

Truncated Milstein method for non-autonomous stochastic differential equations and its modification

Juan Liao Wei Liu* Xiaoyan Wang

Department of Mathematics

Shanghai Normal University, Shanghai, 200234, China

Abstract

The truncated Milstein method, which was initially proposed in (Guo, Liu, Mao and Yue 2018), is extended to the non-autonomous stochastic differential equations with the super-linear state variable and the Hölder continuous time variable. The convergence rate is proved. Compared with the initial work, the requirements on the step-size is also significantly released. In addition, the technique of the randomized step-size is employed to raise the convergence rate of the truncated Milstein method.

Key words: non-autonomous stochastic differential equations, truncated Milstein method, randomized step-size, super-linear state variable, Hölder continuous time variable.

1 Introduction

Numerical methods for stochastic differential equations (SDEs) with super-linear coefficients have been attracting lots of attention in recent years. Due to that the classical Euler-Maruyama (EM) method fails to converge for those types of SDEs [16], different new methods have been proposed.

Implicit methods are natural alternatives, since they have been successful in handling the stiffness in ordinary differential equations (ODEs) [11]. The implicit methods of the Euler's type for SDEs were studied in [2, 15, 28, 34, 35]. The Milstein-type implicit methods for SDEs were discussed in [13, 20, 31, 39]. The multi-stage implicit methods were investigated in [1, 3, 5]. We just mention some

*Corresponding author, Email: weiliu@shnu.edu.cn; lwbvb@hotmail.com

of the works on implicit methods here and refer the readers to the references therein.

Compared with implicit methods, explicit methods still have their advantages, such as simple structures, easy to implement, cheap to simulate large numbers of paths [12]. Recently, many different explicit methods have been proposed to approximate SDEs with super-linear coefficients. The tamed Euler method was initially proposed in [17]. By using the idea of taming the coefficients, different types of tamed methods have been proposed [7, 8, 29, 32, 33, 36, 38]. The truncated EM method is another modification of the classical EM, which was initialized in [26, 27]. Afterwards, the truncating technique has been employed to develop different kinds of truncated methods [6, 9, 18, 22, 23, 37].

Most of the works mentioned above dealt with autonomous SDEs, where the time variable does not appear explicitly in the coefficients. Meanwhile, it is well-known that the non-smoothness of the time variable in non-autonomous SDEs leads to significant difference in the convergence rate of numerical methods for both ODEs [4, 19] and SDEs [21, 24].

Motivated by all the issues mentioned above, we investigate the truncated Milstein method for non-autonomous SDEs with the time variable satisfying Hölder's continuity and the state variable containing super-linear terms. The finite time convergence of the proposed method is proved and the convergence rate is discussed. This result could be regarded as an extension of [10], where autonomous SDEs were considered. We also propose the randomized truncated Milstein method to overcome the low convergence rate due to the Hölder continuous time variable.

The main contribution of this paper are twofold.

- Compared with the existing work [10], our results cover the non-autonomous case and release the requirements on the step-size significantly.
- The randomized truncated Milstein method is proposed. By using numerical simulations, this new method is demonstrated to outperform the truncated Milstein method for non-autonomous SDEs.

This paper is constructed in the following way. Notations, assumptions and the structure of the numerical methods are presented in Section 2. Section 3 contains the main results and their proofs. The randomized truncated Milstein method is proposed in Section 4. Numerical simulations are conducted in Section 5. Section 6 sees the conclusion and some discussions on the future research.

2 Mathematical preliminary

Notations, assumptions and the truncated Milstein method for non-autonomous SDEs are introduced in the section.

Through out this paper, unless otherwise specified, we let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t \geq 0}$ satisfying the usual condition (that is, it is right continuous and increasing while \mathcal{F}_0 contains all p-null sets). Let $B(t)$ be an one-dimensional Brownian motion defined in the probability space and is \mathcal{F}_t -adopted. And let $|\cdot|$ denote both the Euclidean norm in \mathbb{R}^n and the trace norm in $\mathbb{R}^{n \times m}$; Moreover, for two real numbers a and b, we use $a \vee b = \max(a, b)$ and $a \wedge b = \min(a, b)$. For a given set G, its indicator function is denoted by I_G , namely $I_G(x) = 1$ if $x \in G$ and 0 otherwise.

We are concerned with the d-dimension SDEs

$$dy(t) = \mu(t, y(t))dt + \sigma(t, y(t))dB(t) \quad (2.1)$$

on $t \in [t_0, T]$ for any $T > t_0$ with the initial value $y(t_0) = y_0 \in \mathbb{R}^d$, where the drift coefficient function $\mu : [t_0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ and the diffusion coefficient function $\sigma : [t_0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$, and $y(t) = (y^1(t), y^2(t), \dots, y^d(t))^T$.

We define:

$$L\sigma(t, y) = \sum_{l=1}^d \sigma^l(t, y) \frac{\partial \sigma(t, y)}{\partial y^l},$$

where $\sigma = (\sigma^1, \sigma^2, \dots, \sigma^d)^T$, $\sigma^l : \mathbb{R}^d \rightarrow \mathbb{R}$.

And define the derivative of vector $\sigma(t, y)$ with respect to y^l by

$$G^l(t, y) := \left(\frac{\partial \sigma^1(t, y)}{\partial y^l}, \frac{\partial \sigma^2(t, y)}{\partial y^l}, \dots, \frac{\partial \sigma^d(t, y)}{\partial y^l} \right).$$

Moreover, we assume that both σ and μ have a second-order derivative, and we make the following assumptions.

Assumption 2.1. *There exist constants $C_1 > 0$, $\beta > 0$ and $\alpha \in (0, 1]$ such that*

$$|\mu(t, x) - \mu(t, y)| \vee |\sigma(t, x) - \sigma(t, y)| \vee |L\sigma(t, x) - L\sigma(t, y)| \leq C_1(1 + |x|^\beta + |y|^\beta)|x - y|,$$

$$|\mu(t_1, y) - \mu(t_2, y)| \vee |\sigma(t_1, y) - \sigma(t_2, y)| \leq C_4(1 + |y|^{\beta+1})|t_1 - t_2|^\alpha,$$

for all $x, y \in \mathbb{R}^d$, any $t \in (t_0, T]$.

Assumption 2.2. *There exist constants $q \geq 2$ and $C_2 > 0$ such that*

$$\langle x - y, \mu(t, x) - \mu(t, y) \rangle + (q - 1)|\sigma(t, x) - \sigma(t, y)|^2 \leq C_2|x - y|^2,$$

for all $x, y \in \mathbb{R}^d$, any $t \in (t_0, T]$.

Assumption 2.3. *There exist constants $p \geq 2$ and $C_3 > 0$ such that*

$$\langle y, \mu(t, y) \rangle + (p - 1)|\sigma(t, y)|^2 \leq C_3(1 + |y|^2),$$

where C_3 does not depend on t or y , for all $t \in [t_0, T]$, any $y \in \mathbb{R}^d$.

Assumptions 2.1 and 2.2 guarantee a unique global solution of SDE (2.1). In addition, we can derive the boundedness of the moment of the true solution from Assumption 2.2 which is proved in [25], that is, there exists a constant M_1 , which is dependent on t and q , such that

$$E|y(t)|^q \leq M_1(1 + |y(0)|^q). \quad (2.2)$$

And it can be observed from Assumption 2.1 that for all $y \in \mathbb{R}^d$ and $t \in [t_0, T]$

$$|\mu(t, y)| \vee |\sigma(t, y)| \vee |L\sigma(t, y)| \leq M_2(1 + |y|^{\beta+1}) \quad (2.3)$$

where M_2 depends on C_1 and $\sup_{t_0 \leq t \leq T} (|\mu(t, 0)| + |\sigma(t, 0)|)$.

We further assume that there exists a positive constant M_3 such that

$$\left| \frac{\partial \mu(t, y)}{\partial y} \right| \vee \left| \frac{\partial^2 \mu(t, y)}{\partial y^2} \right| \vee \left| \frac{\partial \sigma(t, y)}{\partial y} \right| \vee \left| \frac{\partial^2 \sigma(t, y)}{\partial y^2} \right| \leq M_3(1 + |y|^{\beta+1}), \quad (2.4)$$

for all $y \in \mathbb{R}^d$ and $t \in [t_0, T]$.

To make the paper self-contained, let us revisit the truncated Milstein method. Firstly, we choose a strictly increasing function $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ which is the set of all non-negative real numbers such that $f(u) \rightarrow \infty$ as $u \rightarrow \infty$ and

$$\sup_{t \in [t_0, T]} \sup_{|y| \leq u} (|\mu(t, y)| \vee |\sigma(t, y)| \vee |G^l(t, y)|) \leq f(u), \quad \forall u \geq 1, \quad l = 1, 2, \dots, d. \quad (2.5)$$

Then we use f^{-1} denote the inverse function of f . we can easily observe that f^{-1} is also a strictly increasing continuous function from $[f(1), \infty)$ to \mathbb{R}_+ . We choose a constant $\hat{h} \geq 1 \vee f(1)$ and a strictly decreasing function $h : (0, 1] \rightarrow [f(1), \infty)$ such that

$$\lim_{\Delta \rightarrow 0} h(\Delta) = \infty \quad \text{and} \quad \Delta^{\frac{1}{4}} h(\Delta) \leq \hat{h}, \quad \forall \Delta \in (0, 1]. \quad (2.6)$$

For a given step-size $\Delta \in (0, 1]$, any $t \in [t_0, T]$ and all $y \in \mathbb{R}^d$, define the truncated functions by

$$\sigma_\Delta(t, y) = \sigma \left(t, (|y| \wedge f^{-1}(h(\Delta))) \frac{y}{|y|} \right), \quad (2.7)$$

$$\mu_\Delta(t, y) = \mu \left(t, (|y| \wedge f^{-1}(h(\Delta))) \frac{y}{|y|} \right), \quad (2.8)$$

$$G_{\Delta}^l(t, y) = G^l \left(t, (|y| \wedge f^{-1}(h(\Delta))) \frac{y}{|y|} \right), \quad l = 1, 2, \dots, d, \quad (2.9)$$

where we set $y/|y| = 0$ if $y = 0$.

It is clear that for all $t \in [t_0, T]$, any $y \in \mathbb{R}^d$ and $l = 1, 2, \dots, d$,

$$|\sigma_{\Delta}(t, y)| \vee |\mu_{\Delta}(t, y)| \vee |G_{\Delta}^l(t, y)| \leq f(f^{-1}(h(\Delta))) = h(\Delta). \quad (2.10)$$

We can also obtain the fact that there exists a positive constant M such that

$$\left| \frac{\partial \mu_{\Delta}(t, y)}{\partial y} \right| \vee \left| \frac{\partial^2 \mu_{\Delta}(t, y)}{\partial y^2} \right| \vee \left| \frac{\partial \sigma_{\Delta}(t, y)}{\partial y} \right| \vee \left| \frac{\partial^2 \sigma_{\Delta}(t, y)}{\partial y^2} \right| \leq M, \quad (2.11)$$

for any $t \in [t_0, T]$ and $y \in \mathbb{R}^d$.

Remark 2.4. *Since the truncated functions, $\mu_{\Delta}(t, y)$, $\sigma_{\Delta}(t, y)$ and G_{Δ}^l , are not differential at points, $f^{-1}(h(\Delta))$ and $-f^{-1}(h(\Delta))$, here we mean the derivatives of the truncated functions at $f^{-1}(h(\Delta))$ by their left derivatives and the derivatives at $-f^{-1}(h(\Delta))$ by their right derivatives.*

The discrete-time truncated Milstein numerical solution X_i , to approximate $y(t_i)$ for $t_i = i\Delta + t_0$, are formed by setting $X_0 = y_0$ and computing

$$X_{i+1} = X_i + \mu_{\Delta}(t_i, X_i)\Delta + \sigma_{\Delta}(t_i, X_i)\Delta B_i + \frac{1}{2} \sum_{l=1}^d \sigma_{\Delta}^l(t_i, X_i)G_{\Delta}^l(t_i, X_i)(\Delta B_i^2 - \Delta), \quad (2.12)$$

for $i = 0, 1, \dots, N$, where N is the integer part of T/Δ and let $t_{N+1} = T$ while $\Delta B_i = B(t_{i+1}) - B(t_i)$.

To simplify the notation, we set

$$L\sigma_{\Delta}(t, X_i) := \sum_{l=1}^d \sigma_{\Delta}^l(t, X_i)G_{\Delta}^l(t, X_i),$$

for all $t \in [t_0, T]$ and $l \in \{1, 2, \dots, d\}$.

The continuous version of the truncated Milstein method is defined by

$$\begin{aligned} X_{\Delta}(t) = & \bar{X}_{\Delta}(t) + \int_{t_i}^t \mu_{\Delta}(\kappa(s), \bar{X}_{\Delta}(s)) ds + \int_{t_i}^t \sigma_{\Delta}(\kappa(s), \bar{X}_{\Delta}(s)) dB(s) \\ & + \int_{t_i}^t L\sigma_{\Delta}(\kappa(s), \bar{X}_{\Delta}(s)) \Delta B(s) dB(s), \end{aligned} \quad (2.13)$$

where $\bar{X}_\Delta(t)$ is a piecewise constant solution such that $\bar{X}_\Delta(t) = \bar{X}_\Delta(t_i) = X_\Delta(t_i) = X_i$, for $t_i \leq t < t_{i+1}$ and we define $\bar{X}_\Delta(T) = X_\Delta(T)$, $\Delta B(s) = \sum_{i=0}^N I_{\{t_i \leq s < t_{i+1}\}}(B(s) - B(t_i))$ and $\kappa(s) = t_i I_{\{t_i \leq s < t_{i+1}\}}$.

We need the following version of the Taylor expansion.

If a function $\phi : \mathbb{R}_+ \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ is third-order continuous differentiable, by the Taylor formula we have:

$$\phi(\kappa(t), x) - \phi(\kappa(t), x^*) = \phi'(\kappa(t), x) \Big|_{x=x^*} (x - x^*) + R_\phi(t, x, x^*), \quad (2.14)$$

where $R_\phi(t, x, x^*) = \int_0^1 (1 - \tau) \phi''(\kappa(t), x) \Big|_{x=x^* + \tau(x-x^*)} (x - x^*, x - x^*) d\tau$, for any fixed $t \in [t_0, T]$.

For any $y, j_1, j_2 \in \mathbb{R}^d$, the expressions of the derivatives are as follows:

$$\phi'(\kappa(t), y)(j_1) = \sum_{i=1}^d \frac{\partial \phi}{\partial y^i} j_1^i, \quad \phi'(\kappa(t), y)(j_1, j_2) = \sum_{i,j=1}^d \frac{\partial^2 \phi}{\partial y^i \partial y^j} j_1^i j_2^j,$$

where $\frac{\partial \phi}{\partial y^i} = \left(\frac{\partial \phi_1}{\partial y^i}, \dots, \frac{\partial \phi_d}{\partial y^i} \right)$ and $\phi = (\phi_1, \phi_2, \dots, \phi_d)$, for any fixed $t \in [t_0, T]$.

If we replace x and x^* by $X_\Delta(t)$ and $\bar{X}_\Delta(t)$ respectively from (2.14), for any fixed $t \in [t_0, T]$, we have:

$$\begin{aligned} \phi(\kappa(t), X_\Delta(t)) - \phi(\kappa(t), \bar{X}_\Delta(t)) &= \phi'(\kappa(t), x) \Big|_{x=\bar{X}_\Delta(t)} \int_{t_i}^t \sigma_\Delta(\kappa(s), \bar{X}_\Delta(s)) dB(s) \\ &\quad + \tilde{R}_\phi(t, X_\Delta(t), \bar{X}_\Delta(t)). \end{aligned} \quad (2.15)$$

Here,

$$\begin{aligned} \tilde{R}_\phi(t, X_\Delta(t), \bar{X}_\Delta(t)) &= \phi'(\kappa(t), x) \Big|_{x=\bar{X}_\Delta(t)} \left(\int_{t_i}^t \mu_\Delta(\kappa(s), \bar{X}_\Delta(s)) ds \right. \\ &\quad \left. + \int_{t_i}^t L\sigma_\Delta(\kappa(s), \bar{X}_\Delta(s)) \Delta B(s) dB(s) \right) + R_\phi(t, X_\Delta(t), \bar{X}_\Delta(t)). \end{aligned} \quad (2.16)$$

Thus, replacing ϕ by σ_Δ from (2.15), we obtain

$$R_{\sigma_\Delta}(t, X_\Delta(t), \bar{X}_\Delta(t)) = \sigma_\Delta(\kappa(t), X_\Delta(t)) - \sigma_\Delta(\kappa(t), \bar{X}_\Delta(t)) - L\sigma_\Delta(\kappa(t), \bar{X}_\Delta(t)) \Delta B(t), \quad (2.17)$$

where $\Delta B(t) = B(t) - B(t_i)$, for $t_i \leq t < t_{i+1}$.

3 Main results

This section is divided into three parts. The main theorems of this paper and the comparison with the existing result are presented in Section 3.1. Important

lemmas are proved in Section 3.2. The proofs of the main theorems are postponed to Section 3.3.

3.1 Main theorem

Theorem 3.1. *Let Assumptions 2.1, 2.2 and 2.3 hold and assume that $p > 2(1 + \beta)q$, then for any $\bar{q} \in [2, q)$ and $\Delta \in (0, 1]$, there exists a constant H such that*

$$\sup_{t_0 \leq t \leq T} \mathbb{E}|y(t) - X_\Delta(t)|^{\bar{q}} \leq H \left(\Delta^{\alpha \bar{q}} + \Delta^{\bar{q}} (h(\Delta))^{2\bar{q}} + \left(f^{-1}(h(\Delta)) \right)^{(\beta+1)\bar{q}-p} \right)$$

To see the convergence rate more clearly, we strength the requirement Assumption 2.3 and obtained the following result.

Theorem 3.2. *Let Assumptions 2.1 and 2.2 hold, and Assumption 2.3 hold for any $p > 2(\beta + 1)q$. Then for any $\bar{q} \in [2, q)$, $\varepsilon \in (0, 1/4]$ and $\Delta \in (0, 1]$, there exists a constant H such that*

$$\sup_{t_0 \leq t \leq T} \mathbb{E}|y(t) - X_\Delta(t)|^{\bar{q}} \leq H \left(\Delta^{\min(1-2\varepsilon, \alpha)\bar{q}} \right)$$

Proof. First we define $f(u) = H_4 u^{\beta+2}$, $\forall u \geq 1$. It is easy to get

$$f^{-1}(u) = \left(\frac{u}{H_4} \right)^{\frac{1}{\beta+2}}$$

Then, let

$$h(\Delta) = \Delta^{-\varepsilon}$$

for some $\varepsilon \in (0, 1/4]$ and $\hat{h} > (1 \vee f(1))$.

Applying Theorem 3.1, we can see that

$$\sup_{t_0 \leq t \leq T} \mathbb{E}|y(t) - X_\Delta(t)|^{\bar{q}} \leq H \left(\Delta^{\min\left(\frac{\varepsilon(p-(\beta+1)\bar{q})}{\beta+2}, \alpha\bar{q}, (1-2\varepsilon)\bar{q}\right)} \right).$$

We can get the desired assertions easily by choosing a sufficiently large p . ■

To explain the improvement of the main theorem of this paper, we recall Theorem 3.7 in [10] as follows.

Theorem 3.3. *Let Assumptions 2.1, 2.2 and (2.4) hold. Furthermore, assume that for any given $p \geq 2$, there exists a $q \in (p, \infty)$. In addition, if*

$$h(\Delta) \geq f \left(\left(\Delta^{\frac{p}{2}} (h(\Delta))^p \right)^{-1/(q-\frac{p}{2})} \right) \quad (3.1)$$

holds for all sufficiently small $\Delta \in (0, 1]$, then for any fixed $T \geq t_0$

$$\mathbb{E}|y(T) - X_{N+1}|^p \leq K \Delta^p (h(\Delta))^{2p} \quad (3.2)$$

holds, where K is a positive constant independent of Δ .

Remark 3.4. *Let us demonstrate that compared with the main result in [10] our result in this paper releases the constraint on the step-size.*

Consider the scalar SDE

$$dy(t) = (y(t) - 2y^5(t)) dt + y^2(t)dB(t), \quad t \geq t_0, \quad (3.3)$$

with $t_0 = 0$ and the initial value $y(t_0) = 1$.

Due to the fact that

$$\sup_{|x| \leq u} (|\mu(x)| \vee |\sigma(x)| \vee |L\sigma(x)|) \leq 3u^5, \quad \forall u \geq 1,$$

so we choose $f(u) = 3u^5$ and define $h(u) = u^{-\varepsilon}$ for $\varepsilon \in (0, 1/4]$.

Choose $\varepsilon = 1/4$, then condition (3.1) is satisfied with $p = 1$, $q = 12$ and $\Delta \leq 10^{-21}$. By Theorem 3.3 (Theorem 3.7 in [10]), we can conclude that

$$\mathbb{E}|y(T) - X_{N+1}|^{2p} \leq K\Delta^p,$$

this is to say that the convergence rate is $1/2$.

However, take ε to be $1/4$ and choose p sufficiently large, it can be derived from Theorem 3.2 that

$$\mathbb{E}|y(T) - X_\Delta(T)|^{\bar{q}} \leq K\Delta^{\frac{1}{2}\bar{q}},$$

which means that the convergence rate is $1/2$. It should be noted that we do not put the constraint $\Delta \leq 10^{-21}$ here.

Therefore, compared with the main theorem in [10], the strong requirement on the step-size, $\Delta \leq 10^{-21}$, is not needed for our main result, which shows that our result releases the requirement on the step-size.

Remark 3.5. *Theorem 3.2 tells us that the order of convergence of the truncated Milstein method is $\min(1 - 2\varepsilon, \alpha)$. If α is close to 1, the convergence rate will not very different from that of the traditional truncated Milstein method. Conversely, if α is equal to $1/4$ or more less than 1, then the order of convergence will worse than the traditional sense. This shows that the Hölder-continuous time variable does affect the order of convergence dramatically.*

3.2 Important lemmas

The next lemma shows that the truncated coefficients inherit the same inequality of the original coefficients.

Lemma 3.6. *Assume that Assumption 2.3 holds, then for all $\Delta \in (0, 1]$, there exists a positive constant C_5 such that*

$$\langle y, \mu_\Delta(t, y) \rangle + (p - 1)|\sigma_\Delta(t, y)|^2 \leq C_5(1 + |y|^2), \quad \forall y \in \mathbb{R}^d.$$

The proof of this lemma follows the same idea used in [the proof of Lemma 2.5](#) in [14].

The following lemma shows the difference between the discrete and continuous versions of the truncated Milstein method in the moment sense.

Lemma 3.7. *For any $\Delta \in (0, 1]$, $t \in [t_0, T]$ and all $p \geq 2$*

$$\mathbb{E}|X_\Delta(t) - \bar{X}_\Delta(t)|^p \leq C\Delta^{\frac{p}{2}}(h(\Delta))^p,$$

where C is a constant independent of Δ , consequently,

$$\lim_{\Delta \rightarrow 0} \mathbb{E}|X_\Delta(t) - \bar{X}_\Delta(t)|^p = 0, \quad \forall t \in [t_0, T].$$

Proof. Fix the step size $\Delta \in (0, 1]$ arbitrarily, for any $t \geq t_0$, there exists a constant $i \geq 0$ such that $t_i \leq t < t_{i+1}$. We derive from (2.13) that

$$\begin{aligned} \mathbb{E}|X_\Delta(t) - \bar{X}_\Delta(t)|^p &\leq C\mathbb{E}\left(\left|\int_{t_i}^t \mu_\Delta(\kappa(s), \bar{X}_\Delta(s)) ds\right|^p + \left|\int_{t_i}^t \sigma_\Delta(\kappa(s), \bar{X}_\Delta(s)) dB(s)\right|^p\right. \\ &\quad \left.+ \left|\int_{t_i}^t L\sigma_\Delta(\kappa(s), \bar{X}_\Delta(s)) \Delta B(s) dB(s)\right|^p\right), \end{aligned}$$

where the elementary inequality $|\sum_{i=1}^m a_i|^p \leq m^{p-1} \sum_{i=1}^m |a_i|^p$ has been used, and C is a positive constant independent of Δ that may change from line to line. We derive from the elementary inequality, the Hölder inequality and [the Burkholder-Davis-Gundy inequality](#) (Theorem 1.7.1 in [25]) that

$$\begin{aligned} \mathbb{E}|X_\Delta(t) - \bar{X}_\Delta(t)|^p &\leq C\left(\Delta^{p-1}\mathbb{E}\int_{t_i}^t |\mu_\Delta(\kappa(s), \bar{X}_\Delta(s))|^p ds + \Delta^{\frac{p-2}{2}}\mathbb{E}\int_{t_i}^t |\sigma_\Delta(\kappa(s), \bar{X}_\Delta(s))|^p ds\right. \\ &\quad \left.+ \Delta^{\frac{p-2}{2}}\mathbb{E}\int_{t_i}^t |L\sigma_\Delta(\kappa(s), \bar{X}_\Delta(s)) \Delta B(s)|^p ds\right). \end{aligned}$$

By using (2.10) and the inequality of $\mathbb{E}|\Delta B(s)|^p \leq C\Delta^{p/2}$ for $s \in [t_i, t_{i+1})$, we get

$$\mathbb{E}|X_\Delta(t) - \bar{X}_\Delta(t)|^p \leq C\left(\Delta^p h(\Delta)^p + \Delta^{\frac{p}{2}} h(\Delta)^p + \Delta^p h(\Delta)^{2p}\right).$$

Applying (2.6), the desired assertion holds. \blacksquare

The next lemma shows the moment boundedness of the truncated Milstein method.

Lemma 3.8. *Let Assumption 2.3 holds. Then, for any $\Delta \in (0, 1]$ and any $T \geq t_0$*

$$\sup_{0 < \Delta \leq 1} \sup_{t_0 \leq t \leq T} \mathbb{E}|X_\Delta(t)|^p \leq K(1 + \mathbb{E}|y_0|^p), \quad (3.4)$$

where K is a positive constant dependent on T but independent of Δ .

Proof. For any real number $R > X_0$, we define the stopping time

$$\rho_R := \inf\{t \geq t_0 : |X_\Delta(t)| \geq R\}.$$

Applying the Itô formula, we derive from (2.13), for any $t \in [t_0, \rho_R \wedge T]$

$$\begin{aligned} \mathbb{E}|X_\Delta(t)|^p &= \mathbb{E}|X_\Delta(t_0)|^p + p\mathbb{E} \int_{t_0}^t |X_\Delta(s)|^{p-2} \langle X_\Delta(s), \mu_\Delta(\kappa(s), \bar{X}_\Delta(s)) \rangle ds \\ &\quad + \frac{p(p-1)}{2} \mathbb{E} \int_{t_0}^t |X_\Delta(s)|^{p-2} \left| \sigma_\Delta(\kappa(s), \bar{X}_\Delta(s)) + L\sigma_\Delta(\kappa(s), \bar{X}_\Delta(s)) \Delta B(s) \right|^2 ds, \end{aligned}$$

where the fact that

$$\mathbb{E} \left(\int_{t_0}^t p|X_\Delta(s)|^{p-2} \langle X_\Delta(s), \sigma_\Delta(\kappa(s), \bar{X}_\Delta(s)) + L\sigma_\Delta(\kappa(s), \bar{X}_\Delta(s)) \Delta B(s) \rangle dB(s) \right) = 0$$

is used. Since $p|X_\Delta(s)|^{p-2} \langle X_\Delta(s), \sigma_\Delta(\kappa(s), \bar{X}_\Delta(s)) + L\sigma_\Delta(\kappa(s), \bar{X}_\Delta(s)) \Delta B(s) \rangle$ is \mathcal{F}_s -measurable, by Theorem 3.2.1 in [30] we see the fact above is true.

We rewrite the inequality as

$$\begin{aligned} \mathbb{E}|X_\Delta(t)|^p &\leq \mathbb{E}|X_\Delta(t_0)|^p + p\mathbb{E} \int_{t_0}^t |X_\Delta(s)|^{p-2} \left(\langle \bar{X}_\Delta(s), \mu_\Delta(\kappa(s), \bar{X}_\Delta(s)) \rangle + (p-1) \left| \sigma_\Delta(\kappa(s), \bar{X}_\Delta(s)) \right|^2 \right) ds \\ &\quad + p(p-1) \mathbb{E} \int_{t_0}^t |X_\Delta(s)|^{p-2} \left| L\sigma_\Delta(\kappa(s), \bar{X}_\Delta(s)) \Delta B(s) \right|^2 ds \\ &\quad + p\mathbb{E} \int_{t_0}^t |X_\Delta(s)|^{p-2} \langle X_\Delta(s) - \bar{X}_\Delta(s), \mu_\Delta(\kappa(s), \bar{X}_\Delta(s)) \rangle ds. \end{aligned}$$

By Lemma 3.6, (2.10) and Assumption 2.3, we get

$$\begin{aligned} \mathbb{E}|X_\Delta(t)|^p &\leq \mathbb{E}|y_0|^p + K\mathbb{E} \int_{t_0}^t |X_\Delta(s)|^{p-2} (1 + |\bar{X}_\Delta(s)|^2) ds + K\mathbb{E} \int_{t_0}^t |X_\Delta(s)|^{p-2} |h(\Delta)|^4 \Delta ds \\ &\quad + p\mathbb{E} \int_{t_0}^t |X_\Delta(s)|^{p-2} \langle X_\Delta(s) - \bar{X}_\Delta(s), \mu_\Delta(\kappa(s), \bar{X}_\Delta(s)) \rangle ds, \end{aligned}$$

where K is a constant independent of Δ that may change from line to line. By using the Young inequality

$$a^{p-2}b \leq \frac{p-2}{p}a^p + \frac{2}{p}b^{\frac{p}{2}},$$

we have

$$\begin{aligned} \mathbb{E}|X_\Delta(t)|^p &\leq \mathbb{E}|y_0|^p + K\mathbb{E} \int_{t_0}^t |X_\Delta(s)|^p ds + K\mathbb{E} \int_{t_0}^t |\bar{X}_\Delta(s)|^p ds \\ &\quad + K\mathbb{E} \int_{t_0}^t |h(\Delta)|^{2p} \Delta^{\frac{p}{2}} ds + K\mathbb{E} \int_{t_0}^t |X_\Delta(s) - \bar{X}_\Delta(s)|^{\frac{p}{2}} |\mu_\Delta(\kappa(s), \bar{X}_\Delta(s))|^{\frac{p}{2}} ds. \end{aligned} \tag{3.5}$$

By (2.6), (2.10) and Lemma 3.7, we obtain

$$\mathbb{E} \int_{t_0}^t |X_\Delta(s) - \bar{X}_\Delta(s)|^{\frac{p}{2}} |\mu_\Delta(\kappa(s), \bar{X}_\Delta(s))|^{\frac{p}{2}} ds \leq C \int_{t_0}^t (h(\Delta))^p \Delta^{\frac{p}{4}} ds \leq C(t - t_0). \quad (3.6)$$

Substituting (3.6) into (3.5), by using (2.6) we see that

$$\mathbb{E}|X_\Delta(t)|^p \leq \mathbb{E}|y_0|^p + K(t - t_0) + KC(t - t_0) + K \int_{t_0}^t \left(\sup_{t_0 \leq u \leq s} \mathbb{E}|X_\Delta(u \wedge \rho_R)|^p \right) ds.$$

Under the fact that the sum of the right-hand-side in the above inequality is a increasing function of t , we obtain

$$\sup_{t_0 \leq r \leq t} \mathbb{E}|X_\Delta(r \wedge \rho_R)|^p \leq \mathbb{E}|y_0|^p + K(t - t_0) + KC(t - t_0) + K \int_{t_0}^t \left(\sup_{t_0 \leq u \leq s} \mathbb{E}|X_\Delta(u \wedge \rho_R)|^p \right) ds.$$

Now the Gronwall inequality yields that

$$\sup_{t_0 \leq r \leq T} \mathbb{E}|X_\Delta(r \wedge \rho_R)|^p \leq K(1 + \mathbb{E}|y_0|^p).$$

Finally, the desired assertion follows by letting $R \rightarrow \infty$. \blacksquare

Lemma 3.9. *If Assumptions 2.1, 2.3 and (2.4) hold, and assume that $p \geq 2(1 + \beta)q$, then for any $\bar{q} \in [2, q]$*

$$\sup_{0 < \Delta \leq 1} \sup_{t_0 \leq t \leq T} \left[\mathbb{E}|\mu(t, X_\Delta(t))|^{2\bar{q}} \vee \mathbb{E}|\sigma(t, X_\Delta(t))|^{2\bar{q}} \vee \mathbb{E}|\mu'(t, x)|_{x=X_\Delta(t)}|^{2\bar{q}} \vee \mathbb{E}|\sigma'(t, x)|_{x=X_\Delta(t)}|^{2\bar{q}} \right] < \infty.$$

we can derive it from (2.4) and Lemma 3.8.

Lemma 3.10. *If Assumptions 2.1, 2.2, 2.3 and (2.4) hold and assume that $p \geq 2(1 + \beta)q$, then for any $\bar{q} \in (2, q)$ and $\Delta \in (0, 1]$,*

$$\mathbb{E}|\tilde{R}_\mu(t, X_\Delta(t), \bar{X}_\Delta(t))|^{\bar{q}} \vee \mathbb{E}|\tilde{R}_\sigma(t, X_\Delta(t), \bar{X}_\Delta(t))|^{\bar{q}} \vee \mathbb{E}|\tilde{R}_{\sigma_\Delta}(t, X_\Delta(t), \bar{X}_\Delta(t))|^{\bar{q}} \leq C\Delta^{\bar{q}}(h(\Delta))^{2\bar{q}},$$

where C is a positive constant independent of Δ .

Proof. Firstly, we give an estimate on $|R_\mu(t, X_\Delta(t), \bar{X}_\Delta(t))|^{\bar{q}}$, by Lemmas 3.7, 3.8 and (2.4), we obtain a constant C such that

$$\begin{aligned} & \mathbb{E}|R_\mu(t, X_\Delta(t), \bar{X}_\Delta(t))|^{\bar{q}} \\ & \leq \int_0^1 (1 - \tau)^{\bar{q}} \mathbb{E} \left| \mu''(\kappa(t), x) \Big|_{x=\bar{X}_\Delta(t) + \tau(X_\Delta(t) - \bar{X}_\Delta(t))} (X_\Delta(t) - \bar{X}_\Delta(t), X_\Delta(t) - \bar{X}_\Delta(t)) \right|^{\bar{q}} d\tau \\ & \leq \int_0^1 \left[\mathbb{E} \left| \mu''(\kappa(t), x) \Big|_{x=\bar{X}_\Delta(t) + \tau(X_\Delta(t) - \bar{X}_\Delta(t))} \right|^{2\bar{q}} \mathbb{E} |X_\Delta(t) - \bar{X}_\Delta(t)|^{4\bar{q}} \right]^{\frac{1}{2}} d\tau \\ & \leq C \left(1 + \mathbb{E}|X_\Delta(t)|^{2(1+\beta)\bar{q}} + \mathbb{E}|\bar{X}_\Delta(t)|^{2(1+\beta)\bar{q}} \right)^{\frac{1}{2}} \left(\mathbb{E} |X_\Delta(t) - \bar{X}_\Delta(t)|^{4\bar{q}} \right)^{\frac{1}{2}} \\ & \leq C\Delta^{\bar{q}}h(\Delta)^{2\bar{q}} \end{aligned} \quad (3.7)$$

where the Hölder inequality and the Jensen's inequality are used. Then we can observe from (2.16) and the Hölder inequality that

$$\begin{aligned}
\mathbb{E}|\tilde{R}_\mu(t, X_\Delta(t), \bar{X}_\Delta(t))|^{\bar{q}} &\leq C \left[\Delta^{\bar{q}} \mathbb{E} \left| \mu'(\kappa(t), x) \Big|_{x=\bar{X}_\Delta(t)} \mu_\Delta(\kappa(t), \bar{X}_\Delta(t)) \right|^{\bar{q}} \right. \\
&\quad + \frac{1}{2} \mathbb{E} \left| \mu'(\kappa(t), x) \Big|_{x=\bar{X}_\Delta(t)} L\sigma_\Delta(\kappa(t), \bar{X}_\Delta(t)) (\Delta B(t)^2 - \Delta) \right|^{\bar{q}} \\
&\quad \left. + \mathbb{E} |R_\mu(t, X_\Delta(t), \bar{X}_\Delta(t))|^{\bar{q}} \right] \\
&\leq C \left[\Delta^{\bar{q}} \mathbb{E} \left| \mu'(\kappa(t), x) \Big|_{x=\bar{X}_\Delta(t)} \mu_\Delta(\kappa(t), \bar{X}_\Delta(t)) \right|^{\bar{q}} \right. \\
&\quad + \frac{1}{2} \left(\mathbb{E} \left| \mu'(\kappa(t), x) \Big|_{x=\bar{X}_\Delta(t)} L\sigma_\Delta(\kappa(t), \bar{X}_\Delta(t)) \right|^{2\bar{q}} \mathbb{E} |\Delta B(t)^2 - \Delta|^{2\bar{q}} \right)^{\frac{1}{2}} \\
&\quad \left. + \mathbb{E} |R_\mu(t, X_\Delta(t), \bar{X}_\Delta(t))|^{\bar{q}} \right]
\end{aligned} \tag{3.8}$$

for $t_i \leq t < t_{i+1}$.

We can derive from the Hölder inequality that

$$\mathbb{E} |\Delta B(t)^2 - \Delta|^{2\bar{q}} \leq 2^{2\bar{q}-1} \left(\mathbb{E} |\Delta B(t)|^{4\bar{q}} + \Delta^{2\bar{q}} \right) \leq 2^{2\bar{q}-1} (\Delta^{2\bar{q}} + \Delta^{2\bar{q}}) \leq 2^{2\bar{q}} \Delta^{2\bar{q}}. \tag{3.9}$$

By using Lemma 3.9 and (2.10), we can see that for $t_0 \leq t \leq T$,

$$\begin{aligned}
\mathbb{E} \left| \mu'(\kappa(t), x) \Big|_{x=\bar{X}_\Delta(t)} \mu_\Delta(\kappa(t), \bar{X}_\Delta(t)) \right|^{\bar{q}} &\leq (h(\Delta))^{\bar{q}} \mathbb{E} \left| \mu'(\kappa(t), x) \Big|_{x=\bar{X}_\Delta(t)} \right|^{\bar{q}} \\
&\leq C (h(\Delta))^{\bar{q}},
\end{aligned} \tag{3.10}$$

$$\begin{aligned}
\mathbb{E} \left| \mu'(\kappa(t), x) \Big|_{x=\bar{X}_\Delta(t)} L\sigma_\Delta(\kappa(t), \bar{X}_\Delta(t)) \right|^{2\bar{q}} &\leq (h(\Delta))^{4\bar{q}} \mathbb{E} \left| \mu'(\kappa(t), x) \Big|_{x=\bar{X}_\Delta(t)} \right|^{2\bar{q}} \\
&\leq C (h(\Delta))^{4\bar{q}}.
\end{aligned} \tag{3.11}$$

Substituting (3.7), (3.9), (3.10) and (3.11) into (3.8) and using the independence between $\bar{X}(t)$ and $\Delta B(t)$, we have

$$\mathbb{E} |\tilde{R}_\mu(t, X_\Delta(t), \bar{X}_\Delta(t))|^{\bar{q}} \leq C \Delta^{\bar{q}} (h(\Delta))^{2\bar{q}}.$$

We obtain the desired result.

Similarly, we can show

$$\mathbb{E} |\tilde{R}_\sigma(t, X_\Delta(t), \bar{X}_\Delta(t))|^{\bar{q}} \vee \mathbb{E} |\tilde{R}_{\sigma_\Delta}(t, X_\Delta(t), \bar{X}_\Delta(t))|^{\bar{q}} \leq C \Delta^{\bar{q}} (h(\Delta))^{2\bar{q}}.$$

The proof is complete. \blacksquare

3.3 Proof of Theorem 3.1

Proof. Fix $\bar{q} \in [2, q)$ and $\Delta \in (0, 1]$ arbitrarily, let $e(t) = y(t) - X(t)$ for $t > t_0$, we define the stopping time for each integer $n > |X_0|$

$$\theta_n = \inf \left\{ t \geq t_0 : |X(t)| \vee |y(t)| \geq n \right\}.$$

We can derive from the Itô formula that for any $t_0 \leq t \leq T$,

$$\begin{aligned} \mathbb{E}|e(t \wedge \theta_n)|^{\bar{q}} &= \bar{q} \mathbb{E} \int_{t_0}^{t \wedge \theta_n} |e(s)|^{\bar{q}-2} \left\langle y(s) - X_\Delta(s), \mu(s, y(s)) - \mu_\Delta(\kappa(s), \bar{X}_\Delta(s)) \right\rangle ds \\ &\quad + \bar{q} \mathbb{E} \int_{t_0}^{t \wedge \theta_n} \frac{\bar{q}-1}{2} |e(s)|^{\bar{q}-2} \left| \sigma(s, y(s)) - \sigma_\Delta(\kappa(s), \bar{X}_\Delta(s)) \right. \\ &\quad \left. - L\sigma_\Delta(\kappa(s), \bar{X}_\Delta(s)) \Delta B(s) \right|^2 ds. \end{aligned} \tag{3.12}$$

Substituting (2.17) into (3.12), we have

$$\begin{aligned} \mathbb{E}|e(t \wedge \theta_n)|^{\bar{q}} &\leq \bar{q} \mathbb{E} \int_{t_0}^{t \wedge \theta_n} |e(s)|^{\bar{q}-2} \left\langle y(s) - X_\Delta(s), \mu(s, y(s)) - \mu_\Delta(\kappa(s), \bar{X}_\Delta(s)) \right\rangle ds \\ &\quad + \bar{q} \mathbb{E} \int_{t_0}^{t \wedge \theta_n} \frac{\bar{q}-1}{2} |e(s)|^{\bar{q}-2} \left| \sigma(s, y(s)) - \sigma_\Delta(\kappa(s), X_\Delta(s)) + \tilde{R}_{\sigma_\Delta}(t, X_\Delta(t), \bar{X}_\Delta(t)) \right|^2 ds. \end{aligned}$$

By the Young inequality $2ab \leq \varepsilon a^2 + b^2/\varepsilon$ for any $a, b \geq 0$ and ε arbitrarily, we choose $\varepsilon = \frac{q-\bar{q}}{q-1}$ here.

$$\begin{aligned} &(\bar{q}-1) \left| \sigma(s, y(s)) - \sigma_\Delta(\kappa(s), X_\Delta(s)) \right|^2 \\ &= (\bar{q}-1) \left| \sigma(s, y(s)) - \sigma(s, X_\Delta(s)) + \sigma(s, X_\Delta(s)) - \sigma_\Delta(\kappa(s), X_\Delta(s)) \right|^2 \\ &\leq (\bar{q}-1) \left[\left(1 + \frac{q-\bar{q}}{q-1} \right) \left| \sigma(s, y(s)) - \sigma(s, X_\Delta(s)) \right|^2 \right. \\ &\quad \left. + \left(1 + \frac{\bar{q}-1}{q-\bar{q}} \right) \left| \sigma(s, X_\Delta(s)) - \sigma_\Delta(\kappa(s), X_\Delta(s)) \right|^2 \right] \\ &= (q-1) \left| \sigma(s, y(s)) - \sigma(s, X_\Delta(s)) \right|^2 + \frac{(\bar{q}-1)(q-1)}{q-\bar{q}} \left| \sigma(s, X_\Delta(s)) - \sigma_\Delta(\kappa(s), X_\Delta(s)) \right|^2. \end{aligned}$$

Then

$$\mathbb{E}|e(t \wedge \theta_n)|^{\bar{q}} = J_1 + J_2 + J_3, \tag{3.13}$$

where

$$J_1 = \mathbb{E} \int_{t_0}^{t \wedge \theta_n} \bar{q} |e(s)|^{\bar{q}-2} \left(\left\langle e(s), \mu(s, y(s)) - \mu(s, X_\Delta(s)) \right\rangle + (q-1) \left| \sigma(s, y(s)) - \sigma(s, X_\Delta(s)) \right|^2 \right) ds,$$

$$\begin{aligned}
J_2 &= \mathbb{E} \int_{t_0}^{t \wedge \theta_n} \bar{q} |e(s)|^{\bar{q}-2} \left(\left\langle e(s), \mu(s, X_\Delta(s)) - \mu_\Delta(\kappa(s), \bar{X}_\Delta(s)) \right\rangle \right. \\
&\quad \left. + \frac{(\bar{q}-1)(q-1)}{q-\bar{q}} \left| \sigma(s, X_\Delta(s)) - \sigma_\Delta(\kappa(s), X_\Delta(s)) \right|^2 \right) ds, \\
J_3 &\leq \mathbb{E} \int_{t_0}^{t \wedge \theta_n} \bar{q}(\bar{q}-1) |e(s)|^{\bar{q}-2} |\tilde{R}_{\sigma_\Delta}(t, X_\Delta(t), \bar{X}_\Delta(t))|^2 ds.
\end{aligned}$$

By Assumption 2.2, we have

$$J_1 \leq H_1 \mathbb{E} \int_{t_0}^{t \wedge \theta_n} |e(s)|^{\bar{q}} ds. \quad (3.14)$$

Rearranging J_2 , we get

$$\begin{aligned}
J_2 &\leq \mathbb{E} \int_{t_0}^{t \wedge \theta_n} \bar{q} |e(s)|^{\bar{q}-2} \left(\left\langle e(s), \mu(s, X_\Delta(s)) - \mu(\kappa(s), X_\Delta(s)) \right\rangle \right. \\
&\quad \left. + \frac{2(\bar{q}-1)(q-1)}{q-\bar{q}} \left| \sigma(s, X_\Delta(s)) - \sigma(\kappa(s), X_\Delta(s)) \right|^2 \right) ds \\
&\quad + \mathbb{E} \int_{t_0}^{t \wedge \theta_n} \bar{q} |e(s)|^{\bar{q}-2} \left(\left\langle e(s), \mu(\kappa(s), X_\Delta(s)) - \mu_\Delta(\kappa(s), \bar{X}_\Delta(s)) \right\rangle \right. \\
&\quad \left. + \frac{2(\bar{q}-1)(q-1)}{q-\bar{q}} \left| \sigma(\kappa(s), X_\Delta(s)) - \sigma_\Delta(\kappa(s), X_\Delta(s)) \right|^2 \right) ds \\
&=: J_{21} + J_{22}.
\end{aligned} \quad (3.15)$$

We estimate J_{21} first. Applying the Young inequality $a^{p-2}b^2 \leq (p-2)a^p/p + 2b^p/p$ for any $a, b \geq 0$ and $t_0 \leq t \wedge \theta_n \leq t \leq T$, we obtain

$$\begin{aligned}
J_{21} &\leq \mathbb{E} \int_{t_0}^{t \wedge \theta_n} \bar{q} |e(s)|^{\bar{q}-2} \left(\frac{1}{2} |e(s)|^2 + \frac{1}{2} \left| \mu(s, X_\Delta(s)) - \mu(\kappa(s), X_\Delta(s)) \right|^2 \right. \\
&\quad \left. + \frac{2(\bar{q}-1)(q-1)}{q-\bar{q}} \left| \sigma(s, X_\Delta(s)) - \sigma(\kappa(s), X_\Delta(s)) \right|^2 \right) ds \\
&\leq H_2 \left(\mathbb{E} \int_{t_0}^{t \wedge \theta_n} |e(s)|^{\bar{q}} ds + \mathbb{E} \int_{t_0}^{t \wedge \theta_n} \left| \mu(s, X_\Delta(s)) - \mu(\kappa(s), X_\Delta(s)) \right|^{\bar{q}} ds \right. \\
&\quad \left. + \mathbb{E} \int_{t_0}^{t \wedge \theta_n} \left| \sigma(s, X_\Delta(s)) - \sigma(\kappa(s), X_\Delta(s)) \right|^{\bar{q}} ds \right) \\
&\leq H_2 \left(\mathbb{E} \int_{t_0}^{t \wedge \theta_n} |e(s)|^{\bar{q}} ds + 2C_4 \mathbb{E} \int_{t_0}^T \left(1 + |X_\Delta(s)|^{(1+\beta)\bar{q}} \right) \Delta^{\alpha\bar{q}} ds \right) \\
&\leq H_2 \left(\mathbb{E} \int_{t_0}^{t \wedge \theta_n} |e(s)|^{\bar{q}} ds + \int_{t_0}^T \left(1 + \mathbb{E} |X_\Delta(s)|^{(1+\beta)\bar{q}} \right) \Delta^{\alpha\bar{q}} ds \right) \\
&\leq H_2 \left(\mathbb{E} \int_{t_0}^{t \wedge \theta_n} |e(s)|^{\bar{q}} ds + \Delta^{\alpha\bar{q}} \right),
\end{aligned} \quad (3.16)$$

where the Assumption 2.1 and Lemma 3.8 are used. Rearranging J_{22} shows that

$$\begin{aligned}
J_{22} &= \mathbb{E} \int_{t_0}^{t \wedge \theta_n} \bar{q} |e(s)|^{\bar{q}-2} \left\langle e(s), \mu(\kappa(s), X_\Delta(s)) - \mu(\kappa(s), \bar{X}_\Delta(s)) \right\rangle ds \\
&\quad + \mathbb{E} \int_{t_0}^{t \wedge \theta_n} \bar{q} |e(s)|^{\bar{q}-2} \left(\left\langle e(s), \mu(\kappa(s), \bar{X}_\Delta(s)) - \mu_\Delta(\kappa(s), \bar{X}_\Delta(s)) \right\rangle \right. \\
&\quad \left. + \frac{2(\bar{q}-1)(q-1)}{q-\bar{q}} \left| \sigma(\kappa(s), X_\Delta(s)) - \sigma_\Delta(\kappa(s), X_\Delta(s)) \right|^2 \right) ds \\
&= I_1 + I_2,
\end{aligned} \tag{3.17}$$

where

$$\begin{aligned}
I_1 &= \mathbb{E} \int_{t_0}^{t \wedge \theta_n} \bar{q} |e(s)|^{\bar{q}-2} \left\langle e(s), \mu(\kappa(s), X_\Delta(s)) - \mu(\kappa(s), \bar{X}_\Delta(s)) \right\rangle ds, \\
I_2 &= \mathbb{E} \int_{t_0}^{t \wedge \theta_n} \bar{q} |e(s)|^{\bar{q}-2} \left(\left\langle e(s), \mu(\kappa(s), \bar{X}_\Delta(s)) - \mu_\Delta(\kappa(s), \bar{X}_\Delta(s)) \right\rangle \right. \\
&\quad \left. + \frac{2(\bar{q}-1)(q-1)}{q-\bar{q}} \left| \sigma(\kappa(s), X_\Delta(s)) - \sigma_\Delta(\kappa(s), X_\Delta(s)) \right|^2 \right) ds.
\end{aligned}$$

We can derive from the Young inequality and (2.15) that

$$\begin{aligned}
I_1 &= \mathbb{E} \int_{t_0}^{t \wedge \theta_n} \bar{q} |e(s)|^{\bar{q}-2} \left\langle e(s), \mu'(\kappa(s), x) \Big|_{x=\bar{X}_\Delta(s)} \int_{t_0}^s \sigma_\Delta(\kappa(s_1), \bar{X}_\Delta(s_1)) dB(s_1) + \tilde{R}_\mu(t, X_\Delta(t), \bar{X}_\Delta(t)) \right\rangle ds \\
&\leq H_{21} \mathbb{E} \int_{t_0}^{t \wedge \theta_n} \left(|e(s)|^{\bar{q}} + \left| e(s)^T \mu'(\kappa(s), x) \Big|_{x=\bar{X}_\Delta(s)} \int_{t_0}^s \sigma_\Delta(\kappa(s_1), \bar{X}_\Delta(s_1)) dB(s_1) \right|^{\frac{\bar{q}}{2}} \right. \\
&\quad \left. + |e(s)^T \tilde{R}_\mu(t, X_\Delta(t), \bar{X}_\Delta(t))|^{\frac{\bar{q}}{2}} \right) ds \\
&\leq \left(H_{21} \mathbb{E} \int_{t_0}^{t \wedge \theta_n} \left(|e(s)|^{\bar{q}} + |\tilde{R}_\mu(t, X_\Delta(t), \bar{X}_\Delta(t))|^{\bar{q}} \right) ds + I_{11} \right),
\end{aligned} \tag{3.18}$$

where

$$I_{11} := \mathbb{E} \int_{t_0}^{t \wedge \theta_n} \left| e(s)^T \mu'(\kappa(s), x) \Big|_{x=\bar{X}_\Delta(s)} \int_{t_0}^s \sigma_\Delta(\kappa(s_1), \bar{X}_\Delta(s_1)) dB(s_1) \right|^{\frac{\bar{q}}{2}} ds.$$

Following a very similar approach used for (3.35) in [33], we get

$$I_{11} \leq H_{21} \Delta^{\bar{q}}. \tag{3.19}$$

Combining (3.18), (3.19) and Lemma 3.10, we obtain

$$\begin{aligned}
I_1 &\leq H_{21} \left(\mathbb{E} \int_{t_0}^{t \wedge \theta_n} |e(s)|^{\bar{q}} ds + \mathbb{E} \int_{t_0}^T |\tilde{R}_\mu(t, X_\Delta(t), \bar{X}_\Delta(t))|^{\bar{q}} ds + \Delta^{\bar{q}} \right) \\
&\leq H_{21} \left(\mathbb{E} \int_{t_0}^{t \wedge \theta_n} |e(s)|^{\bar{q}} ds + \int_{t_0}^T \mathbb{E} |\tilde{R}_\mu(t, X_\Delta(t), \bar{X}_\Delta(t))|^{\bar{q}} ds + \Delta^{\bar{q}} \right) \quad (3.20) \\
&\leq H_{21} \left(\mathbb{E} \int_{t_0}^{t \wedge \theta_n} |e(s)|^{\bar{q}} ds + \Delta^{\bar{q}} (h(\Delta))^{2\bar{q}} + \Delta^{\bar{q}} \right).
\end{aligned}$$

And applying the Young inequality and Assumption 2.1, we can show that

$$\begin{aligned}
I_2 &\leq H_{22} \left(\mathbb{E} \int_{t_0}^{t \wedge \theta_n} |e(s)|^{\bar{q}} ds + \mathbb{E} \int_{t_0}^{t \wedge \theta_n} \left(\left| \mu(\kappa(s), \bar{X}_\Delta(s)) - \mu_\Delta(\kappa(s), \bar{X}_\Delta(s)) \right|^{\bar{q}} \right. \right. \\
&\quad \left. \left. + \left| \sigma(\kappa(s), X_\Delta(s)) - \sigma_\Delta(\kappa(s), X_\Delta(s)) \right|^{\bar{q}} \right) ds \right) \\
&\leq H_{22} \left(\mathbb{E} \int_{t_0}^{t \wedge \theta_n} |e(s)|^{\bar{q}} ds + \mathbb{E} \int_{t_0}^{t \wedge \theta_n} \left(1 + |\bar{X}_\Delta(s)|^{\beta\bar{q}} + \left| |\bar{X}_\Delta(s)| \wedge f^{-1}(h(\Delta)) \right|^{\beta\bar{q}} \right) \right. \\
&\quad \times \left| \bar{X}_\Delta(s) - \left(|\bar{X}_\Delta(s)| \wedge f^{-1}(h(\Delta)) \right) \frac{\bar{X}_\Delta(s)}{|\bar{X}_\Delta(s)|} \right|^{\bar{q}} ds + \mathbb{E} \int_{t_0}^{t \wedge \theta_n} \left(1 + |X_\Delta(s)|^{\beta\bar{q}} \right. \\
&\quad \left. + \left| |X_\Delta(s)| \wedge f^{-1}(h(\Delta)) \right|^{\beta\bar{q}} \right) \left| X_\Delta(s) - \left(|X_\Delta(s)| \wedge f^{-1}(h(\Delta)) \right) \frac{X_\Delta(s)}{|X_\Delta(s)|} \right|^{\bar{q}} ds \Big) \\
&\leq H_{22} \left(\mathbb{E} \int_{t_0}^{t \wedge \theta_n} |e(s)|^{\bar{q}} ds + \int_{t_0}^T \left(\mathbb{E} \left[1 + |\bar{X}_\Delta(s)|^p + \left| |\bar{X}_\Delta(s)| \wedge f^{-1}(h(\Delta)) \right|^p \right] \right)^{\frac{\beta\bar{q}}{p}} \right. \\
&\quad \times \left[\mathbb{E} \left| \bar{X}_\Delta(s) - \left(|\bar{X}_\Delta(s)| \wedge f^{-1}(h(\Delta)) \right) \frac{\bar{X}_\Delta(s)}{|\bar{X}_\Delta(s)|} \right|^{\frac{p\bar{q}}{p-\beta\bar{q}}} \right]^{\frac{p-\beta\bar{q}}{p}} ds + \int_{t_0}^T \left(\mathbb{E} \left[1 + |X_\Delta(s)|^p \right. \right. \\
&\quad \left. \left. + \left| |X_\Delta(s)| \wedge f^{-1}(h(\Delta)) \right|^p \right] \right)^{\frac{\beta\bar{q}}{p}} \left[\mathbb{E} \left| X_\Delta(s) - \left(|X_\Delta(s)| \wedge f^{-1}(h(\Delta)) \right) \frac{X_\Delta(s)}{|X_\Delta(s)|} \right|^{\frac{p\bar{q}}{p-\beta\bar{q}}} \right]^{\frac{p-\beta\bar{q}}{p}} ds \Big) \\
&\leq H_{22} \left(\mathbb{E} \int_{t_0}^{t \wedge \theta_n} |e(s)|^{\bar{q}} ds + \int_{t_0}^T \left(\mathbb{E} \left| I \{ |\bar{X}_\Delta(s)| > f^{-1}(h(\Delta)) \} \right| \bar{X}_\Delta(s) \right|^{\frac{p\bar{q}}{p-\beta\bar{q}}} \right)^{\frac{p-\beta\bar{q}}{p}} ds \right. \\
&\quad \left. + \int_{t_0}^T \left(\mathbb{E} \left| I \{ |X_\Delta(s)| > f^{-1}(h(\Delta)) \} \right| X_\Delta(s) \right|^{\frac{p\bar{q}}{p-\beta\bar{q}}} \right)^{\frac{p-\beta\bar{q}}{p}} ds \Big),
\end{aligned}$$

where the Hölder inequality and Lemma 3.8 are used above, and using the Chebyshev inequality yields

$$\begin{aligned}
I_2 &\leq H_{22} \left(\mathbb{E} \int_{t_0}^{t \wedge \theta_n} |e(s)|^{\bar{q}} ds + \int_{t_0}^T \left(\left[P \left\{ |\bar{X}_\Delta(s)| > f^{-1}(h(\Delta)) \right\} \right]^{\frac{p-\beta\bar{q}-\bar{q}}{p-\beta\bar{q}}} \left[\mathbb{E} |\bar{X}_\Delta(s)|^p \right]^{\frac{\bar{q}}{p-\beta\bar{q}}} \right)^{\frac{p-\beta\bar{q}}{p}} ds \right. \\
&\quad \left. + \int_{t_0}^T \left(\left[P \left\{ |X_\Delta(s)| > f^{-1}(h(\Delta)) \right\} \right]^{\frac{p-\beta\bar{q}-\bar{q}}{p-\beta\bar{q}}} \left[\mathbb{E} |X_\Delta(s)|^p \right]^{\frac{\bar{q}}{p-\beta\bar{q}}} \right)^{\frac{p-\beta\bar{q}}{p}} ds \right) \\
&\leq H_{22} \left(\mathbb{E} \int_{t_0}^{t \wedge \theta_n} |e(s)|^{\bar{q}} ds + \int_{t_0}^T \left(\frac{\mathbb{E} |\bar{X}_\Delta(s)|^p}{|f^{-1}(h(\Delta))|^p} \right)^{\frac{p-\beta\bar{q}-\bar{q}}{p}} ds + \int_{t_0}^T \left(\frac{\mathbb{E} |X_\Delta(s)|^p}{|f^{-1}(h(\Delta))|^p} \right)^{\frac{p-\beta\bar{q}-\bar{q}}{p}} ds \right) \\
&\leq H_{22} \left(\mathbb{E} \int_{t_0}^{t \wedge \theta_n} |e(s)|^{\bar{q}} ds + \left(f^{-1}(h(\Delta)) \right)^{(\beta+1)\bar{q}-p} \right). \tag{3.21}
\end{aligned}$$

Substituting (3.20) and (3.21) into (3.17) gives

$$J_{22} \leq H_2 \left(\mathbb{E} \int_{t_0}^{t \wedge \theta_n} |e(s)|^{\bar{q}} ds + \left(f^{-1}(h(\Delta)) \right)^{(\beta+1)\bar{q}-p} + \Delta^q (h(\Delta))^{2\bar{q}} + \Delta^{\bar{q}} \right). \tag{3.22}$$

Due to the Young inequality and Lemma 3.10, we derive that

$$\begin{aligned}
J_3 &\leq H_3 \mathbb{E} \int_{t_0}^{t \wedge \theta_n} \left(|e(s)|^{\bar{q}} + |\tilde{R}_{\sigma_\Delta}(t, X_\Delta(t), \bar{X}_\Delta(t))|^{\bar{q}} \right) ds \\
&\leq H_3 \left(\mathbb{E} \int_{t_0}^{t \wedge \theta_n} |e(s)|^{\bar{q}} ds + \int_{t_0}^T \mathbb{E} |\tilde{R}_{\sigma_\Delta}(t, X_\Delta(t), \bar{X}_\Delta(t))|^{\bar{q}} ds \right) \tag{3.23} \\
&\leq H_3 \left(\mathbb{E} \int_{t_0}^{t \wedge \theta_n} |e(s)|^{\bar{q}} ds + \Delta^{\bar{q}} (h(\Delta))^{2\bar{q}} \right),
\end{aligned}$$

where H_{21} , H_{22} , H_3 and the following H are generic constants independent of Δ that may change from line to line.

Combining (3.13), (3.14), (3.15), (3.16), (3.22) and (3.23) together, we can see that

$$\begin{aligned}
\mathbb{E} |e(t \wedge \theta_n)|^{\bar{q}} &\leq H \left(\mathbb{E} \int_{t_0}^{t \wedge \theta_n} |e(s)|^{\bar{q}} ds + \Delta^{\alpha\bar{q}} + \Delta^{\bar{q}} (h(\Delta))^{2\bar{q}} + \Delta^{\bar{q}} + \left(f^{-1}(h(\Delta)) \right)^{(\beta+1)\bar{q}-p} \right) \\
&\leq H \left(\int_{t_0}^t \sup_{t_0 \leq u \leq s} \mathbb{E} |e(u \wedge \theta_n)|^{\bar{q}} ds + \Delta^{\alpha\bar{q}} + \Delta^{\bar{q}} (h(\Delta))^{2\bar{q}} + \Delta^{\bar{q}} + \left(f^{-1}(h(\Delta)) \right)^{(\beta+1)\bar{q}-p} \right).
\end{aligned}$$

An application of the Gronwall inequality yields that

$$\sup_{t_0 \leq r \leq T} \mathbb{E} |e(r \wedge \theta_n)|^{\bar{q}} \leq H \left(\Delta^{\alpha\bar{q}} + \Delta^{\bar{q}} (h(\Delta))^{2\bar{q}} + \left(f^{-1}(h(\Delta)) \right)^{(\beta+1)\bar{q}-p} \right).$$

Due to the existence and uniqueness of the global solution to SDE (2.1) in $[t_0, T]$, we have $T \wedge \theta_n \rightarrow T$ as $n \rightarrow \infty$ (see, for example, the proof of Lemma 2.3.2 in [25], where the similar argument was used). Using Fatou Lemma and letting $n \rightarrow \infty$, the desired assertion is obtained. \blacksquare

4 Randomized Truncated Milstein method

To define the randomized truncated Milstein method, let $(\tau_j)_{j \in \mathbb{N}}$ be a i.i.d family of $\mathcal{U}(0, 1)$ -distributed random variables on a filtered probability space $(\omega_\tau, \mathcal{F}^\tau, (\mathcal{F}_j^\tau)_{j \in \mathbb{N}}, \mathbb{P}_\tau)$, the space is generated by $\{\tau_1, \dots, \tau_j\}$. Besides we define $\mathcal{U}(0, 1)$ as a unique distribution on the interval $(0, 1)$. Furthermore, $(\tau_j)_{j \in \mathbb{N}}$ is assumed to be independent of the randomness in SDE (2.1).

As already observed in Remark 3.5, the convergence rate of the truncated Milstein method is dominated by the Hölder index α . The purpose of this section is to propose some new method to improve the convergence rate.

Inspired by [21], we embed the randomized time step into (2.12) and propose the following randomized truncated Milstein method.

Given a step-size $\Delta \in (0, 1)$, the randomized truncated Milstein numerical solution X_{i+1} to approximate of SDE (2.1) for $t_i = i\Delta$ is given by the recursion

$$X_{i+1}^\tau = X_i + \tau_i \Delta \mu_\Delta(t_i, X_i) + \sigma_\Delta(t_i, X_i) (B(t_i + \tau_i \Delta) - B(t_i)),$$

$$X_{i+1} = X_i + \Delta \mu_\Delta(t_i + \tau_i \Delta, X_{i+1}^\tau) + \sigma_\Delta(t_i, X_i) \Delta B_i + \frac{1}{2} \sum_{l=1}^d \sigma_\Delta^l(t_i, X_i) G_\Delta^l(t_i, X_i) (\Delta B_i^2 - \Delta),$$

where $X_0 = y_0$, $t_{N+1} = T$ for N is the integer part of T/Δ and $\Delta B_i = B(t_{i+1} - B(t_i))$ for $i \in \{0, 1, \dots, N\}$.

Based on [21], we have the following conjecture on the convergence rate. Briefly speaking, with the employment of the randomized technique the convergence is improved from $\min(1 - 2\varepsilon, \alpha)$ to $\min(1 - 2\varepsilon, \alpha + 1/2)$.

Since we have still been working on the proof of it, we will demonstrate this conjecture by using numerical simulation in the next section.

Conjecture 4.1. *Suppose Assumptions 2.1, 2.2 and 2.3 hold for any $p > 2$, then for any $\bar{q} > 0, \varepsilon \in (0, 1/4)$ and $\Delta \in (0, 1]$,*

$$\sup_{t_0 \leq t \leq T} \mathbb{E}|y(t) - X_\Delta(t)|^{\bar{q}} \leq H \left(\Delta^{\min(1-2\varepsilon, \alpha + \frac{1}{2})\bar{q}} \right),$$

where H is a constant independent from Δ .

5 Numerical examples

The purpose of the example discussed in this section is twofold. On one side, it is used to illustrate Theorem 3.2. On the other side, it demonstrates that the convergence rate in Conjecture 4.1 is promising.

Example 5.1. Consider the scalar SDE

$$\begin{cases} dX(t) = \left([t(1-t)]^{\frac{1}{4}} X^2(t) - X^5(t) \right) dt + [t(1-t)]^{\frac{3}{4}} X(t) dB(t), \\ X(t_0) = 2, \end{cases} \quad (5.1)$$

where $t_0 = 0$, $T = 1$ and $B(t)$ is a scalar Brownian motion.

For any $q > 2$, $t \in [0, 1]$ we can see that

$$\begin{aligned} & (x - y) (\mu(t, x) - \mu(t, y)) + (q - 1) |\sigma(t, x) - \sigma(t, y)|^2 \\ = & (x - y)^2 \left([t(1-t)]^{\frac{1}{4}} (x + y) - (x^4 + x^3y + x^2y^2 + xy^3 + y^4) + (q - 1) [t(1-t)]^{\frac{3}{2}} \right). \end{aligned}$$

But

$$-(x^3y + xy^3) = -xy(x^2 + y^2) \leq 0.5(x^2 + y^2)^2 = 0.5(x^4 + y^4) + x^2y^2.$$

Hence

$$\begin{aligned} & (x - y)^T (\mu(t, x) - \mu(t, y)) + (q - 1) |\sigma(t, x) - \sigma(t, y)|^2 \\ \leq & (x - y)^2 \left([t(1-t)]^{\frac{1}{4}} (x + y) - 0.5(x^4 + y^4) + (q - 1) [t(1-t)]^{\frac{3}{2}} \right) \\ \leq & C(x - y)^2. \end{aligned}$$

Under the fact that polynomials with negative coefficient for the highest order term can always be bounded, we can obtain the assertion above. It means that Assumption 2.2 is satisfied.

Similarly, for any $p > 2$ and any $t \in [0, 1]$, we have

$$x^T \mu(t, x) + (p-1) |\sigma(t, x)|^2 = [t(1-t)]^{\frac{1}{4}} x^3 - x^6 + (p-1) [t(1-t)]^{\frac{3}{2}} x^2 \leq C(1 + |x|^2),$$

which [shows](#) that Assumption 2.3 holds.

Applying the mean value theorem for the temporal variable, Assumption 2.1 are satisfied with $\alpha = 1/4$ and $\beta = 4$. Due to the fact that

$$\sup_{0 < t \leq 1} \sup_{|x| \leq u} (|\mu(t, x)| \vee |\sigma(t, x)| \vee |L\sigma(t, x)|) \leq 2u^5, \quad \forall u \geq 1,$$

Let $f(u) = 2u^5$ and $h(\Delta) = \Delta^{-\varepsilon}$, for any $\varepsilon \in (0, 1/4)$. Choose ϵ sufficiently [small](#), we can derive from Theorem 3.2 that

$$\sup_{0 \leq t \leq 1} \mathbb{E} |y(t) - X_{\Delta}(t)|^{\bar{q}} \leq H \Delta^{\frac{1}{4}\bar{q}},$$

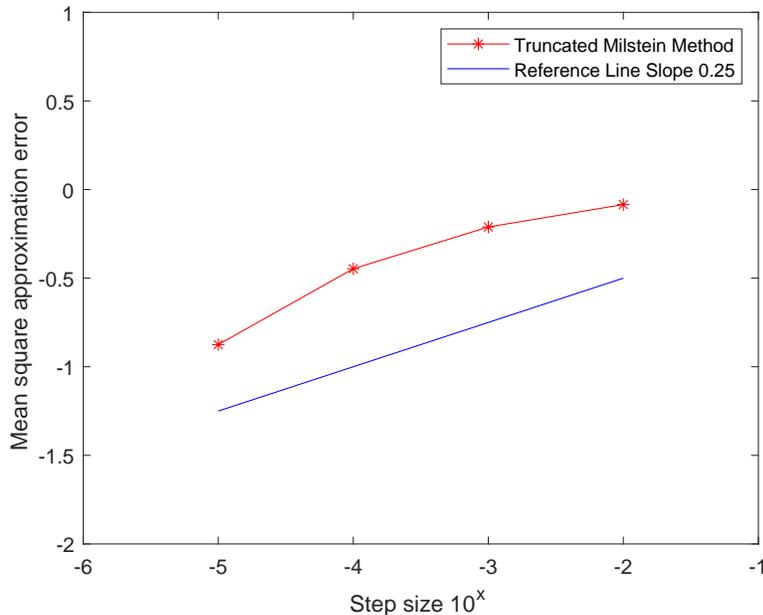


Figure 1: Convergence rate of truncated Milstein method in Example

This shows that the convergence rate of truncated Milstein method for the SDE (5.1) is $1/4$. To approximate the mean square error, we run $M = 1000$ independent trajectories for 5 different step sizes. And we regard the numerical solution as the step-size 10^{-7} as the true solution for the SDE. By numerical simulation we can see in Figure 1 that the slope of the error against the step sizes is about [0.2562](#).

Let us turn to the discussion on the randomized truncated Milstein method. We would expect

$$\sup_{0 \leq t \leq 1} \mathbb{E}|y(t) - X_{\Delta}(t)|^{\bar{q}} \leq H \Delta^{\frac{3}{4}\bar{q}},$$

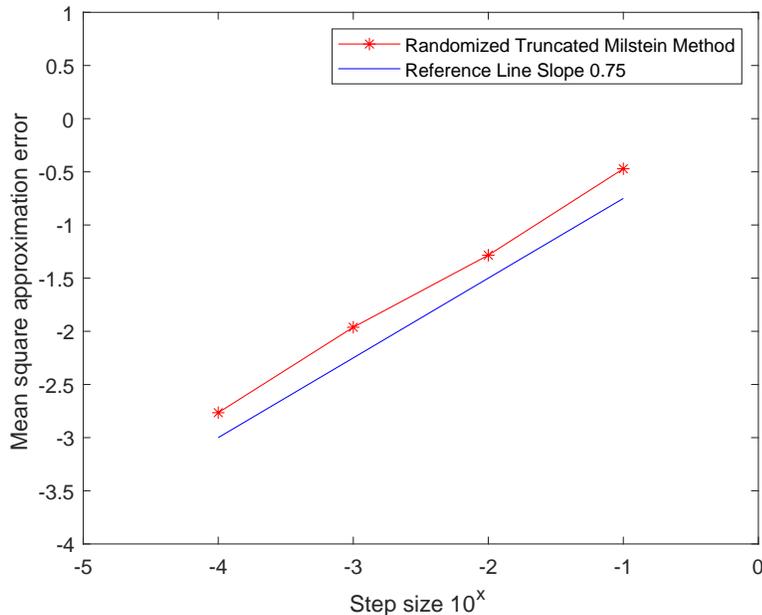


Figure 2: Convergence rate of randomized truncated Milstein method

In numerical simulations, we use the step-size 10^{-6} to approximate the true solution. In Figure 2, we run $M=1000$ independent paths with step-sizes 10^{-4} , 10^{-3} , 10^{-2} , 10^{-1} .

It is clearly to see from Figure 2 that the convergence rate of the randomized truncated Milstein method is indeed improved to be 0.7548. This shows that Conjecture 4.1 is reasonable.

6 Conclusion and future research

This paper revisited the truncated Milstein method and proved the strong convergence of the method for non-autonomous SDEs, which extended and improved the existing result.

With the observation that the convergence rate could be very low due to the Hölder continuous time variable, the randomized truncated Milstein method was proposed. The conjecture on the improvement of the convergence rate is reported. Numerical simulations demonstrate the conjecture is promising.

One of the main future works is to prove the conjecture. In addition, we are working on the stability of the truncated Milstein method and the randomized truncated Milstein method in different senses.

Acknowledgements

We deeply appreciate the detailed comments and suggestions from the reviewers and the editor, which helped to improve this work significantly.

The authors would like to thank the Natural Science Foundation of China (11701378, 11871343, 11971316), and Science and Technology Innovation Plan of Shanghai (20JC1414200) for their financial support.

References

- [1] S. Amiri and S.M. Hosseini, Stochastic Runge-Kutta Rosenbrock type methods for SDE systems, *Applied Numerical Mathematics*, [115](#), 2017, 1-15.
- [2] W.-J. Beyn and R. Kruse, Two-sided error estimates for the stochastic theta method, *Discrete and Continuous Dynamical Systems - Series B*, [14\(2\)](#), 2010, 389-407.
- [3] K. Burrage and T. Tian, Implicit stochastic Runge-Kutta methods for stochastic differential equations, *BIT Numerical Mathematics*, [44\(1\)](#), 2004, 21-39.
- [4] T. Daun, On the randomized solution of initial value problems, *Journal of Complexity*, [27\(3-4\)](#), 2011, 300-311.
- [5] K. Debrabant and A. Rößler, Diagonally drift-implicit Runge-Kutta methods of weak order one and two for Itô SDEs and stability analysis, *Applied Numerical Mathematics*, [59\(3-4\)](#), 2009, 595-607.
- [6] S. Deng, W. Fei, W. Liu and X. Mao, The truncated EM method for stochastic differential equations with Poisson jumps, *Journal of Computational and Applied Mathematics*, [355](#), 2019, 232-257.
- [7] K. Dareiotis, C. Kumar and S. Sabanis, On tamed Euler approximations of SDEs driven by Lévy noise with applications to delay equations, *SIAM Journal on Numerical Analysis*, [54\(3\)](#), 2016, 1840-1872.
- [8] S. Gan, Y. He and X. Wang, Tamed Runge-Kutta methods for SDEs with super-linearly growing drift and diffusion coefficients, *Applied Numerical Mathematics*, [152](#), 2020, 379-402.
- [9] Q. Guo, W. Liu, X. Mao and R. Yue, The partially truncated Euler-Maruyama method and its stability and boundedness, *Applied Numerical Mathematics*, [115](#), 2017, 235-251.

- [10] Q. Guo, W. Liu, X. Mao and R. Yue, The truncated Milstein method for stochastic differential equations with commutative noise, *Journal of Computational and Applied Mathematics*, **338**, 2018, 298-310.
- [11] E. Hairer and G. Wanner, Solving ordinary differential equations. II. Stiff and differential-algebraic problems. Second revised edition. Springer Series in Computational Mathematics, **14**. Springer-Verlag, Berlin, 2010.
- [12] D.J. Higham, Stochastic ordinary differential equations in applied and computational mathematics, *IMA Journal of Applied Mathematics*, **76(3)**, 2011, 449-474.
- [13] D.J. Higham, X. Mao and 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.
- [14] L. Hu, X. Li and X. Mao, Convergence rate and stability of the truncated Euler-Maruyama method for stochastic differential equations, *Journal of Computational and Applied Mathematics*, **337**, 2018, 274-289.
- [15] Y. Hu, Semi-implicit Euler-Maruyama scheme for stiff stochastic equations. *Stochastic analysis and related topics*, V (Silivri, 1994), 183-202, *Progr. Probab.*, 38, Birkhäuser Boston, Boston, MA, 1996.
- [16] 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, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **467(2130)**, 2011, 1563-1576.
- [17] 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.
- [18] Y. Jiang, Z. Huang and W. Liu, Equivalence of the mean square stability between the partially truncated Euler-Maruyama method and stochastic differential equations with super-linear growing coefficients, *Advances in Difference Equations*, 2018, Article No. 355.
- [19] B.Z. Kacwicz, Optimal solution of ordinary differential equations, *Journal of Complexity*, **3(4)**, 1987, 451-465.
- [20] C. Kahl and H. Schurz, Balanced Milstein methods for ordinary SDEs, *Monte Carlo Methods and Applications*, **12(2)**, 2006, 143-170.

- [21] R. Kruse and Y. Wu, A randomized milstein method for stochastic differential equations with non-differentiable drift coefficients, *Discrete and Continuous Dynamical Systems - Series B*, **24(8)**, 2019, 3475-3502.
- [22] G. Lan and F. Xia, Strong convergence rates of modified truncated EM method for stochastic differential equations, *Journal of Computational and Applied Mathematics*, **334**, 2018, 1-17.
- [23] 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 Journal of Numerical Analysis*, **39(2)**, 2019, 847-892.
- [24] W. Liu, X. Mao, J. Tang and Y. Wu, Truncated Euler-Maruyama method for classical and time-changed non-autonomous stochastic differential equations, *Applied Numerical Mathematics*, **153**, 2020, 66-81.
- [25] X. Mao, *Stochastic Differential Equations and Applications*, 2nd Edition, Horwood Publishing, Chichester, 2007.
- [26] X. Mao, The truncated Euler-Maruyama method for stochastic differential equations, *Journal of Computational and Applied Mathematics*, **290**, 2015, 370-384.
- [27] X. Mao, Convergence rates of the truncated Euler-Maruyama method for stochastic differential equations, *Journal of Computational and Applied Mathematics*, **296**, 2016, 362-375.
- [28] X. Mao and L. Szpruch, Strong convergence rates for backward Euler-Maruyama method for non-linear dissipative-type stochastic differential equations with super-linear diffusion coefficients, *Stochastics*, **85(1)**, 2013, 144-171.
- [29] H.L. Ngo and D.T. Luong, Tamed Euler-Maruyama approximation for stochastic differential equations with locally Hölder continuous diffusion coefficients, *Statistics and Probability Letters*, **145**, 2019, 133-140.
- [30] B. Oksendal, *Stochastic Differential Equations*, 6th Edition, Springer, Berlin, Heidelberg, 2003.
- [31] V. Reshniak, A.Q.M. Khaliq, D.A. Voss and G. Zhang, Split-step Milstein methods for multi-channel stiff stochastic differential systems, *Applied Numerical Mathematics*, **89**, 2015, 1-23.
- [32] M. Song, Y. Lu and M. Liu, Convergence of the Tamed Euler Method for Stochastic Differential Equations with Piecewise Continuous Arguments

Under Non-global Lipschitz Continuous Coefficients, *Numerical Functional Analysis and Optimization*, [39\(5\)](#), 2018, 517-536.

- [33] X. Wang and 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.
- [34] X. Wang, J. Wu and B. Dong, Mean-square convergence rates of stochastic theta methods for SDEs under a coupled monotonicity condition, *BIT Numerical Mathematics*, [60\(3\)](#), 2020, 759-790.
- [35] Z. Yan, A. Xiao and X. Tang, Strong convergence of the split-step theta method for neutral stochastic delay differential equations, *Applied Numerical Mathematics*, [120](#), 2017, 215-232.
- [36] Z. Zhang and H. Ma, Order-preserving strong schemes for SDEs with locally Lipschitz coefficients, *Applied Numerical Mathematics*, [112](#), 2017, 1-16.
- [37] W. Zhang, M. Song and M. Liu, Strong convergence of the partially truncated Euler–Maruyama method for a class of stochastic differential delay equations, *Journal of Computational and Applied Mathematics*, [335](#), 2018, 114-128.
- [38] X. Zong, F. Wu and 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.
- [39] X. Zong, F. Wu and G. Xu, Convergence and stability of two classes of theta-Milstein schemes for stochastic differential equations, *Journal of Computational and Applied Mathematics*, [336](#), 2018, 8-29.