Almost sure exponential stability of backward Euler–Maruyama discretizations for hybrid stochastic differential equations

Xuerong Mao, Yi Shen, Alison Gray

PII: S0377-0427(10)00444-9
DOI: 10.1016/j.cam.2010.08.006
Reference: CAM 8008

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

Received date: 15 June 2009
Revised date: 3 August 2010

Almost Sure Exponential Stability of Backward Euler–Maruyama Discretizations for Hybrid Stochastic Differential Equations

Xuerong Mao[1,*], **Yi Shen**[2], **Alison Gray**[1]
[1] Department of Mathematics and Statistics,
University of Strathclyde, Glasgow G1 1XH, U.K.
[2] Department of Control Science and Engineering,
Huazhong University of Science and Technology,
Wuhan, Hubei, 430074, China.

Abstract

This is a continuation of the first author's earlier paper [17] jointly with Pang and Deng, in which the authors established some sufficient conditions under which the Euler–Maruyama (EM) method can reproduce the almost sure exponential stability of the test hybrid SDEs. The key condition imposed in [17] is the *global Lipschitz condition*. However, we will show in this paper that without this global Lipschitz condition the EM method may not preserve the almost sure exponential stability. We will then show that the backward EM method can capture almost sure exponential stability for a certain class of highly nonlinear hybrid SDEs.

Key words: Brownian motion, backward Euler-Maruyama, Markov chain, almost sure exponential stability.

1 Introduction

This is a continuation of the first author's earlier paper [17] jointly with Pang and Deng. It is concerned with the long time dynamics of numerical simulations of hybrid stochastic differential equations (SDEs). The research in this direction is motivated by the question "for what choices of step size does a numerical method reproduce the characteristics of a test SDE?" One of the important characteristics of the test SDE is the stability. Indeed, the stability analysis of numerical methods for SDEs has recently received a great deal of attention (see e.g. [5, 7, 8, 10, 18, 19]). More recently, the authors in [17] established some sufficient conditions under which the Euler–Maruyama (EM) method can reproduce the almost sure exponential stability of the test hybrid SDEs. The key condition imposed in [17] is the *global Lipschitz condition*. However, there are many hybrid SDEs which do

*Corresponding author. E-mail: x.mao@strath.ac.uk

not obey the global Lipschitz condition but are almost surely exponentially stable. For example, consider the scalar hybrid cubic SDE

$$dx(t) = [\alpha(r(t))x(t) - x^3(t)]dt + \beta(r(t))x(t)dB(t), \tag{1.1}$$

where $B(t)$ is a scalar Brownian motion, $r(t)$ is a Markov chain and the parameters $\alpha(\cdot)$ and $\beta(\cdot)$ will be specified in Section 3. We will show that this SDE is almost surely exponentially stable. Further such examples can be found in [16].

Two questions arise:

- Can the EM method preserve the almost sure exponential stability without the global Lipschitz condition?
- If not, what other numerical method can preserve the almost sure exponential stability?

The answer to the first question is not positive. In fact we shall show in Section 3 that the EM method CANNOT reproduce the stability characteristic of the SDE (1.1). Our aim in this paper is to seek a positive answer to the second question. We look for conditions under which positive results of the backward Euler–Maruyama (BEM) method can be derived in the small step size setting. Our work therefore builds on the well known and highly informative analysis for deterministic problems and its more recent extension to SDEs [2, 3, 4, 5, 7, 9, 10, 11, 18, 19].

2 Notation

Throughout this paper, we let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions (i.e. it is increasing and right continuous while $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets) and we let $B(t)$ be a scalar Brownian motion defined on the probability space.

We let $r(t)$, $t \geq 0$, be a right-continuous Markov chain on the probability space taking values in a finite state space $\mathbb{S} = \{1, 2, \dots, N\}$ and independent of the Brownian motion $B(\cdot)$, where N is a positive integer. The corresponding generator is denoted by $\Gamma = (\gamma_{ij})_{N\times N}$, so that

$$\mathbb{P}\{r(t+\delta) = j \mid r(t) = i\} = \begin{cases} \gamma_{ij}\delta + o(\delta) & \text{if } i \neq j, \\ 1 + \gamma_{ij}\delta + o(\delta) & \text{if } i = j, \end{cases}$$

where $\delta > 0$. Here γ_{ij} is the transition rate from i to j and $\gamma_{ij} > 0$ if $i \neq j$ while $\gamma_{ii} = -\sum_{j\neq i}\gamma_{ij}$. We note that almost every sample path of $r(\cdot)$ is a right continuous step function with a finite number of sample jumps in any finite subinterval of $\mathbb{R}_+ := [0, \infty)$ (see e.g. [1]). As a standing hypothesis, we assume moreover in this paper that the Markov chain is *irreducible*. This is equivalent to the condition that for any $i, j \in \mathbb{S}$, we can find $i_1, i_2, \dots, i_k \in \mathbb{S}$ such that

$$\gamma_{i,i_1}\gamma_{i_1,i_2}\cdots\gamma_{i_k,j} > 0.$$

Note that Γ always has an eigenvalue 0. The algebraic interpretation of irreducibility is that $\text{rank}(\Gamma) = N - 1$. Under this condition, the Markov chain has a unique stationary (probability) distribution $\pi = (\pi_1, \pi_2, \ldots, \pi_N) \in \mathbb{R}^{1\times N}$ which can be determined by solving

$$\begin{cases} \pi\Gamma = 0 \\ \text{subject to } \sum_{j=1}^{N} \pi_j = 1 \text{ and } \pi_j > 0 \text{ for all } j \in \mathbb{S}. \end{cases}$$

Let $|\cdot|$ denote both the Euclidean norm in $\mathbb{R}^n$ and the trace (or Frobenius) norm in $\mathbb{R}^{n\times m}$. If A is a vector or matrix, its transpose is denoted by A^T. The inner product of x, y in $\mathbb{R}^n$ is denoted by $\langle x, y\rangle$. We use a.s. to mean almost surely. We will denote the indicator function of a set G by I_G.

We are concerned with the n-dimensional nonlinear hybrid SDE

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

on $t \geq 0$, given $x(0) = x_0 \neq 0$ in $\mathbb{R}^n$ and $r(0) = i_0 \in \mathbb{S}$. As a standing hypothesis, we assume that $f, g : \mathbb{R}^n \times \mathbb{S} \to \mathbb{R}^n$ are smooth enough for the SDE (2.1) to have a unique global solution $x(t)$ on $[0, \infty)$ (see, for example, [16], for sufficient conditions). We make two remarks.

- Scalar Brownian motion $B(t)$ is used to make the analysis in Sections 4 and 5 more accessible. In Section 6 we state how our results can be extended to the case of multidimensional noise.
- The restriction to a deterministic initial condition is convenient and does not lose any generality when almost sure asymptotic stability is studied; see, for example, [13, 14, 15].

We now introduce the EM method and the BEM method. The methods make use of the following lemma (see [1]).

Lemma 2.1 *Given a time step $\Delta > 0$, let $r_k^\Delta = r(k\Delta)$ for $k \geq 0$. Then $\{r_k^\Delta,\ k = 0, 1, 2, \ldots\}$ is a discrete time N-state Markov chain with the one-step transition probability matrix*

$$P(\Delta) = (P_{ij}(\Delta))_{N\times N} = e^{\Delta\Gamma}. \tag{2.2}$$

Given a fixed step size $\Delta \in (0, 1)$ and the one-step transition probability matrix $P(\Delta)$ in (2.2), the discrete Markov chain $\{r_k^\Delta,\ k = 0, 1, 2, \ldots\}$ can be simulated as follows: Let $r_0^\Delta = i_0$ and compute a pseudo-random number ξ_1 from the uniform $(0, 1)$ distribution. Define

$$r_1^\Delta = \begin{cases} i_1 & \text{if } i_1 \in \mathbb{S} - \{N\} \text{ such that } \sum_{j=1}^{i_1-1} P_{i_0,j}(\Delta) \leq \xi_1 < \sum_{j=1}^{i_1} P_{i_0,j}(\Delta), \\ N & \text{if } \sum_{j=1}^{N-1} P_{i_0,j}(\Delta) \leq \xi_1, \end{cases}$$

where we set $\sum_{j=1}^{0} P_{i_0,j}(\Delta) = 0$ as usual. In other words, we ensure that the probability of state s being chosen is given by $\mathbb{P}(r_1^\Delta = s) = P_{i_0,s}(\Delta)$. Generally, having computed

$r_0, r_1, r_2, \ldots, r_k$, we compute r_{k+1} by drawing a uniform $(0,1)$ pseudo-random number ξ_{k+1} and setting

$$r_{k+1}^{\Delta} = \begin{cases} i_{k+1} & \text{if } i_{k+1} \in \mathbb{S} - \{N\} \text{ such that } \sum_{j=1}^{i_{k+1}-1} P_{r_k^{\Delta},j}(\Delta) \leq \xi_{k+1} < \sum_{j=1}^{i_{k+1}} P_{r_k^{\Delta},j}(\Delta), \\ N & \text{if } \sum_{j=1}^{N-1} P_{r_k^{\Delta},j}(\Delta) \leq \xi_{k+1}. \end{cases}$$

This procedure can be carried out independently to obtain more trajectories.

Having explained how to simulate the discrete Markov chain, we now define the EM approximation for the hybrid SDE (2.1). The discrete approximations $X_k \approx x(t_k)$, with $t_k = k\Delta$, are formed by setting $X_0 = x_0,\ r_0^{\Delta} = i_0$ and, generally,

$$X_{k+1} = X_k + f(X_k, r_k^{\Delta})\Delta + g(X_k, r_k^{\Delta})\Delta B_k, \quad k \geq 0, \tag{2.3}$$

where $\Delta B_k = B(t_{k+1}) - B(t_k)$. In words, r_k^{Δ} defines which of the N SDEs is currently active, and we apply the EM to this SDE. Compared with the numerical analysis of standard SDEs, a new source of error arises in the method (2.3): the switching can only occur at discrete time points $\{t_k\}$, whereas for the underlying continuous-time problem (2.1) the Markov chain can produce a switch at any point in time.

Similarly, the BEM method applied to (2.1) produces approximations $X_k \approx x(t_k)$, where $X_0 = x_0,\ r_0^{\Delta} = i_0$ and, generally,

$$X_{k+1} = X_k + f(X_{k+1}, r_k^{\Delta})\Delta + g(X_k, r_k^{\Delta})\Delta B_k, \quad k \geq 0. \tag{2.4}$$

For the BEM method to be well-defined, we will impose condition (4.1) below and will explain in Section 5 that under (4.1) the BEM method is well-defined for sufficiently small step size Δ.

3 Motivating Example

Let $r(t)$ be a Markov chain with the state space $\mathbb{S} = \{1, 2\}$ and the generator

$$\Gamma = \begin{bmatrix} -\gamma_{12} & \gamma_{12} \\ \gamma_{21} & -\gamma_{21} \end{bmatrix},$$

where $\gamma_{12} > 0$ and $\gamma_{21} > 0$. It is easy to see that its unique stationary distribution $\pi = (\pi_1, \pi_2) \in \mathbb{R}^{1\times 2}$ is given by

$$\pi_1 = \frac{\gamma_{21}}{\gamma_{12} + \gamma_{21}}, \quad \pi_2 = \frac{\gamma_{12}}{\gamma_{12} + \gamma_{21}}.$$

Consider the scalar hybrid cubic SDE

$$dx(t) = [\alpha(r(t))x(t) - x^3(t)]dt + \beta(r(t))x(t)dB(t), \tag{3.1}$$

where

$$\alpha(1) = 1, \quad \beta(1) = 2, \quad \alpha(2) = 0.5, \quad \beta(2) = 1.$$

It follows from Corollary 4.5 in Section 4 below that

$$\begin{aligned}\limsup_{t\to\infty}\frac{1}{t}\log|x(t)| &\leq \pi_1[\alpha(1)-0.5\beta^2(1)]+\pi_2[\alpha(2)-0.5\beta^2(2)]\\ &= -\frac{\gamma_{21}}{\gamma_{12}+\gamma_{21}} \quad \text{a.s.}\end{aligned} \tag{3.2}$$

In other words, the SDE (3.1) is almost surely exponentially stable.

The question is: Can the EM method preserve this almost sure exponential stability? The answer is no. To show this, recall that the EM method (2.3) applied to (3.1) produces

$$X_{k+1} = X_k\left(1+\Delta\alpha(r_k^\Delta)-\Delta X_k^2+\beta(r_k^\Delta)\Delta B_k\right). \tag{3.3}$$

Lemma 3.1 *Let $\{X_k\}_{k\geq 1}$ be defined by (3.3). Suppose $0<\Delta<1$. Then the conditional probability*

$$\mathbb{P}\left(|X_{k+1}|\geq \frac{2^{k+2}}{\sqrt{\Delta}},\ \forall\, k\geq 1 \;\Big|\; |X_1|\geq\frac{2^2}{\sqrt{\Delta}}\right)\geq \exp\left(-\frac{2}{2e^2-1}\right).$$

Proof. First, we show that, for $k\geq 1$,

$$|X_k|\geq\frac{2^{k+1}}{\sqrt{\Delta}} \text{ and } |\Delta B_k|\leq 2^k \quad\Rightarrow\quad |X_{k+1}|\geq\frac{2^{k+2}}{\sqrt{\Delta}}. \tag{3.4}$$

To see this, we compute

$$\begin{aligned}|X_{k+1}| &\geq |X_k|\left|\Delta X_k^2-1-|\alpha(r_k^\Delta)|\Delta-|\beta(r_k^\Delta)||\Delta B_k|\right|\\ &\geq \frac{2^{k+1}}{\sqrt{\Delta}}\left|2^{2k+2}-1-\Delta-2|\Delta B_k|\right|\\ &\geq \frac{2^{k+1}}{\sqrt{\Delta}}\left(2^{k+2}-2-2^{k+1}\right)\\ &= \frac{2^{k+2}}{\sqrt{\Delta}}\left(2^k-1\right)\\ &\geq \frac{2^{k+2}}{\sqrt{\Delta}}.\end{aligned}$$

It then follows from (3.4) that

$$\left(|X_1|\geq\frac{2^2}{\sqrt{\Delta}} \text{ and } |\Delta B_k|\leq 2^k,\ \forall\, k\geq 1\right)\subset\left(|X_k|\geq\frac{2^{k+1}}{\sqrt{\Delta}},\ \forall\, k\geq 1\right).$$

Since X_1 and ΔB_k ($k\geq 1$) are all independent,

$$\mathbb{P}\left(|X_1|\geq\frac{2^2}{\sqrt{\Delta}}\right)\mathbb{P}\left(|\Delta B_k|\leq 2^k,\ \forall\, k\geq 1\right)\leq\mathbb{P}\left(|X_k|\geq\frac{2^{k+1}}{\sqrt{\Delta}},\ \forall\, k\geq 1\right).$$

This implies

$$\begin{aligned}\mathbb{P}\left(|X_{k+1}|\geq\frac{2^{k+2}}{\sqrt{\Delta}},\ \forall\, k\geq 1 \;\Big|\; |X_1|\geq\frac{2^2}{\sqrt{\Delta}}\right) &\geq \mathbb{P}\left(|\Delta B_k|\leq 2^k\ \forall\, k\geq 1\right)\\ &= \prod_{k=1}^{\infty}\mathbb{P}\left(|\Delta B_k|\leq 2^k\right).\end{aligned} \tag{3.5}$$

Now, because $\Delta B_k \sim N(0,\Delta)$ and $\Delta < 1$, we have

$$\begin{aligned}
\mathbb{P}\left(|\Delta B_k| \geq 2^k\right) &= \mathbb{P}\left(|\Delta B_k|/\sqrt{\Delta} \geq 2^k/\sqrt{\Delta}\right) \\
&\leq \mathbb{P}\left(|\Delta B_k|/\sqrt{\Delta} \geq 2^k\right) \\
&= \frac{2}{\sqrt{2\pi}} \int_{2^k}^{\infty} e^{-x^2/2}\, dx \\
&\leq \frac{2}{2^k\sqrt{2\pi}} \int_{2^k}^{\infty} x e^{-x^2/2}\, dx \\
&\leq \frac{1}{2^k} \exp\left(-2^{2k-1}\right).
\end{aligned}$$

Hence, in (3.5)

$$\mathbb{P}\left(|X_{k+1}| \geq \frac{2^{k+2}}{\sqrt{\Delta}},\ \forall\, k \geq 1 \,\middle|\, |X_1| \geq \frac{2^2}{\sqrt{\Delta}}\right) \geq \prod_{k=1}^{\infty}\left(1 - \frac{1}{2^k}\exp\left(-2^{2k-1}\right)\right).$$

Since

$$\log(1-u) \geq -2u, \qquad \text{for } 0 < u < \tfrac{1}{2},$$

we then have

$$\begin{aligned}
\log\left(\mathbb{P}\left(|X_{k+1}| \geq \frac{2^{k+2}}{\sqrt{\Delta}},\ \forall\, k \geq 1 \,\middle|\, |X_1| \geq \frac{2^2}{\sqrt{\Delta}}\right)\right) &\geq \sum_{k=1}^{\infty} \log\left(1 - \frac{1}{2^k}\exp\left(-2^{2k-1}\right)\right) \\
&\geq -\sum_{k=1}^{\infty} \frac{1}{2^{k-1}} \exp\left(-2^{2k-1}\right). \qquad (3.6)
\end{aligned}$$

But, using $2^{2k-1} \geq 2k$,

$$\sum_{k=1}^{\infty} \frac{1}{2^{k-1}} \exp\left(-2^{2k-1}\right) \leq \sum_{k=1}^{\infty} \frac{1}{2^{k-1}} e^{-2k} = e^{-2} \sum_{k=1}^{\infty} \frac{1}{(2e^2)^{k-1}} = \frac{2}{2e^2-1}.$$

Hence, in (3.6),

$$\log\left(\mathbb{P}\left(|X_{k+1}| \geq \frac{2^{k+2}}{\sqrt{\Delta}},\ \forall\, k \geq 1 \,\middle|\, |X_1| \geq \frac{2^2}{\sqrt{\Delta}}\right)\right) \geq -\frac{2}{2e^2-1}$$

and the result follows.

Lemma 3.1 shows that

$$\mathbb{P}\left(|X_k| \geq \frac{2^{k+1}}{\sqrt{\Delta}},\ \forall\, k \geq 1\right) \geq \exp\left(-\frac{2}{2e^2-1}\right) \mathbb{P}\left(|X_1| \geq \frac{2^2}{\sqrt{\Delta}}\right).$$

But we note that given any $x(0) \neq 0$ and any $\Delta > 0$, there is a non-zero probability that the first Brownian increment, ΔB_1, will cause $|X_1| \geq 2^2/\sqrt{\Delta}$, namely

$$\mathbb{P}\left(|X_1| \geq \frac{2^2}{\sqrt{\Delta}}\right) > 0.$$

Hence

$$\mathbb{P}\left(|X_k| \geq \frac{2^{k+1}}{\sqrt{\Delta}},\ \forall\, k \geq 1\right) > 0.$$

In other words, there is a non-zero probability that EM will produce a numerical solution that blows up at a geometric rate. This contrasts with the almost sure exponential stability of the underlying SDE, shown by (3.2).

For example, we set $\gamma_{12} = 1$ and $\gamma_{21} = 4$. Figures 3.1 and 3.2 show the results of two computer simulations based on the EM method with step size $\Delta = 0.001$ and initial values $(x(0), r(0)) = (20, 1)$ and $(50, 1)$, respectively. Both simulations show that the EM method does not capture the stability property of the underlying SDE (3.1), while the second simulation shows that the EM method can blow up very quickly.

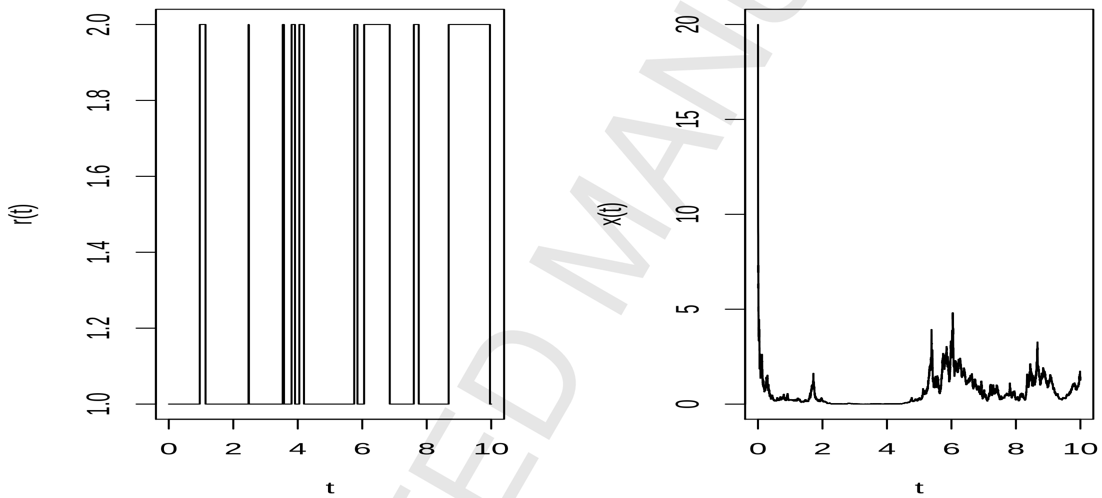

Figure 3.1: Computer simulation of the path $x(t)$ and corresponding state $r(t)$, using the EM method with step size $\Delta = 0.001$, and initial values $(x(0), r(0)) = (20, 1)$. The generator parameters are $\gamma_{12} = 1$ and $\gamma_{21} = 4$, and the SDE parameters are $\alpha(1) = 1, \beta(1) = 2, \alpha(2) = 0.5, \beta(2) = 1$.

4 Stability Criteria

The motivating example above shows that in the case of general nonlinear hybrid SDEs the EM method cannot guarantee to preserve the almost sure exponential stability, even for arbitrarily small step sizes. However, we will show that the BEM method can preserve the stability for a class of nonlinear hybrid SDEs. We look for conditions under which positive results of the BEM method can be derived in the small step size setting. In this section we will establish some sufficient conditions under which the hybrid SDE (2.1) is almost surely exponentially stable, while in the next section we will show that under these conditions the BEM method can preserve this stability. Let us first state the conditions.

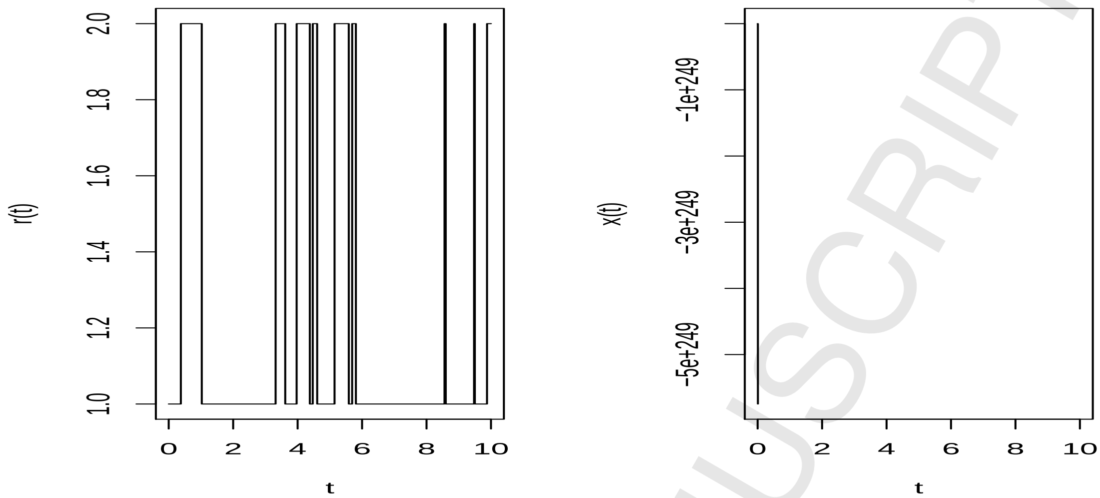

Figure 3.2: Computer simulation of the path $x(t)$ and state $r(t)$, using the EM method with step size $\Delta = 0.001$, and initial values $(x(0), r(0)) = (50, 1)$. The generator parameters are $\gamma_{12} = 1$ and $\gamma_{21} = 4$, and the SDE parameters are $\alpha(1) = 1, \beta(1) = 2, \alpha(2) = 0.5, \beta(2) = 1$.

Assumption 4.1 *Assume that for each $k = 1, 2, \ldots$, there is an $h_k > 0$ such that*

$$|f(x, i)| \leq h_k |x|$$

for all $i \in \mathbb{S}$ and those $x \in \mathbb{R}^n$ with $|x| \leq k$ and, moreover, there is an $h > 0$ such that

$$|g(x, i)| \leq h|x| \quad \forall (x, i) \in \mathbb{R}^n \times \mathbb{S}.$$

Assumption 4.2 *Assume that there is a symmetric positive-definite matrix $Q \in \mathbb{R}^{n\times n}$ and constants μ_i $(i \in \mathbb{S})$ such that*

$$(x-y)^T Q(f(x,i) - f(y,i)) \leq \mu_i (x-y)^T Q(x-y), \quad \forall x, y \in \mathbb{R}^n \tag{4.1}$$

and

$$\sigma_i := \sup_{x\in\mathbb{R}^n, x\neq 0} \left(\frac{g^T(x,i)Qg(x,i)}{x^T Qx} - \frac{2|x^T Qg(x,i)|^2}{(x^T Qx)^2} \right) < \infty. \tag{4.2}$$

We first note that Assumption 4.1 implies that

$$f(0, i) = 0 \quad \text{and} \quad g(0, i) = 0 \quad \text{for all } i \in \mathbb{S}. \tag{4.3}$$

It is therefore easy to observe that the solution of equation (2.1) will remain zero if it starts from zero ($x_0 = 0$). In other words, zero is an *equilibrium* or *stationary state*. This solution $x(t) \equiv 0$ is often called a *trivial solution*. However it is not so obvious to observe from Assumption 4.1 that any solution of equation (2.1) starting from a non-zero state will remain non-zero. In the sequel we will need this non-zero property so we cite a result from [16, Lemma 5.1 on page 164].

Lemma 4.3 *Under Assumption 4.1, if $x_0 \neq 0$, then the solution of equation (2.1) obeys*

$$\mathbb{P}\{x(t) \neq 0 \text{ on } t \geq 0\} = 1.$$

That is, almost all the sample paths of any solution of equation (2.1) starting from a non-zero state will never be zero.

We next observe from (4.1) and (4.3) that

$$x^T Q f(x,i) \leq \mu_i x^T Q x, \quad \forall x \in \mathbb{R}^n. \tag{4.4}$$

We will see from the proof of the next theorem that this is the condition we need for stability. The reason why we impose the stronger condition (4.1) instead of this (4.4) is to guarantee that the BEM (2.4) is well defined (see the details in the beginning of Section 5).

Theorem 4.4 *Let Assumptions 4.1 and 4.2 hold. If*

$$\sum_{i \in \mathbb{S}} \pi_i(\mu_i + 0.5\sigma_i) < 0, \tag{4.5}$$

then the solution of equation (2.1) obeys

$$\limsup \frac{1}{t} \log(|x(t)|) \leq \sum_{i \in \mathbb{S}} \pi_i(\mu_i + 0.5\sigma_i) \quad a.s. \tag{4.6}$$

for all $x_0 \in \mathbb{R}^n$. In other words, (the trivial solution of) equation (2.1) is almost surely exponentially stable.

To highlight the numerical analysis in this paper we defer the proof of this theorem to the Appendix.

If we let Q in Assumption 4.2 be the identity matrix, we obtain the following useful corollary.

Corollary 4.5 *Let Assumption 4.1 hold. Assume that there are constants μ_i ($i \in \mathbb{S}$) such that*

$$(x-y)^T(f(x,i) - f(y,i)) \leq \mu_i |x-y|^2, \quad \forall x, y \in \mathbb{R}^n \tag{4.7}$$

and

$$\sigma_i := \sup_{x \in \mathbb{R}^n, x \neq 0} \left(\frac{|g(x,i)|^2}{|x|^2} - \frac{2|x^T g(x,i)|^2}{|x|^4} \right) < \infty. \tag{4.8}$$

If $\sum_{i \in \mathbb{S}} \pi_i(\mu_i + 0.5\sigma_i) < 0$, then the solution of equation (2.1) obeys (4.6).

As an example, let us verify (3.2). It follows from (3.1) that

$$f(x,i) = \alpha(i)x - x^3 \quad \text{and} \quad g(x,i) = \beta(i)x$$

for $(x,i) \in \mathbb{R} \times \{1,2\}$. Hence

$$(x-y)(f(x,i) - f(y,i)) = \alpha(i)|x-y|^2 - (x-y)(x^3 - y^3) \leq \alpha(i)|x-y|^2$$

and

$$\sup_{x\in\mathbb{R},x\neq 0}\left(\frac{|g(x,i)|^2}{|x|^2}-\frac{2|x^Tg(x,i)|^2}{|x|^2}\right)=-\beta^2(i).$$

In other words, (4.7) and (4.8) hold with $\mu_i=\alpha(i)$ and $\sigma_i=-\beta^2(i)$. Compute

$$\sum_{i=1}^{2}\pi_i(\mu_i+0.5\sigma_i)=\sum_{i=1}^{2}\pi_i(\alpha(i)-0.5\beta^2(i))=-\frac{\gamma_{21}}{\gamma_{12}+\gamma_{21}}.$$

Hence (3.2) follows from Corollary 4.5.

5 Stability of the BEM Method

In this section we will show that the BEM method can preserve the almost sure exponential stability described in Theorem 4.4. First of all, let us explain that the BEM method (2.4) is well-defined under condition (4.1) and this follows from the following lemma.

Lemma 5.1 *Let (4.1) hold and $\Delta<(\max_{i\in\mathbb{S}}|\mu_i|)^{-1}$. Then for any $i\in\mathbb{S}$ and $b\in\mathbb{R}^n$, there is a unique root $x\in\mathbb{R}^n$ of the equation*

$$x=b+f(x,i)\Delta. \tag{5.1}$$

Proof. Since Q is a symmetric positive-definite matrix, we can define its square root matrix $Q^{\frac{1}{2}}$ in the sense

$$(Q^{\frac{1}{2}})^TQ^{\frac{1}{2}}=Q.$$

Define $F:\mathbb{R}^n\times\mathbb{S}\to\mathbb{R}^n$ by

$$F(y,i)=Q^{\frac{1}{2}}f((Q^{\frac{1}{2}})^{-1}y,i),\quad (y,i)\in\mathbb{R}^n\times\mathbb{S}.$$

Then, for any $y_1,\ y_2\in\mathbb{R}^n$, setting $x_1=(Q^{\frac{1}{2}})^{-1}y_1$ and $x_2=(Q^{\frac{1}{2}})^{-1}y_2$, we compute, by condition (4.1), that

$$\begin{aligned}\langle y_1-y_2,F(y_1,i)-F(y_2,i)\rangle &= \langle Q^{\frac{1}{2}}(x_1-x_2),Q^{\frac{1}{2}}(f(x_1,i)-f(x_2,i))\rangle\\ &=(x_1-x_2)^TQ(f(x_1,i)-f(x_2,i))\le\mu_i(x_1-x_2)^TQ(x_1-x_2)\\ &=\mu_i|Q^{\frac{1}{2}}(x_1-x_2)|^2=\mu_i|y_1-y_2|^2.\end{aligned}$$

In other words, for each $i\in\mathbb{S}$, $F(\cdot,i)$ obeys the one-sided Lipschitz condition ([6, 9]). It is therefore known (see e.g. [16, 20]) that if $\Delta<(\max_{i\in\mathbb{S}}|\mu_i|)^{-1}$, then for any $b\in\mathbb{R}^n$, there is a unique root $y\in\mathbb{R}^n$ of the equation

$$y=Q^{\frac{1}{2}}b+F(y,i)\Delta. \tag{5.2}$$

This is equivalent to

$$y=Q^{\frac{1}{2}}b+Q^{\frac{1}{2}}f((Q^{\frac{1}{2}})^{-1}y,i)\Delta$$

or

$$(Q^{\frac{1}{2}})^{-1}y = b + f((Q^{\frac{1}{2}})^{-1}y, i)\Delta.$$

By setting $x = (Q^{\frac{1}{2}})^{-1}y$, this becomes

$$x = b + f(x, i)\Delta.$$

In other words, we have shown that there is a root x of equation (5.1) which is given by $x = (Q^{\frac{1}{2}})^{-1}y$.

To show the uniqueness, we let $\bar{x}$ be another root of equation (5.1), namely $\bar{x} = b + f(\bar{x}, i)\Delta$. Then

$$Q^{\frac{1}{2}}\bar{x} = Q^{\frac{1}{2}}b + Q^{\frac{1}{2}}f(\bar{x}, i)\Delta = Q^{\frac{1}{2}}b + F(Q^{\frac{1}{2}}\bar{x}, i)\Delta.$$

Hence $Q^{\frac{1}{2}}\bar{x} = y$ because y is the unique root of equation (5.2). We therefore must have $\bar{x} = x$. This completes the proof.

Let us now show that the BEM method can preserve the almost sure exponential stability described in Theorem 4.4.

Theorem 5.2 *Let Assumptions 4.1 and 4.2 hold, as well as condition (4.5). Then for any* $\varepsilon \in (0, \lambda)$, *where* $\lambda = |\sum_{i\in\mathbb{S}} \pi_i(\mu_i + 0.5\sigma_i)|$, *there is* $\Delta^\star \in (0, 1)$ *with* $2\Delta^\star(\max_{i\in\mathbb{S}} |\mu_i|) < 1$ *such that for any* $\Delta < \Delta^\star$, *the BEM method (2.4) has the property that*

$$\limsup_{k\to\infty} \frac{1}{k\Delta} \log |X_k| \le \sum_{i\in\mathbb{S}} \pi_i(\mu_i + 0.5\sigma_i) + \varepsilon < 0 \quad a.s. \tag{5.3}$$

The proof of this theorem is based on the following lemma, which is a simple version of the theorem.

Lemma 5.3 *Let Assumption 4.1 hold, as well as conditions (4.7), (4.8) and (4.5). Then the conclusion of Theorem 5.2 holds.*

We defer the proof of this lemma to the Appendix in order to highlight the following proof of our main result.

Proof of Theorem 5.2. We will use the notation defined in the proof of Lemma 5.1. We observe from the proof there that the function F defined there obeys condition (4.7). Define, moreover, $G : \mathbb{R}^n \times \mathbb{S} \to \mathbb{R}^n$ by

$$G(y, i) = Q^{\frac{1}{2}}g((Q^{\frac{1}{2}})^{-1}y, i), \quad (y, i) \in \mathbb{R}^n \times \mathbb{S}.$$

For any $y \in \mathbb{R}^n$ and $y \neq 0$, setting $x = (Q^{\frac{1}{2}})^{-1}y$, we compute

$$\frac{|G(y, i)|^2}{|y|^2} = \frac{|Q^{\frac{1}{2}}g(x, i)|^2}{|Q^{\frac{1}{2}}x|^2} = \frac{g^T(x, i)(Q^{\frac{1}{2}})^T Q^{\frac{1}{2}}g(x, i)}{x^T(Q^{\frac{1}{2}})^T Q^{\frac{1}{2}}x} = \frac{g^T(x, i)Qg(x, i)}{x^T Qx}$$

and

$$\frac{|y^TG(y,i)|^2}{|y|^4} = \frac{|x^T(Q^{\frac{1}{2}})^TQ^{\frac{1}{2}}g(x,i)|^2}{|x^T(Q^{\frac{1}{2}})^TQ^{\frac{1}{2}}x|^2} = \frac{|x^TQg(x,i)|^2}{|x^TQx|^2}.$$

Hence, by Assumption 4.2,

$$\begin{aligned}
&\sup_{y\in\mathbb{R}^n, y\neq 0}\left(\frac{|G(y,i)|^2}{|y|^2} - \frac{2|y^TG(y,i)|^2}{|y|^4}\right)\\
=&\sup_{x\in\mathbb{R}^n, x\neq 0}\left(\frac{g^T(x,i)Qg(x,i)}{x^TQx} - \frac{2|x^TQg(x,i)|^2}{(x^TQx)^2}\right)\\
=&\sigma_i < \infty.
\end{aligned}$$

In other words, the function G obeys condition (4.8). It is also easy to see that the functions F and G satisfy Assumption 4.1.

Now, for the BEM approximations X_k defined by (2.4), set $Y_k = Q^{\frac{1}{2}}X_k$. It follows from (2.4) that

$$Q^{\frac{1}{2}}X_{k+1} = Q^{\frac{1}{2}}X_k + Q^{\frac{1}{2}}f(X_{k+1}, r_k^\Delta)\Delta + Q^{\frac{1}{2}}g(X_k, r_k^\Delta)\Delta B_k,$$

whence

$$\begin{aligned}
Y_{k+1} &= Y_k + Q^{\frac{1}{2}}f((Q^{\frac{1}{2}})^{-1}Y_{k+1}, r_k^\Delta)\Delta + Q^{\frac{1}{2}}g((Q^{\frac{1}{2}})^{-1}Y_k, r_k^\Delta)\Delta B_k\\
&= Y_k + F(Y_{k+1}, r_k^\Delta)\Delta + G(Y_k, r_k^\Delta)\Delta B_k. \qquad (5.4)
\end{aligned}$$

In other words, Y_k ($k \geq 0$) are the approximations when the BEM method is applied to the following hybrid SDE

$$dy(t) = F(y(t), t)dt + G(y(t), t)dB(t) \qquad (5.5)$$

with initial value $y(0) = Q^{\frac{1}{2}}x_0$. By Lemma 5.3, for any $\varepsilon \in (0, \lambda)$, there is $\Delta^\star \in (0, 1)$ with $2\Delta^\star(\max_{i\in\mathbb{S}} |\mu_i|) < 1$ such that for any $\Delta < \Delta^\star$, the approximations Y_k ($k \geq 0$) defined by (5.4) obey

$$\limsup_{k\to\infty} \frac{1}{k\Delta}\log|Y_k| \leq \sum_{i\in\mathbb{S}} \pi_i(\mu_i + 0.5\sigma_i) + \varepsilon < 0 \quad a.s.$$

This implies that the BEM method (2.4) obeys

$$\limsup_{k\to\infty} \frac{1}{k\Delta}\log|X_k| \leq \sum_{i\in\mathbb{S}} \pi_i(\mu_i + 0.5\sigma_i) + \varepsilon < 0 \quad a.s.$$

as required. The proof is therefore complete.

6 Generalisation

So far, in order to streamline the presentation, we have only considered scalar noise. In this section we state, without proof, how the results generalize to the multidimensional noise case, as follows:

$$dx(t) = f(x(t), r(t))dt + \sum_{j=1}^{d} g_j(x(t), r(t))dB_j(t) \tag{6.1}$$

on $t \geq 0$, given $x(0) = x_0 \neq 0$ in $\mathbb{R}^n$ and $r(0) = i_0 \in \mathbb{S}$. Here $(B_1(t), \dots, B_d(t))$ is a d-dimensional Brownian motion and is independent of the Markov chain $r(t)$. As before, we assume, as a standing hypothesis, that $f, g_1, \dots, g_d : \mathbb{R}^n \times \mathbb{S} \to \mathbb{R}^n$ are smooth enough for the hybrid SDE (6.1) to have a unique global solution $x(t)$ on $[0, \infty)$. For the purpose of stability, we impose the following assumptions.

Assumption 6.1 *Assume that for each $k = 1, 2, \cdots$, there is an $h_k > 0$ such that*

$$|f(x, i)| \leq h_k|x|$$

for all $i \in \mathbb{S}$ and those $x \in \mathbb{R}^n$ with $|x| \leq k$ and, moreover, there is an $h > 0$ such that

$$|g_j(x, i)| \leq h|x| \quad \forall (x, i) \in \mathbb{R}^n \times \mathbb{S},\ 1 \leq j \leq d.$$

Assumption 6.2 *Assume that there is a symmetric positive-definite matrix $Q \in \mathbb{R}^{n\times n}$ and constants μ_i $(i \in \mathbb{S})$ such that*

$$(x - y)^T Q(f(x, i) - f(y, i)) \leq \mu_i(x - y)^T Q(x - y), \quad \forall x, y \in \mathbb{R}^n \tag{6.2}$$

and

$$\sigma_i := \sup_{x\in\mathbb{R}^n, x\neq 0} \left\{ \sum_{j=1}^{d} \left(\frac{g_j^T(x, i)Qg_j(x, i)}{x^T Qx} - \frac{2|x^T Qg_j(x, i)|^2}{(x^T Qx)^2} \right) \right\} < \infty. \tag{6.3}$$

The following generalization of Theorem 4.4 gives a criterion for the almost sure moment exponential stability of the SDE.

Theorem 6.3 *Let Assumptions 6.1 and 6.2 hold and assume that*

$$\sum_{i\in\mathbb{S}} \pi_i(\mu_i + 0.5\sigma_i) < 0. \tag{6.4}$$

Then the solution of equation (6.1) obeys

$$\limsup \frac{1}{t} \log(|x(t)|) \leq \sum_{i\in\mathbb{S}} \pi_i(\mu_i + 0.5\sigma_i) \quad a.s. \tag{6.5}$$

for all $x_0 \in \mathbb{R}^n$.

This theorem can be proved in a similar way to Theorem 4.4.

The BEM method applied to (6.1) produces approximations $X_k \approx x(k\Delta)$ with $X_0 = x(0)$ and

$$X_{k+1} = X_k + f(X_{k+1}, r_k^\Delta) + \sum_{j=1}^{d} g_j(X_k, r_k^\Delta)\Delta B_{jk}, \tag{6.6}$$

where $\Delta B_{jk} := B_j((k+1)\Delta) - B_j(k\Delta)$.

Theorem 6.4 *Under the same conditions of Theorem 6.3, for any* $\varepsilon \in (0, \lambda)$*, where* $\lambda = |\sum_{i\in\mathbb{S}} \pi_i(\mu_i + 0.5\sigma_i)|$*, there is* $\Delta^\star \in (0,1)$ *with* $2\Delta^\star(\max_{i\in\mathbb{S}} |\mu_i|) < 1$ *such that for any* $\Delta < \Delta^\star$*, the BEM method (6.6) has the property that*

$$\limsup_{k\to\infty} \frac{1}{k\Delta} \log |X_k| \le \sum_{i\in\mathbb{S}} \pi_i(\mu_i + 0.5\sigma_i) + \varepsilon < 0 \quad a.s. \tag{6.7}$$

This theorem can be proved in the same way as the scalar noise version, Theorem 5.2, was proved.

7 Example and Simulations

Let us now return to the SDE (3.1). In Section 3 we have shown that the EM method cannot reproduce the almost sure exponential stability of the SDE. However, our theory established in the previous sections shows that the BEM method can reproduce the stability. To illustrate our theory, as well as to compare to the simulations in Section 3, we set the system parameters $\gamma_{12} = 1$ and $\gamma_{21} = 4$ as before and use the same step size $\Delta = 0.001$, and the same SDE parameters $\alpha(1) = 1, \beta(1) = 2, \alpha(2) = 0.5, \beta(2) = 1$. The two simulations shown in Figures 7.1 and 7.2 are based on the BEM method with initial values set as before to $(x(0), r(0)) = (20, 1)$ and $(50, 1)$, respectively. Both figures show clearly that the BEM method reproduces the almost sure exponential stability of the underlying SDE (3.1).

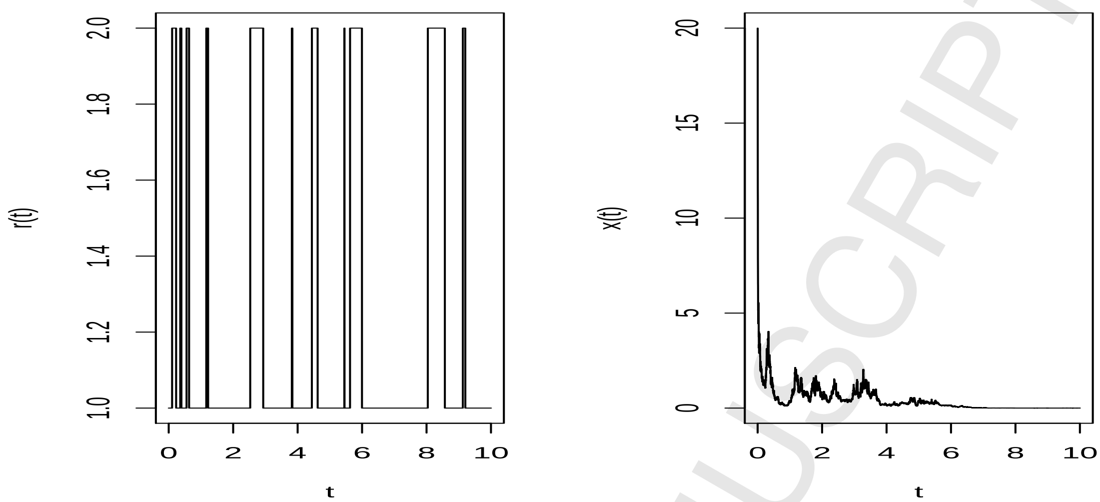

Figure 7.1: Computer simulation of the path $x(t)$ and corresponding state $r(t)$, using the BEM method with step size $\Delta = 0.001$, and initial values $(x(0), r(0)) = (20, 1)$. The generator parameters are $\gamma_{12} = 1$ and $\gamma_{21} = 4$, and the SDE parameters are $\alpha(1) = 1, \beta(1) = 2, \alpha(2) = 0.5, \beta(2) = 1$.

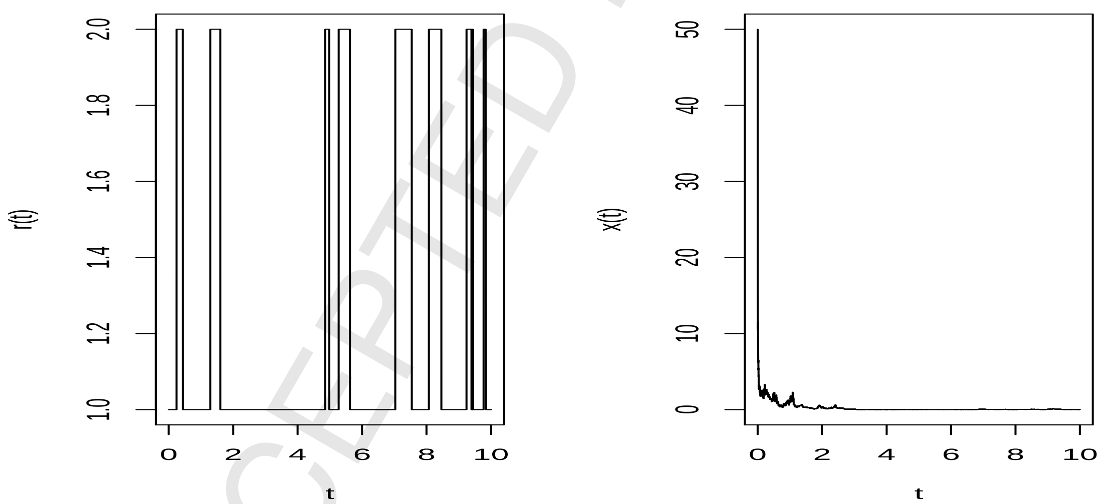

Figure 7.2: Computer simulation of the path $x(t)$ and state $r(t)$, using the BEM method with step size $\Delta = 0.001$, and initial values $(x(0), r(0)) = (50, 1)$. The generator parameters are $\gamma_{12} = 1$ and $\gamma_{21} = 4$, and the SDE parameters are $\alpha(1) = 1, \beta(1) = 2, \alpha(2) = 0.5, \beta(2) = 1$.

A Appendix: Proof of Theorem 4.4

In this appendix we prove Theorem 4.4. Clearly, assertion (4.6) holds when $x_0 = 0$ since in this case the solution $x(t) \equiv 0$. Fix any initial value $x_0 \neq 0$. By Lemma 4.3, with probability one this solution $x(t)$ will never reach zero. We can then apply the Itô formula to $\log(x^T(t)Qx(t))$ to obtain that

$$\begin{aligned} d[\log(x^T(t)Qx(t))] = &\Big[\frac{2x^T(t)Qf(x(t),r(t))}{x^T(t)Qx(t)} + \frac{g^T(x(t),r(t))Qg(x(t),r(t))}{x^T(t)Qx(t)} \\ &- \frac{2|x^T(t)Qg(x(t),r(t))|^2}{(x^T(t)Qx(t))^2}\Big]\,dt \\ &+ \frac{2x^T(t)Qg(x(t),r(t))}{x^T(t)Qx(t)}\,dB(t). \end{aligned}$$

By Assumption 4.2 (namely (4.2) and (4.4)), we obtain that

$$\log(x^T(t)Qx(t)) \le \log(x_0^TQx_0)) + \int_0^t \big[2\mu_{r(s)} + \sigma_{r(s)}\big]\,ds + M(t), \tag{A.1}$$

where

$$M(t) = \int_0^t \frac{2x^T(s)Qg(x(s),r(s))}{x^T(s)Qx(s)}\,dB(s)$$

is a continuous martingale vanishing at $t = 0$. The quadratic variation of the martingale is given by

$$\langle M(t)\rangle = \int_0^t \frac{4|x^T(s)Qg(x(s),r(s))|^2}{(x^T(s)Qx(s))^2}ds.$$

By Assumption 4.1, we see that

$$\langle M(t)\rangle \le \frac{4h^2\|Q\|^2 t}{(\lambda_{\min}(Q))^2}.$$

Hence, by the strong law of large numbers for martingales (see e.g. [15, Theorem 3.4 on page 12]), we have

$$\lim_{t\to\infty} \frac{M(t)}{t} = 0 \quad a.s.$$

We can therefore divide both sides of (A.1) by t and then let $t \to \infty$ to obtain

$$\limsup_{t\to\infty} \frac{1}{t}\,\log(x^T(t)Qx(t)) \le \limsup_{t\to\infty} \frac{1}{t}\int_0^t \big[2\mu_{r(s)} + \sigma_{r(s)}\big]\,ds. \tag{A.2}$$

But, by the ergodic property of the Markov chain (see e.g. [1]) we have

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \big[2\mu_{r(s)} + \sigma_{r(s)}\big]\,ds = \sum_{i\in\mathbb{S}} \pi_i(2\mu_i + \sigma_i) \quad a.s.$$

Hence

$$\limsup_{t\to\infty} \frac{1}{t}\,\log(x^T(t)Qx(t)) \le \sum_{i\in\mathbb{S}} \pi_i(2\mu_i + \sigma_i) \quad a.s.$$

This implies immediately that

$$\limsup_{t\to\infty} \frac{1}{t}\log(|x(t)|) \le \frac{1}{2}\sum_{i\in\mathbb{S}} \pi_i(2\mu_i + \sigma_i) \quad a.s.,$$

which is the required assertion (4.6).

B Appendix: Proof of Lemma 5.3

The proof borrows heavily from [17] but for the completeness of the paper we present the details here. The proof is very technical so we divide it into three steps.

Step 1. From (2.4), we have

$$|X_{k+1}|^2 = \langle X_{k+1}, f(X_{k+1}, r_k^\Delta)\rangle\Delta + \langle X_{k+1}, X_k + g(X_k, r_k^\Delta)\Delta B_k\rangle.$$

By Assumption 4.1 and condition (4.7), we have

$$\langle X_{k+1}, f(X_{k+1}, r_k^\Delta)\rangle \le \mu_{r_k^\Delta}|X_{k+1}|^2.$$

But,

$$\langle X_{k+1}, X_k + g(X_k, r_k^\Delta)\Delta B_k\rangle \le \tfrac{1}{2}|X_{k+1}|^2 + \tfrac{1}{2}|X_k + g(X_k, r_k^\Delta)\Delta B_k|^2.$$

We hence obtain

$$\begin{aligned} |X_{k+1}|^2 &\le \frac{1}{1-2\mu_{r_k^\Delta}\Delta}|X_k + g(X_k, r_k^\Delta)\Delta B_k|^2 \\ &= \frac{1}{1-2\mu_{r_k^\Delta}\Delta}\left(|X_k|^2 + 2\langle X_k, g(X_k, r_k^\Delta)\rangle\Delta B_k + |g(X_k, r_k^\Delta)|^2|\Delta B_k|^2\right) \\ &= \frac{|X_k|^2}{1-2\mu_{r_k^\Delta}\Delta}(1+\zeta_k(r_k^\Delta)), \end{aligned}$$

where

$$\zeta_k(r_k^\Delta) = \frac{1}{|X_k|^2}\left(2\langle X_k, g(X_k, r_k^\Delta)\rangle\Delta B_k + |g(X_k, r_k^\Delta)|^2|\Delta B_k|^2\right)$$

if $X_k \neq 0$, otherwise it is set to -1. Clearly, $\zeta_k \ge -1$. Let $\mathcal{G}_t = \sigma(\{r(u)\}_{u\ge 0}, \{B(s)\}_{0\le s\le t})$, namely the σ-algebra generated by $\{r(u)\}_{u\ge0}$ and $\{B(s)\}_{0\le s\le t}$. For any $p \in (0,1)$, recall the fundamental inequality

$$(1+u)^{p/2} \le 1 + \frac{p}{2}u + \frac{p(p-2)}{8}u^2 + \frac{p(p-2)(p-4)}{48}u^3, \qquad u \ge -1. \tag{B.1}$$

We then have

$$\begin{aligned} \mathbb{E}(|X_{k+1}|^p|\mathcal{G}_{k\Delta}) \le{}& \frac{|X_k|^p}{(1-2\mu_{r_k^\Delta}\Delta)^{p/2}}I_{\{X_k\neq0\}} \\ &\times \mathbb{E}\left(1 + \frac{p}{2}\zeta_k(r_k^\Delta) + \frac{p(p-2)}{8}\zeta_k^2(r_k^\Delta) + \frac{p(p-4)(p-2)}{48}\zeta_k^3(r_k^\Delta)\Big|\mathcal{G}_{k\Delta}\right). \end{aligned} \tag{B.2}$$

Now,

$$\begin{aligned} &I_{\{X_k\neq0\}}\mathbb{E}(\zeta_k(r_k^\Delta)|\mathcal{G}_{k\Delta}) \\ =&I_{\{X_k\neq0\}}\mathbb{E}\Big(\frac{1}{|X_k|^2}\big(2\langle X_k, g(X_k, r_k^\Delta)\rangle\Delta B_k + |g(X_k, r_k^\Delta)|^2|\Delta B_k|^2\big)\Big|\mathcal{G}_{k\Delta}\Big) \\ =&I_{\{X_k\neq0\}}\frac{1}{|X_k|^2}\Big(2\langle X_k, g(X_k, r_k^\Delta)\rangle\mathbb{E}(\Delta B_k|\mathcal{G}_{k\Delta}) + |g(X_k, r_k^\Delta)|^2\mathbb{E}(|\Delta B_k|^2|\mathcal{G}_{k\Delta})\Big). \end{aligned}$$

Since ΔB_k is independent of $\mathcal{G}_{k\Delta}$, we have

$$\mathbb{E}(\Delta B_k|\mathcal{G}_{k\Delta}) = \mathbb{E}(\Delta B_k) = 0 \quad \text{and} \quad \mathbb{E}(|\Delta B_k|^2|\mathcal{G}_{k\Delta}) = \mathbb{E}(|\Delta B_k|^2) = \Delta.$$

Hence

$$I_{\{X_k \neq 0\}}\mathbb{E}(\zeta_k(r_k^\Delta)|\mathcal{G}_{k\Delta}) = I_{\{X_k \neq 0\}}\frac{|g(X_k, r_k^\Delta)|^2}{|X_k|^2}\Delta. \tag{B.3}$$

Making use of the properties

$$\mathbb{E}((\Delta B_k)^{2i}) = (2i-1)!!\Delta^i \quad \text{and} \quad \mathbb{E}((\Delta B_k)^{2i-1}) = 0$$

for $i = 1, 2, \cdots$, where $(2i-1)!! = (2i-1)(2i-3)\cdots 3 \cdot 1$, we compute similarly that

$$\begin{aligned} I_{\{X_k \neq 0\}}\mathbb{E}(\zeta_k^2(r_k^\Delta)|\mathcal{G}_{k\Delta}) =& I_{\{X_k \neq 0\}}\Big(\frac{4\langle X_k, g(X_k, r_k^\Delta)\rangle^2}{|X_k|^4}\Delta + \frac{|g(X_k, r_k^\Delta)|^4}{|X_k|^4}3\Delta^2\Big) \\ \geq& I_{\{X_k \neq 0\}}\frac{4\langle X_k, g(X_k, r_k^\Delta)\rangle^2}{|X_k|^4}\Delta \end{aligned} \tag{B.4}$$

and

$$\begin{aligned} & I_{\{X_k \neq 0\}}\mathbb{E}(\zeta_k^3(r_k^\Delta)|\mathcal{G}_{k\Delta}) \\ =& I_{\{X_k \neq 0\}}\frac{1}{|X_k|^6}\Big(36\langle X_k, g(X_k, r_k^\Delta)\rangle^2|g(X_k, r_k^\Delta)|^2\Delta^2 + 15|g(X_k, r_k^\Delta)|^6\Delta^3\Big) \\ \leq& I_{\{X_k \neq 0\}}c_1\Delta^2, \end{aligned} \tag{B.5}$$

where $c_1 = c_1(h) = 36h^4 + 15h^6$ and the h is specified in Assumption 4.1. Substituting (B.3)-(B.5) into (B.2) and then using (4.8) and Assumption 4.1, we derive that

$$\begin{aligned} \mathbb{E}(|X_{k+1}|^p\big|\mathcal{G}_{k\Delta}) \leq& \frac{|X_k|^p}{(1-2\mu_{r_k^\Delta}\Delta)^{p/2}}I_{\{X_k \neq 0\}} \\ & \times\Big(1 + \frac{p}{2}\frac{|g(X_k, r_k^\Delta)|^2}{|X_k|^2}\Delta + \frac{p(p-2)}{8}\frac{4\langle X_k, g(X_k, r_k^\Delta)\rangle^2}{|X_k|^4}\Delta + c_2\Delta^2\Big) \\ \leq& \frac{|X_k|^p}{(1-2\mu_{r_k^\Delta}\Delta)^{p/2}}\Big(1 + \tfrac{1}{2}p\sigma_{r_k^\Delta}\Delta + \tfrac{1}{2}p^2h^2\Delta + c_2\Delta^2\Big), \end{aligned} \tag{B.6}$$

where $c_2 = c_2(h, p) = c_1(h)p(p-2)(p-4)/48$.

Step 2. Now, for any $\varepsilon \in (0, \lambda)$, we may choose p sufficiently small for $ph^2 \leq \frac{1}{4}\varepsilon$. Then we have

$$(1-2\mu_{r_k^\Delta}\Delta)^{p/2} \geq 1 - p\mu_{r_k^\Delta}\Delta - c_3\Delta^2 > 0, \tag{B.7}$$

for sufficiently small Δ, where $c_3 = c_3(p)$ is a positive constant. By further reducing Δ, if necessary, we may ensure that

$$c_2\Delta < \tfrac{1}{8}p\varepsilon, \qquad c_3\Delta < \tfrac{1}{4}p\varepsilon, \qquad |p(\mu_{r_k^\Delta} + \tfrac{1}{4}\varepsilon)\Delta| \leq \tfrac{1}{2}. \tag{B.8}$$

Using (B.7) and (B.8) in (B.6) gives

$$\mathbb{E}(|X_{k+1}|^p\big|\mathcal{G}_{k\Delta}) \leq \frac{1 + \frac{1}{2}p(\sigma_{r_k^\Delta} + \frac{1}{2}\varepsilon)\Delta}{1 - p(\mu_{r_k^\Delta} + \frac{1}{4}\varepsilon)\Delta}|X_k|^p. \tag{B.9}$$

Note that for any $u \in [-\frac{1}{2}, \frac{1}{2}]$

$$\frac{1}{1-u} = 1 + u + u^2 \sum_{i=0}^{\infty} u^i \le 1 + u + u^2 \sum_{i=0}^{\infty} (\tfrac{1}{2})^i = 1 + u + 2u^2.$$

By further reducing Δ, if necessary, so that

$$2p(\mu_{r_k^\Delta} + \tfrac{1}{4}\varepsilon)^2\Delta + \tfrac{1}{2}(\sigma_{r_k^\Delta} + \tfrac{1}{2}\varepsilon)\left(p(\mu_{r_k^\Delta} + \tfrac{1}{4}\varepsilon)\Delta + 2[p(\mu_{r_k^\Delta} + \tfrac{1}{4}\varepsilon)\Delta]^2\right) \le \tfrac{1}{4}\varepsilon,$$

and using (B.9), we compute that

$$\begin{aligned}\mathbb{E}(|X_{k+1}|^p \big| \mathcal{G}_{k\Delta}) &\le \left(1 + \tfrac{1}{2}p(\sigma_{r_k^\Delta} + \tfrac{1}{2}\varepsilon)\Delta\right)\left(1 + p(\mu_{r_k^\Delta} + \tfrac{1}{4}\varepsilon)\Delta + 2[p(\mu_{r_k^\Delta} + \tfrac{1}{4}\varepsilon)\Delta]^2\right)|X_k|^p \\ &\le [1 + p(\mu_{r_k^\Delta} + \tfrac{1}{2}\sigma_{r_k^\Delta} + \tfrac{3}{4}\varepsilon)\Delta]|X_k|^p. \qquad \text{(B.10)}\end{aligned}$$

Since this holds for all $k \ge 0$, we further compute

$$\begin{aligned}\mathbb{E}(|X_{k+1}|^p \big| \mathcal{G}_{(k-1)\Delta}) &\le \mathbb{E}(|X_k|^p \big| \mathcal{G}_{(k-1)\Delta})[1 + p(\mu_{r_k^\Delta} + \tfrac{1}{2}\sigma_{r_k^\Delta} + \tfrac{3}{4}\varepsilon)\Delta] \\ &\le |X_{k-1}|^p \prod_{z=k-1}^{k} [1 + p(\mu_{r_z^\Delta} + \tfrac{1}{2}\sigma_{r_z^\Delta} + \tfrac{3}{4}\varepsilon)\Delta].\end{aligned}$$

Repeating this procedure we obtain

$$\mathbb{E}(|X_{k+1}|^p \big| \mathcal{G}_0) \le |x_0|^p \prod_{z=0}^{k} [1 + p(\mu_{r_z^\Delta} + \tfrac{1}{2}\sigma_{r_z^\Delta} + \tfrac{3}{4}\varepsilon)\Delta].$$

Taking expectations on both sides yields

$$\mathbb{E}(|X_{k+1}|^p) \le |x_0|^p \, \mathbb{E}\exp\Big(\sum_{z=0}^{k} \log[1 + p(\mu_{r_z^\Delta} + \tfrac{1}{2}\sigma_{r_z^\Delta} + \tfrac{3}{4}\varepsilon)\Delta]\Big). \qquad \text{(B.11)}$$

If necessary, we can further reduce Δ to ensure that

$$p(\mu_i + \tfrac{1}{2}\sigma_i + \tfrac{3}{4}\varepsilon)\Delta > -1, \quad i \in \mathbb{S}.$$

Then, by the ergodic property of the Markov chain and inequality

$$\log(1 + x) \le x, \quad x > -1,$$

we compute

$$\begin{aligned}\lim_{k\to\infty} \frac{1}{1+k} \sum_{z=0}^{k} \log[1 + p(\mu_{r_z^\Delta} + \tfrac{1}{2}\sigma_{r_z^\Delta} + \tfrac{3}{4}\varepsilon)\Delta] &= \sum_{i\in\mathbb{S}} \pi_i \log[1 + p(\mu_i + \tfrac{1}{2}\sigma_i + \tfrac{3}{4}\varepsilon)\Delta] \\ &\le p\Delta \sum_{i\in\mathbb{S}} \pi_i(\mu_i + \tfrac{1}{2}\sigma_i + \tfrac{3}{4}\varepsilon) \\ &= p\Delta(-\lambda + \tfrac{3}{4}\varepsilon) \quad a.s.\end{aligned}$$

This yields

$$\lim_{k\to\infty}\Big(p\Delta(\lambda-\varepsilon)(k+1)+\sum_{z=0}^{k}\log[1+p(\mu_{r_z^\Delta}+\tfrac{1}{2}\sigma_{r_z^\Delta}+\tfrac{3}{4}\varepsilon)\Delta]\Big)=-\infty \quad a.s. \tag{B.12}$$

However, it follows from (B.11)

$$\begin{aligned} & e^{p\Delta(\lambda-\varepsilon)(k+1)}\mathbb{E}(|X_{k+1}|^p) \\ \le\ & |x_0|^p\mathbb{E}\exp\Big(p\Delta(\lambda-\varepsilon)(k+1)+\sum_{z=0}^{k}\log[1+p(\mu_{r_z^\Delta}+\tfrac{1}{2}\sigma_{r_z^\Delta}+\tfrac{3}{4}\varepsilon)\Delta]\Big). \end{aligned}$$

By the Fatou lemma (see e.g. [12]) and property (B.12), we hence derive that

$$\begin{aligned} & \limsup_{k\to\infty}\Big[e^{p\Delta(\lambda-\varepsilon)(k+1)}\mathbb{E}(|X_{k+1}|^p)\Big] \\ \le\ & |x_0|^p\mathbb{E}\Big[\limsup_{k\to\infty}\exp\Big(p\Delta(\lambda-\varepsilon)(k+1)+\sum_{z=0}^{k}\log[1+p(\mu_{r_z^\Delta}+\tfrac{1}{2}\sigma_{r_z^\Delta}+\tfrac{3}{4}\varepsilon)\Delta]\Big)\Big] \\ =\ & 0. \end{aligned} \tag{B.13}$$

Step 3. It follows from (B.13) that there is an integer k_0 such that

$$\mathbb{E}(|X_k|^p)\le e^{-pk\Delta(\lambda-\varepsilon)}, \quad \forall k\ge k_0.$$

This implies, by the Chebycheff inequality, that

$$\mathbb{P}\{|X_k|^p>k^2e^{-pk\Delta(\lambda-\varepsilon)}\}\le\frac{1}{k^2}, \quad \forall k\ge k_0.$$

Applying the Borel–Cantelli lemma (see e.g. [15, p17]), we obtain that for almost all $\omega\in\Omega$,

$$|X_k|^p\le k^2e^{-pk\Delta(\lambda-\varepsilon)} \tag{B.14}$$

holds for all but finitely many $k\ge k_0$. Hence, there exists a $k_1(\omega)\ge k_0$, for almost all $\omega\in\Omega$, for which (B.14) holds whenever $k\ge k_1$. Consequently, for almost all $\omega\in\Omega$,

$$\frac{1}{k\Delta}\log(|X_k|)\le\frac{2\log(k)}{pk\Delta}-(\lambda-\varepsilon)$$

whenever $k\ge k_1$. Therefore

$$\limsup_{k\to\infty}\frac{1}{k\Delta}\log(|X_k|)\le-\lambda+\varepsilon \quad a.s.,$$

which is the desired assertion (5.3). The proof is complete.

Acknowledgements

The authors would like to thank the referees and the Editor for their very helpful suggestions and comments. The authors would also like to thank the Royal Society of Edinburgh, the Scottish Government, the British Council Shanghai and the National Natural Science Foundation of China (grants 60874031 and 60740430664) for their financial support.

References

[1] Anderson, W.J., *Continuous-Time Markov Chains*, Springer-Verlag, 1991.

[2] Appleby, J.A.D., Berkolaiko, G. and Rodkina, A., On the asymptotic behaviour of the moments of solutions of stochastic difference equations, *Mathematical Sciences Preprint* **MS-05-21** (2005), Dublin City University.

[3] Baker, C.T.H.B. and Buckwar, E., Exponential stability in p-th mean of solutions, and of convergent Euler-type solutions, of stochastic delay differential equations, *J. Comp. Appl. Math.* **184** (2005), 404–427.

[4] Buckwar, E., Horvath, R. and Winkler, R., Asymptotic mean-square stability of two-step methods for stochastic ordinary differential equations, *Preprint* **04-25** (2004), Institut für Mathematik, Humboldt-Universität Berlin.

[5] Higham. D.J., Mean-square and asymptotic stability of the stochastic theta method, *SIAM J. Numer. Anal.* **38** (2000), 753–769.

[6] Higham, D.J., Mao X. and Stuart, A.M., Strong convergence of numerical methods for nonlinear stochastic differential equations, *SIAM J. Numer. Anal.* **40** (2002), 1041–1063.

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

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

[9] Kloeden, P.E. and Platen, E., *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, 1992.

[10] Komori, Y. and Mitsui, T., Stable ROW-type weak scheme for stochastic differential equations, *Monte Carlo Methods Appl.* **1** (1995), 279-300.

[11] Lamba H. and Seaman, T. L., Mean-square stability of an adaptive SDE solver, to appear in *J. Comp. App. Math.*

[12] Loève, M., *Probability Theory*, D. Van Nostrand Company Inc., 1963.

[13] Mao, X., *Stability of Stochastic Differential Equations with Respect to Semimartingales*, Longman Scientific and Technical, 1991.

[14] Mao, X., *Exponential Stability of Stochastic Differential Equations*, Marcel Dekker, 1994.

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

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

[17] Pang, S., Deng, F. and Mao, X., Almost sure and moment exponential stability of Euler–Maruyama discretizations for hybrid stochastic differential equations *J. Comp. App. Math.* **213** (2008), 127–141.

[18] Saito, Y. and Mitsui, T., T-stability of numerical scheme for stochastic differential equations, *World Sci. Ser. Appl. Anal.* **2** (1993), 333-344.

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

[20] Stuart, A.M. and Humphries, A.R., *Dynamical Systems and Numerical Analysis*, Cambridge University Press, Cambridge, 1996.