mu, c$  are constants and  $w^0(t)$  is a scalar Brownian motion. Here, we consider the single Brownian motion and since the linear combination of independent Brownian motions is still a Brownian motion. It is known that mean-square stability for (5.23) is equivalent to

$$2\mu + c^2 < 0. \quad (5.24)$$

Moreover, it is easy to observe that the two classes of Milstein schemes are equivalents for the linear SDE (5.23). Here we rewrite this scheme as follows:

$$y_{k+1} = y_k + \theta\mu\Delta y_{k+1} + (1-\theta)\mu\Delta y_k + cy_k\Delta w_k^0 + \frac{c^2}{2}y_k[|\Delta w_k^0|^2 - \Delta]. \quad (5.25)$$

For the linear scalar SDE (5.23), we have the following stability theorem.

**Theorem 5.3.** *Let condition (5.24) hold. Then theta Milstein scheme (5.25) is exponentially mean-square stable if and only if*

$$(2\mu + c^2) + \frac{1}{2}c^4\Delta + (1-2\theta)\mu^2\Delta < 0. \quad (5.26)$$

Moreover, if (5.26) holds, then theta Milstein scheme (5.25) holds

$$\mathbb{E}|y_k|^2 = e^{-\gamma_\Delta k\Delta} \mathbb{E}|x_0|^2,$$

$$\text{where } \gamma_\Delta = -\frac{1}{\Delta} \log \left( 1 + \frac{[2\mu + c^2 + (1-2\theta)\mu^2\Delta + 1/2c^4\Delta]\Delta}{(1-\theta\mu\Delta)^2} \right).$$

*Proof.* Rearranging equation (5.25) gives

$$y_{k+1} = \frac{1}{1-\theta\mu\Delta} \left( 1 + (1-\theta)\mu\Delta + c\Delta w_k^0 + \frac{c^2}{2} [|\Delta w_k^0|^2 - \Delta] \right) y_k.$$

Note that  $y_k$  is  $\mathfrak{F}_{k\Delta}$ -measurable and  $\Delta w_k^0$  is independent of  $\mathfrak{F}_{k\Delta}$ . We therefore have

$$\begin{aligned} \mathbb{E}|y_{k+1}|^2 &= \frac{[1 + (1-\theta)\mu\Delta]^2 + c^2\Delta + \frac{c^4}{2}\Delta^2}{(1-\theta\mu\Delta)^2} \mathbb{E}|y_k|^2 \\ &= \left[ 1 + \frac{[2\mu + c^2 + (1-2\theta)\mu^2\Delta + 1/2c^4\Delta]\Delta}{(1-\theta\mu\Delta)^2} \right] \mathbb{E}|y_k|^2, \end{aligned}$$

where we used  $\mathbb{E}\Delta w_k^0 = 0$ ,  $\mathbb{E}|\Delta w_k^0|^2 = \Delta$ ,  $\mathbb{E}(\Delta w_k^0)^3 = 0$  and  $\mathbb{E}|\Delta w_k^0|^4 = 3\Delta^2$ . Hence, theta Milstein scheme (5.25) is exponentially mean-square stable if and only if (5.26) holds. Moreover, if (5.26) holds, then

$$\mathbb{E}|y_{k+1}|^2 = e^{-\gamma_\Delta(k+1)\Delta}\mathbb{E}|x_0|^2,$$

and the desired assertion follows.  $\square$

Theorem 5.3 shows that (i) for  $\theta \in [0, 1/2]$ , theta Milstein scheme (5.25) shares the exponential mean-square stability of the exact solution if and only if the stepsize  $\Delta < \Delta^* := \frac{-2\mu - c^2}{0.5c^4 + (1-2\theta)\mu^2}$ . (ii) For  $\theta \in (1/2, 1]$ , if  $\mu^2 < \frac{c^4}{2(2\theta-1)}$  (the diffusion term plays a crucial role), then for  $\Delta < \Delta^*$ , theta Milstein scheme (5.25) is exponentially mean-square stable, and if  $\mu^2 \geq \frac{c^4}{2(2\theta-1)}$  (the drift term plays a crucial role), then theta Milstein scheme (5.25) is exponentially mean-square stable unconditionally. These results are coincident with Theorem 2.1 in [38]. Theorem 5.3 also presents the exponential decay rate  $\gamma_\Delta$  of theta-Milstein scheme

## 6 Simulations

In this section, we introduce some numerical experiments to illustrate our theoretical results. The STM method has been well studied under the global Lipschitz conditions. Here, we are concerned about the SDEs with non-global Lipschitz condition. Let us consider the scalar stochastic differential equation:

$$dx(t) = (\mu x(t) - x(t)^3)dt + \sum_{j=1}^2 \sigma_j \sin(x(t))dw^j(t), t > 0 \quad (6.1)$$

with the initial data  $x_0 = 0.25$ , where  $\mu, \sigma_1, \sigma_2$  are constants. Then the drift  $f(x) = \mu x - x^3$  does not satisfy the global Lipschitz condition, but obeys the one-sided Lipschitz condition (2.2). It is easy to see that the diffusion  $g_i(x) = \sigma_i \sin(x)$  satisfies the Lipschitz condition (2.3).

**Convergence Analysis:** We first investigate the convergence of the theta-Milstein methods for the case  $\mu = \sigma_1 = 1, \sigma_2 = 2$ . By Theorem 4.1, we know that SSTM and STM with  $\theta \in (1/2, 1]$  are convergent with the order one.

Now, we resort to MATLAB and follow the algorithms in [41] to make the numerical experiments. We compute  $10^4$  different discretized Brownian paths over  $[0, 1]$  with  $\Delta_0 = 2^{-13}$ . For each path, SSTM and STM with  $\theta = 0.7$  are applied with 5 different stepsizes:  $\Delta = 2^p \Delta_0$  for  $p = 0, 4, \dots, 7$ . Figure 1 depicts the root mean-square errors as a function of the stepsize  $\Delta$  in log-log plot, where the expectation is approximated by the mean of  $10^4$  independent realizations. As expected, the two classes of Milstein methods give the errors that decreases proportional to  $\Delta^1$ . From the numerical experiments, we obtain the simulated convergence orders about 1.0756 for STM and 1.0244 for SSTM with the least squares residuals ([41]) 0.0198 for SSM and 0.0081 for SSTM. In fact, we conduct the experiments many times and find that the least squares residual of SSTM method is smaller than that of STM method.

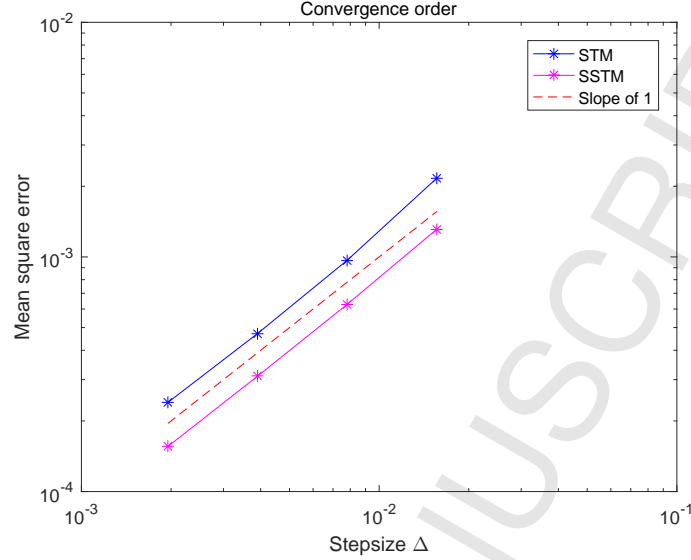


Figure 1: Mean-square approximation error versus stepsize  $\Delta$  to approximate (6.1)

**Stability Analysis:** Then we investigate the numerical stability of the theta Milstein schemes for SDE (6.1) with  $\mu = -2, \sigma_1 = \sigma_2 = 1/\sqrt{2}$ . It is easy to deduce from Theorem 5.1 that the solution (6.1) is exponentially mean-square stable with the Lyapunov exponent not greater than  $-3$ , since

$$2x^T f(x) + \sum_{i=1}^2 |g_i(x)|^2 = -4x^2 - 2x^4 + x^2 \leq -3x^2.$$

We test the stability of SSTM and STM schemes with different stepsizes. We take  $\Delta = 2, 1$  and  $1/2$  ( $\Delta < \Delta_2^* = 3$ ) and generate  $10^4$  sample paths for each numerical method. Under the choices of  $\theta = 0.6$  and  $\theta = 0.9$ , the mean-square stability of SSTM and STM are plotted in Figure 2 and Figure 3, respectively. The two figures show that the numerical schemes are mean-square asymptotically stable under the stepsizes  $\Delta = 2, 1$  and  $1/2$ , but SSTM method decay faster than STM method under the same stepsizes.

## 7 Conclusion

The work examines the convergence and stability of SSTM proposed in current work and STM for approximating SDEs with non-Lipschitz continuous coefficients. By establishing the uniform boundedness estimation of  $p$ th moments, we firstly prove that under some appropriate conditions, all the two classes methods are strongly convergent with the order one. Under a coupled condition, we show that SSTM and STM can reproduce the exponential mean-square stability of the exact solution. The numerical experiments confirm the theoretical results and reveal some interesting phenomena that SSTM method may work better than STM method since SSTM method has smaller least squares residual and decays faster than STM method.

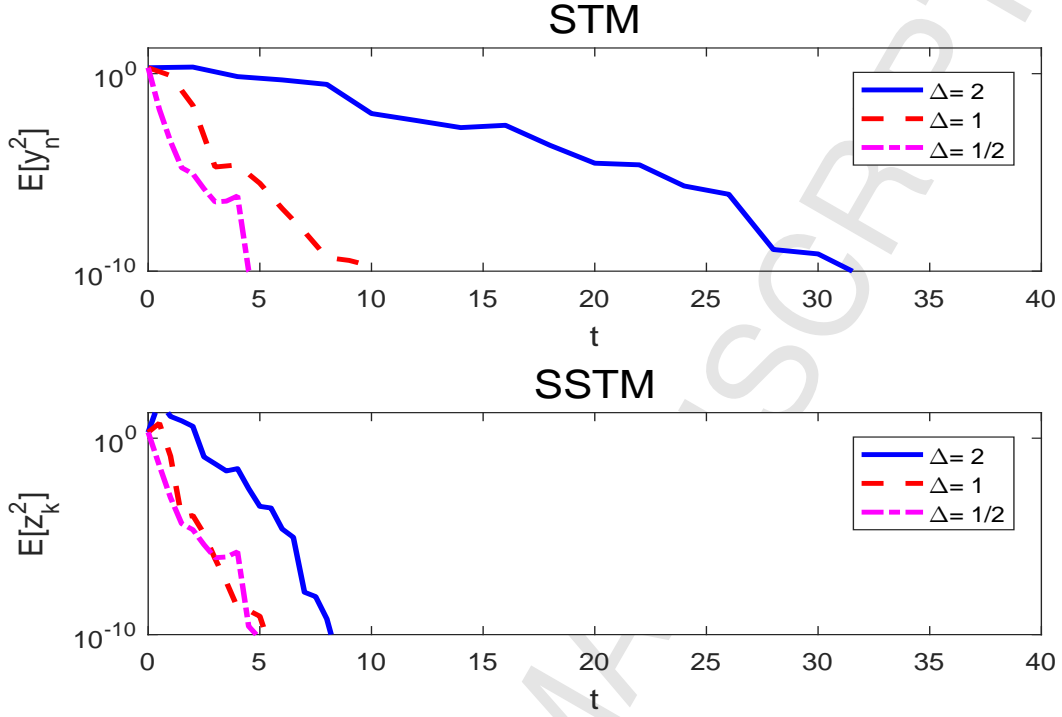


Figure 2: Simulation of  $\mathbb{E}|x(t)|^2$ , using STM and SSTM with different stepsizes under  $\theta = 0.6$

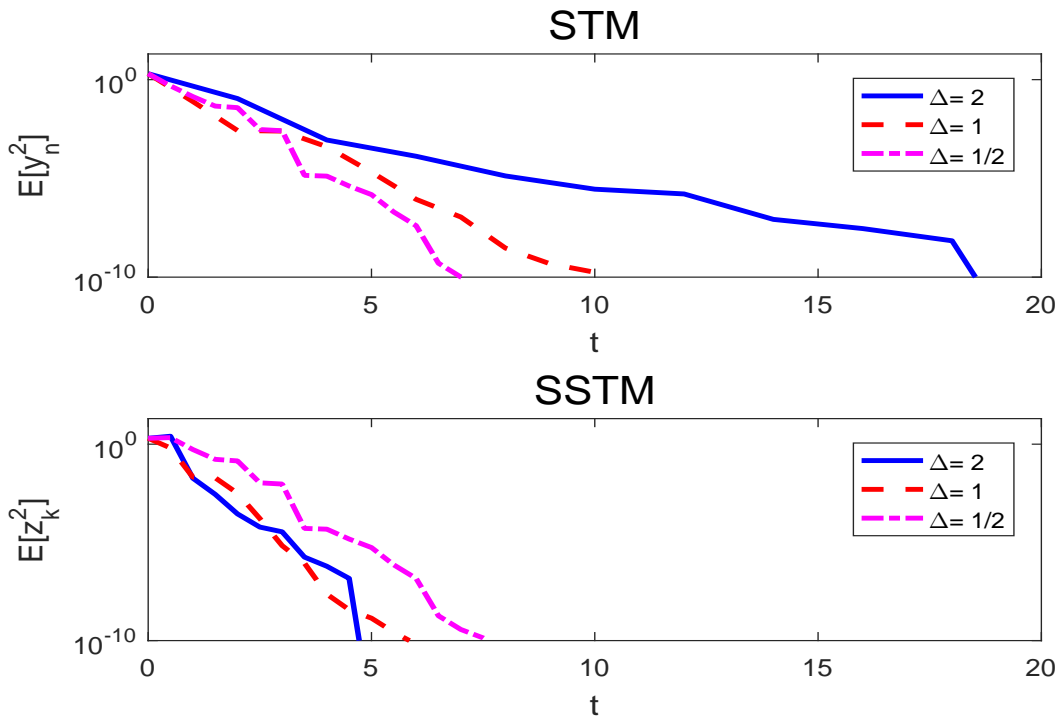


Figure 3: Simulation of  $\mathbb{E}|x(t)|^2$ , using STM and SSTM with different stepsizes under  $\theta = 0.9$

In this paper, the diffusions are assumed to be Lipschitz. In the future, we hope to discuss more complex stochastic differential equations without such Lipschitz condition on diffusion or with general monotone conditions, and extend current work to the delayed version.

## References

- [1] P. E. Kloeden, E. Platen, Numerical solution of stochastic differential equations, Vol. 2 of Applications of Mathematics (New York), Springer-Verlag, Berlin, 1992.
- [2] G. N. Milstein, Numerical Integration of Stochastic Differential Equations., Vol. 313 of Mathematics and Its Applications, Kluwer Academic, Dordrecht, 1995.
- [3] A Rathinasamy, K. Balachandran Mean-square stability of second-order RungeCKutta methods for multi-dimensional linear stochastic differential systems[J]. Journal of Computational and Applied Mathematics, 219 (1) (2008) 170–197.
- [4] G. N. Milstein, M. V. Tretyakov, Stochastic Numerics for Mathematical Physics, Scientific Computation, Springer, New York, 2004.
- [5] L. Arnold, Random dynamical systems, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 1998.
- [6] M. Hutzenthaler, A. Jentzen, Convergence of the stochastic Euler scheme for locally Lipschitz coefficients, Foundations of Computational Mathematics 11 (6) (2011) 657–706.
- [7] B. Schmalfuss, The random attractor of the stochastic Lorenz system, Zeitschrift für angewandte Mathematik und Physik ZAMP 48 (1997) 951–975.
- [8] M. Arato, A famous nonlinear stochastic equation (Lotka-Volterra model with diffusion), Mathematical and Computer Modelling, 38 (7) (2003) 709–726.
- [9] M. Hutzenthaler, A. Jentzen, Non-globally Lipschitz counterexamples for the stochastic Euler scheme, arXiv:0905.0273v1.
- [10] M. Hutzenthaler, A. Jentzen, P. E. Kloeden, Strong and weak divergence in finite time of Eulers method for stochastic differential equations with non-globally Lipschitz continuous coefficients, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences 467 (2011) 1563–1576.
- [11] Y. Hu, Semi-implicit Euler-Maruyama scheme for stiff stochastic equations, in: H. Körezlioğlu, B. Øksendal, A. Üstünel (Eds.), Stochastic Analysis and Related Topics V, Vol. 38 of Progress in Probability, Birkhäuser Boston, 1996, pp. 183–202.
- [12] D. J. Higham, X. Mao, A. M. Stuart, Strong convergence of Euler-type methods for nonlinear stochastic differential equations, SIAM Journal on Numerical Analysis 40 (2) (2002) 1041–1063.

- [13] D. J. Higham, P. E. Kloeden, Strong convergence rates for backward Euler on a class of non-linear jump-diffusion problems, *Journal of Computational and Applied Mathematics* 205 (2) (2007) 949–956.
- [14] A. F. Bastani, M. Tahmasebi, Strong convergence of split-step backward Euler method for stochastic differential equations with non-smooth drift, *Journal of Computational and Applied Mathematics* 236 (7) (2012) 1903–1918.
- [15] X. Mao, L. Szpruch, Strong convergence rates for backward Euler-Maruyama method for non-linear dissipative-type stochastic differential equations with super-linear diffusion coefficients, *Stochastics An International Journal of Probability and Stochastic Processes* 85 (1) (2013) 144–171.
- [16] L. Szpruch, X. Mao, Strong convergence of numerical methods for nonlinear stochastic differential equations under monotone conditions, working paper.
- [17] X. Mao, L. Szpruch, Strong convergence and stability of implicit numerical methods for stochastic differential equations with non-globally Lipschitz continuous coefficients, *Journal of Computational and Applied Mathematics* 238 (0) (2013) 14–28.
- [18] M. Hutzenthaler, A. Jentzen, On a perturbation theory and on strong convergence rates for stochastic ordinary and partial differential equations with non-globally monotone coefficients. *arXiv:1401.0295*, 2014.
- [19] M. Hutzenthaler, A. Jentzen, and P. E. Kloeden, Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients, *Annals of Applied Probability* 22(4) (2012) 1611–1641.
- [20] S. Sabanis, Euler approximations with varying coefficients: the case of superlinearly growing diffusion coefficients. *arXiv:1308.1796*, 2013.
- [21] X. Zong, F. Wu, C. Huang, Convergence and stability of the semi-tamed Euler scheme for stochastic differential equations with non-Lipschitz continuous coefficients, *Applied Mathematics and Computation* 228 (2014) 240–250.
- [22] Y. Hu, S.-E. A. Mohammed, F. Yan, Discrete-time approximations of stochastic delay equations: The Milstein scheme, *The Annals of Probability* 32 (1) (2004) 265–314.
- [23] T. Tian, K. Burrage, Implicit Taylor methods for stiff stochastic differential equations, *Applied Numerical Mathematics* 38 (1-2) (2001) 167–185.
- [24] X. Wang, S. Gan, D. Wang, A family of fully implicit Milstein methods for stiff stochastic differential equations with multiplicative noise, *BIT Numerical Mathematics* 52 (3) (2012) 741–772.

- [25] J. Alcock, K. Burrage, Stable strong order 1.0 schemes for solving stochastic ordinary differential equations, *BIT Numerical Mathematics* 52 (3) (2012) 539–557.
- [26] X. Wang, S. Gan, The tamed Milstein method for commutative stochastic differential equations with non-globally lipschitz continuous coefficients, *Journal of Difference Equations and Applications* 19 (3) (2013) 466–490.
- [27] D. Higham, X. Mao, L. Szpruch, Convergence, non-negativity and stability of a new Milstein scheme with applications to finance, *Discrete and Continuous Dynamical Systems - Series B* 18 (8) (2013) 2083–2100.
- [28] D. J. Higham, Mean-square and asymptotic stability of the stochastic theta method, *SIAM Journal on Numerical Analysis* 38 (2000) 753–769.
- [29] Y. Saito, T. Mitsui, Stability analysis of numerical schemes for stochastic differential equations, *SIAM Journal on Numerical Analysis* 33 (6) (1996) 2254–2267.
- [30] Y. Saito, T. Mitsui, Mean-square stability of numerical schemes for stochastic differential systems, *Vietnam Journal of Mathematics* 30 (2002) 551–560.
- [31] D. J. Higham, X. Mao, A. Stuart, Exponential mean-square stability of numerical solutions to stochastic differential equations, *LMS Journal of Computation and Mathematics* 6 (2003) 297–313.
- [32] C. Huang, Exponential mean square stability of numerical methods for systems of stochastic differential equations, *Journal of Computational and Applied Mathematics* 236 (2012) 4016–4026.
- [33] X. Zong, F. Wu, Choice of  $\theta$  and mean-square exponential stability in the stochastic theta method of stochastic differential equations, *Journal of Computational and Applied Mathematics* 255 (0) (2014) 837–847.
- [34] X. Zong, F. Wu, C. Huang, Preserving exponential mean square stability and decay rates in two classes of theta approximations of stochastic differential equations, *Journal of Difference Equations and Applications* 20 (7) (2014) 1091–1111.
- [35] X. Zong, F. Wu, C. Huang, Theta schemes for SDDEs with non-globally lipschitz continuous coefficients. *Journal of Computational and Applied Mathematics*, 278(1) (2015) 258–277.
- [36] X. Mao, C. Yuan, *Stochastic Differential Equations with Markovian Switching*, Imperial College Press, London, 2006.
- [37] X. Mao, *Stochastic Differential Equations and their Applications*, Horwood Publishing Series in Mathematics & Applications, Horwood Publishing Limited, Chichester, 1997.

- [38] D. J. Higham, A-stability and stochastic mean-square stability, *BIT Numerical Mathematics* 40 (2) (2000) 404–409.
- [39] A. Neuenkirch, L. Szpruch, First order strong approximations of scalar SDEs defined in a domain, *Numerische Mathematik* 128 (2014) 103–136.
- [40] P. E. Kloeden, A. Neuenkirch, Convergence of numerical methods for stochastic differential equations in mathematical finance, in “Recent Developments in Computational Finance: Foundations, Algorithms and Applications”, T. Gerstner and P.E. Kloeden (Editors), *Interdisciplinary Mathematical Sciences Series*, Vol. 14, World Scientific Publishing Co. Inc, Singapur, 2013; pp. 49-80.
- [41] D. J. Higham, An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Review* 43 (3) (2001) 525–546.