Positivity and Boundedness Preserving Numerical Scheme for the Stochastic Epidemic Model with Square-root Diffusion Term

Yongmei Cai[1], Junhao Hu[2]*, Xuerong Mao[3]

[1]School of Mathematical Sciences, University of Nottingham Ningbo China,
Ningbo 315100, China

[2]School of Mathematics and Statistics, South-Central University for Nationalities,
Wuhan 430074, China

[3]Department of Mathematics and Statistics, University of Strathclyde,
Glasgow G1 1XH, UK

Abstract

This work concerns about the numerical solution to the stochastic epidemic model proposed by Cai et al. [2]. The typical features of the model including the positivity and boundedness of the solution and the presence of the square-root diffusion term make this an interesting and challenging work. By modifying the classical Euler-Maruyama (EM) scheme, we generate a positivity and boundedness preserving numerical scheme, which is proved to have a strong convergence to the true solution over finite time intervals. We also demonstrate that the principle of this method is applicable to a bunch of popular stochastic differential equation (SDE) models, e.g. the mean-reverting square-root process, an important financial model, and the multi-dimensional SDE SIR epidemic model.

Keywords: Stochastic differential equation, square-root process, positivity and boundedness preserving numerical method, strong convergence

1 Introduction

In 2011, Gray et al. [5] explored the SIS epidemic model incorporating environmental variability

$$dx(t) = \big[\beta x(t)(N - x(t)) - (\mu + \gamma)x(t)\big]dt + \sigma_1 x(t)(N - x(t))dB_1(t) \tag{1.1}$$

with initial value $x(0) \in (0, N)$, where $x(t)$ represents the infected population at time t, N is the total amount of population where the disease spreads, β is the disease transmission coefficient, μ is the per capita death rate, γ is the recovery rate of the infected population, σ_1 is a positive constant and $B_1(t)$ is a scalar Brownian motion. Since then, the stochastic epidemic modelling has been extensively investigated. On the basis of this work, Cai et al. [2, 3] modified the SIS model further by incorporating an additional external noise, which is of the form

$$\begin{aligned} dx(t) = \big[\beta x(t)(N - x(t)) - (\mu + \gamma)x(t)\big]dt + \sigma_1 x(t)(N - x(t))dB_1(t) \\ - \sigma_2 x(t)\sqrt{N - x(t)}dB_2(t) \end{aligned} \tag{1.2}$$

*Corresponding author and e-mail: junhaohu74@163.com

with initial value $x(0) \in (0, N)$, where σ_2 and $B_2(t)$ have similar meanings as σ_1 and $B_1(t)$. By defining

$$F(x) := \beta x(N-x) - (\mu+\gamma)x \quad \text{and} \quad G(x) := (\sigma_1 x(N-x), -\sigma_2 x\sqrt{N-x}),$$

SDE (1.2) can be rewritten as

$$dx(t) = F(x(t))dt + G(x(t))dB(t),$$

where $B(t) = (B_1(t), B_2(t))^T$.

Due to the nonlinearity, the solution to SDE (1.2) cannot be explicitly represented. In [5], a complete analytical study has been carried out, however, numerical analysis is a more efficient and intuitive way of understanding the dynamical behaviours of such epidemic models in the real life, as well as allowing us to conduct further study such as parameter estimations.

In the past few decades, numerical analysis has been intensively developed. As an explicit numerical method, the Euler-Maruyama (EM) scheme is easily implementable and hence is widely used in practice. Since the EM method is not strongly convergent without global Lipchitz continuous coefficients, several modified EM methods have been proposed to address this issue, including the tamed EM method [9], the stopped EM method [14], the truncated EM method [16] and the multilevel EM method [1], etc.

Recently, Cai et al.[2] has shown that the solution to SDE (1.2) satisfies $x(t) \in (0, N)$ for all time almost surely. However, the condition obtained is rather restrictive. So one contribution in this paper is to verify the presence of a unique global positive solution under much weaker condition with the Ikeda-Watanabe technique [10] as well as the Feller test [12].

From the numerical viewpoint, however, the positivity and boundedness of SDE (1.2) cannot be guaranteed by most of the existing explicit methods including the tamed/truncated EM, though there are some implicit ones [18–20]. Recently, Chen et al. [4] addressed this issue for SDE (1.1) by using the Lamperti smoothing truncation scheme. However, this technique has its limitation when coping with SDE (1.2) due to the presence of an additional square-root term as well as other multi-dimensional models (see Section 5 below please).

To fill this gap, there are generally two main issues to overcome. Firstly, to the best of our knowledge, there exists few explicit numerical scheme that preserves the positivity and boundedness. Secondly, compared to SDE (1.1), the existence of the non-Lipschitzian square-root diffusion term in SDE (1.2) makes our analysis more challenging.

In fact, the numerical study on the square-root process can be traced back to 2005, when Higham and Mao [8] considered the following mean-reverting square-root process

$$ds(t) = \lambda(\xi - s(t))dt + \sigma\sqrt{s(t)}dW(t), \tag{1.3}$$

where λ, ξ and σ are positive constants and $W(t)$ is a scalar Brownian motion. To avoid the computational issues due to the presence of the square-root function, (1.3) is replaced by the following SDE

$$ds(t) = \lambda(\xi - s(t))dt + \sigma\sqrt{|s(t)|}dW(t). \tag{1.4}$$

The EM applied to (1.4) is to form the discrete-time EM solutions $S_k \approx s(t_k)$ by setting $S_0 = s(t_0)$ and computing

$$S_{k+1} = S_k(1 - \lambda\Delta) + \lambda\xi\Delta + \sigma\sqrt{|S_k|}\Delta W_k$$

for $k \geqslant 0$, where $\Delta W_k = W(t_{k+1}) - W(t_k)$. They found the EM is a strong convergent approximation of $s(t)$, though the positivity is not yet guaranteed.

In this paper, we formulate an explicit numerical algorithm that preserves the positivity and boundedness of a bunch of popular SDE models including the square-root process. The method can natually be applied to the multi-dimensional scenario. To make the arguments concise, we interpret our scheme using the 1-dimensional SIS model (1.2).

2 Generalized Existence-and-Uniqueness Theorem

Let $(\Omega, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P})$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$ satisfying the usual conditions. Let $B(t) = (B_1(t), B_2(t))^T$ be a two-dimensional Brownian motion defined on the probability space. For a value b, define

$$[b]^+ = \begin{cases} -b, & b < 0, \\ 0, & \text{otherwise.} \end{cases}$$

In addition, we let T and θ be two fixed positive values. Let C represent generic positive constants dependent on $x(0)$, T and θ but independent of the step size Δ which will be used later on.

Before getting into the numerical part, we first study the presence of a unique global positive solution to (1.2) with the Ikeda-Watanabe technique [10] and the Feller test [12].

Assumption 2.1. $N > 1 \vee (\mu+\gamma)/\beta$.

Remark 2.2. *In many real world systems, $(\mu+\gamma)/\beta$ might be a large value (see, e.g., [7]), while N can be even larger as in the pandemic, e.g., SARS, MERS and COVID-19, where the virus spreads across a fairly huge amount of population. In other words, $N > (\mu+\gamma)/\beta$ holds naturally, not mentioning $N > 1$. We state $N > 1$ explicitly to make it strictly clear from the mathematical point of view.*

Theorem 2.3. *Under Assumption 2.1, for any initial value $x(0) \in (0, N)$, there is a unique global solution $x(t)$ on $t \geqslant 0$ which has the properties*

$$\mathbb{P}\big(0 < \min_{0\leqslant t\leqslant T} x(t) < \max_{0\leqslant t\leqslant T} x(t) < N\big) = 1, \quad \forall T > 0 \tag{2.1}$$

and

$$\mathbb{P}\big(\inf_{0\leqslant t<\infty} x(t) = 0\big) = \mathbb{P}\big(\sup_{0\leqslant t<\infty} x(t) = N\big) = 1.$$

Proof. We first show that there exists a unique solution $\bar{x}(t) > 0$ to

$$\begin{aligned} d\bar{x}(t) = \big[\beta\bar{x}(t)(N-\bar{x}(t)) - (\mu+\gamma)\bar{x}(t)\big]dt + \sigma_1\bar{x}(t)(N-\bar{x}(t))dB_1(t) \\ - \sigma_2\bar{x}(t)\sqrt{|N-\bar{x}(t)|}dB_2(t) \end{aligned} \tag{2.2}$$

on $t \geqslant 0$ for any initial value $\bar{x}(0) > 0$ almost surely. Let $m_0 > 0$ satisfy $1/m_0 < \bar{x}(0) < m_0$. For any integer $m > m_0$, consider the stopping time

$$\tau_m = \inf\{t \geqslant 0 : x(t) \notin (1/m, m)\}.$$

We then need to show that $\tau_\infty = \infty$ almost surely. To this end, we use the classical method as adopted in e.g. [15, 17]. Considering a C^2-function $V : \mathbb{R}_+ \to \mathbb{R}_+$ by

$$V(x) = x^{1/2} - \log x,$$

we have

$$\begin{aligned} LV(\bar{x}(t)) &\leqslant \big(\frac{1}{2}\bar{x}^{1/2}(t) - 1\big)\big[\beta(N-\bar{x}(t)) - (\mu+\gamma)\big] + \frac{1}{2}\big(-\frac{1}{4}\bar{x}^{1/2}(t) + 1\big)\big[\sigma_1^2(N-\bar{x}(t))^2 + \sigma_2^2|N-\bar{x}(t)|\big] \\ &\leqslant -\beta N + \mu + \gamma + \frac{1}{2}\sigma_1^2N^2 + \frac{1}{2}\sigma_2^2N + \Big[\frac{1}{2}\big(\beta N - \mu - \gamma\big) - \frac{1}{8}\sigma_1^2N^2 + \frac{1}{8}\sigma_2^2N\Big]\bar{x}^{1/2}(t) \\ &+ \Big(\beta - N\sigma_1^2 + \frac{1}{2}\sigma_2^2\Big)\bar{x}(t) + \Big[-\frac{1}{2}\beta + \frac{1}{4}N\sigma_1^2 + \frac{1}{8}\sigma_2^2\Big]\bar{x}^{3/2}(t) + \frac{1}{2}\sigma_1^2\bar{x}^2(t) - \frac{1}{8}\sigma_1^2\bar{x}^{5/2}(t) \\ &\leqslant C \end{aligned}$$

and further analysis shows that the solution $\bar{x}(t)$ to SDE (2.2) is unique and positive for any initial value $\bar{x}(0) > 0$.

Next, we aim to verify that for any initial value $\bar{x}(0) \leqslant N$, the solution $\bar{x}(t)$ to SDE (1.2) does not exceed N for $t \geqslant 0$ with probability 1, and thus SDE (2.2) can be rewritten as (1.2).

Now, choose a strictly decreasing sequence c_k for $k \geqslant 0$ such that $c_0 = 1$ and

$$\int_{c_k}^{c_{k-1}} \frac{1}{(u+u^2)(N+u)^2} du = k.$$

Denote C^1-functions $\zeta_k : (c_k, c_{k-1}) \to \mathbb{R}$ such that

$$0 \leqslant \zeta_k(u) \leqslant \frac{2}{k(u+u^2)(N+u)^2}$$

and moreover,

$$\int_{c_k}^{c_{k-1}} \zeta_k(u) du = 1.$$

Next let

$$\rho_k(x) = \begin{cases} \int_0^{-(N-x)} dy \int_0^y \zeta_k(u) du & x > N, \\ 0 & \text{otherwise.} \end{cases}$$

It can be revealed that $\rho_k \in C^2(\mathbb{R}, \mathbb{R})$ and

$$0 \leqslant \rho_k'(x) \leqslant 1 \text{ for } c_k + N < x < \infty \text{ and } \rho_k'(x) = 0 \text{ otherwise.}$$
$$\rho_k''(x) \leqslant \frac{2}{kx^2[-(N-x) + (N-x)^2]} \text{ for } N + c_k < x < N + c_{k-1} \text{ and } \rho_k''(x) = 0 \text{ otherwise.}$$

Also, we have

$$[N-x]^+ - c_{k-1} \leqslant \rho_k(x) \leqslant [N-x]^+ \text{ for } x \in \mathbb{R}.$$

Applying the Itô formula on $\rho_k(\bar{x}(t))$ for $t \geqslant 0$ yields

$$\begin{aligned} \rho_k(\bar{x}(t)) = \rho_k(\bar{x}(0)) &+ \int_0^t \Big[\rho_k'(\bar{x}(v))\big(\beta\bar{x}(v)(N - x(v)) - (\mu+\gamma)\bar{x}(v)\big) + \frac{1}{2}\rho_k''(\bar{x}(v))\Big(\sigma_1^2\bar{x}^2(v)(N-\bar{x}(v))^2 \\ &+ \sigma_2^2\bar{x}^2(v)|N-\bar{x}(v)|\Big)\Big] dv + \sigma_1 \int_0^t \bar{x}(v)(N-\bar{x}(v)) dB_1(v) - \sigma_2 \int_0^t \bar{x}(v)\sqrt{|N-\bar{x}(v)|} dB_2(v) \\ &\leqslant \frac{1}{k}(\sigma_1^2 + \sigma_2^2)t + \sigma_1 \int_0^t \bar{x}(v)(N-\bar{x}(v)) dB_1(v) - \sigma_2 \int_0^t \bar{x}(v)\sqrt{|N-\bar{x}(v)|} dB_2(v). \end{aligned}$$

Taking expectation gives

$$\mathbb{E}[N - \bar{x}(t \wedge \tau_m)]^+ - c_{k-1} \leqslant \mathbb{E}\rho_k(\bar{x}(t \wedge \tau_m)) \leqslant \frac{1}{k}(\sigma_1^2 + \sigma_2^2)(t \wedge \tau_m).$$

Letting $m \to \infty$ and then $k \to \infty$ gives

$$\mathbb{E}[N - \bar{x}(t)]^+ \leqslant 0$$

and this suggests

$$\mathbb{E}[N - \bar{x}(t)]^+ = 0 \text{ for all } t \geqslant 0.$$

Finally we obtain

$$\mathbb{P}(\bar{x}(t) > N) = 0 \text{ for all } t \geqslant 0.$$

Since $\bar{x}(t)$ is continuous, we must have $\bar{x}(t) \leqslant N$ for all $t \geqslant 0$ almost surely. Namely, SDE (2.2) is equivalent to (1.2).

Now, we will take a further step by justifying the solution $x(t)$ is actually strictly less than N for all $t \geqslant 0$ almost surely. This is achieved using the classical Feller test for explosions [12].

Firstly, it is known that for $x \in (0, N)$,

$$F(x) := \beta x(N - x) - (\mu + \gamma)x$$

is continuous while

$$G(x) := (\sigma_1 x(N - x), -\sigma_2 x\sqrt{N - x})$$

is continuous with $|G(x)|^2 > 0$. From the standard result of ordinary differential equations, for any given numbers p and q satisfying $0 < p < x(0) < q < N$, there is a unique solution $L(v)$ to the equation

$$F(v)L'(v) + 0.5|G(v)|^2 L''(v) = -1, \quad p < v < q$$

with the boundary condition $L(p) = L(q) = 0$. Consider the two stopping times

$$\eta_p = \inf\{t \geqslant 0 : x(t) \leqslant p\} \quad \text{and} \quad \eta_q = \inf\{t \geqslant 0 : x(t) \geqslant q\}.$$

From Itô's formula,

$$\mathbb{E}L(x(t \wedge \eta_p \wedge \eta_q)) = L(x(0)) - \mathbb{E}(t \wedge \eta_p \wedge \eta_q),$$

which implies

$$\mathbb{E}(t \wedge \eta_p \wedge \eta_q) \leqslant L(x(0)).$$

Letting $t \to \infty$, we have

$$\mathbb{E}(\eta_p \wedge \eta_q) \leqslant L(x(0)) < \infty.$$

Namely, $x(t)$ moves across each compact subinterval of $(0, N)$ over finite expected time. So

$$\mathbb{P}(\eta_p \wedge \eta_q < \infty) = 1.$$

This and the boundary condition suggest

$$\lim_{t\to\infty} \mathbb{E}L(x(t \wedge \eta_p \wedge \eta_q)) = 0$$

and therefore

$$\mathbb{E}(\eta_p \wedge \eta_q) = L(x(0)).$$

Now define

$$J(x) = \int_1^x \exp\Big\{ - \int_1^y \frac{2F(z)}{|G(z)|^2} dz \Big\} dy, \quad x \in (0, N).$$

The Itô formula leads to

$$J(x(t \wedge \eta_p \wedge \eta_q)) = J(x(0)) + \int_0^{t\wedge\eta_p\wedge\eta_q} J'(x(r))G(x(r))dB(r), \quad t > 0.$$

Taking expectation and then letting $t \to \infty$ infers

$$J(x(0)) = \mathbb{E}J(x(\eta_p \wedge \eta_q)) = J(p)\mathbb{P}(\eta_p < \eta_q) + J(q)\mathbb{P}(\eta_p > \eta_q).$$

This combines with the fact that $\mathbb{P}(\eta_q > \eta_p) + \mathbb{P}(\eta_q < \eta_p) = 1$ leads to

$$\mathbb{P}(\eta_p < \eta_q) = \frac{J(q) - J(x(0))}{J(q) - J(p)} \quad \text{and} \quad \mathbb{P}(\eta_p > \eta_q) = \frac{J(x(0)) - J(p)}{J(q) - J(p)}. \tag{2.3}$$

Compute

$$J(x) = \int_1^x \exp\Big\{ -2\int_1^y \frac{\beta z(N-z) - (\mu+\gamma)z}{\sigma_1^2 z^2(N-z)^2 + \sigma_2^2 z^2(N-z)} dz \Big\} dy$$
$$= \int_1^x \exp\Big\{ -2\int_1^y \frac{\beta(N-z) - (\mu+\gamma)}{\sigma_1^2 z(N-z)^2 + \sigma_2^2 z(N-z)} dz \Big\} dy.$$

Under Assumption 2.1, it is easy to see that when $N > 1$,

$$\lim_{x\uparrow N} J(x) = \infty$$

while when $N > (\mu+\gamma)/\beta$,

$$\lim_{x\downarrow 0} J(x) = -\infty.$$

Define

$$\eta_0 = \lim_{p\downarrow 0} \eta_p, \quad \eta_N = \lim_{q\uparrow N} \eta_q \quad \text{and then} \quad \eta = \eta_0 \wedge \eta_N.$$

From (2.3), we see

$$\mathbb{P}\big(\inf_{0\leqslant t<\eta} x(t) \leqslant p \big) \geqslant \mathbb{P}(\eta_p < \eta_q) = \frac{1 - J(x(0))/J(q)}{1 - J(p)/J(q)}.$$

Letting $q \uparrow N$ yields

$$\mathbb{P}\big(\inf_{0\leqslant t<\eta} x(t) \leqslant p \big) = 1.$$

As this holds for all $0 < p < N$, we obtain

$$\mathbb{P}\big(\inf_{0\leqslant t<\eta} x(t) = 0 \big) = 1. \tag{2.4}$$

On the other hand, we have

$$\mathbb{P}\big(\sup_{0\leqslant t<\eta} x(t) \geqslant q \big) \geqslant \mathbb{P}(\eta_p > \eta_q) = \frac{J(x(0))/J(p) - 1}{J(q)/J(p) - 1}.$$

Letting $p \downarrow 0$ gives

$$\mathbb{P}\big(\sup_{0\leqslant t<\eta} x(t) \geqslant q \big) = 1.$$

and thus

$$\mathbb{P}\big(\sup_{0\leqslant t<\eta} x(t) = N \big) = 1. \tag{2.5}$$

Now we claim that $\mathbb{P}(\eta < \infty) = 0$, since if this is not true, we will not have (2.4) and (2.5) hold simultaneously. As a result, we have

$$\mathbb{P}(\eta = \infty) = \mathbb{P}\big(\inf_{0\leqslant t<\infty} x(t) = 0 \big) = \mathbb{P}\big(\sup_{0\leqslant t<\infty} x(t) = N \big) = 1.$$

This completes our proof. □

Remark 2.4. *One may wonder what happens when the disease spreads on a small scale, namely,* $1 < N \leqslant (\mu+\gamma)/\beta$. *In this case, we still have*

$$\lim_{x\uparrow N} J(x) = \infty$$

when $N > 1$, *while when* $N \leqslant (\mu+\gamma)/\beta$, *we see*

$$\lim_{x\downarrow 0} J(x) > -\infty.$$

We thus still have (2.4) *being hold, while* (2.5) *is no longer true. Recall the second equality of* (2.3)

$$\mathbb{P}(\eta_p > \eta_q) = \frac{J(x(0))/J(p) - 1}{J(q)/J(p) - 1}.$$

Letting $p \downarrow 0$ *and* $q \uparrow N$ *leads to*

$$\mathbb{P}(\eta_0 > \eta_N) = 0.$$

In all,

$$\mathbb{P}\big(\inf_{0\leqslant t<\eta} x(t) = 0\big) = \mathbb{P}\big(\sup_{0\leqslant t<\eta} x(t) < N\big) = 1.$$

Based on the results, we are unable to conclude that $x(t) \in (0, N)$ *for all* $t \geqslant 0$.

3 Positivity and Boundedness Preserving EM Method

In this section, we will set up the positivity and boundedness preserving Euler-Maruyama (PBPEM) method which guarantees our numerical solution to SDE (1.2) remains positive and bounded within the open interval $(0, N)$. To this end, we need to extend the domain of our model from $(0, N)$ to $\mathbb{R}$. Recall the two functions we have defined in section 1, $F : (0, N) \to \mathbb{R}$ and $G : (0, N) \to \mathbb{R}^{1\times 2}$ by

$$F(x) = \beta x(N - x) - (\mu + \gamma)x \quad \text{and} \quad G(x) = (\sigma_1 x(N - x), -\sigma_2 x\sqrt{N - x}).$$

Hence our SDE (1.2) can be rewritten as

$$dx(t) = f(x(t))dt + g(x(t))dB(t),$$

where $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}^{1\times 2}$ are defined as

$$f(x) = F((x \vee 0) \wedge N) \quad \text{and} \quad g(x) = G((x \vee 0) \wedge N).$$

In this way the domain of (1.2) has been extended from $(0, N)$ to $\mathbb{R}$. Moreover, it is easy to see that

$$|f(x)| \leqslant C \quad \text{and} \quad |g(x)| \leqslant C \quad \text{for} \quad x \in \mathbb{R}. \tag{3.1}$$

We form the discrete-time PBPEM solution $X_\Delta(t_k) \approx x(t_k)$ for $t_k = k\Delta$ by setting $\bar{X}_\Delta(0) = X_\Delta(0) = x(0)$ and computing

$$\begin{aligned} \bar{X}_\Delta(t_{k+1}) &= \bar{X}_\Delta(t_k) + f(X_\Delta(t_k))\Delta + g(X_\Delta(t_k))\Delta B_k, \\ X_\Delta(t_{k+1}) &= (\Delta \vee \bar{X}_\Delta(t_{k+1})) \wedge (N - \Delta) \end{aligned} \tag{3.2}$$

for $k = 0, 1, 2, \cdots$ and $\Delta B_k = B(t_{k+1}) - B(t_k)$. Then we define

$$X_\Delta(t) = X_\Delta(t_k), \quad t \in [t_k, t_{k+1})$$

and

$$\bar{X}_\Delta(t) = \bar{X}_\Delta(t_k), \quad t \in [t_k, t_{k+1}).$$

Thus we see

$$X_\Delta(t) = (\Delta \vee \bar{X}_\Delta(t)) \wedge (N - \Delta) \tag{3.3}$$

For the convenience of further analysis, we also define an Itô process

$$x_\Delta(t) = x(0) + \int_0^t f(X_\Delta(s))ds + \int_0^t g(X_\Delta(s))dB(s) \quad \text{for } t \geqslant 0. \tag{3.4}$$

with

$$dx_\Delta(t) = f(X_\Delta(t))dt + g(X_\Delta(t))dB(t).$$

Clearly, $x_\Delta(t_k) = \bar{X}_\Delta(t_k)$ for $k = 0, 1, 2, \cdots$.

Lemma 3.1. *For any $\theta > 0$,*

$$\sup_{0<\Delta\leqslant 1} \mathbb{E}\big(\sup_{0\leqslant t\leqslant T} |x_\Delta(t)|^\theta \big) \leqslant C.$$

This is a direct result as both $f(\cdot)$ and $g(\cdot)$ are bounded.

Now three key lemmas are presented before showing our final results.

Lemma 3.2. *For any $\theta > 0$,*

$$\lim_{\Delta\to 0} \mathbb{E}\big(\sup_{0\leqslant t\leqslant T} \big|x_\Delta(t) - \bar{X}_\Delta(t)\big|^\theta \big) = 0.$$

Proof. Set e as the integer part of T/Δ. From (3.1),

$$\begin{aligned}
\mathbb{E}\big(\sup_{0\leqslant t\leqslant T} \big|x_\Delta(t) - \bar{X}_\Delta(t)\big|^\theta \big) &\leqslant \mathbb{E}\Big(\max_{0\leqslant k\leqslant e} \sup_{t_k\leqslant t\leqslant t_{k+1}} \Big|f(X_\Delta(t_k))(t-t_k) + g(X_\Delta(t_k))(B(t)-B(t_k))\Big|^\theta \Big) \\
&\leqslant C\Delta^\theta + C\mathbb{E}\big(\max_{0\leqslant k\leqslant e} \sup_{t_k\leqslant t\leqslant t_{k+1}} |B(t)-B(t_k)|^\theta \big) \\
&\leqslant C\Delta^\theta + C\Big(\sum_{k=0}^{e} \mathbb{E}\big(\sup_{t_k\leqslant t\leqslant t_{k+1}} |B(t)-B(t_k)|^{2d} \big) \Big)^{\theta/2d}
\end{aligned}$$

by the Hölder's inequality, where d is an integer satisfying

$$d \geqslant 3 \vee \frac{\theta}{2} \quad \text{and} \quad \Big(\frac{2d}{2d-1}\Big)^\theta (T+1)^{\theta/2d} \leqslant 2. \tag{3.5}$$

Then the Doob martingale inequality yields

$$\begin{aligned}
\mathbb{E}\big(\sup_{0\leqslant t\leqslant T} \big|x_\Delta(t) - \bar{X}_\Delta(t)\big|^\theta \big) &\leqslant C\Delta^\theta + C\Big(\sum_{k=0}^{e} \Big(\frac{2d}{2d-1}\Big)^{2d} \mathbb{E}\big|B(t_{k+1}) - B(t_k)\big|^{2d} \Big)^{\theta/2d} \\
&\leqslant C\Delta^\theta + C\Big(\sum_{k=0}^{e} \Big(\frac{2d}{2d-1}\Big)^{2d} (2d-1)!!\Delta^d \Big)^{\theta/2d} \\
&\leqslant C\Delta^\theta + C\Big(\Big(\frac{2d}{2d-1}\Big)^{2d} (T+1)(2d-1)!!\Delta^{d-1} \Big)^{\theta/2d},
\end{aligned}$$

where $(2d-1)!! = 1\cdot 3\cdots(2d-3)\cdot(2d-1)$. Note that

$$[(2d-1)!!]^{1/d} \leqslant \frac{1}{d}\sum_{k=1}^{d}(2k-1) = d.$$

Hence under (3.5),

$$\begin{aligned}
\mathbb{E}\big(\sup_{0\leqslant t\leqslant T} \big|x_\Delta(t) - \bar{X}_\Delta(t)\big|^\theta \big) &\leqslant C\Delta^\theta + C\Big(\frac{2d}{2d-1}\Big)^\theta d^{\theta/2}(T+1)^{\theta/2d}\Delta^{(d-1)\theta/2d} \\
&\leqslant C\Delta^\theta + C\Delta^{\theta/3} \leqslant C\Delta^{\theta/3}
\end{aligned}$$

and the required assertion follows. □

Lemma 3.3. *Under Assumption 2.1, for any $\theta > 0$,*

$$\lim_{\Delta\to 0} \mathbb{E}\big[\sup_{0\leqslant t\leqslant T} \big|x(t) - x_\Delta(t)\big|^\theta \big] = 0 \tag{3.6}$$

and

$$\lim_{\Delta\to 0} \mathbb{E}\big[\sup_{0\leqslant t\leqslant T} \big|X_\Delta(t) - \bar{X}_\Delta(t)\big|^\theta \big] = 0. \tag{3.7}$$

Proof. Let ϵ be arbitrary. From (2.1), one can find a $\delta = \delta(\epsilon) \in (0,1)$ so small that

$$\mathbb{P}(\Omega_1) \geqslant 1 - \frac{\epsilon}{2}, \tag{3.8}$$

where

$$\Omega_1 = \{\delta < \min_{0\leqslant t\leqslant T} x(t) \leqslant \max_{0\leqslant t\leqslant T} x(t) < N - \delta\}. \tag{3.9}$$

With the chosen δ, define

$$f_\delta(x) = F((x \vee \delta/2) \wedge (N - \delta/2)) \quad \text{and} \quad g_\delta(x) = G((x \vee \delta/2) \wedge (N - \delta/2)) \quad \text{for } x \in \mathbb{R}.$$

Since both $f_\delta(\cdot)$ and $g_\delta(\cdot)$ are globally Lipschitz continuous, the SDE

$$dy(t) = f_\delta(y(t))dt + g_\delta(y(t))dB(t) \tag{3.10}$$

on $t \geqslant 0$ with initial value $y(0) = x(0)$ has a unique global solution $y(t)$ on $y \geqslant 0$. For each step size $\Delta \in (0,1]$, apply EM method to (3.10). More precise, we form the EM solution $Y_k \approx y(t_k)$ for $t_k = k\Delta$ by setting $Y_0 = x(0)$ and computing

$$Y_{k+1} = Y_k + f_\delta(Y_k)\Delta + g_\delta(Y_k)\Delta B_k \quad \text{for } k = 0, 1, \cdots. \tag{3.11}$$

Then we define

$$Y(t) = Y_k, \quad t \in [t_k, t_{k+1})$$

and hence the Itô process

$$y_\Delta(t) = x(0) + \int_0^t f_\delta(Y(s))ds + \int_0^t g_\delta(Y(s))dB(s). \tag{3.12}$$

It then follows from the theory in [15] that

$$\mathbb{E}\big(\sup_{0\leqslant t\leqslant T} |y(t) - y_\Delta(t)|^\theta\big) \leqslant C_\delta \Delta^{\theta/2}, \quad \forall \Delta \in (0,1],$$

where C_δ is a positive constant dependent on δ but independent of Δ. And thus, of course,

$$\mathbb{E}\big(\sup_{0\leqslant t\leqslant T} |y(t) - y_\Delta(t)|^\theta \mathbb{I}_{\Omega_1}(\omega)\big) \leqslant C_\delta \Delta^{\theta/2}. \tag{3.13}$$

This and Chebyshev's inequality infer that

$$\mathbb{P}\big(\sup_{0\leqslant t\leqslant T} |y(t) - y_\Delta(t)|\mathbb{I}_{\Omega_1}(\omega) \geqslant \delta/2\big) \leqslant C_\delta \Delta^{\theta/2}\Big(\frac{2}{\delta}\Big)^\theta, \quad \forall \Delta \in (0,1].$$

Defining

$$\Omega_2 = \{\sup_{0\leqslant t\leqslant T} |y(t) - y_\Delta(t)| < \delta/2\} \cap \Omega_1,$$

one can find a $\Delta_1 \in (0, \delta/2]$ such that

$$\mathbb{P}\big(\Omega_2\big) \geqslant 1 - \epsilon, \quad \forall \Delta \in (0, \Delta_1].$$

Note that whenever $\omega \in \Omega_2 \subset \Omega_1$, we always have

$$y(t) = x(t), \quad t \in [0, T]. \tag{3.14}$$

Hence for any $\omega \in \Omega_2$, it implies

$$\sup_{0\leqslant t\leqslant T} |x(t) - y_\Delta(t)| < \delta/2. \tag{3.15}$$

This together with (3.9) yields

$$\delta/2 < \min_{0\leqslant t\leqslant T} y_\Delta(t) \leqslant \max_{0\leqslant t\leqslant T} y_\Delta(t) < N - \delta/2, \quad \forall \omega \in \Omega_2.$$

It then follows by comparing (3.11) and (3.12) with (3.2) and (3.4) that for any $\omega \in \Omega_2$,

$$y_\Delta(t) = x_\Delta(t) \quad \text{for } t \in [0,T] \text{ and } \Delta \in (0, \Delta_1]. \tag{3.16}$$

Combining (3.14) and (3.16) with (3.13) leads to

$$\mathbb{E}\big(\sup_{0\leqslant t\leqslant T} |x(t) - x_\Delta(t)|^\theta \mathbb{I}_{\Omega_2}(\omega)\big) \leqslant C_\delta \Delta^{\theta/2}, \quad \forall \Delta \in (0, \Delta_1]. \tag{3.17}$$

Furthermore, substituting (3.16) into (3.15) indicates that for any $\omega \in \Omega_2$,

$$\sup_{0\leqslant t\leqslant T} |x(t) - x_\Delta(t)| < \delta/2, \quad \forall \Delta \in (0, \Delta_1]. \tag{3.18}$$

We are now ready to show the required two assertions (3.6) and (3.7). It follows from (3.17) that for any $\Delta \in (0, \Delta_1]$

$$\begin{aligned}
\mathbb{E}\big[\sup_{0\leqslant t\leqslant T} \big|x(t) - x_\Delta(t)\big|^\theta\big] &= \mathbb{E}\big[\sup_{0\leqslant t\leqslant T} \big|x(t) - x_\Delta(t)\big|^\theta \mathbb{I}_{\Omega_2}\big] + \mathbb{E}\big[\sup_{0\leqslant t\leqslant T} \big|x(t) - x_\Delta(t)\big|^\theta \mathbb{I}_{\Omega_2^c}\big] \\
&\leqslant C_\delta \Delta^{\theta/2} + \big[\mathbb{P}(\Omega_2^c)\big]^{1/2} \Big[\mathbb{E}\big(\sup_{0\leqslant t\leqslant T} \big|x(t) - x_\Delta(t)\big|^{2\theta}\big)\Big]^{1/2} \\
&\leqslant C_\delta \Delta^{\theta/2} + 2^\theta \epsilon^{1/2} \Big[\mathbb{E}\big(\sup_{0\leqslant t\leqslant T} |x(t)|^{2\theta}\big) + \mathbb{E}\big(\sup_{0\leqslant t\leqslant T} |x_\Delta(t)|^{2\theta}\big)\Big]^{1/2} \\
&\leqslant C_\delta \Delta^{\theta/2} + C\epsilon^{1/2},
\end{aligned}$$

where Lemma 3.1 is used in the final step. Let $\Delta \to 0$ and then assertion (3.6) follows as ϵ is arbitrary.

On the other hand, for any $\Delta \in (0, \Delta_1]$ and $\omega \in \Omega_2$, it follows from (3.9) and (3.18) that

$$\sup_{0\leqslant t\leqslant T} \bar{X}_\Delta(t) \leqslant \sup_{0\leqslant t\leqslant T} x_\Delta(t) \leqslant \sup_{0\leqslant t\leqslant T} x(t) + \sup_{0\leqslant t\leqslant T} \big|x_\Delta(t) - x(t)\big| < N - \delta + \frac{\delta}{2} = N - \frac{\delta}{2},$$

and

$$\inf_{0\leqslant t\leqslant T} \bar{X}_\Delta(t) \geqslant \inf_{0\leqslant t\leqslant T} x_\Delta(t) \geqslant \inf_{0\leqslant t\leqslant T} x(t) - \sup_{0\leqslant t\leqslant T} \big|x(t) - x_\Delta(t)\big| > \delta - \frac{\delta}{2} = \frac{\delta}{2}.$$

Fix any $\Delta \in (0, \Delta_1]$. Recalling (3.3) and noting that $\Delta_1 \leqslant \delta/2$, we see that for any $\omega \in \Omega_2$,

$$X_\Delta(t) = \bar{X}_\Delta(t), \quad \forall t \in [0,T].$$

Now consider

$$\begin{aligned}
\mathbb{E}\big(\sup_{0\leqslant t\leqslant T} \big|X_\Delta(t) - \bar{X}_\Delta(t)\big|^\theta\big) &= \mathbb{E}\Big(\mathbb{I}_{\Omega_2^c} \sup_{0\leqslant t\leqslant T} \big|X_\Delta(t) - \bar{X}_\Delta(t)\big|^\theta\Big) \\
&\leqslant \big(\mathbb{P}(\Omega_2^c)\big)^{1/2} \Big[\mathbb{E}\big(\sup_{0\leqslant t\leqslant T} \big|X_\Delta(t) - \bar{X}_\Delta(t)\big|^{2\theta}\big)\Big]^{1/2} \\
&\leqslant 2^\theta \epsilon^{1/2} \Big[\mathbb{E}\big(\sup_{0\leqslant t\leqslant T} |X_\Delta(t)|^{2\theta}\big) + \mathbb{E}\big(\sup_{0\leqslant t\leqslant T} |\bar{X}_\Delta(t)|^{2\theta}\big)\Big]^{1/2} \\
&\leqslant C\epsilon^{1/2},
\end{aligned}$$

where Lemma 3.1 is used in the final step.

Since $\epsilon \in (0,1)$ is arbitrary, assertion (3.7) follows clearly.

□

Theorem 3.4. *Let Assumption 2.1 hold. Then for any $\theta > 0$,*

$$\lim_{\Delta\to 0} \mathbb{E}\big[\sup_{0\leqslant t\leqslant T} \big|X_\Delta(t) - x(t)\big|^\theta\big] = 0.$$

This follows immediately from Lemma 3.2 and 3.3

4 Numerical Study

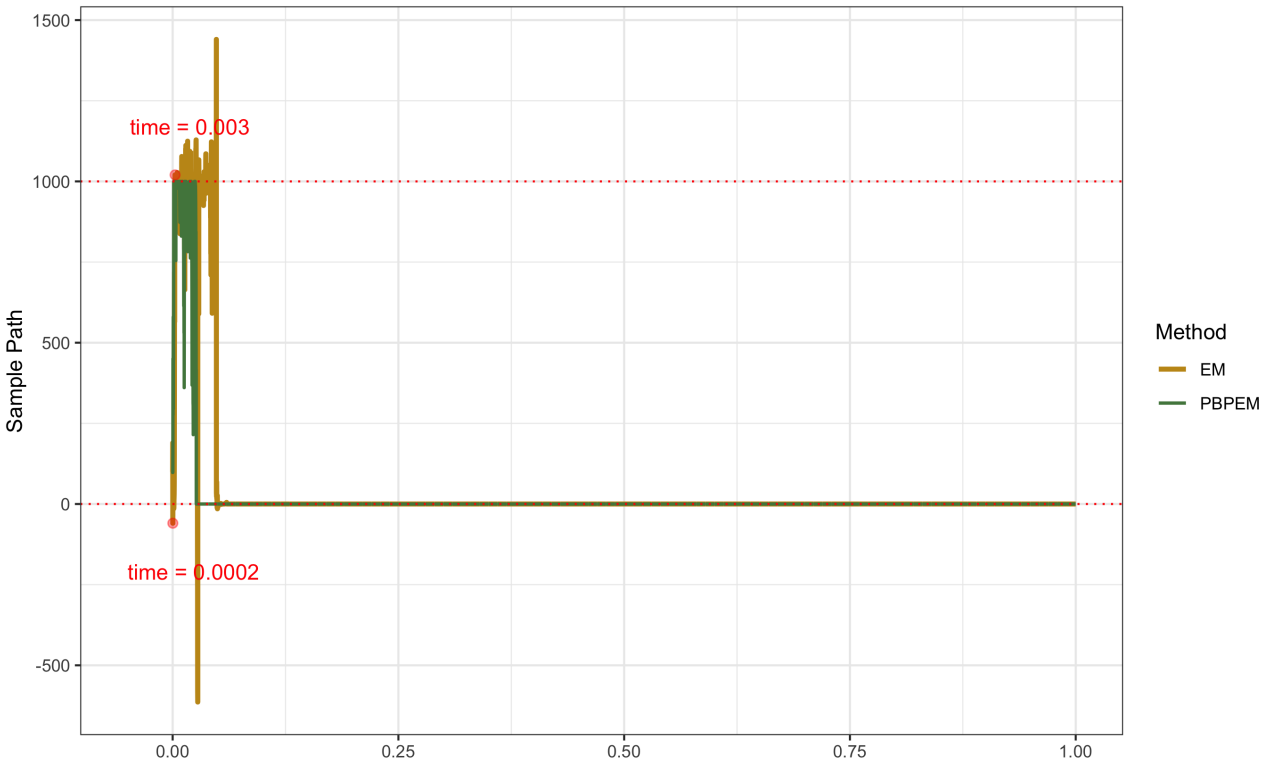

Figure 1: Simulated paths of 10^4 iterations using the PBPEM and the EM, both with $\Delta = 10^{-4}$ and $x(0) = 100$ under the group of parameters (4.2). The two red points respectively show the first time the EM sample path crosses 0 and N.

As stated in Section 1, Higham and Mao's earlier work [8] has proved the strong convergence of the EM algorithm for the mean-reverting square-root process, though the positivity is not guaranteed. In this section, we will reveal the advantages of our newly proposed PBPEM scheme by comparing the performances of these two methods. Notice that when adopting the EM method, SDE (1.2) is replaced by

$$dx(t) = \big[\beta x(t)(N - x(t)) - (\mu + \gamma)x(t)\big]dt + \sigma_1 x(t)(N - x(t))dB_1(t) - \sigma_2 x(t)\sqrt{|N - x(t)|}dB_2(t) \tag{4.1}$$

to avoid computational failure.

Example 4.1. We conduct computer simulations of 10^4 iterations using the PBPEM and the EM method, both with $x(0) = 100$, $\Delta = 10^{-4}$ and parameters

$$\beta = 0.6, N = 10^3, \mu = 21, \gamma = 26, \text{ and } \sigma_1 = \sigma_2 = 0.1. \tag{4.2}$$

Figure 1 generally shows similar patterns developed by the two methods. It has been proved by Cai et al.[2] that the system dies out under this group of parameters, which has been detected by both methods. However, we see that the EM simulated path crosses over the upper and lower boundaries frequently during the early time. In particular, the first drop below 0 occurs at $t = 0.0002$ while the first jump beyond N at $t = 0.003$.

By comparison, the PBPEM scheme prevents the simulated data from overstepping the boundaries for all time. After waving below N for a while, the path quickly approaches 0 but never hit 0.

Example 4.2. We conduct another pair of simulations of 10^4 iterations with parameters

$$\beta = 5, N = 10^3, \mu = 21, \gamma = 26, \sigma_1 = 0.01, \text{ and } \sigma_2 = 0.1. \tag{4.3}$$

From figure 2, the two paths generated by the two methods are both persistent, which is consistent with the result in [2]. However, we spot that the EM simulated path first crosses over N at $t = 0.03$. Since then it happens all the way.

In contrast, the PBPEM scheme prevents the path from passing through N for all time. As a result, the sample path keeps fluctuating below N.

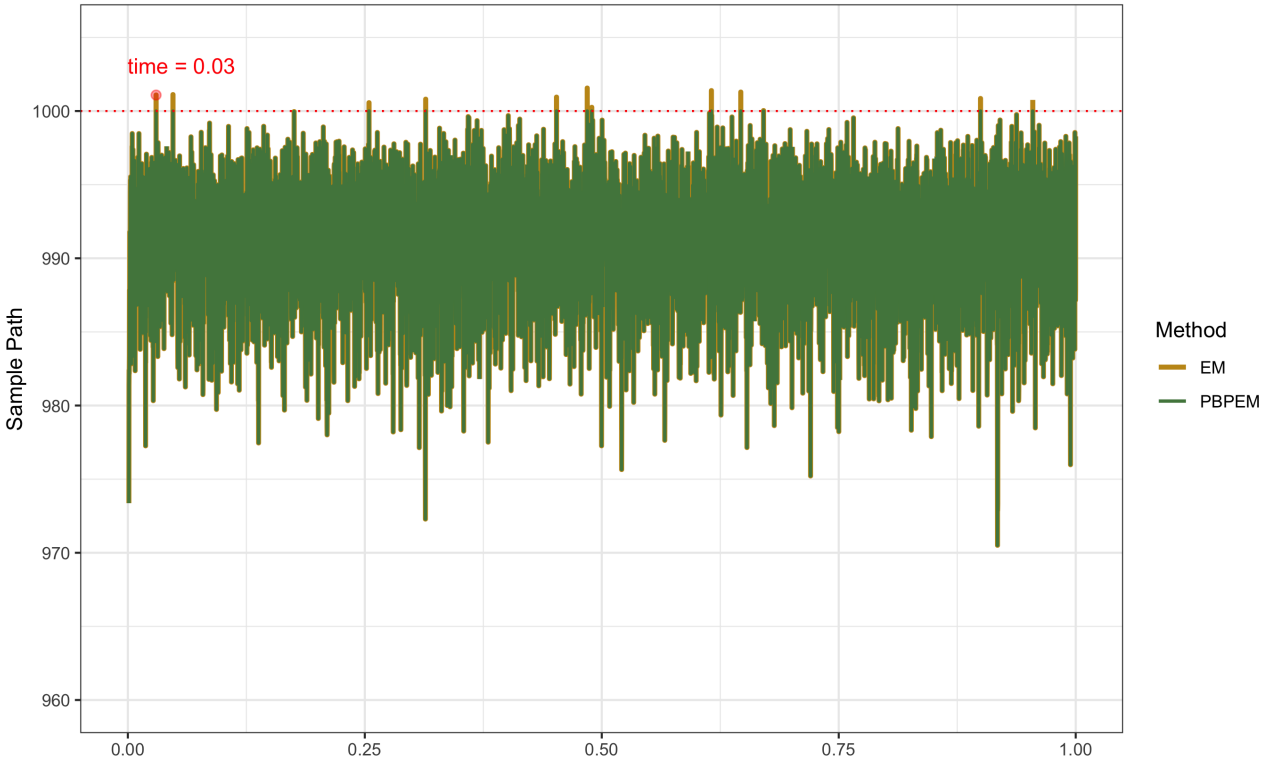

Figure 2: Simulated paths of 10^4 iterations using the PBPEM and the EM, both with $\Delta = 10^{-4}$ and $x(0) = 100$ under the group of parameters (4.3). The red point shows the first time the EM sample path crosses N.

The above two examples have intuitively justified the power of our PBPEM method in preserving the positivity and boundedness of the stochastic SIS model, though both schemes have captured the general long-time behaviour of system (1.2).

5 Useful Comments

Although our newly proposed PBPEM scheme is constructed in the context of the 1-dimensional SIS model (1.2), it can natually be applied to a bunch of popular SDE models which are widely used in finance, biology and engineering. In this section, we focus on two typical models: the mean-reverting square-root process and the SIR epidemic model.

5.1 The mean-reverting square-root process

The mean-reverting square-root process is an important financial model that has come to the fore recently. In 2005, Higham and Mao [8] have shown the EM is a strong convergent approximation

of the mean-reverting square-root process. However, the positivity preserving was left as an open problem. So far, this issue has been considered by several authors, see e.g. [6, 18, 19]. We would like to show now that the idea of our new scheme can also deal with it.

Assumption 5.1. $\sigma^2 \leqslant 2\lambda\xi$

According to [12, 15], under Assumption 5.1, the mean-reverting square-root process (1.3) has a unique global positive solution $s(t)$ on $t \geqslant 0$ with probability 1. We can hence form a positivity preserving EM (PPEM) scheme by modifying the existing PBPEM sheme as follows.

We first define two functions $F_1 : (0,\infty) \to \mathbb{R}$ and $G_1 : (0,\infty) \to \mathbb{R}$ by

$$F_1(s) = \lambda(\xi - s) \quad \text{and} \quad G_1(s) = \sigma\sqrt{s}.$$

Hence SDE (1.3) can be rewritten as

$$ds(t) = f_1(s(t))dt + g_1(s(t))dB(t),$$

where $f_1, g_1 : \mathbb{R} \to \mathbb{R}$ are defined as

$$f_1(s) = F_1(s \vee 0) \quad \text{and} \quad g_1(s) = G_1(s \vee 0).$$

Note that $f_1(\cdot)$ and $g_1(\cdot)$ are linearly bounded. We then formulate the discrete-time PPEM solution $S_\Delta(t_k) \approx s(t_k)$ for $t_k = k\Delta$ by setting $\bar{S}_\Delta(0) = S_\Delta(0) = s(0)$ and computing

$$\begin{aligned}\bar{S}_\Delta(t_{k+1}) &= \bar{S}_\Delta(t_k) + f_1(S_\Delta(t_k))\Delta + g_1(S_\Delta(t_k))\Delta W_k,\\ S_\Delta(t_{k+1}) &= \Delta \vee \bar{S}_\Delta(t_{k+1})\end{aligned}$$

for $k = 0, 1, 2, \cdots$ and $\Delta W_k = W(t_{k+1}) - W(t_k)$. The definition of $S(\cdot)$ is then extended to the whole $t \geqslant 0$ by

$$S_\Delta(t) = S_\Delta(t_k), \quad t \in [t_k, t_{k+1}).$$

We also define a new process

$$s_\Delta(t) = s(0) + \int_0^t f_1(S_\Delta(v))dv + \int_0^t g_1(S_\Delta(v))dW(v) \quad \text{for } t \geqslant 0.$$

Using the key idea of Theorem 3.4, we see that for any $\theta > 0$,

$$\begin{aligned}&\mathbb{E}\big[\sup_{0\leqslant t\leqslant T}\big|S_\Delta(t) - s(t)\big|^\theta\big]\\ &\leqslant \mathbb{E}\big[\sup_{0\leqslant t\leqslant T}\big|S_\Delta(t) - \bar{S}_\Delta(t)\big|^\theta\big] + \mathbb{E}\big[\sup_{0\leqslant t\leqslant T}\big|\bar{S}_\Delta(t) - s_\Delta(t)\big|^\theta\big] + \mathbb{E}\big[\sup_{0\leqslant t\leqslant T}\big|s_\Delta(t) - s(t)\big|^\theta\big]\\ &\to 0 \quad \text{as} \quad \Delta \to 0.\end{aligned}$$

under Assumption 5.1. Thus the PPEM not only converges strongly to the true solution of (1.3), but also preserves the positivity. This is a new advance in the numerical analysis of the mean-reverting square-root process. A pair of simulations of 10^5 iterations is performed using the EM (see [8]) and our new PPTEM with the system parameters given by

$$\lambda = \xi = 1 \quad \text{and} \quad \sigma = 2. \tag{5.1}$$

From Figure 3, at $t = 7.04$ the EM solution drops down to 0, while the PBPEM keeps waving above 0. This supports our theory clearly.

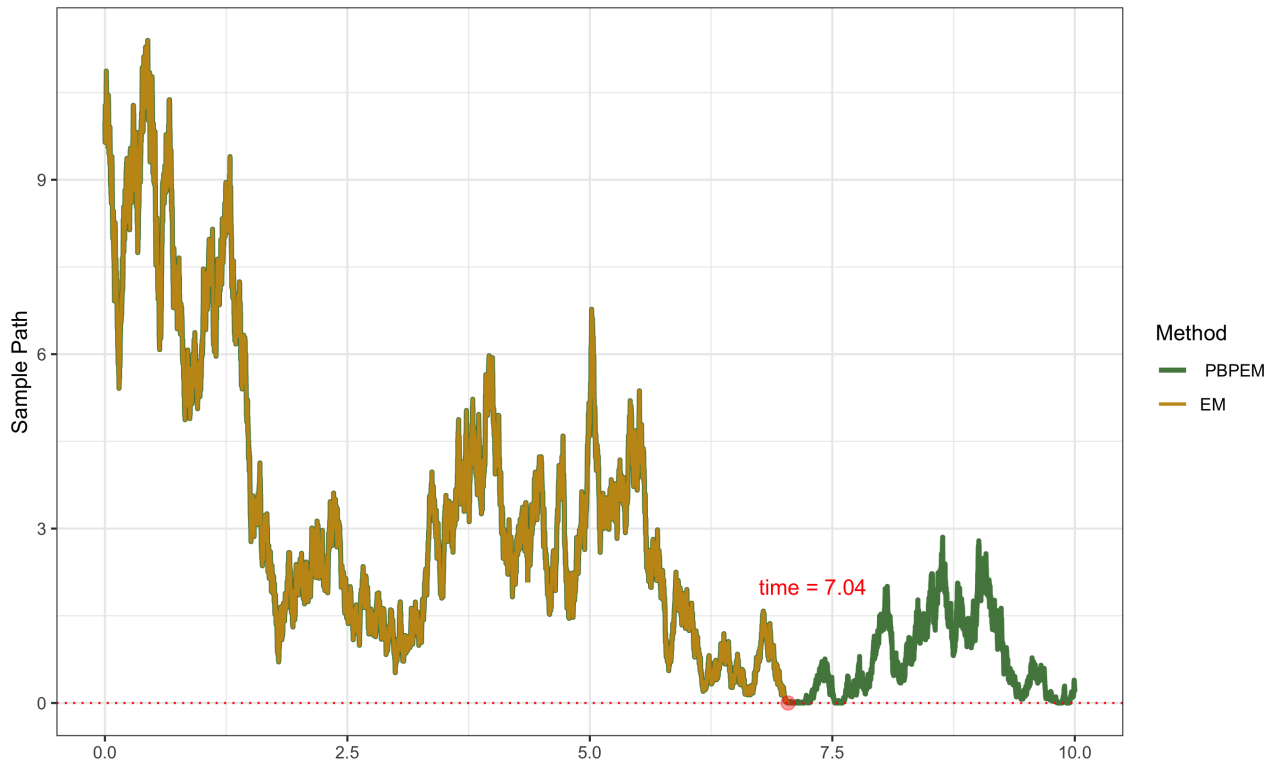

Figure 3: Simulated paths of 10^5 iterations using the PBPEM and the EM, both with $\Delta = 10^{-4}$ and $x(0) = 10$ under the group of parameters (5.1). The red point shows the time the EM sample path crosses 0.

5.2 The SIR epidemic model

Moreover, our PBPEM scheme can easily be extended to the multi-dimensional systems, e.g. the 3-dimensional stochastic SIR model [21] and the stochastic SIRS model [13], each of which has a positive and bounded solution. The stochastic SIR model discussed in [21] has the following form

$$\begin{aligned} dz_1(t) &= (\mu - \beta z_1(t)z_2(t) - \mu z_1(t))dt - \sigma z_1(t)z_2(t)dW(t), \\ dz_2(t) &= (\beta z_1(t)z_2(t) - (\lambda+\mu)z_2(t))dt + \sigma z_1(t)z_2(t)dW(t), \\ dz_3(t) &= (\lambda z_2(t) - \mu z_3(t))dt, \end{aligned} \tag{5.2}$$

where $z_1(t), z_2(t)$ and $z_3(t)$ respectively represent the susceptible group to the disease, the infected group and the recovered one, and $\lambda, \beta, \mu, \sigma$ are positive constants. From [11], SDE (5.2) has a unique solution $(z_1(t), z_2(t), z_3(t))^T$ on $t \geqslant 0$ that remains in $\mathbb{R}^3_+$ almost surely.

Moreover, given that the total amount of population is a constant N, we then normalize the variables to $N = 1$, namely, $z_1(t) + z_2(t) + z_3(t) = 1$ for all $t \geqslant 0$. It is then necessary to study the following 2-dimensional SDE

$$\begin{aligned} dz_1(t) &= (\mu - \beta z_1(t)z_2(t) - \mu z_1(t))dt - \sigma z_1(t)z_2(t)dW(t), \\ dz_2(t) &= (\beta z_1(t)z_2(t) - (\lambda+\mu)z_2(t))dt + \sigma z_1(t)z_2(t)dW(t). \end{aligned} \tag{5.3}$$

Let $z(t) = (z_1(t), z_2(t))^T$ be the solution to (5.3). We see that the solution is positive and bounded, however, this is not guaranteed by the existing numerical methods. Now we show that this problem can also be fixed by the PBPEM.

The PBPEM scheme is set up as below. Let $F_2, G_2 : \mathbb{R}^2_+ \to \mathbb{R}^2$

$$F_2(z) = \begin{bmatrix} -\beta z_1 z_2 - \mu z_1 + \mu \\ \beta z_1 z_2 - (\lambda+\mu)z_2 \end{bmatrix} \quad \text{and} \quad G_2(z) = \begin{bmatrix} -\sigma z_1 z_2 \\ \sigma z_1 z_2 \end{bmatrix}.$$

Then define $f_2, g_2 : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$f_2(z) = F_2(\pi_{01}(z)) \quad \text{and} \quad g_2(z) = G_2(\pi_{01}(z)),$$

where

$$\pi_{01}(z) = \begin{cases} (z_1^0, z_2^0)^T & z_1^0 + z_2^0 \leqslant 1, \\ (\bar{z}_1^0, \bar{z}_2^0)^T & \text{otherwise}, \end{cases}$$

and

$$z_i^0 = z_i \vee 0, \quad \bar{z}_i^0 = \frac{z_i^0}{z_1^0 + z_2^0} \quad \text{for } i = 1, 2.$$

We can then rewrite (5.3) as

$$dz(t) = f_2(z(t))dt + g_2(z(t))dW(t).$$

Notice that $f_2(\cdot)$ and $g_2(\cdot)$ are both globally Lipschitz continuous. Defining

$$\pi_\Delta(z) = \begin{cases} (z_1^\Delta, z_2^\Delta)^T & z_1^\Delta + z_2^\Delta \leqslant 1 - \Delta, \\ (1-\Delta)(\bar{z}_1^0, \bar{z}_2^0)^T & z_1^\Delta + z_2^\Delta > 1 - \Delta \ \& \ \bar{z}_1^0, \bar{z}_2^0 \geqslant \Delta/(1-\Delta), \\ (\Delta, 1-2\Delta)^T & z_1^\Delta + z_2^\Delta > 1 - \Delta \ \& \ \bar{z}_1^0 < \Delta/(1-\Delta), \\ (1-2\Delta, \Delta)^T & z_1^\Delta + z_2^\Delta > 1 - \Delta \ \& \ \bar{z}_2^0 < \Delta/(1-\Delta), \end{cases}$$

where

$$z_i^\Delta = z_i \vee \Delta \quad \text{for } i = 1, 2,$$

we form the discrete-time PBPEM solution $Z_\Delta(t_k) \approx z(t_k)$ for $t_k = k\Delta$ by setting $\bar{Z}_\Delta(0) = Z_\Delta(0) = z(0)$ and computing

$$\begin{aligned} \bar{Z}_\Delta(t_{k+1}) &= \bar{Z}_\Delta(t_k) + f_2(Z_\Delta(t_k))\Delta + g_2(Z_\Delta(t_k))\Delta W_k, \\ Z_\Delta(t_{k+1}) &= \pi_\Delta(\bar{Z}_\Delta(t_{k+1})) \end{aligned}$$

for $k = 0, 1, 2, \cdots$ and $\Delta W_k = W(t_{k+1}) - W(t_k)$. We then extend the definition of $Z_\Delta(\cdot)$ to the whole $t \geqslant 0$ by

$$Z_\Delta(t) = Z_\Delta(t_k), \quad t \in [t_k, t_{k+1}).$$

We also define a new process

$$z_\Delta(t) = z(0) + \int_0^t f_2(Z_\Delta(s))ds + \int_0^t g_2(Z_\Delta(s))dW(s) \quad \text{for } t \geqslant 0.$$

Denote the ith component of $z_\Delta(t), \bar{Z}(t)$ and $Z(t)$ by $z_{\Delta,i}(t), \bar{Z}_{\Delta,i}(t)$ and $Z_{\Delta,i}(t)$ for $i = 1, 2$. Similar to Lemma 3.2, we can deduce that for any $\theta > 0$

$$\lim_{\Delta \to 0} \mathbb{E}\big[\sup_{0 \leqslant t \leqslant T} \big|\bar{Z}_\Delta(t) - z_\Delta(t)\big|^\theta \big] = 0. \tag{5.4}$$

Also, using the key idea of Lemma 3.3, we see that for any $\theta > 0$,

$$\lim_{\Delta \to 0} \mathbb{E}\big[\sup_{0 \leqslant t \leqslant T} \big|z_\Delta(t) - z(t)\big|^\theta \big] = 0 \tag{5.5}$$

and

$$\lim_{\Delta \to 0} \mathbb{E}\big[\sup_{0 \leqslant t \leqslant T} \big|Z_\Delta(t) - \bar{Z}_\Delta(t)\big|^\theta \big] = 0, \tag{5.6}$$

Here gives the sketch of proof.

Sketch of proof. Let ϵ be arbitrary. Find a $\delta = \delta(\epsilon) \in (0,1)$ so small that

$$\mathbb{P}(\Omega_{01}) \geqslant 1 - \epsilon/2$$

with

$$\Omega_{01} = \{ \min_{0\leqslant t\leqslant T} z_1(t) > \delta, \min_{0\leqslant t\leqslant T} z_2(t) > \delta, \max_{0\leqslant t\leqslant T} \big(z_1(t) + z_2(t)\big) < 1 - \delta\}$$

With the chosen δ, define

$$f_{2\delta}(z) = F_2(\pi_\delta(z)) \quad \text{and} \quad g_{2\delta}(z) = G_2(\pi_\delta(z)) \quad \text{for } z \in \mathbb{R}^2,$$

where

$$\pi_\delta = \begin{cases} (z_1^\delta, z_2^\delta)^T & z_1^\delta + z_2^\delta \leqslant 1 - \delta/2, \\ (1-\delta/2)(\bar{z}_1^0, \bar{z}_2^0)^T & z_1^\delta + z_2^\delta > 1 - \delta/2 \ \&\ \bar{z}_1^0, \bar{z}_2^0 \geqslant \delta/(2-\delta), \\ (\delta/2, 1-\delta)^T & z_1^\delta + z_2^\delta > 1 - \delta/2 \ \&\ \bar{z}_1^0 < \delta/(2-\delta), \\ (1-\delta, \delta/2)^T & z_1^\delta + z_2^\delta > 1 - \delta/2 \ \&\ \bar{z}_2^0 < \delta/(2-\delta), \end{cases}$$

where

$$z_i^\delta = z_i \vee \delta/2 \quad \text{for } i = 1, 2.$$

Hence the SDE

$$du(t) = f_{2\delta}(u(t))dt + g_{2\delta}(u(t))dW(t)$$

on $t \geqslant 0$ with initial value $u(0) = z(0)$ has a unique global solution $u(t)$ on $y \geqslant 0$. For any $\Delta \in (0,1]$, we form the EM solution $U_k \approx u(t_k)$ for $t_k = k\Delta$ by setting $U_0 = u(0)$ and computing

$$U_{k+1} = U_k + f_{2\delta}(U_k)\Delta + g_{2\delta}(U_k)\Delta W_k \quad \text{for } k = 0, 1, \cdots.$$

Then we define

$$U(t) = U_k, \quad t \in [t_k, t_{k+1}).$$

and hence the Itô process

$$u_\Delta(t) = z(0) + \int_0^t f_{2\delta}(U(s))ds + \int_0^t g_{2\delta}(U(s))dW(s).$$

Let the ith component of $u_\Delta(t), \bar{U}(t)$ and $U(t)$ be $u_{\Delta,i}(t), \bar{U}_{\Delta,i}(t)$ and $U_{\Delta,i}(t)$ for $i = 1, 2$. We can then deduce

$$\mathbb{E}\big(\sup_{0\leqslant t\leqslant T} |u(t) - u_\Delta(t)|^\theta \big) \leqslant C_\delta \Delta^{\theta/2}, \quad \forall \Delta \in (0,1].$$

Defining

$$\Omega_{02} = \{ \sup_{0\leqslant t\leqslant T} |u(t) - u_\Delta(t)| < \delta/4\} \cap \Omega_{01},$$

one can find a $\Delta_{01} \in (0, \delta/2]$ such that

$$\mathbb{P}\big(\Omega_{02}\big) \geqslant 1 - \epsilon, \quad \forall \Delta \in (0, \Delta_{01}].$$

Note that whenever $\omega \in \Omega_{02} \subset \Omega_{01}$, we always have

$$u(t) = z(t), \quad t \in [0, T],$$

and therefore

$$\sup_{0\leqslant t\leqslant T}|z(t)-u_\Delta(t)|<\delta/4.$$

This implies

$$\min_{0\leqslant t\leqslant T}u_{\Delta,1}(t)>3\delta/4,\quad \min_{0\leqslant t\leqslant T}u_{\Delta,2}(t)>3\delta/4\quad\text{and}\quad \max_{0\leqslant t\leqslant T}\big(u_{\Delta,1}(t)+u_{\Delta,2}(t)\big)<1-\delta/2,\quad \forall\omega\in\Omega_{02}.$$

We then conclude that for any $\omega\in\Omega_{02}$

$$u_\Delta(t)=z_\Delta(t)\quad\text{for } t\in[0,T] \text{ and } \Delta\in(0,\Delta_{01}].$$

As a result, for any $\Delta\in(0,\Delta_{01}]$

$$\mathbb{E}\big(\sup_{0\leqslant t\leqslant T}|z(t)-z_\Delta(t)|^\theta\mathbb{I}_{\Omega_{02}}(\omega)\big)\leqslant C_\delta\Delta^{\theta/2},\tag{5.7}$$

and for $\omega\in\Omega_{02}$,

$$\sup_{0\leqslant t\leqslant T}|z(t)-z_\Delta(t)|<\delta/4.\tag{5.8}$$

Inequality (5.7) immediately yields assertion (5.6), while (5.8) indicates that for any $\Delta\in(0,\Delta_{01}]$ and $\omega\in\Omega_{02}$,

$$\begin{aligned}&\sup_{0\leqslant t\leqslant T}\big(\bar Z_{\Delta,1}(t)+\bar Z_{\Delta,2}(t)\big)\\ \leqslant&\sup_{0\leqslant t\leqslant T}\big(z_{\Delta,1}(t)+z_{\Delta,2}(t)\big)\\ \leqslant&\sup_{0\leqslant t\leqslant T}\big(z_1(t)+z_2(t)+|z_1(t)-z_{\Delta,1}(t)|+|z_2(t)-z_{\Delta,2}(t)|\big)\\ \leqslant&1-\delta+\delta/2=1-\delta/2\end{aligned}$$

and

$$\inf_{0\leqslant t\leqslant T}\bar Z_{\Delta,i}(t)\geqslant\inf_{0\leqslant t\leqslant T}z_{\Delta,i}(t)\geqslant\inf_{0\leqslant t\leqslant T}z_i(t)-\sup_{0\leqslant t\leqslant T}|z_i(t)-z_{\Delta,i}(t)|>\delta-\delta/4=3\delta/4\quad\text{for } i=1,2,$$

which suggests

$$Z_\Delta(t)=\bar Z_\Delta(t),\quad t\in[0,T].$$

This implies the required assertion (5.6) immediately. Combining (5.4)-(5.6) leads to the strong convergence of PBPEM. □

A pair of simulations of 10^4 iterations is performed using the EM and PPTEM with parameters given by

$$\beta=3,\quad \mu=30,\quad \lambda=2,\quad\text{and}\quad \sigma=5.5.\tag{5.9}$$

From Figure 4, at $t=0.49$ the EM solution of the susceptible individuals touches 1, and the infective and recovered population hit 0. However, the PBPEM keeps positive and bounded for all time.

We have demonstrated via these two well-known SDE models that the numerical method established in this paper can be modified easily to fit into a large class of SDE models and hence is of full value.

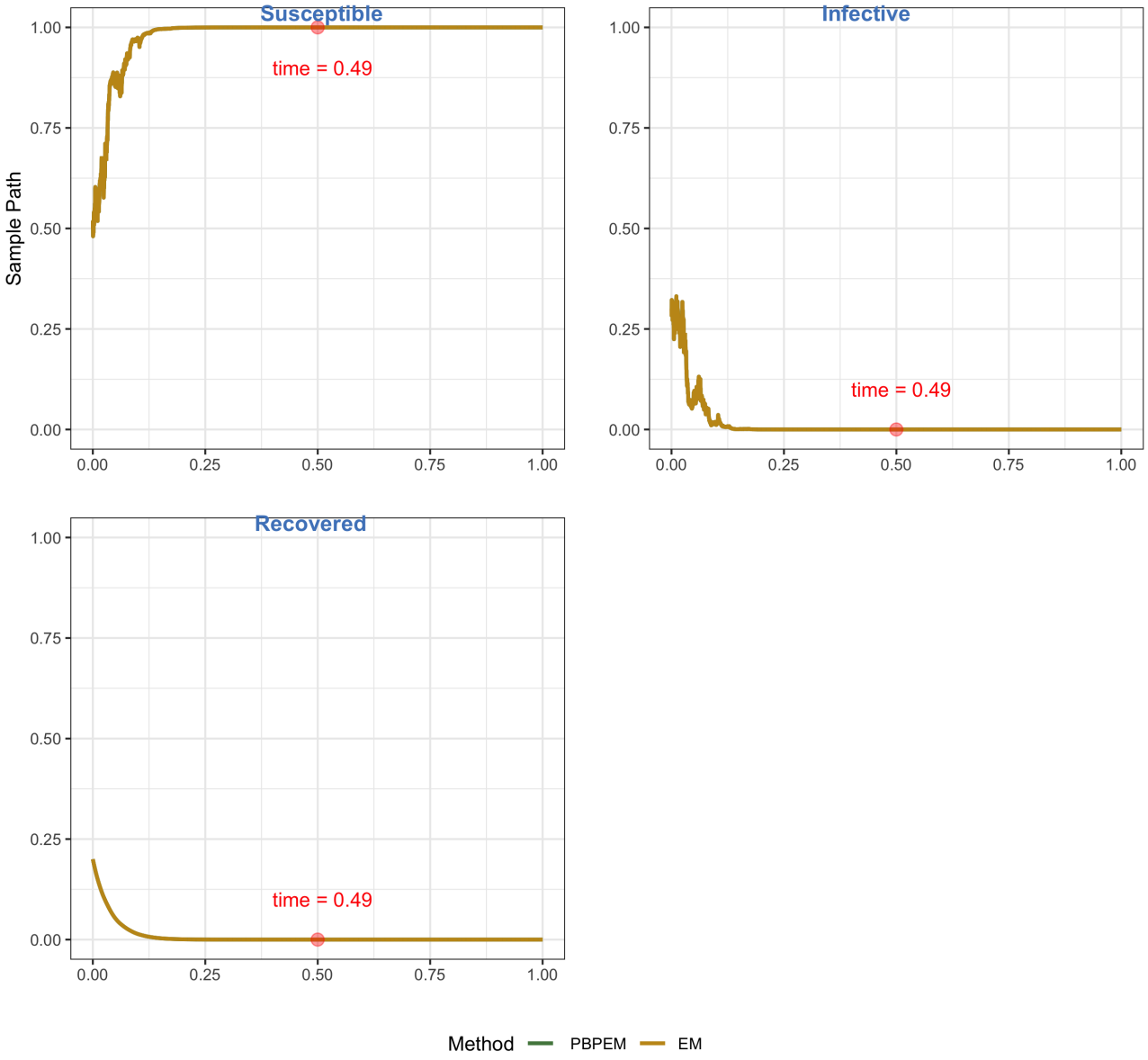

Figure 4: Simulated paths of 10^4 iterations using the PBPEM and the EM, both with $\Delta = 10^{-4}$ and $z(0) = (0.5, 0.3, 0.2)^T$ under the group of parameters (5.9). The red points show the time the EM sample paths cross 0 or 1.

6 Conclusion

In this paper we established a much better condition under which the stochastic epidemic model proposed by Cai et al.[2] has a unique positive and bounded solution. We then modified the classical Euler-Maruyama (EM) scheme to create the positivity and boundedness preserving Euler-Maruyama (PBPEM) scheme. We showed that the PBPEM convergences strongly to the true solution over finite time intervals. We have also demonstrated that the key of the method is applicable to a bunch of popular SDE models including the mean-reverting square-root process and the SDE SIR epidemic model. The advantages of this method were fully illustrated by the numerical simulations.

Acknowledgements

The authors would like to thank the editors and the referees for their professional comments and suggestions. Y. Cai acknowledges the financial support from Zhejiang Natural Science Foundation

(LQ22A010009). J. Hu would like to thank the National Natural Science Foundation of China (61876192), the Fundamental Research Funds for the Central Universities (CZT20020), and Academic Team in Universities (KTZ20051) for their financial support. X. Mao would like to thank the Royal Society (WM160014, Royal Society Wolfson Research Merit Award), the Royal Society of Edinburgh (RSE1832), Shanghai Administration of Foreign Experts Affairs (21WZ2503700, the Foreign Expert Program) for their financial support.

References

[1] D. F. Anderson, D. J. Higham, and Y. Sun. Multilevel Monte Carlo for stochastic differential equations with small noise. *SIAM Journal on Numerical Analysis*, 54(2):505–529, 2016.

[2] S. Cai, Y. Cai, and X. Mao. A stochastic differential equation SIS epidemic model with two independent Brownian motions. *Journal of Mathematical Analysis and Applications*, 474(2):1536 – 1550, 2019.

[3] S. Cai, Y. Cai, and X. Mao. A stochastic differential equation SIS epidemic model with regime switching. *Discrete & Continuous Dynamical Systems-B*, 26(9):4887, 2021.

[4] L. Chen, S. Gan, and X. Wang. First order strong convergence of an explicit scheme for the stochastic SIS epidemic model. *Journal of Computational and Applied Mathematics*, 392:113482, 2021.

[5] A. Gray, D. Greenhalgh, L. Hu, X. Mao, and J. Pan. A stochastic differential equation SIS epidemic model. *SIAM Journal on Applied Mathematics*, 71(3):876–902, 2011.

[6] N. Halidias. A new numerical scheme for the CIR process. *Monte Carlo Methods and Applications*, 21(3):245–253, 2015.

[7] H. W. Hethcote and J. A. Yorke. *Gonorrhea transmission dynamics and control*, volume 56. Springer, 2014.

[8] D. Higham and X. Mao. Convergence of Monte Carlo simulations involving the mean-reverting square root process. *Journal of Computational Finance*, 8(3):35–62, 2005.

[9] M. Hutzenthaler, A. Jentzen, P. E. Kloeden, et al. Strong convergence of an explicit numerical method for sdes with nonglobally lipschitz continuous coefficients. *Annals of Applied Probability*, 22(4):1611–1641, 2012.

[10] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes*. Elsevier, 2014.

[11] C. Ji, D. Jiang, and N. Shi. The behavior of an SIR epidemic model with stochastic perturbation. *Stochastic analysis and applications*, 30(5):755–773, 2012.

[12] I. Karatzas and S. Shreve. *Brownian motion and stochastic calculus*, volume 113. springer, 2014.

[13] A. Lahrouz, L. Omari, and D. Kiouach. Global analysis of a deterministic and stochastic nonlinear SIRS epidemic model. *Nonlinear Analysis: Modelling and Control*, 16(1):59–76, 2011.

[14] W. Liu and X. Mao. Strong convergence of the stopped Euler–Maruyama method for nonlinear stochastic differential equations. *Applied Mathematics and Computation*, 223:389–400, 2013.

[15] X. Mao. *Stochastic differential equations and applications*. Elsevier, Horwood, Chichester, 2007.

[16] X. Mao. The truncated Euler–Maruyama method for stochastic differential equations. *Journal of Computational and Applied Mathematics*, 290:370–384, 2015.

[17] X. Mao, G. Marion, and E. Renshaw. Environmental Brownian noise suppresses explosions in population dynamics. *Stochastic Processes and their Applications*, 97(1):95 – 110, 2002.

[18] A. Neuenkirch and L. Szpruch. First order strong approximations of scalar SDEs defined in a domain. *Numerische Mathematik*, 128(1):103–136, 2014.

[19] J. Tan, Y. Chen, W. Men, and Y. Guo. Positivity and convergence of the balanced implicit method for the nonlinear jump-extended CIR model. *Mathematics and Computers in Simulation*, 182:195–210, 2021.

[20] J. Tan, W. Men, Y. Pei, and Y. Guo. Construction of positivity preserving numerical method for stochastic age-dependent population equations. *Applied Mathematics and Computation*, 293:57–64, 2017.

[21] E. Tornatore, S. M. Buccellato, and P Vetro. Stability of a stochastic SIR system. *Physica A: Statistical Mechanics and its Applications*, 354:111–126, 2005.