The Partially Truncated Euler–Maruyama Method and its Stability and Boundedness

Qian Guo[a], Wei Liu[a,*], Xuerong Mao[b], Rongxian Yue[a]

[a]Department of Mathematics, Shanghai Normal University, Shanghai, 200234, China.
[b]Department of Mathematics and Statistics, University of Strathclyde, Glasgow, G1 1XH, U.K.

Abstract

The partially truncated Euler–Maruyama (EM) method is proposed in this paper for highly nonlinear stochastic differential equations (SDEs). We will not only establish the finite-time strong L^r-convergence theory for the partially truncated EM method, but also demonstrate the real benefit of the method by showing that the method can preserve the asymptotic stability and boundedness of the underlying SDEs.

Keywords: Stochastic differential equation, local Lipschitz condition, Khasminskii-type condition, partially truncated Euler-Maruyama method, stability

1. Motivation

It is known (see, e.g., [13, 16, 17]) that the scalar stochastic differential equation (SDE)

$$dx(t) = -\left(x(t) + x^5(t)\right) dt + x^2(t) dB(t), \quad t \geq 0, \tag{1.1}$$

is exponentially stable in the mean square sense, where $B(t)$ is a scalar Brownian motion. More precisely, the solution satisfies

$$\mathbb{E}|x(t)|^2 \leq |x_0|^2 e^{-\frac{15t}{8}}, \quad t \geq 0, \tag{1.2}$$

*Corresponding author

Email addresses: qguo@shnu.edu.cn (Qian Guo), weiliu@shnu.edu.cn (Wei Liu), x.mao@strath.ac.uk (Xuerong Mao), yue2@shnu.edu.cn (Rongxian Yue)

for any initial value $x(0) = x_0 \in \mathbb{R}$ (see Example 4.4 below). It is also known (see, e.g., [9, 11]) that the (classical) Euler–Maruyama (EM) method may not preserve the exponential stability in the mean square sense (see, e.g., [14, 18] for the EM method).

Recently, the truncated EM method was developed in [20, 21], where the finite-time strong convergence theory was established and the order of L^q-convergence was shown to be arbitrarily close to $q/2$ for a class of SDEs including the underlying SDE (1.1). We therefore wonder if the truncated EM method can preserve the mean square exponential stability of the underlying SDE (1.1).

To apply the truncated EM method for a given step size Δ, we need to truncate the drift coefficient $f(x) = -x - x^5$ and the diffusion coefficient $g(x) = x^2$ into

$$f_\Delta(x) = f(\pi_\Delta(x)) \quad \text{and} \quad g_\Delta(x) = g(\pi_\Delta(x)),$$

where $\pi_\Delta(x) = (|x| \wedge \mu^{-1}(h(\Delta)))x/|x|$ and both functions μ^{-1} and h will be explained in the next section. The truncated EM solution is then obtained by applying the EM method to the truncated SDE

$$dx(t) = f_\Delta(x(t))dt + g_\Delta(x(t))dB(t).$$

In other words, the truncated EM solution is formed by setting $X_0 = x_0$ and computing

$$X_{k+1} = X_k + f_\Delta(X_k)\Delta + g_\Delta(X_k)\Delta B_k, \quad k \geq 0.$$

When we try to show if this truncated EM solution is exponentially stable in the mean square sense for all sufficiently small step size Δ, we note the following factor: the drift coefficient contains the fifth power term $-x^5$ and the linear term $-x$ while the diffusion coefficient contains the square term x^2 but all these terms are truncated. We realise that it is necessary to truncate the the fifth power term $-x^5$ and the square term x^2; otherwise the EM solution will not converge to the true solution in the moment sense at a finite time (see, e.g., [9, 11]). However, we feel that it is unnecessary to truncate the linear term. In fact, from the finite-time-convergence point of view, the linear term does not cause any problem to the EM method and hence there is no point to truncate it. Moreover, from the stability point of view, it is this linear term that plays a key role for the mean square exponential stability of the underlying SDE (1.1). In other words, truncating the linear term spoils the stability feature of the underlying SDE (1.1). Based on these observations, we feel it is better to partially truncate the underlying SDE (1.1) into the following form

$$dx(t) = -(x(t) + (\pi_\Delta(x(t)))^5)dt + (\pi_\Delta(x(t)))^2 dB(t), \tag{1.3}$$

and then apply the EM method to this SDE to form the numerical solution: $X_0 = x_0$ and

$$X_{k+1} = X_k - (X_k + (\pi_\Delta(X_k))^5)\Delta + (\pi_\Delta(X_k))^2 \Delta B_k, \quad k \geq 0. \tag{1.4}$$

We shall see that this numerical solution does not only converge to the true solution at a finite time but it is also exponentially stable in the mean square sense for sufficiently small step size Δ. This example motivates us to propose the the partially truncated EM method in the next section.

It turns out that the partially truncated EM method can preserve the asymptotic boundedness of the SDEs. For example, consider the scalar stochastic Ginzburg–Landau equation (see, e.g., [5, 14])

$$dx(t) = (ax(t) - bx^3(t))dt + cx(t)dB(t), \tag{1.5}$$

where a, b, c are three positive numbers. It is known (see [22] or Example 5.4 below) that the second moment of the solution of this SDE is asymptotically bounded. It is also known (see, e.g., [9, 11]) that the EM method may not preserve this asymptotic boundedness. However, we will show that our partially truncated EM method can preserve this boundedness very well.

It needs to mention that several nice explicit methods have been developed recently for SDEs with both drift and diffusion coefficients growing super-linearly. The fully tamed Euler method is developed in [12]. A new explicit balanced scheme using sine functions to control the highly nonlinear terms is developed in [28] and the strong convergence order of 1/2 is obtained. The two-step BDF-Maruyama scheme of order 1/2 is proposed in [1]. The projected Euler scheme that uses a different truncating strategy is developed in [26]. Some general criteria on the convergence and the asymptotic stability of numerical methods are discussed in [10, 26, 27].

The convergence of numerical methods in other senses are interesting and important as well. In [2], the authors propose a new algorithm to approximate the laws of the solutions to a class of SDEs with irregular coefficients. The pathwise convergences of numerical methods with constant and adaptive step sizes for some highly non-linear SDEs are studied in [6] and [24], respectively. It is also interesting to see if these methods could preserve asymptotic properties of the underlying SDEs in their corresponding senses.

The main contribution of this paper is to prove that the partially truncated EM method is able to preserve the mean square exponential stability and asymptotic boundedness of underlying SDEs, both of whose drift and diffusion coefficients are allowed to grow super-linearly.

Let us begin to develop our partially truncated EM method and demonstrate its real benefits.

2. The partially truncated EM method

Throughout this paper, unless otherwise specified, we will use the following notation. If A is a vector or matrix, its transpose is denoted by A^T. If $x \in \mathbb{R}^d$, then $|x|$ is the Euclidean norm. If A is a matrix, we let $|A| = \sqrt{\text{trace}(A^T A)}$ be its trace norm. If A is a symmetric matrix, denote by $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ its largest and smallest eigenvalue, respectively. For two real numbers a and b, we use $a \vee b = \max(a, b)$ and $a \wedge b = \min(a, b)$. If D is a set, its indicator function is denoted by I_D, namely $I_D(x) = 1$ if $x \in D$ and 0 otherwise. Moreover, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions (that is, it is right continuous and increasing while $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets), and let $\mathbb{E}$ denote the expectation corresponding to $\mathbb{P}$. Let $B(t)$

be an m-dimensional Brownian motion defined on the space.

Consider a d-dimensional SDE

$$dx(t) = f(x(t))dt + g(x(t))dB(t) \tag{2.1}$$

on $t \geq 0$ with the initial value $x(0) = x_0 \in \mathbb{R}^d$, where

$$f : \mathbb{R}^d \to \mathbb{R}^d \quad \text{and} \quad g : \mathbb{R}^d \to \mathbb{R}^{d\times m}.$$

We assume that f and g can be decomposed as

$$f(x) = F_1(x) + F(x) \quad \text{and} \quad g(x) = G_1(x) + G(x), \tag{2.2}$$

where $F_1, F : \mathbb{R}^d \to \mathbb{R}^d$ and $G_1, G : \mathbb{R}^d \to \mathbb{R}^{d\times m}$. We also impose three standing hypotheses.

Assumption 2.1. *Assume that the coefficients $F_1,\ F,\ G_1,\ G$ satisfy the following conditions: there are constants $L_1 > 0$ and $r \geq 0$ such that*

$$|F_1(x) - F_1(y)| \vee |G_1(x) - G_1(y)| \leq L_1|x - y| \tag{2.3}$$

and

$$|F(x) - F(y)| \vee |G(x) - G(y)| \leq L_1(1 + |x|^\gamma + |y|^\gamma)|x - y| \tag{2.4}$$

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

We can derive from (2.3) that the coefficients F_1 and G_1 satisfy the linear growth condition that there exists a constant $K_1 > 0$ such that

$$|F_1(x)| \vee |G_1(x)| \leq K_1(1 + |x|) \tag{2.5}$$

for all $x \in \mathbb{R}^d$.

Assumption 2.2. *Assume that the coefficients F and G satisfy the following condition: there is a pair of constants $\bar{r} > 2$ and L_2 such that*

$$(x - y)^T(F(x) - F(y)) + \frac{\bar{r} - 1}{2}|G(x) - G(y)|^2 \leq L_2|x - y|^2 \tag{2.6}$$

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

Assumption 2.3. *Assume that the coefficients F and G satisfy the Khasminskii-type condition: there is a pair of constants $\bar{p} > \bar{r}$ and $K_2 > 0$ such that*

$$x^T F(x) + \frac{\bar{p}-1}{2}|G(x)|^2 \leq K_2(1+|x|^2) \tag{2.7}$$

for all $x \in \mathbb{R}^d$.

Indeed, (2.7) can be indicated by (2.6). But this approach may force $\bar{p}$ to be less than $\bar{r}$, which is not necessary. We will see it by the example in Section 3.2.

We derive from (2.5) and (2.7) that for any $p \in (2, \bar{p})$,

$$\begin{aligned} & x^T f(x) + \frac{p-1}{2}|g(x)|^2 \\ \leq\ & x^T(F_1(x) + F(x)) + \frac{p-1}{2}(|G_1(x)|^2 + 2|G_1(x)||G(x)| + |G(x)|^2) \\ \leq\ & |x||F_1(x)| + x^T F(x) + \frac{p-1}{2}\Big(|G_1(x)|^2 + \frac{p-1}{\bar{p}-p}|G_1(x)|^2 + \frac{\bar{p}-p}{p-1}|G(x)|^2 + |G(x)|^2\Big) \\ =\ & |x||F_1(x)| + \frac{(p-1)(\bar{p}-1)}{2(\bar{p}-p)}|G_1(x)|^2 + x^T F(x) + \frac{\bar{p}-1}{2}|G(x)|^2 \\ \leq\ & K_3(1+|x|^2), \end{aligned} \tag{2.8}$$

where

$$K_3 = 2K_1 + K_2 + \frac{K_1^2(p-1)(\bar{p}-1)}{\bar{p}-p}.$$

In a similar manner, we can derive from (2.3) and (2.6) that for any $r \in (2, \bar{r})$

$$(x-y)^T(f(x)-f(y)) + \frac{r-1}{2}|g(x)-g(y)|^2 \leq L_3|x-y|^2, \tag{2.9}$$

where

$$L_3 = 2L_1 + L_2 + \frac{L_1^2(r-1)(\bar{r}-1)}{\bar{r}-r}.$$

We can therefore state a known result (see, e.g., [18, 25]) as a lemma for the use of this paper.

Lemma 2.4. *Under Assumptions 2.1, 2.2 and 2.3, the SDE (2.1) has a unique global solution $x(t)$ and, moreover, for any $p \in (2, \bar{p})$,*

$$\sup_{0 \leq t \leq T} \mathbb{E}|x(t)|^p < C, \quad \forall T > 0, \tag{2.10}$$

where, and from now on, C stands for generic positive real constants dependent on $T, \bar{p}, p, K_1,\ K_2, x_0$ but independent of the step size Δ (and R later) and its values may change between occurrences.

To define the partially truncated EM numerical solutions, we first choose a strictly increasing continuous function $\mu : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\mu(r) \to \infty$ as $r \to \infty$ and

$$\sup_{|x|\le r} \big(|F(x)| \vee |G(x)|\big) \le \mu(r), \quad \forall r \ge 1. \tag{2.11}$$

Denote by μ^{-1} the inverse function of μ and we see that μ^{-1} is a strictly increasing continuous function from $[\mu(0), \infty)$ to $\mathbb{R}_+$. We also choose a number $\Delta^* \in (0, 1]$ and a strictly decreasing function $h : (0, \Delta^*] \to (0, \infty)$ such that

$$h(\Delta^*) \ge \mu(1), \quad \lim_{\Delta\to 0} h(\Delta) = \infty \quad \text{and} \quad \Delta^{1/4} h(\Delta) \le 1, \quad \forall \Delta \in (0, 1). \tag{2.12}$$

For a given step size $\Delta \in (0, 1)$, let us define the mapping $\pi_\Delta : \mathbb{R}^d \to \mathbb{R}^d$ by

$$\pi_\Delta(x) = (|x| \wedge \mu^{-1}(h(\Delta)))\frac{x}{|x|},$$

where we set $x/|x| = 0$ when $x = 0$. We then define the truncated functions

$$F_\Delta(x) = F(\pi_\Delta(x)) \quad \text{and} \quad G_\Delta(x) = G(\pi_\Delta(x)) \tag{2.13}$$

for $x \in \mathbb{R}^d$. It is easy to see that

$$|F_\Delta(x)| \vee |G_\Delta(x)| \le \mu(\mu^{-1}(h(\Delta))) = h(\Delta) \quad \forall x \in \mathbb{R}^d. \tag{2.14}$$

That is, both truncated functions F_Δ and G_Δ are bounded. Moreover, these truncated functions preserve the Khasminskii-type condition (2.7) for all $\Delta \in (0, \Delta^*]$ as shown in [20] and we state it here as a lemma for the use of this paper.

Lemma 2.5. *Let Assumption 2.3 hold. Then, for all $\Delta \in (0, \Delta^*]$, we have*

$$x^T F_\Delta(x) + \frac{\bar{p}-1}{2}|G_\Delta(x)|^2 \le 2K_2(1 + |x|^2), \quad \forall x \in \mathbb{R}^d. \tag{2.15}$$

In the same way as (2.8) was proved, we can show that for any $p \in (2, \bar{p})$,

$$x^T(F_1(x) + F_\Delta(x)) + \frac{p-1}{2}|G_1(x) + G_\Delta(x)|^2 \le K_4(1 + |x|^2) \tag{2.16}$$

for all $x \in \mathbb{R}^d$, where

$$K_4 = 2K_1 + 2K_2 + \frac{K_1^2(p-1)(\bar{p}-1)}{\bar{p}-p}.$$

The discrete-time partially truncated EM numerical solutions $X_\Delta(t_k) \approx x(t_k)$ for $t_k = k\Delta$ are formed by setting $X_\Delta(0) = x_0$ and computing

$$X_\Delta(t_{k+1}) = X_\Delta(t_k)+[F_1(X_\Delta(t_k))+F_\Delta(X_\Delta(t_k))]\Delta+[G_1(X_\Delta(t_k))+G_\Delta(X_\Delta(t_k))]\Delta B_k, \tag{2.17}$$

for $k = 0, 1, \cdots$, where $\Delta B_k = B(t_{k+1}) - B(t_k)$. There are two versions of the continuous-time truncated EM solutions. The first one is defined by

$$\bar{x}_\Delta(t) = \sum_{k=0}^{\infty} X_\Delta(t_k) I_{[t_k, t_{k+1})}(t), \quad t \geq 0. \tag{2.18}$$

This is a simple step process so its sample paths are not continuous. We will refer this as the continuous-time step-process partially truncated EM solution. The other one is defined by

$$x_\Delta(t) = x_0 + \int_0^t [F_1(\bar{x}_\Delta(s)) + F_\Delta(\bar{x}_\Delta(s))]ds + \int_0^t [G_1(\bar{x}_\Delta(s)) + G_\Delta(\bar{x}_\Delta(s))]dB(s) \tag{2.19}$$

for $t \geq 0$. We will refer this as the continuous-time continuous-sample partially truncated EM solution. We observe that $x_\Delta(t_k) = \bar{x}_\Delta(t_k) = X_\Delta(t_k)$ for all $k \geq 0$. Moreover, $x_\Delta(t)$ is an Itô process with its Itô differential

$$dx_\Delta(t) = [F_1(\bar{x}_\Delta(t)) + F_\Delta(\bar{x}_\Delta(t))]dt + [G_1(\bar{x}_\Delta(t)) + G_\Delta(\bar{x}_\Delta(t))]dB(t). \tag{2.20}$$

3. Finite-Time L^r-Convergence

This section is divided into two parts. The theoretical results of the strong convergence are proved in the first subsection and a manual of the method is presented in the second one.

3.1. Theoretical Results

In this part, we will fix $T > 0$ arbitrarily. The following theorem shows the strong L^r-convergence of the partially truncated EM method.

Theorem 3.1. *Let Assumptions 2.1, 2.2 and 2.3 hold. If $p > \bar{r}$, $2p > \bar{r}\gamma$ and for any $r \in [2, \bar{r})$*

$$h(\Delta) \geq \mu((\Delta^{r/2}(h(\Delta))^r)^{-1/(p-r)}), \tag{3.1}$$

then there is a $\bar{\Delta} \in (0, \Delta^*]$ such that for all $\Delta \in (0, \bar{\Delta}]$

$$\mathbb{E}|x_\Delta(T) - x(T)|^r \leq C\Delta^{r/2}(h(\Delta))^r \tag{3.2}$$

and

$$\mathbb{E}|\bar{x}_\Delta(T) - x(T)|^r \leq C\Delta^{r/2}(h(\Delta))^r. \tag{3.3}$$

We will prove this theorem in a similar fashion as [21, Theorem 3.8], so we need to establish a number of lemmas as in [21].

Lemma 3.2. *Let Assumptions 2.1, 2.2 and 2.3 hold and let $p \in (2, \bar{p})$ be arbitrary. Then*

$$\sup_{0<\Delta\leq\Delta^*} \sup_{0\leq t\leq T} \mathbb{E}|x_\Delta(t)|^p \leq C. \tag{3.4}$$

Proof. Fix any $\Delta \in (0, \Delta^*]$. By the Itô formula, we derive from (2.19) that, for $0 \leq t \leq T$,

$$\begin{aligned} & \mathbb{E}|x_\Delta(t)|^p - |x_0|^p \\ \leq\ & \mathbb{E}\int_0^t p|x_\Delta(s)|^{p-2}\Big(x_\Delta^T(s)[F_1(\bar{x}_\Delta(s)) + F_\Delta(\bar{x}_\Delta(s))] + \frac{p-1}{2}|G_1(\bar{x}_\Delta(s)) + G_\Delta(\bar{x}_\Delta(s))|^2\Big)ds \\ =\ & \mathbb{E}\int_0^t p|x_\Delta(s)|^{p-2}\Big(\bar{x}_\Delta^T(s)[F_1(\bar{x}_\Delta(s)) + F_\Delta(\bar{x}_\Delta(s))] + \frac{p-1}{2}|G_1(\bar{x}_\Delta(s)) + G_\Delta(\bar{x}_\Delta(s))|^2\Big)ds \\ +\ & \mathbb{E}\int_0^t p|x_\Delta(s)|^{p-2}(x_\Delta(s) - \bar{x}_\Delta(s))^T[F_1(\bar{x}_\Delta(s)) + F_\Delta(\bar{x}_\Delta(s))]ds. \end{aligned} \tag{3.5}$$

By (2.16), we then have

$$\mathbb{E}|x_\Delta(t)|^p - |x_0|^p \leq J_1 + J_2 + J_3, \tag{3.6}$$

where

$$J_1 = \mathbb{E}\int_0^t pK_4|x_\Delta(s)|^{p-2}(1 + |\bar{x}_\Delta(s)|^2)ds, \tag{3.7}$$

$$J_2 = \mathbb{E}\int_0^t p|x_\Delta(s)|^{p-2}|x_\Delta(s) - \bar{x}_\Delta(s)||F_1(\bar{x}_\Delta(s))|ds, \tag{3.8}$$

and

$$J_3 = \mathbb{E}\int_0^t p|x_\Delta(s)|^{p-2}|x_\Delta(s) - \bar{x}_\Delta(s)||F_\Delta(\bar{x}_\Delta(s))|ds. \tag{3.9}$$

By the Young inequality $a^\beta b^{1-\beta} \leq \beta a + (1-\beta)b$ for $a, b \geq 0$ and $\beta \in (0,1)$ as well as the elementary inequality $|x|^{p-2} \leq 1 + |x|^p$, we can show easily that

$$J_1 \leq C\Big(1 + \int_0^t (\mathbb{E}|x_\Delta(s)|^p + \mathbb{E}|\bar{x}_\Delta(s)|^p)ds\Big). \tag{3.10}$$

Similarly, by Assumption 2.1, we can show that

$$J_2 \leq C\Big(1 + \int_0^t (\mathbb{E}|x_\Delta(s)|^p + \mathbb{E}|\bar{x}_\Delta(s)|^p)ds\Big). \tag{3.11}$$

Moreover, by the Young inequality and (2.14), we derive

$$\begin{aligned} J_3 &\leq (p-2)\mathbb{E}\int_0^t |x_\Delta(s)|^p ds + 2\mathbb{E}\int_0^t |x_\Delta(s) - \bar{x}_\Delta(s)|^{p/2}|F_\Delta(\bar{x}_\Delta(s))|^{p/2}ds \\ &\leq (p-2)\int_0^t \mathbb{E}|x_\Delta(s)|^p ds + 2(h(\Delta))^{p/2}\int_0^t \mathbb{E}|x_\Delta(s) - \bar{x}_\Delta(s)|^{p/2}ds. \end{aligned} \tag{3.12}$$

On the other hand, for any $s \in [0, T]$, there is a unique $k \geq 0$ such that $t_k \leq s \leq t_{k+1}$. By Assumption 2.1, (2.14) and the properties of the Itô integral (see, e.g., [18]), we then derive from (2.19) that

$$\begin{aligned} &\mathbb{E}|x_\Delta(s) - \bar{x}_\Delta(s)|^{p/2} = \mathbb{E}|x_\Delta(s) - x_\Delta(t_k)|^{p/2} \\ = &\ \mathbb{E}\Big|\int_{t_k}^s [F_1(\bar{x}_\Delta(t_k)) + F_\Delta(\bar{x}_\Delta(t_k))]du + \int_{t_k}^t [G_1(\bar{x}_\Delta(t_k)) + G_\Delta(\bar{x}_\Delta(t_k))]dB(u)\Big|^{p/2} \\ \leq &\ C\Delta^{p/4}\Big(1 + \mathbb{E}|\bar{x}_\Delta(t_k)|^{p/2} + (h(\Delta))^{p/2}\Big) \\ = &\ C\Delta^{p/4}\Big(1 + \mathbb{E}|\bar{x}_\Delta(s)|^{p/2} + (h(\Delta))^{p/2}\Big). \end{aligned} \tag{3.13}$$

Substituting this into (3.12) and recalling (2.12), we get

$$\begin{aligned} J_3 &\leq (p-2)\int_0^t \mathbb{E}|x_\Delta(s)|^p ds + 2C(h(\Delta))^{p/2}\Delta^{p/4}\int_0^t \Big(1 + \mathbb{E}|\bar{x}_\Delta(s)|^{p/2} + (h(\Delta))^{p/2}\Big)ds \\ &\leq C\Big(1 + \int_0^t (\mathbb{E}|x_\Delta(s)|^p + \mathbb{E}|\bar{x}_\Delta(s)|^p)ds\Big). \end{aligned} \tag{3.14}$$

Substituting (3.10), (3.11) and (3.14) into (3.6), we have

$$\mathbb{E}|x_\Delta(t)|^p \leq C\Big(1 + \int_0^t \sup_{0\leq u\leq s} \mathbb{E}|x_\Delta(u)|^p ds\Big).$$

As this holds for any $t \in [0, T]$ while the right-hand side is non-decreasing in t, we then see

$$\sup_{0 \le u \le t} \mathbb{E}|x_\Delta(u)|^p \le C\Big(1 + \int_0^t \sup_{0 \le u \le s} \mathbb{E}|x_\Delta(u)|^p ds\Big).$$

The well-known Gronwall inequality yields that

$$\sup_{0 \le u \le T} \mathbb{E}|x_\Delta(u)|^p \le C.$$

As this holds for any $\Delta \in (0, \Delta^*]$ while C is independent of Δ, we see the required assertion (3.4). □

The following lemma shows that $x_\Delta(t)$ and $\bar{x}_\Delta(t)$ are close to each other in the sense of L^p.

Lemma 3.3. *Let Assumptions 2.1, 2.2 and 2.3 hold and let $p \in (2, \bar{p})$ be arbitrary. Then there is a $\bar{\Delta} \in (0, \Delta^*]$ such that for all $\Delta \in (0, \bar{\Delta}]$,*

$$\mathbb{E}|x_\Delta(t) - \bar{x}_\Delta(t)|^p \le C\Delta^{p/2}(h(\Delta))^p, \quad \forall t \in [0, T]. \tag{3.15}$$

Consequently

$$\lim_{\Delta \to 0} \mathbb{E}|x_\Delta(t) - \bar{x}_\Delta(t)|^p = 0. \tag{3.16}$$

Proof. By Lemma 3.2, there is a $\bar{\Delta} \in (0, \Delta^*]$ such that

$$\sup_{0 < \Delta \le \bar{\Delta}} \sup_{0 \le t \le T} \mathbb{E}|x_\Delta(t)|^p \le C. \tag{3.17}$$

Now, fix any $\Delta \in (0, \bar{\Delta}]$. For any $t \in [0, T]$, there is a unique $k \ge 0$ such that $t_k \le t \le t_{k+1}$. In the same way as (3.13) was proved, we can then show

$$\mathbb{E}|x_\Delta(t) - \bar{x}_\Delta(t)|^p \le C\Delta^{p/2}\Big(1 + \mathbb{E}|\bar{x}_\Delta(t)|^p + (h(\Delta))^p\Big).$$

By (3.17), we therefore have

$$\mathbb{E}|x_\Delta(t) - \bar{x}_\Delta(t)|^p \le C\Delta^{p/2}(h(\Delta))^p,$$

which is (3.15). Noting from (2.12) that $\Delta^{p/2}(h(\Delta))^p \le \Delta^{p/4}$, we obtain (3.16) from (3.15) immediately. □

Let us now cite another lemma from [20, Lemma 3.3].

Lemma 3.4. *Let Assumptions 2.1, 2.2 and 2.3 hold. For any real number $R > |x_0|$, define the stopping time*

$$\tau_R = \inf\{t \geq 0 : |x(t)| \geq R\},$$

where throughout this paper we set $\inf \emptyset = \infty$ *(and as usual* $\emptyset$ *denotes the empty set). Then*

$$\mathbb{P}(\tau_R \leq T) \leq \frac{C}{R^p}. \tag{3.18}$$

(Recall that C stands for generic positive real constants independent of Δ and R.)

The following lemma can be proved in the same way as [20, Lemma 3.4] was proved.

Lemma 3.5. *Let Assumptions 2.1, 2.2 and 2.3 hold. For any real number $R > |x_0|$ and $\Delta \in (0, \bar{\Delta}]$ (the same $\bar{\Delta}$ as in Lemma 3.3), define the stopping time*

$$\rho_{\Delta,R} = \inf\{t \geq 0 : |x_\Delta(t)| \geq R\}.$$

Then

$$\mathbb{P}(\rho_{\Delta,R} \leq T) \leq \frac{C}{R^p}. \tag{3.19}$$

We can now prove Theorem 3.1. As the proof is in a similar fashion as [20, Theorem 3.5] was proved so we only highlight the different parts.

Proof of Theorem 3.1.

Let $\varepsilon > 0$ be arbitrary. Let τ_R and $\rho_{\Delta,R}$ be the same as the definitions in Lemmas 3.4 and 3.5. Set

$$\theta_{\Delta,R} = \tau_R \wedge \rho_{\Delta,R} \quad \text{and} \quad e_\Delta(T) = x_\Delta(T) - x(T).$$

For a sufficiently large $R > |x(0)|$, we have that

$$\mathbb{E}|e_\Delta(T)|^r = \mathbb{E}\Big(|e_\Delta(T)|^r I_{\{\theta_{\Delta,R} > T\}}\Big) + \mathbb{E}\Big(|e_\Delta(T)|^r I_{\{\theta_{\Delta,R} \leq T\}}\Big). \tag{3.20}$$

For any $\delta > 0$, using the Young inequality we obtain that

$$\mathbb{E}\Big(|e_\Delta(T)|^r I_{\{\theta_{\Delta,R} \leq T\}}\Big) \leq \frac{r\delta}{p}\mathbb{E}|e_\Delta(T)|^p + \frac{p-r}{p\delta^{r/(p-r)}}\mathbb{P}(\theta_{\Delta,R} \leq T). \tag{3.21}$$

Applying Lemmas 2.4 and 3.2, we can see that

$$\mathbb{E}|e_\Delta(T)|^p \le 2^{p-1}\mathbb{E}|x(T)|^p + 2^{p-1}\mathbb{E}|x_\Delta(T)|^p \le C.$$

Using Lemmas 3.4 and 3.5, we obtain that

$$\mathbb{P}(\theta_{\Delta,R} \le T) \le \mathbb{P}(\tau_R \le T) + \mathbb{P}(\rho_{\Delta,R} \le T) \le \frac{C}{R^p}.$$

Substituting the two estimates above back into (3.21), and choosing $\delta = \Delta^{r/2}(h(\Delta))^r$ and $R = (\Delta^{r/2}(h(\Delta))^r)^{-1/(p-r)}$ we have that

$$\mathbb{E}\Big(|e_\Delta(T)|^r I_{\{\theta_{\Delta,R} \le T\}}\Big) \le C\Delta^{r/2}(h(\Delta))^r. \tag{3.22}$$

In the same way as the proof of Lemma 3.7 in [21], we can show that

$$\mathbb{E}\Big(|e_\Delta(T \wedge \theta_{\Delta,R})|^r\Big) \le C\Delta^{r/2}(h(\Delta))^r. \tag{3.23}$$

By (3.1), we can see that

$$\mu^{-1}(h(\Delta)) \ge (\Delta^{r/2}(h(\Delta))^r)^{-1/(p-r)} = R.$$

Therefore, substituting (3.22) and (3.23) into (3.20) yields (3.2). In addition, (3.2) together with Lemma 3.3 indicates (3.3). □

3.2. A Manual of the Method

We demonstrate the process of implementing the partially truncated EM by the following example.

Example 3.6. Consider a nonlinear test scalar SDE

$$dx(t) = (x(t) - x^5(t))dt + x^2(t)dB(t), \quad t \ge 0,$$

with the initial value $x(0) = 1$. It can be seen that $F_1(x) = x$, $F(x) = -x^5$, $G_1(x) = 0$ and $G(x) = x^2$.

Step 1. Check the assumptions

Assumption 2.1 holds clearly. For Assumption 2.2, it is straightforward to see that

$$
\begin{aligned}
&(x-y)(F(x)-F(y)) + \frac{\bar r - 1}{2}|G(x)-G(y)|^2 \\
=& (x-y)\big[-(x-y)(x^4+x^3y+x^2y^2+xy^3+y^4)\big] + \frac{\bar r - 1}{2}(x+y)^2(x-y)^2 \\
=& \big[-(x^4+x^3y+x^2y^2+xy^3+y^4) + \frac{\bar r - 1}{2}(x+y)^2\big]|x-y|^2.
\end{aligned}
$$

But

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

Hence

$$
\begin{aligned}
&(x-y)(F(x)-F(y)) + \frac{\bar r - 1}{2}|G(x)-G(y)|^2 \\
\le& \big[-0.5(x^4+y^4) + \frac{\bar r - 1}{2}(x^2+y^2)\big]|x-y|^2 \\
\le& \big[1 + \frac{(\bar r - 1)^2}{4}\big]|x-y|^2.
\end{aligned}
$$

In other words, Assumption 2.2 is also fulfilled for any $\bar r$. Moreover,

$$
\begin{aligned}
xF(x) + \frac{\bar p - 1}{2}|G(x)|^2 &= -x^6 + \frac{\bar p - 1}{2}|x^2|^2 \\
&= -x^2(x^2 - \frac{\bar p - 1}{4})^2 + \frac{(\bar p - 1)^2}{16}x^2 \le \frac{(\bar p - 1)^2}{16}x^2,
\end{aligned}
$$

i.e. Assumption 2.3 is satisfied for any $\bar p$.

Step 2. Choose $\mu(\cdot)$ and $h(\cdot)$

According to (2.11), we set $\mu(r) = r^5$ such that

$$
\sup_{|x|\le r}\big(|F(x)| \vee |G(x)|\big) = \sup_{|x|\le r}\big(|x|^5 \vee |x|^2\big) \le r^5, \quad \forall r \ge 1.
$$

We set $h(\Delta) = \Delta^{-1/10}$, then all the conditions in (2.12) hold for all $\Delta^* \in (0,1]$.[1]

Step 3. Define $F_\Delta(x)$ and $G_\Delta(x)$

[1] One may notice that the choices of both $\mu(\cdot)$ and $h(\cdot)$ are not unique and we do not know if there are optimal choices currently.

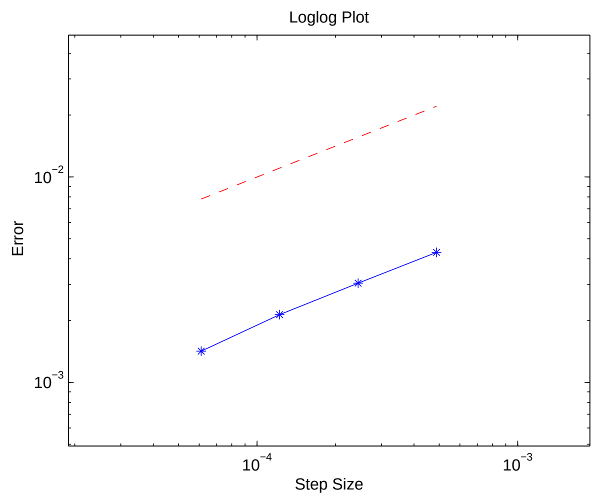

Figure 1: The strong convergence order at the terminal time $T = 2$. The red dashed line is the reference line with the slope of $1/2$.

From *Step 2*, we can see the truncating factor is defined as $\mu^{-1}(h(\Delta)) = \Delta^{-1/50}$. Then according to (2.13), $F_\Delta(x)$ and $G_\Delta(x)$ are defined as

$$F_\Delta(x) = F((|x| \wedge \Delta^{-1/50})\frac{x}{|x|}) \quad \text{and} \quad G_\Delta(x) = G((|x| \wedge \Delta^{-1/50})\frac{x}{|x|}).$$

Step 4. Calculation in each iteration

For the given step size Δ and X_k, we compare $|X_k|$ and $\Delta^{-1/50}$. Then substituting the product of the smaller one and $X_k/|X_k|$ into $F(\cdot)$ and $G(\cdot)$ yields $F_\Delta(X_k)$ and $G_\Delta(X_k)$. The X_{k+1} is calculated by

$$X_{k+1} = X_k + (F_1(X_k) + F_\Delta(X_k))\Delta + G_\Delta(X_k)\Delta B_k.$$

Figure 1 displays the L^1 errors at the time $T = 2$ with step sizes 2^{-14}, 2^{-13}, 2^{-12} and 2^{-11}. The simulations with step size 2^{-17} are regarded as the true solutions. For each step size, 1000 paths are simulated. Compared with the red dashed reference line, strong convergence order of the partially truncated Euler-Maruyama method is approximately $1/2$, which is in line with the theoretical result.

4. Stability

Finite-time convergence is a fundamental property for a numerical method. However, a nice numerical method for an SDE should also preserve some asymptotic properties of the underlying SDE, for example, stability and boundedness (see, e.g., [4, 8, 9, 15, 19, 23]).

In this section we will show that the partially truncated EM method can preserve the mean square exponential stability of the underlying SDE (2.1). We will let Assumptions 2.1–2.3 be the standing hypotheses so we will not mention them explicitly in the theorems in this section. Moreover, for the stability purpose, we also assume in this section that

$$F_1(0) = F(0) = 0, \quad G_1(0) = G(0) = 0. \tag{4.1}$$

So the linear growth condition (2.5) becomes

$$|F_1(x)| \vee |G_1(x)| \le K_1|x|. \tag{4.2}$$

Our main assumption in this section is the following one.

Assumption 4.1. *Assume that there are constants $\theta \in [0, \infty]$ and $\lambda_1 > \lambda_2 \ge 0$ such that*

$$2x^T F_1(x) + (1+\theta)|G_1(x)|^2 \le -\lambda_1|x|^2 \tag{4.3}$$

and

$$2x^T F(x) + (1+\theta^{-1})|G(x)|^2 \le \lambda_2|x|^2 \tag{4.4}$$

for all $x \in \mathbb{R}^d$, where throughout the remaining part of this paper we choose $\theta = 0$ and set $\theta^{-1}|G(x)|^2 = 0$ when there is no $G(x)$ term in $g(x)$, while choose $\theta = \infty$ and set $\theta|G_1(x)|^2 = 0$ when there is no $G_1(x)$ term in $g(x)$.

This assumption implies

$$2x^T f(x) + |g(x)|^2 \le -(\lambda_1 - \lambda_2)|x|^2, \quad x \in \mathbb{R}^d. \tag{4.5}$$

It is therefore known (see, e.g., [13, 16, 17]) that the SDE (2.1) is exponentially stable in the mean square sense. To be precise, we state it as a theorem.

Theorem 4.2. *Let Assumption 4.1 hold. Then for any initial value $x_0 \in \mathbb{R}^d$, the solution of the SDE (2.1) satisfies*

$$\mathbb{E}|x(t)|^2 \le |x_0|^2 e^{-(\lambda_1-\lambda_2)t}, \quad \forall t \ge 0. \tag{4.6}$$

The following theorem shows that the partially truncated EM method can preserve this mean square exponential stability perfectly.

Theorem 4.3. *Let Assumption 4.1 hold. Then for any $\varepsilon \in (0, \lambda_1 - \lambda_2)$, there is a $\hat{\Delta} \in (0, \Delta^*)$ such that for every $\Delta \in (0, \hat{\Delta})$ and any initial value $x_0 \in \mathbb{R}^d$, the solution of the partially truncated EM method (2.17) satisfies*

$$\mathbb{E}|X_\Delta(t_k)|^2 \le |x_0|^2 e^{-(\lambda_1-\lambda_2-\varepsilon)t_k}, \quad \forall k \ge 0. \tag{4.7}$$

Proof. To simplify the notation, we define, in the remaining part of this paper,

$$f_\Delta(x) = F_1(x) + F_\Delta(x) \quad \text{and} \quad g_\Delta(x) = G_1(x) + G_\Delta(x), \quad x \in \mathbb{R}^d,$$

for every $\Delta \in (0, \Delta^*]$. We first show that these functions preserve property (4.5) perfectly in the sense that

$$2x^T f_\Delta(x) + |g_\Delta(x)|^2 \le -(\lambda_1 - \lambda_2)|x|^2, \quad x \in \mathbb{R}^d. \tag{4.8}$$

In fact, this holds obviously for $x \in \mathbb{R}^d$ with $|x| \le \mu^{-1}(h(\Delta))$. For $x \in \mathbb{R}^d$ with $|x| > \mu^{-1}(h(\Delta))$, we derive, by Assumption 4.1,

$$\begin{aligned}
& 2x^T f_\Delta(x) + |g_\Delta(x)|^2 \\
\le\ & 2x^T F_1(x) + (1+\theta)|G_1(x)|^2 + 2x^T F(\pi_\Delta(x)) + (1+\theta^{-1})|G(\pi_\Delta(x))|^2 \\
\le\ & -\lambda_1|x|^2 + 2(x-\pi_\Delta(x))^T F(\pi_\Delta(x)) + 2(\pi_\Delta(x))^T F(\pi_\Delta(x)) + (1+\theta^{-1})|G(\pi_\Delta(x))|^2 \\
\le\ & -\lambda_1|x|^2 + \lambda_2|\pi_\Delta(x)|^2 + 2(x-\pi_\Delta(x))^T F(\pi_\Delta(x)).
\end{aligned} \tag{4.9}$$

But, by Assumption 4.1 again,

$$\begin{aligned}
2(x-\pi_\Delta(x))^T F(\pi_\Delta(x)) &= 2[|x|/\mu^{-1}(h(\Delta)) - 1](\pi_\Delta(x))^T F(\pi_\Delta(x)) \\
&\le [|x|/\mu^{-1}(h(\Delta)) - 1]\lambda_2|\pi_\Delta(x)|^2.
\end{aligned}$$

Substituting this into (4.9) and noting that $|\pi_\Delta(x)| = \mu^{-1}(h(\Delta))$, we get

$$2x^T f_\Delta(x) + |g_\Delta(x)|^2 \le -\lambda_1|x|^2 + \lambda_2|x||\pi_\Delta(x)| \le -(\lambda_1 - \lambda_2)|x|^2. \tag{4.10}$$

Fix $x_0 \in \mathbb{R}^d$ arbitrarily. For any $\Delta \in (0, \Delta^*]$, we can easily obtain from (2.17) that

$$\begin{aligned} \mathbb{E}|X_\Delta(t_{k+1})|^2 &= \mathbb{E}\Big(|X_\Delta(t_k)|^2 + |f_\Delta(X_\Delta(t_k))|^2\Delta^2 + |g_\Delta(X_\Delta(t_k))\Delta B_k|^2 \\ &+ 2X_\Delta(t_k)^T f_\Delta(X_\Delta(t_k))\Delta\Big) \end{aligned} \tag{4.11}$$

for $k = 0, 1, \cdots$. But

$$\begin{aligned} \mathbb{E}(|g_\Delta(X_\Delta(t_k))\Delta B_k|^2) &= \mathbb{E}\big(\text{trace}\big[g_\Delta(X_\Delta(t_k))\Delta B_k \Delta B_k^T g_\Delta(X_\Delta(t_k))^T\big]\big) \\ &= \mathbb{E}\Big(\mathbb{E}\big(\text{trace}\big[g_\Delta(X_\Delta(t_k))\Delta B_k \Delta B_k^T g_\Delta(X_\Delta(t_k))^T\big]\big|\mathcal{F}_{t_k}\big)\Big) \\ &= \mathbb{E}\Big(\text{trace}\big[g_\Delta(X_\Delta(t_k))\mathbb{E}\big(\Delta B_k \Delta B_k^T\big|\mathcal{F}_{t_k}\big) g_\Delta(X_\Delta(t_k))^T\big]\Big) \\ &= \mathbb{E}\Big(\text{trace}\big[g_\Delta(X_\Delta(t_k))\Delta I_m g_\Delta(X_\Delta(t_k))^T\big]\Big) \\ &= \Delta\mathbb{E}|g_\Delta(X_\Delta(t_k))|^2, \end{aligned}$$

where I_m denotes the $m \times m$ identity matrix. Substituting this into (4.11) yields

$$\begin{aligned} \mathbb{E}|X_\Delta(t_{k+1})|^2 &= \mathbb{E}\Big(|X_\Delta(t_k)|^2 + |f_\Delta(X_\Delta(t_k))|^2\Delta^2 + |g_\Delta(X_\Delta(t_k))|^2\Delta \\ &+ 2X_\Delta(t_k)^T f_\Delta(X_\Delta(t_k))\Delta\Big). \end{aligned} \tag{4.12}$$

Using (4.8), we get

$$\mathbb{E}|X_\Delta(t_{k+1})|^2 \le (1 - (\lambda_1 - \lambda_2)\Delta)\mathbb{E}|X_\Delta(t_k)|^2 + \Delta^2\mathbb{E}|f_\Delta(X_\Delta(t_k))|^2. \tag{4.13}$$

Now, by (4.2), we have

$$|f_\Delta(x)|^2 \le 2K_1^2|x|^2 + 2|F_\Delta(x)|^2, \quad \forall x \in \mathbb{R}^d.$$

But, by (2.4) and (4.1), we have

$$|F_\Delta(x)|^2 \le 4L_1^2|x|^2 \quad \text{if } |x| \le 1$$

while

$$|F_\Delta(x)|^2 \le h^2(\Delta) \le h^2(\Delta)|x|^2 \quad \text{if } |x| > 1.$$

We hence always have

$$|f_\Delta(x)|^2 \le 2(K_1^2 + 4L_1^2 + h^2(\Delta))|x|^2, \quad \forall x \in \mathbb{R}^d.$$

Recalling (2.12), we see that for any $\varepsilon \in (0, \lambda_1 - \lambda_2)$, there is a $\hat{\Delta} \in (0, \Delta^*)$ sufficiently small such that for all $\Delta \in (0, \hat{\Delta})$, $(\lambda_1 - \lambda_2 - \varepsilon)\Delta < 1$ and

$$\Delta |f_\Delta(x)|^2 \le \varepsilon |x|^2, \quad \forall x \in \mathbb{R}^d. \tag{4.14}$$

For each such Δ, we hence obtain from (4.13) and (4.14) that

$$\mathbb{E}|X_\Delta(t_{k+1})|^2 \le (1 - (\lambda_1 - \lambda_2 - \varepsilon)\Delta)\mathbb{E}|X_\Delta(t_k)|^2 \le |x_0|^2 (1 - (\lambda_1 - \lambda_2 - \varepsilon)\Delta)^{k+1}. \tag{4.15}$$

By the elementary inequality

$$1 - (\lambda_1 - \lambda_2 - \varepsilon)\Delta \le e^{-(\lambda_1 - \lambda_2 - \varepsilon)\Delta},$$

we further have

$$\mathbb{E}|X_\Delta(t_{k+1})|^2 \le |x_0|^2 e^{-(\lambda_1 - \lambda_2 - \varepsilon)t_{k+1}}, \tag{4.16}$$

which is the desired assertion (4.7). The proof is complete. $\Box$

Example 4.4. Let us return to the scalar SDE (1.1), namely

$$dx(t) = -(x(t) + x^5(t))dt + x^2(t)dB(t), \quad t \ge 0, \tag{4.17}$$

with the initial value $x(0) = x_0 \in \mathbb{R}$, where $B(t)$ is a scalar Brownian motion. We decompose the coefficients $f(x)$ and $g(x)$ in the form of (2.2) with

$$F_1(x) = -x, \quad F(x) = -x^5, \quad G_1(x) = 0, \quad G(x) = x^2$$

for $x \in \mathbb{R}$. Choosing $\theta = \infty$, we then have

$$2x^T F_1(x) + (1 + \theta)|G_1(x)|^2 = -2|x|^2$$

and

$$2x^T F(x) + (1 + \theta^{-1})|G(x)|^2 = -2x^6 + x^4.$$

But

$$-2x^6 + x^4 = -\Big(2x^6 - x^4 + \frac{1}{8}x^2\Big) + \frac{1}{8}x^2 = -2x^2\Big(x^2 - \frac{1}{4}\Big)^2 + \frac{1}{8}x^2 \le \frac{1}{8}x^2.$$

In other words, Assumption 4.1 is satisfied with $\lambda_1 = 2$ and $\lambda_2 = 1/8$. By Theorem 4.2, the SDE (4.17) is exponentially stable in the mean square sense, namely, for any initial value $x_0 \in \mathbb{R}$, the solution of the SDE (4.17) satisfies

$$\mathbb{E}|x(t)|^2 \le |x_0|^2 e^{-\frac{15t}{8}}, \quad \forall t \ge 0. \tag{4.18}$$

It is also known (see, e.g., [9, 11]) that the EM method might not preserve this mean square exponential stability. However, our new partially truncated EM method does preserve this stability perfectly. In fact, it is easy to see that our standing hypotheses, Assumption 2.1 is satisfied. Assumption 2.2 can be verified in the same way as that in Example 3.6. Moreover, for any $\bar{p} > 2$,

$$x^T F(x) + \frac{\bar{p}-1}{2}|G(x)|^2 = -x^6 + \frac{\bar{p}-1}{2}x^4$$

which is bounded above in $x \in \mathbb{R}$. In other words, Assumption 2.3 is also satisfied for any $\bar{p} > 2$. We can choose $\mu(r) = r^5$ and $h(\Delta) = \Delta^{-1/4}$ to define the numerical solution $X_\Delta(t_k)$ by the partially truncated EM method (2.17). By Theorem 3.1, this numerical solution will converge to the true solution in L^r for any $r \ge 2$ at any finite time. Moreover, by Theorem 4.3, we can also conclude that for any $\varepsilon \in (0, 15/8)$, there is a positive number $\hat{\Delta}$ such that for every $\Delta \in (0, \hat{\Delta})$ and any initial value $x_0 \in \mathbb{R}^d$, this numerical solution satisfies

$$\mathbb{E}|X_\Delta(t_k)|^2 \le |x_0|^2 e^{-(15/8-\varepsilon)t_k}, \quad \forall k \ge 0. \tag{4.19}$$

Figure 2 displays the asymptotic behaviour of the equation (4.17). The lower plot shows that the second moment of the partially truncated Euler-Maruyama method tends to zero as the time advances. In addition, the behaviour of the pathwise asymptotic stability can also be observed from the upper plot.

5. Boundedness

Although the stability of numerical methods for SDEs has been studied intensively (see, e.g., [4, 8, 9, 19, 23]), there are only a few papers on the

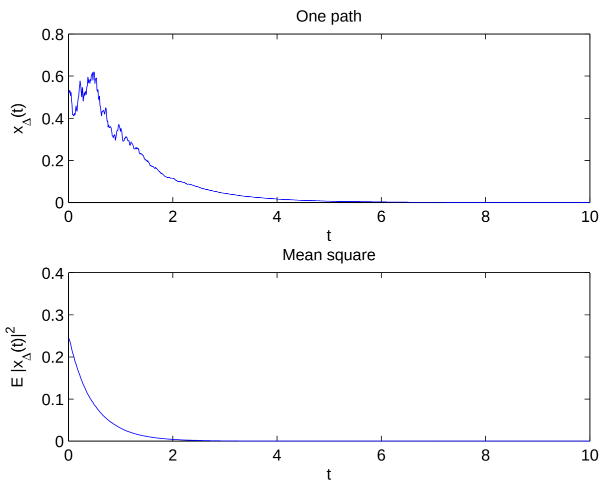

Figure 2: The upper plot is the simulation of one path and the lower one is the mean square of 1000 paths.

asymptotic boundedness of numerical methods (see, e.g., [15]).

In this section we will show that the partially truncated EM method can preserve the asymptotic boundedness of the underlying SDE (2.1). As in the previous section, we let Assumptions 2.1–2.3 be the standing hypotheses so we will not mention them explicitly in the theorems in this section. Of course we will no longer need condition (4.1) and Assumption 4.1 in this section. The main assumption in this section is the following one.

Assumption 5.1. *Assume that there are constants $\theta \in [0,\infty]$, $\alpha_1, \alpha_2 \ge 0$ and $\beta_1 > \beta_2 \ge 0$ such that*

$$2x^T F_1(x) + (1+\theta)|G_1(x)|^2 \le \alpha_1 - \beta_1 |x|^2 \tag{5.1}$$

and

$$2x^T F(x) + (1+\theta^{-1})|G(x)|^2 \le \alpha_2 + \beta_2 |x|^2 \tag{5.2}$$

for all $x \in \mathbb{R}^d$.

This assumption implies

$$2x^T f(x) + |g(x)|^2 \le \alpha_1 + \alpha_2 - (\beta_1 - \beta_2)|x|^2, \quad x \in \mathbb{R}^d. \tag{5.3}$$

We can hence state a theorem which follows easily from [22, Theorem 5.2 on page 157].

Theorem 5.2. *Let Assumption 5.1 hold. Then for any initial value $x_0 \in \mathbb{R}^d$, the solution of the SDE (2.1) satisfies*

$$\limsup_{t\to\infty} \mathbb{E}|x(t)|^2 \le \frac{\alpha_1+\alpha_2}{\beta_1-\beta_2}. \tag{5.4}$$

The following theorem shows that the partially truncated EM method can preserve this asymptotic boundedness perfectly.

Theorem 5.3. *Let Assumption 5.1 hold. Then for any $\varepsilon \in (0, \beta_1-\beta_2)$, there is a $\hat\Delta \in (0,\Delta^*)$ such that for every $\Delta \in (0,\hat\Delta)$ and any initial value $x_0 \in \mathbb{R}^d$, the solution of the partially truncated EM method (2.17) satisfies*

$$\limsup_{k\to\infty} \mathbb{E}|X_\Delta(t_k)|^2 \le \frac{\alpha_1+\alpha_2+\varepsilon}{\beta_1-\beta_2-\varepsilon}. \tag{5.5}$$

Proof. Fix $\varepsilon \in (0,\lambda_1-\lambda_2)$ arbitrarily. We first show that the functions f_Δ and g_Δ defined in the previous section preserve property (5.3) almost perfectly in the sense that

$$2x^T f_\Delta(x) + |g_\Delta(x)|^2 \le \alpha_1+\alpha_2-(\beta_1-\beta_2-0.5\varepsilon)|x|^2, \quad x\in\mathbb{R}^d, \tag{5.6}$$

as long as $\Delta \in (0,\hat\Delta_1)$, where $\hat\Delta_1 \in (0,\Delta^*)$ is sufficiently small for which

$$\frac{\alpha_2}{(\mu^{-1}(h(\hat\Delta_1)))^2} \le 0.5\varepsilon. \tag{5.7}$$

In fact, fix any $\Delta \in (0,\hat\Delta_1)$ and it is obvious that (5.6) holds for $x\in\mathbb{R}^d$ with $|x| \le \mu^{-1}(h(\Delta))$. For $x \in \mathbb{R}^d$ with $|x| > \mu^{-1}(h(\Delta))$, we derive, by Assumption 5.1,

$$\begin{aligned}
& 2x^T f_\Delta(x) + |g_\Delta(x)|^2 \\
\le\ & 2x^T F_1(x) + (1+\theta)|G_1(x)|^2 + 2x^T F(\pi_\Delta(x)) + (1+\theta^{-1})|G(\pi_\Delta(x))|^2 \\
\le\ & \alpha_1 - \beta_1|x|^2 + 2(x-\pi_\Delta(x))^T F(\pi_\Delta(x)) \\
+\ & 2(\pi_\Delta(x))^T F(\pi_\Delta(x)) + (1+\theta^{-1})|G(\pi_\Delta(x))|^2 \\
\le\ & \alpha_1 - \beta_1|x|^2 + 2(x-\pi_\Delta(x))^T F(\pi_\Delta(x)) + \alpha_2 + \beta_2(\mu^{-1}(h(\Delta)))^2.
\end{aligned} \tag{5.8}$$

But, by Assumption 5.1 again,

$$
\begin{aligned}
2(x-\pi_\Delta(x))^T F(\pi_\Delta(x)) &= 2[|x|/\mu^{-1}(h(\Delta))-1](\pi_\Delta(x))^T F(\pi_\Delta(x)) \\
&\le [|x|/\mu^{-1}(h(\Delta))-1]\Big(\alpha_2+\beta_2(\mu^{-1}(h(\Delta)))^2\Big)
\end{aligned}
$$

Substituting this into (5.8) yields

$$
\begin{aligned}
2x^T f_\Delta(x)+|g_\Delta(x)|^2 &\le \alpha_1-\beta_1|x|^2+\frac{|x|}{\mu^{-1}(h(\Delta))}\Big(\alpha_2+\beta_2(\mu^{-1}(h(\Delta)))^2\Big) \\
&\le \alpha_1-\beta_1|x|^2+\alpha_2\Big(\frac{|x|}{\mu^{-1}(h(\Delta))}\Big)^2+\beta_2|x|^2 \\
&\le \alpha_1-(\beta_1-\beta_2-0.5\varepsilon)|x|^2, \qquad (5.9)
\end{aligned}
$$

where (5.7) have been used. In other words, (5.6) holds for any $x\in\mathbb{R}^d$ with $|x|>\mu^{-1}(h(\Delta))$ too so it holds for all $x\in\mathbb{R}^d$ as claimed.

Fix $x_0\in\mathbb{R}^d$ arbitrarily. For any $\Delta\in(0,\hat\Delta_1)$, it follows from (4.12) and (5.6) that

$$
\begin{aligned}
\mathbb{E}|X_\Delta(t_{k+1})|^2 &\le \Delta(\alpha_1+\alpha_2)+\Delta^2\mathbb{E}|f_\Delta(X_\Delta(t_k))|^2 \\
&+ \big[1-\Delta(\beta_1-\beta_2-0.5\varepsilon)\big]\mathbb{E}|X_\Delta(t_k)|^2. \qquad (5.10)
\end{aligned}
$$

But, by (2.5) and (2.14),

$$|f_\Delta(X_\Delta(t_k))|^2\le 2|F_1(X_\Delta(t_k))|^2+2|F_\Delta(X_\Delta(t_k))|^2\le 4K_1(1+|X_\Delta(t_k)|^2)+2(h(\Delta))^2.$$

Hence, by (2.12),

$$\Delta|f_\Delta(X_\Delta(t_k))|^2\le 4\Delta K_1(1+|X_\Delta(t_k)|^2)+2\sqrt{\Delta}.$$

Consequently, there is a $\hat\Delta\in(0,\hat\Delta_1]$ sufficiently small such that for any $\Delta\in(0,\hat\Delta)$, $\Delta(\beta_1-\beta_2-\varepsilon)<1$ and

$$\Delta|f_\Delta(X_\Delta(t_k))|^2\le\varepsilon+0.5\varepsilon|X_\Delta(t_k)|^2. \qquad (5.11)$$

Now, fix any $\Delta\in(0,\hat\Delta)$. Substituting (5.11) into (5.10) yields

$$\mathbb{E}|X_\Delta(t_{k+1})|^2\le\Delta(\alpha_1+\alpha_2+\varepsilon)+\big[1-\Delta(\beta_1-\beta_2-\varepsilon)\big]\mathbb{E}|X_\Delta(t_k)|^2. \qquad (5.12)$$

This implies

$$\begin{aligned}
\mathbb{E}|X_\Delta(t_{k+1})|^2 &\le \Delta(\alpha_1+\alpha_2+\varepsilon)\Big(1+\big[1-\Delta(\beta_1-\beta_2-\varepsilon)\big]\Big) \\
&+ \big[1-\Delta(\beta_1-\beta_2-\varepsilon)\big]^2\,\mathbb{E}|X_\Delta(t_{k-1})|^2 \\
&\le \cdots \\
&\le \Delta(\alpha_1+\alpha_2+\varepsilon)\Big(1+\sum_{i=1}^{k}\big[1-\Delta(\beta_1-\beta_2-\varepsilon)\big]^i\Big) \\
&+ \big[1-\Delta(\beta_1-\beta_2-\varepsilon)\big]^{k+1}|x_0|^2 \\
&= \frac{\alpha_1+\alpha_2+\varepsilon}{\beta_1-\beta_2-\varepsilon}\Big(1-\big[1-\Delta(\beta_1-\beta_2-\varepsilon)\big]^{k+1}\Big) \\
&+ \big[1-\Delta(\beta_1-\beta_2-\varepsilon)\big]^{k+1}|x_0|^2. \qquad (5.13)
\end{aligned}$$

Letting $k \to \infty$, we obtain the required assertion (5.5). The proof is complete. $\square$

Example 5.4. Let us return to the SDE (1.5), namely consider the scalar stochastic Ginzburg–Landau equation (see, e.g., [5, 14])

$$dx(t) = (ax(t) - bx^3(t))dt + cx(t)dB(t), \qquad (5.14)$$

where $B(t)$ is a scalar Brownian motion and a, b, c are three positive numbers. We decompose the coefficients $f(x)$ and $g(x)$ in the form of (2.2) with

$$F_1(x) = -(a+c^2)x, \quad F(x) = (2a+c^2)x - bx^3, \quad G_1(x) = cx, \quad G(x) = 0 \qquad (5.15)$$

for $x \in \mathbb{R}$. Choosing $\theta = 0$, we then have

$$2xF_1(x) + (1+\theta)|G_1(x)|^2 = -(2a+c^2)x^2$$

and

$$2xF(x) + (1+\theta^{-1})|G(x)|^2 = 2(2a+c^2)x^2 - 2bx^4 \le \frac{(2a+c^2)^2}{2b}.$$

That is, Assumption 5.1 holds with

$$\alpha_1 = 0, \quad \beta_1 = 2a+c^2, \quad \alpha_2 = \frac{(2a+c^2)^2}{2b}, \quad \beta_2 = 0.$$

By Theorem 5.2, we then see that for any initial value $x_0 \in \mathbb{R}^d$, the solution of the SDE (5.14) satisfies

$$\limsup_{t\to\infty} \mathbb{E}|x(t)|^2 \le \frac{2a+c^2}{2b}. \tag{5.16}$$

It is known (see, e.g., [9, 11]) that the EM method may not preserve this asymptotic boundedness. However, we now show that our partially truncated EM method can preserve this boundedness perfectly. In fact, it is easy to see that the coefficients of the SDE (5.14) with their decompositions in (5.15) satisfy Assumptions 2.1 - 2.3 for any $\bar{p} > 2$. We can choose $\mu(r) = (2a + c^2 + b)r^3$ and $h(\Delta) = \Delta^{-1/4}$ to define the numerical solution $X_\Delta(t_k)$ by the partially truncated EM method (2.17). By Theorem 3.1, this numerical solution will converge to the true solution in L^r for any $r \ge 2$ at any finite time. Moreover, by Theorem 5.3, we can also conclude that for any $\varepsilon \in (0, 2a + c^2)$, there is a positive number $\hat{\Delta}$ such that for every $\Delta \in (0, \hat{\Delta})$ and any initial value $x_0 \in \mathbb{R}^d$, this numerical solution satisfies

$$\limsup_{k\to\infty} \mathbb{E}|X_\Delta(t_k)|^2 \le \frac{\frac{(2a+c^2)^2}{2b} + \varepsilon}{2a + c^2 - \varepsilon}. \tag{5.17}$$

Example 5.5. Let us now discuss a d-dimensional SDE

$$dx(t) = f(x(t))dt + g(x(t))dB(t), \tag{5.18}$$

on $t \ge 0$ with the initial value $x(0) = x_0 \in \mathbb{R}^d$. Here $B(t)$ is a scalar Brownian motion and $f, g : \mathbb{R}^d \to \mathbb{R}^d$ are defined by

$$f(x) = \text{diag}(x_1, x_2, ..., x_d)(b + Ax^2) \quad \text{and} \quad g(x) = \text{diag}(x_1, x_2, ..., x_d)Cx$$

for $x \in \mathbb{R}^d$, where $b \in \mathbb{R}^d$, $A, C \in \mathbb{R}^{d\times d}$ and $x^2 = (x_1^2, \cdots, x_d^2)^T$. If we restrict the state space of this SDE in the positive cone $\mathbb{R}^d_+$, it is known as the stochastic power Lotka-Volterra model (see, e.g., [3]). But we here treat this SDE in the whole $\mathbb{R}^d$-space. Let $\bar{b} = \max_{1\le i\le d} |b_i|$ and decompose the coefficients $f(x)$ and $g(x)$ in the form of (2.2) with

$$F_1(x) = -\bar{b}x, \quad F(x) = \bar{b}x + \text{diag}(x_1, x_2, ..., x_d)(b + Ax^2),$$

and

$$G_1(x) = 0, \quad G(x) = \mathrm{diag}(x_1, x_2, ..., x_d)Cx.$$

It is easy to see that Assumption 2.1 is satisfied. To satisfy Assumption 2.3, we assume that

$$-\lambda_{\max}(A + A^T) > d\,\lambda_{\max}(C^T C). \tag{5.19}$$

We then derive that

$$x^T F(x) = \bar{b}|x|^2 + (x^2)^T b + (x^2)^T A x^2 \le 2\bar{b}|x|^2 + \frac{1}{2}\lambda_{\max}(A + A^T)|x^2|^2.$$

But

$$|x|^4 = \sum_{i,j=1}^{d} x_i^2 x_j^2 \le \sum_{i=1}^{d} x_i^4 + \frac{1}{2}\sum_{i \ne j}(x_i^4 + x_j^4) = d\sum_{i=1}^{d} x_i^4 = d|x^2|^2.$$

So

$$x^T F(x) \le 2\bar{b}|x|^2 + \frac{1}{2d}\lambda_{\max}(A + A^T)|x|^4. \tag{5.20}$$

Moreover,

$$|G(x)|^2 = x^T C^T \mathrm{diag}(x_1^2, x_2^2, ..., x_d^2)Cx \le |x|^2 x^T C^T C x \le \lambda_{\max}(C^T C)|x|^4. \tag{5.21}$$

Set

$$\bar{p} = 1 + \frac{-\lambda_{\max}(A + A^T)}{d\lambda_{\max}(C^T C)}. \tag{5.22}$$

We have $\bar{p} > 2$ by condition (5.19) and, by (5.20) and (5.21),

$$x^T F(x) + \frac{\bar{p} - 1}{2}|G(x)|^2 \le 2\bar{b}|x|^2.$$

In other words, Assumption 2.3 is satisfied. Let us now verify Assumption 5.1. Choosing $\theta = \infty$, we have

$$2x^T F_1(x) + (1 + \theta)|G_1(x)|^2 = -2\bar{b}|x|^2 \tag{5.23}$$

and, by (5.20) and (5.21) again,

$$2x^T F(x) + (1 + \theta^{-1})|G(x)|^2 \le 4\bar{b}|x|^2 - \frac{1}{d}\Big(-\lambda_{\max}(A + A^T) - d\,\lambda_{\max}(C^T C)\Big)|x|^4 \le \alpha_2, \tag{5.24}$$

where

$$\alpha_2 = \frac{4d\bar{b}^2}{-\lambda_{\max}(A+A^T) - d\,\lambda_{\max}(C^TC)}. \tag{5.25}$$

That is, Assumption 5.1 is satisfied with

$$\alpha_1 = 0, \quad \beta_1 = 2\bar{b}, \quad \beta_2 = 0 \quad \text{and } \alpha_2 \text{ as defined above.}$$

By Theorem 5.2, we can therefore conclude that under condition (5.19), for any initial value $x_0 \in \mathbb{R}^d$, the solution of the SDE (5.18) satisfies

$$\limsup_{t\to\infty} \mathbb{E}|x(t)|^2 \le \frac{\alpha_2}{2\bar{b}}. \tag{5.26}$$

It is known (see, e.g., [9, 11]) that the EM method may not preserve this asymptotic boundedness. However, our partially truncated EM method will do. In fact, We can choose $\mu(r) = \delta r^3$, for a sufficiently large positive number δ, and $h(\Delta) = \Delta^{-1/4}$ to define the numerical solution $X_\Delta(t_k)$ by the partially truncated EM method (2.17). By Theorem 3.1, this numerical solution will converge to the true solution in L^r for any $2 \le r < \bar{p}$ at any finite time, where p is defined by (5.22). Moreover, by Theorem 5.3, we can also conclude that for any $\varepsilon \in (0, 2\bar{b})$, there is a positive number $\hat{\Delta}$ such that for every $\Delta \in (0, \hat{\Delta})$ and any initial value $x_0 \in \mathbb{R}^d$, this numerical solution satisfies

$$\limsup_{k\to\infty} \mathbb{E}|X_\Delta(t_k)|^2 \le \frac{\alpha_2 + \varepsilon}{2\bar{b} - \varepsilon}. \tag{5.27}$$

6. Discussions and Conclusions

Motivated by two examples discussed in Section 1, we developed a new explicit numerical scheme, called the partially truncated EM method for nonlinear SDEs under the local Lipschitz condition plus the Khasminskii-type condition. We established the finite-time strong L^r-convergence theory for the partially truncated EM method.

With respect of the finite convergence, we do not claim that our method outperforms those explicit methods, such as [1] [12] [21] [26] [28], that were also designed for SDEs with both drift and diffusion coefficients growing super-linearly. Actually, the finite time strong convergence order of those methods and the partially truncated EM method are 1/2 or arbitrarily close to 1/2.

The real benefits of this new method lie in that the method can preserve the asymptotic stability and boundedness of the underlying SDEs.

It should be noted that the conditions we imposed to guarantee the mean square exponential stability and the mean square asymptotic boundedness are only sufficient, but not necessary. In addition, our assumptions require the drift coefficient to dominate the diffusion coefficient in the negative direction, which may exclude some types of SDEs, such as some driftless SDEs with super-linear diffusion. Therefore, it is interesting to investigate whether the partially truncated EM method can still work if the assumptions in this paper are further released.

Acknowledgements

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

The authors also would like to thank Shanghai Pujiang Program (16PJ1408000), the Natural Science Fund of Shanghai Normal University (SK201603), Young Scholar Training Program of Shanghai's Universities, the EPSRC (EP/K503174/1), the Leverhulme Trust (RF-2015-385), the Royal Society (Wolfson Research Merit Award WM160014), the Natural Science Foundation of China (11471216), the Natural Science Foundation of Shanghai (14ZR1431300), E-Institutes of Shanghai Municipal Education Commission (No. E03004) and the Ministry of Education (MOE) of China (MS2014DHDX020), for their financial support.

The second author also would like to thank Dr Yue Lou for the very useful chat.

References

[1] A. Andersson, R. Kruse, Mean-square convergence of the BDF2-Maruyama and backward Euler schemes for SDE satisfying a global monotonicity condition, arXiv:1509.00609.

[2] S. Ankirchner, T. Kruse, M. Urusov, Numerical approximation of irregular SDEs via Skorokhod embeddings, J. Math. Anal. Appl. 440 (2016), 692–715.

[3] A. Bahar, X. Mao, Stochastic delay population dynamics, J. Int. Appl. Math. 11 (4) (2004), 377–400.

[4] G. Berkolaiko, E. Buckwar, C. Kelly, A. Rodkina, Almost sure asymptotic stability analysis of the Euler-Maruyama method applied to a test system with stabilising and destabilising stochastic perturbations, LMS J. Comput. Math. 15 (2012), 71–83.

[5] V.L. Ginzburg, L.D. Landau, On the theory of superconductivity, Zh. Eksperim. i teor. Fiz. 20 (1950), 1064–1082.

[6] I. Gyöngy, A note on Euler's approximations, Potential Anal. 8 (3) 1998, 205–216.

[7] D.J. Higham, X. Mao, A.M. Stuart, Strong convergence of Euler-type methods for nonlinear stochastic differential equations, SIAM J. Numer. Anal. 40 (3) (2003), 1041–1063.

[8] D.J. Higham, X. Mao, A.M. Stuart, Exponential mean-square stability of numerical solutions to stochastic differential equations, LMS J. Comput. Math. 6 (2003), 297–313.

[9] D.J. Higham, X. Mao, C. Yuan, Almost sure and moment exponential stability in the numerical simulation of stochastic differential equations, SIAM J. Numer. Anal. 45 (2) (2007), 592–609.

[10] 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

[11] M. Hutzenthaler, A. Jentzen, P.E. Kloeden, Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients, Proc. R. Soc. A 467 (2011), 1563-1576.

[12] M. Hutzenthaler, A. Jentzen, Numerical approximations of stochastic differential equations with non-globally Lipschitz continuous coefficients, Mem. Amer. Math. Soc. 236(2) (2015) 99 pages.

[13] R.Z. Khasminskii, Stochastic Stability of Differential Equations, Alphen: Sijtjoff and Noordhoff, 1980. (Translation of the Russian edition, Moscow, Nauka 1969).

[14] P.E. Kloeden, E. Platen, Numerical Solution of Stochastic Differential Equations, Springer-Verlog, Berlin, 1992.

[15] W. Liu, X. Mao, Asymptotic moment boundedness of the numerical solutions of stochastic differential equations, J. Comput. Appl. Math. 251 (2013), 22–32.

[16] X. Mao, Stability of Stochastic Differential Equations with Respect to Semimartingales, Pitman Research Notes in Mathematics Series 251, Longman Scientific and Technical, 1991.

[17] X. Mao, Exponential Stability of Stochastic Differential Equations, Monographs and Textbooks in Pure and Applied Mathematics Series, Marcel Dekker, 1994.

[18] X. Mao, Stochastic Differential Equations and Applications, 2nd Edition, Horwood, Chichester, UK, 2007.

[19] X. Mao, Almost sure exponential stability in the numerical simulation of stochastic differential equations, SIAM J. Numer. Anal. 53 (2015), 370–389.

[20] X. Mao, The truncated Euler–Maruyama method for stochastic differential equations, J. Comput. Appl. Math. 290 (2015), 370–384.

[21] X. Mao, Convergence rates of the truncated Euler–Maruyama method for stochastic differential equations, J. Comput. Appl. Math. 296 (2016), 362–375.

[22] X. Mao, C. Yuan, Stochastic Differential Equations with Markovian Switching, Imperial College Press, London, 2006.

[23] Y. Saito, T. Mitsui, Stability analysis of numerical schemes for stochastic differential equations, SIAM J. Numer. Anal. 33 (1996), 2254–2267.

[24] T. Shardlow, P. Taylor, On the pathwise approximation of stochastic differential equations, BIT 56 (3) (2016), 1101–1129.

[25] M. Song, L. Hu, X. Mao, L. Zhang, Khasminskii-Type theorems for stochastic functional differential equations, Discrete Contin. Dyn. Syst. Ser. B 18 (6) (2013), 1697–1714.

[26] L. Szpruch, X. Zhang, V-Integrability, Asymptotic Stability And Comparison Theorem of Explicit Numerical Schemes for SDEs, arXiv:1310.0785v2

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

[28] Z. Zhang, New explicit balanced schemes for SDEs with locally Lipschitz coefficients, arXiv:1402.3708.