View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Cronfa at Swansea University

Swansea University
Prifysgol Abertawe

Cronfa
Setting Research Free

Cronfa - Swansea University Open Access Repository

This is an author produced version of a paper published in:
SIAM Journal on Numerical Analysis

Cronfa URL for this paper:
http://cronfa.swan.ac.uk/Record/cronfa40078

Paper:

Li, X., Ma, Q., Yang, H. & Yuan, C. (2018). The Numerical Invariant Measure of Stochastic Differential Equations With Markovian Switching. *SIAM Journal on Numerical Analysis, 56*(3), 1435-1455.
http://dx.doi.org/10.1137/17M1143927

This item is brought to you by Swansea University. Any person downloading material is agreeing to abide by the terms of the repository licence. Copies of full text items may be used or reproduced in any format or medium, without prior permission for personal research or study, educational or non-commercial purposes only. The copyright for any work remains with the original author unless otherwise specified. The full-text must not be sold in any format or medium without the formal permission of the copyright holder.

Permission for multiple reproductions should be obtained from the original author.

Authors are personally responsible for adhering to copyright and publisher restrictions when uploading content to the repository.

http://www.swansea.ac.uk/library/researchsupport/ris-support/

THE NUMERICAL INVARIANT MEASURE OF STOCHASTIC DIFFERENTIAL EQUATIONS WITH MARKOVIAN SWITCHING

XIAOYUE LI*, QIANLIN MA†, HONGFU YANG‡, AND CHENGGUI YUAN§

Abstract. The existence and uniqueness of the numerical invariant measure of the backward Euler-Maruyama method for stochastic differential equations with Markovian switching is yielded, and it is revealed that the numerical invariant measure converges to the underlying invariant measure in the Wasserstein metric. The global Lipschitz condition on the drift coefficients required by Bao et al., 2016 and Yuan et al., 2005 is released. Under a polynomial growth condition imposed on drift coefficients we show that the convergence is polynomial. Several examples and numerical experiments are given to verify our theory.

Key words. The backward Euler-Maruyama method·Markovian switching·Numerical invariant measure·Wasserstein metric

AMS subject classifications. 60H10 · 34F05

1. Introduction. As one of the important classes of hybrid systems, stochastic differential equations (SDEs) with Markovian switching have been widely used in biology, control problems, neutral activity, mathematical finance and other sciences (see, e.g., the monographs [13, 32] and the references therein). So far, various dynamical properties including moment boundedness, stability, ergodicity, recurrence and transience on SDEs with Markovian switching have been investigated extensively, refer to [3, 4, 5, 13, 24, 26, 27, 31, 32]. Yin and Zhu [32, pp.181-280], and Mao and Yuan [13, pp.164-190] investigated the stability of SDEs with Markovian switching and showed that the Markov chain facilitates the stochastic stabilization in which the stationary distribution of the Markov chain plays an important role. Pinsky and Scheutzow [24] revealed the fact that the overall system may not to be positive recurrence (resp. transience) even though each subsystem is. So, the dynamical behaviors of SDEs with Markovian switching are significantly different from those of SDEs.

However, solving the SDEs with Markovian switching is still a challenging task that requires using numerical methods or approximation techniques, see, e.g., the monographs [10, 13, 14, 32]. Some long-time behaviors of the SDEs with Markovian switching, for instance, the almost sure stability and the moment stability, have been preserved by the numerical solutions, see, e.g., [7, 13, 15, 23, 32, 35] and the references therein. For deterministic systems, the stability of equilibrium point is among of the interesting topics. However, many stochastic systems don't posses a deterministic equilibrium state. Recently, for stochastic systems with Markovian switching, the stability of the "stochastic equilibrium state"-the existence of the invariant measure has drawn increasing attention [3, 4, 26, 27, 31, 32]. Since the corresponding Kolmogorov-Fokker-Planck equations are always computationally intensive, it is important to be able to approximate the invariant measure numerically. Therefore, approximations of

*School of Mathematics and Statistics, Northeast Normal University, Changchun, Jilin, 130024, China. Research of this author was supported by National Natural Science Foundation of China (11171056, 11471071, 11671072), the Natural Science Foundation of Jilin Province (No.20170101044JC), the Education Department of Jilin Province (No.JJKH20170904KJ).

†School of Mathematics and Statistics, Northeast Normal University, Changchun, Jilin, 130024, China.

‡School of Mathematics and Statistics, Northeast Normal University, Changchun, Jilin, 130024, China.

§Department of Mathematics, Swansea University, Singleton Park, SA2 8PP, UK.

invariant measures for SDEs with Markovian switching have attracted much attention recently. Mao et al. [12], Yuan and Mao [33] and Bao et al. [4] made use of Euler-Maruyama (EM) method with a constant step size to approximate the underlying invariant measure while Yin and Zhu [32, p.159-179] did that using the EM scheme with the decreasing step size. In the mentioned papers, both the drift coefficients and the diffusion coefficients of the SDEs with regime switching are required to be global Lipschitz continuous. Although the classical Euler-Maruyama (EM) method is convenient for computations and implementations, the absolute moments of its approximation for SDEs with super-linear coefficients may diverge to infinity at a finite time (see, e.g. [8]). It is well know, see [16], that the EM numerical solutions fail to be ergodic, even when the underlying SDE is geometrically ergodic. Many implicit methods were used to study the numerical solutions to SDEs with nonlinear coefficients (see, e.g., [6, 22]). Higham et al. [6] proved that the implicit EM numerical solutions converge strongly to the exact solutions of SDEs with globally one-sided Lipschitz continuous drift term and globally Lipschitz continuous diffusion term, but the explicit EM method fails to do that. Mattingly et al. [16] introduced variants of the implicit EM method to preserve the ergodicity for SDEs with additional noises usually established through the use of Foster-Lyapunov conditions in [18, 19, 20] while Liu and Mao [11] took advantage of the implicit EM method to approximate the stability in distribution of non-globally Lipschitz continuous SDEs. For the background on the implicit methods, we refer the reader to the books [10, 21]. Shardlow and Stuart [28] established the perturbation theory of geometrically ergodic Markov chain with an application to numerical approximations.

Motivated by the papers above, this paper focuses on using the backward Euler-Maruyama (BEM) method to approximate the invariant measure of nonlinear SDEs with Markovian switching that the drift coefficients need not to satisfy the global Lipschitz condition. The BEM scheme, which is implicit in the drift term, has been implemented for SDEs with Markovian switching to investigate the strong convergence and the approximation of the almost sure stability as well as the moment stability (see, e.g., [15, 35, 34] and the references therein). The main aim of this paper is to study the existence and uniqueness of the numerical invariant measure of the BEM method and the convergence in the Wasserstein metric to the invariant measure of the corresponding exact solution as well as the convergence rate.

The rest of our paper is organized as follows. Section 2 gives some preliminary results on the existence and uniqueness of the invariant measure for the exact solution. Section 3 focuses on the existence and uniqueness of the numerical invariant measures in BEM scheme. Then we go further to reveal that the numerical invariant measure converges in the Wasserstein distance to the underlying one. Section 4 presents several examples and numerical experiments to illustrate our results.

2. Preliminary. Throughout this paper, let $|\cdot|$ denote the Euclidean norm in $\mathbb{R}^n := \mathbb{R}^{n\times 1}$ and the trace norm in $\mathbb{R}^{n\times m}$. If A is a vector or matrix, its transpose is denoted by A^T and its trace norm is denoted by $|A| = \sqrt{\mathrm{trace}(A^TA)}$. For vectors or matrixes A and B with compatible dimensions, AB denotes the usual matrix multiplication. We denote the indicator function of a set $\mathbb{D}$ by $I_{\mathbb{D}}$, and $\mathbf{0} \in \mathbb{R}^n$ is a zero vector. For any $\xi = (\xi_1, \xi_2, \cdots, \xi_n)^T \in \mathbb{R}^n$, $\xi \gg \mathbf{0}$ means each component $\xi_j > 0, j = 1, 2, \cdots, n$. Define $\hat{\xi} = \min_{1\le j\le n} \xi_j$ and $\check{\xi} = \max_{1\le j\le n} \xi_j$. For any $a, b \in \mathbb{R}$, $a \vee b = \max\{a, b\}$, and $a \wedge b = \min\{a, b\}$. For each $R > 0$, let $B_R(0) = \{x \in \mathbb{R}^n : |x| \le R\}$. Let $\mathscr{B}(\mathbb{R}^n)$ denote the family of all Borel sets in $\mathbb{R}^n$.

Let $(\Omega,\ \mathcal{F},\ \mathbb{P})$ be a complete probability space, and $\mathbb{E}$ denotes the expectation corresponding to $\mathbb{P}$. Let $B(t)$ be an m-dimensional Brownian motion defined on this probability space. Suppose that $\{r(t)\}_{t\geq 0}$ is a right-continuous Markov chain with finite state space $\mathbb{S}=\{1,2,\cdots,N\}$ and independent of the Brownian motion $B(\cdot)$, where N is a positive integer. Suppose $\{\mathcal{F}_t\}_{t\geq 0}$ is a filtration defined on this probability space satisfying the usual conditions (i.e., it is right continuous in t and $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets) such that $B(t)$ and $r(t)$ are $\mathcal{F}_t$ adapted. The generator of $\{r(t)\}_{t\geq 0}$ is denoted by $Q=(q_{lj})_{N\times N}$, so that for a sufficiently small $\epsilon>0$,

$$\mathbb{P}\{r(t+\epsilon)=j|r(t)=l\}=\begin{cases} q_{lj}\epsilon+o(\epsilon), & \text{if } \ l\neq j,\\ 1+q_{ll}\epsilon+o(\epsilon), & \text{if } \ l=j. \end{cases}$$

Here $q_{lj}\geq 0$ is the transition rate from l to j if $l\neq j$ while $q_{ll}=-\sum_{l\neq j}q_{lj}$. It is well known that almost every sample path of $r(t)$ is a right-continuous step function with a finite number of simple jumps in any finite subinterval of $\mathbb{R}_+ := [0,+\infty)$ (see [1, p.17-18]). As a standing hypothesis, we assume that the transition probability matrix Q are *irreducible and conservative*. So Markov chain $\{r(t)\}_{t\geq 0}$ has a unique stationary distribution $\mu:=(\mu_1,\mu_2,\cdots,\mu_N)\gg \mathbf{0}\in\mathbb{R}^{1\times N}$ which can be determined by solving the linear equation

$$\mu Q=0, \qquad \text{subject to } \ \sum_{j=1}^{N}\mu_j=1. \tag{2.1}$$

In this paper, we consider the two-component diffusion process $(Y(t),r(t))$ described by the SDE with Markovian switching

$$dY(t)=f(Y(t),r(t))dt+g(Y(t),r(t))\mathrm{d}B(t) \tag{2.2}$$

on $t\geq 0$ with the initial data $(Y(0),r(0))=(x,i)\in\mathbb{R}^n\times\mathbb{S}$, where

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

For convenience we further impose the following hypothesises on the drift and diffusion coefficients.

Assumption 2.1. *For any $j\in\mathbb{S}$, there exists a constant $\alpha_j\in\mathbb{R}$ such that*

$$(u-v)^T(f(u,j)-f(v,j))\leq\alpha_j|u-v|^2, \qquad \forall u,\ v\in\mathbb{R}^n. \tag{2.3}$$

Moreover, for any $R\geq 0$, there exists a positive constant K_R such that

$$|f(u,j)-f(v,j)|\leq K_R|u-v|,$$

for any $u,\ v\in\mathbb{R}^n$, $|u|\vee|v|\leq R$, $j\in\mathbb{S}$.

Assumption 2.2. *For any $j\in\mathbb{S}$, there exist constants $h_j\in\mathbb{R}$ and $h>0$ such that*

$$|u-v|^2|g(u,j)-g(v,j)|^2-2|(u-v)^T(g(u,j)-g(v,j))|^2\leq h_j|u-v|^4, \tag{2.4}$$

and

$$|g(u,j)-g(v,j)|^2\leq h|u-v|^2, \tag{2.5}$$

for any $u,\ v \in \mathbb{R}^n$.

Next, for convenience, define

$$\beta_j = 2\alpha_j + h_j,\quad \beta = (\beta_1, \cdots, \beta_N)^T,\quad \lambda = |\mu\beta|. \tag{2.6}$$

Assumptions 2.1 and the elementary inequality imply that for any $u \in \mathbb{R}^n$

$$\begin{aligned} u^T f(u,j) &\le \alpha_j |u|^2 + |u^T f(0,j)| \le \alpha_j |u|^2 + \frac{\lambda |u|^2}{8} + \frac{2|f(0,j)|^2}{\lambda} \\ &\le \Big(\alpha_j + \frac{1}{8}\lambda\Big)|u|^2 + \sigma_1, \end{aligned} \tag{2.7}$$

and Assumption 2.2 and the elementary inequality imply that

$$|g(u,j)|^2 \le 2h|u|^2 + \sigma_2, \tag{2.8}$$

where $\sigma_1 = 2\max_{j\in\mathbb{S}}\{|f(0,j)|^2/\lambda\}$, $\sigma_2 = 2\max_{j\in\mathbb{S}}\{|g(0,j)|^2\}$. Moreover, choosing constants $2p \le \varepsilon = \lambda/16h$, we find that

$$\begin{aligned} |u|^2|g(u,j)|^2 + (p-2)|u^T g(u,j)|^2 &\le \big(h_j + (3\varepsilon + 2p)h\big)|u|^4 + \frac{\sigma_2(1 + 2p + 3\varepsilon^{-1})}{2}|u|^2 \\ &\le \big(h_j + \frac{1}{4}\lambda\big)|u|^4 + \sigma_3 |u|^2, \end{aligned} \tag{2.9}$$

where $\sigma_3 = (1 + \lambda/16h + 48h/\lambda)\sigma_2/2$.

Under Assumptions 2.1 and 2.2, the equation (2.2) admits a unique solution $(Y(t), r(t))$ (see, e.g., [13, Theorem 3.17, p.93]). Throughout the paper, we write $(Y_t^{x,i}, r_t^i)$ in lieu of $(Y(t), r(t))$ to highlight the initial data $(Y(0), r(0)) = (x, i)$. Let $\mathcal{P}(\mathbb{R}^n \times \mathbb{S})$ denote the family of all probability measures on $\mathbb{R}^n \times \mathbb{S}$. For any $p \in (0, 1]$, define a metric on $\mathbb{R}^n \times \mathbb{S}$ as the following

$$d_p((u,j),(v,l)) := |u - v|^p + I_{\{j \ne l\}},\quad (u,j),(v,l) \in \mathbb{R}^n \times \mathbb{S},$$

and the corresponding Wasserstein distance between $\nu, \tilde\nu \in \mathcal{P}(\mathbb{R}^n \times \mathbb{S})$ by

$$W_p(\nu, \tilde\nu) := \inf_{\pi \in C(\nu,\tilde\nu)} \int_{(\mathbb{R}^n\times\mathbb{S})\times(\mathbb{R}^n\times\mathbb{S})} d_p(u,v)\pi(\mathrm{d}u, \mathrm{d}v),$$

where $C(\nu, \tilde\nu)$ denotes the set of all couplings of ν and $\tilde\nu$. Let $\mathbf{P}_t(x, i; \mathrm{d}u \times \{l\})$ be the transition probability kenel of the pair $\big(Y_t^{x,i}, r_t^i\big)$, a time homogeneous Markov process (see, e.g, [13, Theorem 3.28, pp.105-106]). Recall that $\pi \in \mathcal{P}(\mathbb{R}^n \times S)$ is called an invariant measure of $\big(Y_t^{x,i}, r_t^i\big)$ if

$$\pi(\Gamma \times \{j\}) = \sum_{l=1}^{N} \int_{\mathbb{R}^n} \mathbf{P}_t(u, l; \Gamma \times \{j\})\pi(\mathrm{d}u \times \{l\}),\quad \forall t \ge 0,\ \Gamma \in \mathscr{B}(\mathbb{R}^n),\ j \in \mathbb{S}$$

holds. For each $p > 0$, define

$$\Lambda = \mathrm{diag}(2\beta_1 + \lambda, \cdots, 2\beta_N + \lambda),\quad Q_p = Q + \frac{p}{4}\Lambda,\quad \eta_p = -\max_{\gamma \in \mathrm{spec}(Q_p)} \mathrm{Re}\gamma, \tag{2.10}$$

where λ and β_j are introduced in (2.6), Q is the generator of $\{r(t)\}_{t\ge0}$, and $\mathrm{spec}(Q_p)$ denotes the spectrum of Q_p.

The following lemma highlights the relationship between the sign of $\mu\beta$ and the sign of η_p.

Lemma 2.1. *Assume that* $\mu\beta < 0$, *then*

(1) $\eta_p > 0$ *if* $2\check{\beta} + \lambda \leq 0$;

(2) $\eta_p > 0$ *for* $p \in \left(0, \min_{j \in \mathbb{S}, 2\beta_j + \lambda > 0} \{-4q_{jj}/(2\beta_j + \lambda)\}\right)$ *if* $2\check{\beta} + \lambda > 0$,

where λ *and* β_j *are introduced in* (2.6).

Proof. According to (2.1) and $\mu\beta < 0$, it is easy to obtain

$$\sum_{j=1}^{N} \mu_j (2\beta_j + \lambda) = 2\mu\beta + \lambda = \mu\beta < 0.$$

Then the desired assertion follows from [5, Proposition 4.2] directly. □

We have the following result on the invariant measure for the exact solution.

Theorem 2.2. *Suppose that Assumptions 2.1, 2.2, and* $\mu\beta < 0$ *hold, then the solutions of the SDE with Markovian switching (2.2) converge to a unique invariant measure* $\pi \in \mathcal{P}(\mathbb{R}^n \times \mathbb{S})$ *with some exponential rate* $\xi > 0$ *in the Wasserstein distance.*

Proof. We shall adopt the approach of [4, Theorem 2.3] to complete the proof. Let

$$p_0 = 1 \wedge \min_{j \in \mathbb{S}, 2\beta_j + \lambda > 0} \{-4q_{jj}/(2\beta_j + \lambda)\} \wedge \lambda/32h. \tag{2.11}$$

Thus, for any $p \in (0, p_0)$, (2.9) holds, and $\eta_p > 0$ follows from Lemma 2.1. One observes that

$$\mathcal{L}\left((1 + |x|^2)^{\frac{p}{2}} \xi_i^{(p)}\right) \leq C - \eta_p \xi_i^{(p)} (1 + |x|^2)^{\frac{p}{2}}, \tag{2.12}$$

for $p \in (0, p_0)$, where $\xi^{(p)} = (\xi_1^{(p)}, \cdots, \xi_N^{(p)}) \gg \mathbf{0}$ is a eigenvector of Q_p corresponding to $-\eta_p$, C is a positive constant. Borrowing the proof method of [4, Theorem 2.3] we can get the result on the existence and uniqueness of the invariant measure but omit the details to avoid duplication. By the similar way to Theorem 3.5, we yield the exponential convergence rate. □

Remark 2.3. By virtue of Theorem 2.2, the solution $(Y(t), r(t))$ is ergodic and the transition probability of $(Y(t), r(t))$ converges to its invariant measure with exponential rate in the Wasserstein distance. Furthermore, due to (2.12) the Foster-Lyapunov criterion [20, Theorem 6.1, p.536] implies that $(Y(t), r(t))$ is exponentially ergodic, provided all compact sets are petite for some skeleton chain. Thus, this pair is strongly mixing since it is positively Harris-recurrent, see details in [2, p.881]. However more conditions should be imposed on the coefficients of the equation in order for all compact sets are petite for some skeleton chain.

3. Numerical Invariant Measure. This section is devoted to the existence and uniqueness of the numerical invariant measure of the BEM method and approximation of the numerical invariant measure to the underlying one in the Wasserstein metric. In order to define the numerical solution, we need to explain how to simulate a discrete Markov chain, which has been formulated in [13, Chapter 4, p.111]. To make the content self-contained, we sketch it here.

Given a stepsize $\Delta > 0$ and let $P(\Delta) = (P_{ij}(\Delta))_{N \times N} = \exp(\Delta Q)$. The discrete Markov chain $\{r_k,\ k = 0, 1, \cdots\}$ can be simulated as follows: let $r(0) = i$ and give a random pseudo number ς_1 obeying the uniform $(0,1)$ distribution. Define

$$r_1 = \begin{cases} i_1, & \text{if } i_1 \in \mathbb{S} - \{N\} \text{ such that } \displaystyle\sum_{j=1}^{i_1-1} P_{ij}(\Delta) \le \varsigma_1 < \sum_{j=1}^{i_1} P_{ij}(\Delta), \\ N, & \text{if } \displaystyle\sum_{j=1}^{N-1} P_{ij}(\Delta) \le \varsigma_1, \end{cases}$$

where $\sum_{j=1}^{N} P_{ij}(\Delta) = 0$ as usual. In other words, the probability of state s being chosen is given by $\mathbb{P}(r_1 = s) = P_{is}(\Delta)$. Generally, after the computations of $r_0, r_1, \cdots, r_k$, give a random pseudo number ς_{k+1} obeying a uniform $(0,1)$ distribution and define r_{k+1} by

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

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

We can now define the BEM scheme for the SDEs with Markovian switching (2.2). Let $X_0 = x$, $r_0 = i$, and define

$$X_{k+1} = X_k + f(X_{k+1}, r_k)\triangle + g(X_k, r_k)\triangle B_k, \quad k \ge 0, \tag{3.1}$$

where $\triangle B_k = B(t_{k+1}) - B(t_k)$. Here $X_k, r_k, k \ge 0$, depend on the step size $\triangle$, we drop it for simplicity. We point out that the BEM method (3.1) is well-defined under Assumption 2.1 based on a known result [15, Lemma 5.1] as follows.

Lemma 3.1. *Let Assumption 2.1 holds and $\triangle < 1/|\breve{\alpha}|$. Then for any $j \in \mathbb{S}$, $b \in \mathbb{R}^n$, there is a unique root $u \in \mathbb{R}^n$ of the equation*

$$u = b + f(u, j)\triangle.$$

It is useful to write (3.1) as

$$X_{k+1} - f(X_{k+1}, r_k)\triangle = X_k + g(X_k, r_k)\triangle B_k. \tag{3.2}$$

For any $j \in \mathbb{S}$, define a function $G_j : \mathbb{R}^n \to \mathbb{R}^n$ satisfying $G_j(u) = u - f(u,j)\triangle$. Then G_j has its inverse function $G_j^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ for any $j \in \mathbb{S}$. Moreover, the BEM method (3.1) can be represented as

$$X_{k+1} = G_{r_k}^{-1}(X_k + g(X_k, r_k)\triangle B_k), \qquad \forall k \ge 0. \tag{3.3}$$

Similar to that of [13, Theorem 6.14, p.250], we can prove the following result.

Lemma 3.2. $\{(X_k, r_k)\}_{k\ge 0}$ *is a time homogeneous Markov chain.*

Let $\mathbf{P}^{\triangle}_{k\triangle}(x, i; \mathrm{d}u \times \{l\})$ be the transition probability kernel of the pair $\left(X_k^{x,i}, r_k^i\right)$, a time homogeneous Markov chain. If $\pi^{\triangle} \in \mathcal{P}(\mathbb{R}^n \times \mathbb{S})$ satisfies

$$\pi^{\triangle}(\Gamma \times \{j\}) = \sum_{l=1}^{N} \int_{\mathbb{R}^n} \mathbf{P}^{\triangle}_{k\triangle}(u, l; \Gamma \times \{j\}) \pi^{\triangle}(\mathrm{d}u \times \{l\}), \forall k \ge 0, \Gamma \in \mathscr{B}(\mathbb{R}^n), j \in \mathbb{S},$$

then $\pi^{\triangle}$ is called an invariant measure of $\big(X_k^{x,i}, r_k^i\big)$. For convenience, Denote by C a generic positive constant which value may be different with different appearance and is independent of the iteration number k and the time stepsize $\triangle$.

In order to show the existence of the numerical invariant measure we prepare the following lemma on the moment boundedness of the numerical solution of the BEM scheme borrowing the idea of [11].

LEMMA 3.3. *Under the conditions of Theorem 2.2, there exists a constant* $\bar{\triangle}$ *such that the numerical solution of BEM scheme with any initial value* $(x,i) \in \mathbb{R}^n \times \mathbb{S}$ *satisfies*

$$\sup_{k\ge 0} \mathbb{E}|X_k|^p \le C(1+|x|^p) \tag{3.4}$$

for any $\triangle \in (0,\bar{\triangle})$ *and any* $p \in (0,p_0)$*, where* p_0 *is defined by* (2.11).

Proof. It follows from (2.7) and (3.1) that

$$\begin{aligned} |X_{k+1}|^2 =& X_{k+1}^T\Big(f(X_{k+1},r_k)\triangle + X_k + g(X_k,r_k)\triangle B_k\Big) \\ \le& \big(\alpha_{r_k} + \frac{1}{8}\lambda\big)|X_{k+1}|^2\triangle + \sigma_1\triangle + \frac{1}{2}|X_{k+1}|^2 + \frac{1}{2}|X_k + g(X_k,r_k)\triangle B_k|^2. \end{aligned}$$

Choosing a constant $0 < \triangle_1 < 1$ such that $(2|\breve{\alpha}| + \frac{1}{4}\lambda)\triangle_1 \le 1/3$ (where $|\breve{\alpha}| := \min_{i\in\mathbb{S}} |\alpha_i|$), we then obtain for any $\triangle \in (0,\triangle_1]$,

$$|X_{k+1}|^2 \le \frac{1}{1-(2\alpha_{r_k}+\frac{1}{4}\lambda)\triangle}|X_k + g(X_k,r_k)\triangle B_k|^2 + \frac{2\sigma_1\triangle}{1-(2\alpha_{r_k}+\frac{1}{4}\lambda)\triangle},$$

which implies

$$\begin{aligned} 1+|X_{k+1}|^2 \le& \frac{1}{1-(2\alpha_{r_k}+\frac{1}{4}\lambda)\triangle}\Big[1+|X_k + g(X_k,r_k)\triangle B_k|^2 + \Big(2\sigma_1 - 2\alpha_{r_k}\Big)\triangle\Big] \\ \le& \frac{(1+|X_k|^2)}{1-(2\alpha_{r_k}+\frac{1}{4}\lambda)\triangle}\Big(1+\upsilon_k(r_k)\Big), \end{aligned}$$

where

$$\upsilon_k(r_k) = \frac{2X_k^T g(X_k,r_k)\triangle B_k + |g(X_k,r_k)\triangle B_k|^2 + c_1\triangle}{1+|X_k|^2}, \quad c_1 = |2\sigma_1 - 2\hat{\alpha}|.$$

For any $p \in (0,p_0)$ where p_0 is defined by (2.11), noting that

$$(1+u)^{\frac{p}{2}} \le 1 + \frac{p}{2}u + \frac{p(p-2)}{8}u^2 + \frac{p(p-2)(p-4)}{48}u^3, \quad u \ge -1 \tag{3.5}$$

and $\upsilon_k(r_k) > -1$, we then have

$$\begin{aligned} \mathbb{E}\Big((1+|X_{k+1}|^2)^{\frac{p}{2}}|\mathcal{F}_{t_k}\Big) \le& \frac{(1+|X_k|^2)^{\frac{p}{2}}}{[1-(2\alpha_{r_k}+\frac{1}{4}\lambda)\triangle]^{\frac{p}{2}}} \\ &\times \mathbb{E}\Big(1+\frac{p}{2}\upsilon_k(r_k) + \frac{p(p-2)}{8}\upsilon_k^2(r_k) + \frac{p(p-2)(p-4)}{48}\upsilon_k^3(r_k)\Big|\mathcal{F}_{t_k}\Big). \end{aligned} \tag{3.6}$$

Since $\triangle B_k$ is independent of $\mathcal{F}_{t_k}$, we have $\mathbb{E}(\triangle B_k|\mathcal{F}_{t_k}) = 0$, $\mathbb{E}(|A\triangle B_k|^2|\mathcal{F}_{t_k}) = |A|^2\triangle$, for any $A \in \mathbb{R}^{n\times m}$. Hence,

$$\mathbb{E}\big(\upsilon_k(r_k)|\mathcal{F}_{t_k}\big) = \frac{|g(X_k,r_k)|^2\triangle + c_1\triangle}{1+|X_k|^2}. \tag{3.7}$$

Using the properties $\mathbb{E}(|\triangle B_k|^{2j}) = C\triangle^j,\ \mathbb{E}(|\triangle B_k|^{2j-1}|\mathcal{F}_{t_k}) \le C\triangle^{j-\frac{1}{2}},\ j=2,3,\cdots,$ we compute

$$\mathbb{E}\big(\upsilon_k^2(r_k)|\mathcal{F}_{t_k}\big) = \frac{1}{(1+|X_k|^2)^2}\Big(4|X_k^Tg(X_k,r_k)|^2\triangle + C\triangle^{\frac{3}{2}}\Big) \ge \frac{4|X_k^Tg(X_k,r_k)|^2\triangle}{(1+|X_k|^2)^2}, \tag{3.8}$$

and

$$\begin{aligned}&\mathbb{E}\big(\upsilon_k^3(r_k)|\mathcal{F}_{t_k}\big)\\ &\le \frac{9}{(1+|X_k|^2)^3}\mathbb{E}\Big[8|X_k^Tg(X_k,r_k)\triangle B_k|^3 + |g(X_k,r_k)\triangle B_k|^6 + c_1^3\triangle^3\Big|\mathcal{F}_{t_k}\Big]\\ &\le C\triangle^{\frac{3}{2}}.\end{aligned} \tag{3.9}$$

Combining (3.6)-(3.9) and using (2.8), for any $k \ge 0$ we obtain,

$$\begin{aligned}&\mathbb{E}\Big((1+|X_{k+1}|^2)^{\frac{p}{2}}|\mathcal{F}_{t_k}\Big)\\ \le&\frac{(1+|X_k|^2)^{\frac{p}{2}}}{[1-(2\alpha_{r_k}+\frac{1}{4}\lambda)\triangle]^{\frac{p}{2}}}\Big\{1+\frac{p}{2}\Big[\frac{|X_k|^2|g(X_k,r_k)|^2+(p-2)|X_k^Tg(X_k,r_k)|^2}{(1+|X_k|^2)^2}\triangle\\ &+\frac{\big(2h+c_1\big)|X_k|^2+\sigma_2+c_1}{(1+|X_k|^2)^2}\triangle\Big]+C\triangle^{\frac{3}{2}}\Big\}.\end{aligned} \tag{3.10}$$

This, together with (2.8) and (2.9), implies

$$\begin{aligned}&\mathbb{E}\Big((1+|X_{k+1}|^2)^{\frac{p}{2}}|\mathcal{F}_{t_k}\Big)\\ \le&\frac{(1+|X_k|^2)^{\frac{p}{2}}}{[1-(2\alpha_{r_k}+\frac{1}{4}\lambda)\triangle]^{\frac{p}{2}}}\Big\{1+\frac{p}{2}\Big[\frac{\big(h_{r_k}+\frac{1}{4}\lambda\big)|X_k|^4+\sigma_3|X_k|^2}{(1+|X_k|^2)^2}\triangle\\ &+\frac{\big(2h+c_1\big)|X_k|^2+\sigma_2+c_1}{(1+|X_k|^2)^2}\triangle\Big]+C\triangle^{\frac{3}{2}}\Big\}\\ \le&\frac{(1+|X_k|^2)^{\frac{p}{2}}}{[1-(2\alpha_{r_k}+\frac{1}{4}\lambda)\triangle]^{\frac{p}{2}}}\Big[1+\frac{p}{2}\big(h_{r_k}+\frac{1}{4}\lambda\big)\triangle+C\triangle^{\frac{3}{2}}\Big]+C\triangle.\end{aligned} \tag{3.11}$$

Choosing a constant $0<\triangle_2\le\triangle_1$ sufficiently small such that $C\triangle_2^{\frac{1}{2}}\le p\lambda/8$, and $27(p+2)\big(2|\breve{\alpha}|+\lambda/4\big)^2\triangle_2\le 2\lambda$, this yields that for any $\triangle\in(0,\triangle_2]$

$$\frac{p}{2}\big(h_{r_k}+\frac{1}{4}\lambda\big)\triangle+C\triangle^{\frac{3}{2}}\le\frac{p}{2}\big(h_{r_k}+\frac{1}{2}\lambda\big)\triangle \tag{3.12}$$

and

$$\begin{aligned}&\Big[1-\big(2\alpha_{r_k}+\frac{\lambda}{4}\big)\triangle\Big]^{-\frac{p}{2}}\\ \le&1+\frac{p}{2}\big(2\alpha_{r_k}+\frac{\lambda}{4}\big)\triangle+\frac{p(p+2)(2|\breve{\alpha}|+\frac{1}{4}\lambda)^2}{8[1-(2|\breve{\alpha}|+\frac{1}{4}\lambda)\triangle_2]^{\frac{p}{2}+2}}\triangle^2\\ \le&1+\frac{p}{2}\big(2\alpha_{r_k}+\frac{5\lambda}{16}\big)\triangle.\end{aligned} \tag{3.13}$$

Then for any $\triangle \in (0, \triangle_2]$, combining (3.11)-(3.13) we obtain

$$
\begin{aligned}
&\mathbb{E}\big((1+|X_{k+1}|^2)^{\frac{p}{2}}|\mathcal{F}_{t_k}\big)\\
\le&(1+|X_k|^2)^{\frac{p}{2}}\Big[1+\frac{p}{2}\big(2\alpha_{r_k}+h_{r_k}+\frac{13\lambda}{16}\big)\triangle+C\triangle^2\Big]+C\triangle.
\end{aligned}
$$

Letting $\bar{\triangle}$ be a constant such that $\bar{\triangle} \in (0, \triangle_2]$, $C\bar{\triangle} \le p\lambda/32$ and $(|\breve{\beta}| + \frac{7}{8}\lambda)\bar{\triangle} < 1$ (where $|\breve{\beta}| = \max_{i\in\mathbb{S}} |\beta_i|$), we arrive at for $\triangle \in (0, \bar{\triangle}]$

$$
\mathbb{E}\big((1+|X_{k+1}|^2)^{\frac{p}{2}}|\mathcal{F}_{t_k}\big) \le \Big[1+\frac{p}{2}\big(\beta_{r_k}+\frac{7}{8}\lambda\big)\triangle\Big](1+|X_k|^2)^{\frac{p}{2}}+C\triangle, \tag{3.14}
$$

where β_i is defined as (2.6) for each $i \in \mathbb{S}$. For any $k \ge 1$, we further compute

$$
\begin{aligned}
\mathbb{E}\big((1+|X_{k+1}|^2)^{\frac{p}{2}}|\mathcal{F}_{t_{k-1}}\big) \le&\Big[1+\frac{p}{2}\big(\beta_{r_k}+\frac{7}{8}\lambda\big)\triangle\Big]\mathbb{E}((1+|X_k|^2)^{\frac{p}{2}}|\mathcal{F}_{t_{k-1}})+C\triangle\\
\le&\prod_{j=k-1}^{k}\Big[1+\frac{p}{2}\big(\beta_{r_j}+\frac{7}{8}\lambda\big)\triangle\Big](1+|X_{k-1}|^2)^{\frac{p}{2}}\\
&+C\triangle\Big[1+\frac{p}{2}\big(\beta_{r_k}+\frac{7}{8}\lambda\big)\triangle\Big]+C\triangle.
\end{aligned}
\tag{3.15}
$$

Repeating (3.15) we obtain

$$
\begin{aligned}
\mathbb{E}\big((1+|X_{k+1}|^2)^{\frac{p}{2}}|\mathcal{F}_0\big) \le&(1+|X_0|^2)^{\frac{p}{2}}\prod_{j=0}^{k}\Big[1+\frac{p}{2}\big(\beta_{r_j}+\frac{7}{8}\lambda\big)\triangle\Big]\\
&+C\triangle\sum_{i=1}^{k}\prod_{j=k-i+1}^{k}\Big[1+\frac{p}{2}\big(\beta_{r_j}+\frac{7}{8}\lambda\big)\triangle\Big]+C\triangle.
\end{aligned}
$$

Hence, for any $k \ge 0$, by virtue of the homogeneous property of the Markov chain, taking expectations on both sides yields

$$
\begin{aligned}
\mathbb{E}\big((1+|X_{k+1}|^2)^{\frac{p}{2}}\big) \le&(1+|x|^2)^{\frac{p}{2}}\mathbb{E}\Big[\prod_{j=0}^{k}\Big(1+\frac{p}{2}\big(\beta_{r_j}+\frac{7}{8}\lambda\big)\triangle\Big)\Big]\\
&+C\triangle\sum_{i=1}^{k}\mathbb{E}\Big[\mathbb{E}\Big(\prod_{j=k-i+1}^{k}\Big(1+\frac{p}{2}\big(\beta_{r_j}+\frac{7}{8}\lambda\big)\triangle\Big)|\mathcal{F}_{k-i}\Big)\Big]+C\triangle\\
\le&(1+|x|^2)^{\frac{p}{2}}\mathbb{E}\Big[\prod_{j=0}^{k}\Big(1+\frac{p}{2}\big(\beta_{r_j}+\frac{7}{8}\lambda\big)\triangle\Big)\Big]\\
&+C\triangle\sum_{i=1}^{k}\mathbb{E}\Big[\prod_{j=1}^{i}\Big(1+\frac{p}{2}\big(\beta_{r_j}+\frac{7}{8}\lambda\big)\triangle\Big)\Big]+C\triangle.
\end{aligned}
$$

Thus, we have

$$
\begin{aligned}
&\mathbb{E}\big((1+|X_{k+1}|^2)^{\frac{p}{2}}\big)\\
&\le (1+|x|^2)^{\frac{p}{2}}\mathbb{E}\bigg[\exp\Big(\sum_{j=0}^{k}\log\Big(1+\frac{p}{2}\big(\beta_{r_j}+\frac{7}{8}\lambda\big)\triangle\Big)\Big)\bigg]\\
&\quad + C\triangle\sum_{i=1}^{k}\mathbb{E}\bigg[\exp\Big(\sum_{j=1}^{i}\log\Big(1+\frac{p}{2}\big(\beta_{r_j}+\frac{7}{8}\lambda\big)\triangle\Big)\Big)\bigg]+C\triangle.\\
&=: H_1+H_2+C\triangle.
\end{aligned}
\tag{3.16}
$$

Then, by the ergodic property of the Markov chain (see, e.g., [15]) and inequality $\log(1+u)\le u,\ \forall u>-1$, we compute

$$
\begin{aligned}
\lim_{i\to\infty}\frac{1}{i}\sum_{j=1}^{i}\log\Big(1+\frac{p}{16}\big(8\beta_{r_j}+7\lambda\big)\triangle\Big) &= \sum_{j\in\mathbb{S}}\mu_j\log\Big(1+\frac{p}{16}\big(8\beta_j+7\lambda\big)\triangle\Big)\\
&\le \frac{p\triangle}{16}\sum_{j\in\mathbb{S}}\mu_j\big(8\beta_j+7\lambda\big) = -\frac{\lambda p\triangle}{16}\ \ a.s.
\end{aligned}
$$

which implies $\lim_{i\to\infty}\exp\Big(\frac{\lambda p\triangle i}{32}+\sum_{j=1}^{i}\log\Big(1+\frac{p}{16}\big(8\beta_{r_j}+7\lambda\big)\triangle\Big)\Big)=0\ a.s.$ By virtue of the Fatou lemma (see, e.g. [30, p.187, Theorem 2]) we have

$$
\limsup_{i\to\infty}\mathbb{E}\bigg[\exp\Big(\frac{\lambda p\triangle i}{32}+\sum_{j=1}^{i}\log\Big(1+\frac{p}{16}\big(8\beta_{r_j}+7\lambda\big)\triangle\Big)\Big)\bigg]=0.
$$

Thus there is a positive integer N such that

$$
\mathbb{E}\bigg[\exp\Big(\sum_{j=1}^{i}\log\Big(1+\frac{p}{16}\big(8\beta_{r_j}+7\lambda\big)\triangle\Big)\Big)\bigg]\le\exp\Big(-\frac{\lambda p\triangle}{32}i\Big),\qquad \forall i>N.
\tag{3.17}
$$

Therefore

$$
H_1\le(1+|x|^2)^{\frac{p}{2}}\Big(1+\frac{p}{2}\big(\beta_{r_0}+\frac{7}{8}\lambda\big)\triangle\Big)\exp\Big(-\frac{\lambda p\triangle}{32}k\Big),\qquad\forall k>N.
\tag{3.18}
$$

For the given N we know

$$
\sum_{i=1}^{N}\mathbb{E}\bigg[\exp\Big(\sum_{j=1}^{i}\log\Big(1+\frac{p}{2}\big(\beta_{r_j}+\frac{7}{8}\lambda\big)\triangle\Big)\Big)\bigg]\le\sum_{i=1}^{N}\Big(1+\frac{p}{2}\big(|\breve{\beta}|+\frac{7}{8}\lambda\big)\triangle\Big)^{i}\le C.
$$

This, together with (3.17), implies

$$
\begin{aligned}
H_2&\le C\triangle+C\triangle\sum_{i=N+1}^{\infty}\mathbb{E}\bigg[\exp\Big(\sum_{j=1}^{i}\log\Big(1+\frac{p}{2}\big(\beta_{r_j}+\frac{7}{8}\lambda\big)\triangle\Big)\Big)\bigg]\\
&\le C\triangle+C\triangle\sum_{i=N+1}^{k}\exp\Big(-\frac{\lambda p\triangle}{32}i\Big),\qquad\forall k>N.
\end{aligned}
\tag{3.19}
$$

Combining (3.18) and (3.19) with (3.16) yields

$$
\begin{aligned}
&\mathbb{E}\big((1+|X_{k+1}|^2)^{\frac{p}{2}}\big) \le C\triangle + C(1+|x|^2)^{\frac{p}{2}} \exp\Big(-\frac{\lambda p\triangle}{32}(k\vee(N+1))\Big)\\
&\qquad + C\triangle \sum_{i=N+1}^{k\vee(N+1)} \exp\Big(-\frac{\lambda p\triangle}{32} i\Big) \le C(1+|x|^p), \qquad \forall k\ge 0.
\end{aligned} \tag{3.20}
$$

Therefore the desired assertion follows. □

Remark 3.1. Recently, the work of [11] gives the the moment boundedness of the BEM numerical solutions for SDEs without globally Lipschitz continuous coefficients. However the proof techniques can't be adopted for SDEs with regime switching directly since their dynamical behaviors are significantly different from those of SDEs. In the proof of Lemma 3.3 we establish the recursion formula (3.14) dependent on the states, and then yield the desired result by making use of the ergodic property of the Markov chain.

To investigate the uniqueness of the invariant measure we provide the asymptotically attractive property of the numerical solutions of BEM scheme. Here we denote the numerical solution of BEM scheme with any given initial value (x,i) by $X_k^{x,i}$.

Lemma 3.4. *Under the conditions of Theorem 2.2, it holds that*

$$
\mathbb{E}|X_k^{x,i} - X_k^{y,j}|^p \le C(1+|x|^p+|y|^p)\mathrm{e}^{-\varsigma k\triangle} \tag{3.21}
$$

for any $\triangle\in(0,\bar{\triangle})$ and for any $p\in(0,p_0)$, $(x,i),(y,j)\in\mathbb{R}^n\times\mathbb{S}$, $\bar{\triangle}$ and p_0 are given in Lemma 3.3, $\varsigma>0$ is a constant.

Proof. Note that

$$
\begin{cases}
X_{k+1}^{x,i} = X_k^{x,i} + f(X_{k+1}^{x,i}, r_k^i)\triangle + g(X_k^{x,i}, r_k^i)\triangle B_k,\\
X_{k+1}^{y,i} = X_k^{y,i} + f(X_{k+1}^{y,i}, r_k^i)\triangle + g(X_k^{y,i}, r_k^i)\triangle B_k.
\end{cases}
$$

It follows from Assumption 2.1 that

$$
\begin{aligned}
&|X_{k+1}^{x,i} - X_{k+1}^{y,i}|^2\\
=&\Big(X_{k+1}^{x,i} - X_{k+1}^{y,i}\Big)^T \Big(f(X_{k+1}^{x,i}, r_k^i) - f(X_{k+1}^{y,i}, r_k^i)\Big)\triangle\\
&+\Big(X_{k+1}^{x,i} - X_{k+1}^{y,i}\Big)^T \Big(X_k^{x,i} - X_k^{y,i} + \big(g(X_k^{x,i}, r_k^i) - g(X_k^{y,i}, r_k^i)\big)\triangle B_k\Big)\\
\le&\alpha_{r_k^i}\big|X_{k+1}^{x,i} - X_{k+1}^{y,i}\big|^2\triangle\\
&+\frac{1}{2}\big|X_{k+1}^{x,i} - X_{k+1}^{y,i}\big|^2 + \frac{1}{2}\big|(X_k^{x,i} - X_k^{y,i}) + \big(g(X_k^{x,i}, r_k^i) - g(X_k^{y,i}, r_k^i)\big)\triangle B_k\big|^2.
\end{aligned}
$$

We hence obtain

$$
\begin{aligned}
|X_{k+1}^{x,i} - X_{k+1}^{y,i}|^2 \le& \frac{1}{1-2\alpha_{r_k^i}\triangle}\big|(X_k^{x,i} - X_k^{y,i}) + (g(X_k^{x,i}, r_k^i) - g(X_k^{y,i}, r_k^i)\triangle B_k)\big|^2\\
=&\frac{|X_k^{x,i} - X_k^{y,i}|^2}{1-2\alpha_{r_k^i}\triangle}\Big(1+\vartheta(r_k^i)\Big),
\end{aligned}
$$

where

$$
\begin{aligned}
&\vartheta_k(r_k^i)\\
=&\frac{2(X_k^{x,i} - X_k^{y,i})^T(g(X_k^{x,i}, r_k^i) - g(X_k^{y,i}, r_k^i))\triangle B_k + |\big(g(X_k^{x,i}, r_k^i) - g(X_k^{y,i}, r_k^i)\big)\triangle B_k|^2}{|X_k^{x,i} - X_k^{y,i}|^2}
\end{aligned}
$$

if $|X_k^{x,i}-X_k^{y,i}|\neq 0$, otherwise it is set to -1. Clear, $\vartheta_k(r_k^i)\geq -1$. For any $p\in(0,p_0)$, then using (3.5) we derive that

$$\begin{aligned}&\mathbb{E}\big(|X_{k+1}^{x,i}-X_{k+1}^{y,i}|^p\big|\mathcal{F}_{t_k}\big)\\ \leq&\frac{|X_k^{x,i}-X_k^{y,i}|^p}{\big(1-2\alpha_{r_k^i}\triangle\big)^{\frac{p}{2}}}I_{\{|X_k^{x,i}-X_k^{y,i}|\neq 0\}}\mathbb{E}\Big[1+\frac{p}{2}\vartheta_k(r_k^i)\\ &+\frac{p(p-2)}{8}\vartheta_k^2(r_k^i)+\frac{p(p-2)(p-4)}{48}\vartheta_k^3(r_k^i)\big|\mathcal{F}_{t_k}\Big].\end{aligned}\tag{3.22}$$

Then following the same way as (3.7)-(3.9), by (2.5) we can show

$$I_{\{|X_k^{x,i}-X_k^{y,i}|\neq 0\}}\mathbb{E}\big(\vartheta_k(r_k^i)|\mathcal{F}_{t_k}\big)=I_{\{|X_k^{x,i}-X_k^{y,i}|\neq 0\}}\frac{|g(X_k^{x,i},r_k^i)-g(X_k^{y,i},r_k^i)|^2\triangle}{|X_k^{x,i}-X_k^{y,i}|^2},\tag{3.23}$$

$$\begin{aligned}&I_{\{|X_k^{x,i}-X_k^{y,i}|\neq 0\}}\mathbb{E}\big(\vartheta_k^2(r_k^i)|\mathcal{F}_{t_k}\big)\\ \geq&I_{\{|X_k^{x,i}-X_k^{y,i}|\neq 0\}}\frac{4|(X_k^{x,i}-X_k^{y,i})^T(g(X_k^{x,i},r_k^i)-g(X_k^{y,i},r_k^i))|^2\triangle}{|X_k^{x,i}-X_k^{y,i}|^4},\end{aligned}\tag{3.24}$$

and

$$I_{\{|X_k^{x,i}-X_k^{y,i}|\neq 0\}}\mathbb{E}\big(\vartheta_k^3(r_k^i)|\mathcal{F}_{t_k}\big)\leq I_{\{|X_k^{x,i}-X_k^{y,i}|\neq 0\}}C\triangle^{\frac{3}{2}}.\tag{3.25}$$

Combining (3.22)-(3.25) and using Assumption 2.2, for any $k\geq 0$ we arrive at

$$\begin{aligned}&\mathbb{E}\big(|X_{k+1}^{x,i}-X_{k+1}^{y,i}|^p|\mathcal{F}_{t_k}\big)\\ \leq&\frac{|X_k^{x,i}-X_k^{y,i}|^p}{(1-2\alpha_{r_k^i}\triangle)^{\frac{p}{2}}}I_{\{|X_k^{x,i}-X_k^{y,i}|\neq 0\}}\Big[1+\frac{p}{2}\Big(\frac{|g(X_k^{x,i},r_k^i)-g(X_k^{y,i},r_k^i)|^2}{|X_k^{x,i}-X_k^{y,i}|^2}\triangle\\ &+(p-2)\frac{|(X_k^{x,i}-X_k^{y,i})^T(g(X_k^{x,i},r_k^i)-g(X_k^{y,i},r_k^i))|^2\triangle}{|X_k^{x,i}-X_k^{y,i}|^4}\Big)+\frac{p(p-2)(p-4)}{48}C\triangle^{\frac{3}{2}}\Big]\\ \leq&\frac{|X_k^{x,i}-X_k^{y,i}|^p}{(1-2\alpha_{r_k^i}\triangle)^{\frac{p}{2}}}\Big[1+\frac{p}{2}\big(h_{r_k^i}+ph\big)\triangle+\frac{p(p-2)(p-4)}{48}C\triangle^{\frac{3}{2}}\Big].\end{aligned}$$

It is easy to find from (2.11) that $4ph<\lambda$ holds for each $p\in(0,p_0)$. Choose a constant $0<\triangle_4\leq\bar{\triangle}$ ($\bar{\triangle}$ is a positive constant given in Lemma 3.3) sufficiently small such that $C\triangle_4^{1/2}\leq 3\lambda/8$, which implies that for any $\triangle\in(0,\triangle_4]$

$$\mathbb{E}\big(|X_{k+1}^{x,i}-X_{k+1}^{y,i}|^p|\mathcal{F}_{t_k}\big)\leq\frac{|X_k^{x,i}-X_k^{y,i}|^p}{\big(1-2\alpha_{r_k^i}\triangle\big)^{\frac{p}{2}}}\Big[1+\frac{p}{2}\big(h_{r_k^i}+\frac{1}{4}\lambda\big)\triangle+\frac{p\lambda}{16}\triangle\Big].\tag{3.26}$$

Further choose $0<\triangle_5\leq\triangle_4$ such that for any $\triangle\in(0,\triangle_5]$, any $i\in\mathbb{S}$, any integer k

$$(1-2\alpha_{r_k^i}\triangle)^{\frac{p}{2}}\geq 1-p\alpha_{r_k^i}\triangle-C\triangle^2\geq 1-\frac{p}{2}(2\alpha_{r_k^i}+\frac{1}{16}\lambda)\triangle\tag{3.27}$$

holds. Substituting this in (3.26) yields

$$\mathbb{E}\big(|X_{k+1}^{x,i}-X_{k+1}^{y,i}|^p|\mathcal{F}_{t_k}\big)\leq\frac{1+\frac{p}{2}(h_{r_k^i}+\frac{3}{8}\lambda)\triangle}{1-\frac{p}{2}(2\alpha_{r_k^i}+\frac{1}{16}\lambda)\triangle}|X_k^{x,i}-X_k^{y,i}|^p.$$

Using inequality $1/(1-u)\le 1+u+2u^2$ for any $u\in(-1/2,1/2)$, we obtain

$$\mathbb{E}\big(|X_{k+1}^{x,i}-X_{k+1}^{y,i}|^p|\mathcal{F}_{t_k}\big)\le\Big(1+\frac{p}{2}(\beta_{r_k^i}+\frac{1}{2}\lambda)\triangle\Big)|X_k^{x,i}-X_k^{y,i}|^p \tag{3.28}$$

for any $\triangle\in(0,\triangle^*)$, $p\in(0,p_0)$, where $0<\triangle^*\le\triangle_5$ satisfying $C\triangle^*\le p\lambda/32$, and $p_0(|\breve{\beta}|+\lambda/2)\triangle^*/2<1$. This implies that

$$\begin{aligned}\mathbb{E}\big(|X_k^{x,i}-X_k^{y,i}|^p\big)&\le|x-y|^p\mathbb{E}\Big[\prod_{j=0}^{k-1}\Big(1+\frac{p}{2}(\beta_{r_j^i}+\frac{1}{2}\lambda)\triangle\Big)\Big]\\&\le|x-y|^p\mathbb{E}\Big[\exp\Big(\sum_{j=0}^{k-1}\log\big(1+\frac{p}{2}(\beta_{r_j^i}+\frac{1}{2}\lambda)\triangle\big)\Big)\Big].\end{aligned}\tag{3.29}$$

This, together with the ergodic property of the Markov chain, yields

$$\lim_{k\to\infty}\frac{1}{k}\sum_{j=0}^{k-1}\log\Big(1+\frac{p}{2}(\beta_{r_j^i}+\frac{1}{2}\lambda)\triangle\Big)=\sum_{j\in\mathbb{S}}\mu_j\log\Big(1+\frac{p}{2}(\beta_j+\frac{1}{2}\lambda)\triangle\Big)\le-\frac{\lambda p\triangle}{4}\quad a.s.$$

We therefore have

$$\lim_{k\to\infty}\Big[\frac{\lambda pk\triangle}{8}+\sum_{j=0}^{k-1}\log\Big(1+\frac{p}{4}(2\beta_{r_j}+\lambda)\triangle\Big)\Big]=-\infty\qquad a.s. \tag{3.30}$$

By virtue of the Fatou lemma, we have

$$\lim_{k\to\infty}\mathbb{E}\Big[\exp\Big(\frac{\lambda pk\triangle}{8}+\sum_{j=0}^{k-1}\log\Big(1+\frac{p}{4}(2\beta_{r_j}+\lambda)\triangle\Big)\Big)\Big]=0.$$

This together with (3.29) implies

$$\mathbb{E}|X_k^{x,i}-X_k^{y,i}|^p\le C|x-y|^p\mathrm{e}^{-\frac{\lambda pk\triangle}{8}},\qquad\forall\ k>0. \tag{3.31}$$

Define $\bar{\tau}=\inf\{k\ge0:r_k^i=r_k^j\}$. Since the state space $\mathbb{S}$ is finite, and Q is irreducible, there exists $\bar{\gamma}>0$ such that

$$\mathbb{P}(\bar{\tau}>k)\le\mathrm{e}^{-\bar{\gamma}k\triangle} \tag{3.32}$$

for any $k>0$. For the fixed $p\in(0,p_0)$, let $q=(p+p_0)/(2p)>1$, then $pq=(p+p_0)/2\in(0,p_0)$. Moreover, Hölder's inequality implies that

(3.33)

$$\begin{aligned}&\mathbb{E}|X_k^{x,i}-X_k^{y,j}|^p\\=&\mathbb{E}\Big(|X_k^{x,i}-X_k^{y,j}|^pI_{\{\bar{\tau}>[\frac{k}{2}]\}}\Big)+\mathbb{E}\Big(|X_k^{x,i}-X_k^{y,j}|^pI_{\{\bar{\tau}\le[\frac{k}{2}]\}}\Big)\\\le&\Big(\mathbb{E}|X_k^{x,i}-X_k^{y,j}|^{pq}\Big)^{\frac{1}{q}}\Big(\mathbb{P}(\bar{\tau}>[\frac{k}{2}])\Big)^{1-\frac{1}{q}}+\mathbb{E}\Big[I_{\{\bar{\tau}\le[\frac{k}{2}]\}}\mathbb{E}\Big(|X_k^{x,i}-X_k^{y,j}|^p\big|\mathcal{F}_{\bar{\tau}\triangle}\Big)\Big]\\\le&\Big(\mathbb{E}|X_k^{x,i}-X_k^{y,j}|^{pq}\Big)^{\frac{1}{q}}\Big(\mathbb{P}(\bar{\tau}>[\frac{k}{2}])\Big)^{1-\frac{1}{q}}+\mathbb{E}\Big[I_{\{\bar{\tau}\le[\frac{k}{2}]\}}\mathbb{E}\Big(|X_{k-\bar{\tau}}^{X_{\bar{\tau}}^{x,i},r_{\bar{\tau}}^i}-X_{k-\bar{\tau}}^{X_{\bar{\tau}}^{y,j},r_{\bar{\tau}}^j}|^p\Big)\Big]\\\le&C\mathrm{e}^{-\frac{q-1}{2q}\bar{\gamma}k\triangle}\Big(\mathbb{E}|X_k^{x,i}-X_k^{y,j}|^{pq}\Big)^{\frac{1}{q}}+C\mathrm{e}^{-\frac{p}{16}\lambda k\triangle}\mathbb{E}\Big[I_{\{\bar{\tau}\le[\frac{k}{2}]\}}\mathbb{E}\Big(|X_{\bar{\tau}}^{x,i}-X_{\bar{\tau}}^{y,j}|^p\Big)\Big]\\\le&C\mathrm{e}^{-\frac{p_0-p}{2(p+p_0)}\bar{\gamma}k\triangle}\Big(\mathbb{E}|X_k^{x,i}-X_k^{y,j}|^{\frac{p+p_0}{2}}\Big)^{\frac{2p}{p+p_0}}+C\mathrm{e}^{-\frac{p}{16}\lambda k\triangle}\mathbb{E}\Big(\big|X_{\bar{\tau}\wedge[\frac{k}{2}]}^{x,i}-X_{\bar{\tau}\wedge[\frac{k}{2}]}^{y,j}\big|^p\Big),\end{aligned}$$

where $[x]$ represents the integer part of x for any $x \in \mathbb{R}$. Applying the elementary inequality $(a+b)^p \le 2^p(a^p+b^p)$ for all $a,b>0$, by (3.20), yields that $\big(\mathbb{E}|X_k^{x,i} - X_k^{y,j}|^{\frac{p+p_0}{2}}\big)^{\frac{2p}{p+p_0}} \le C(1+|x|^p+|y|^p)$, and

$$\begin{aligned}\mathbb{E}\Big(|X^{x,i}_{\bar\tau\wedge[\frac{k}{2}]} - X^{y,j}_{\bar\tau\wedge[\frac{k}{2}]}|^p\Big) \le &\mathbb{E}\Big(|X^{x,i}_{\bar\tau\wedge[\frac{k}{2}]}|^p\Big) + \mathbb{E}\Big(|X^{y,j}_{\bar\tau\wedge[\frac{k}{2}]}|^p\Big)\\ =&\mathbb{E}\Big(\sum_{l=0}^{[\frac{k}{2}]}|X_l^{x,i}|^p I_{\{\bar\tau\wedge[\frac{k}{2}]=l\}}(\omega)\Big) + \mathbb{E}\Big(\sum_{l=0}^{[\frac{k}{2}]}|X_l^{y,j}|^p I_{\{\bar\tau\wedge[\frac{k}{2}]=l\}}(\omega)\Big)\\ \le&\sum_{l=0}^{[\frac{k}{2}]}\Big[\mathbb{E}\big(|X_l^{x,i}|^p\big) + \mathbb{E}\big(|X_l^{y,j}|^p\big)\Big] \le C(1+|x|^p+|y|^p)(k+2).\end{aligned}$$

The desired assertion (3.21) follows by using (3.33). □

Next we give the existence and uniqueness of the numerical invariant measure for SDE (2.2) of BEM method.

THEOREM 3.5. *Under the conditions of Theorem 2.2, there is a positive* $\triangle^*$ *sufficiently small such that for any* $\triangle \in (0,\triangle^*)$*, the solutions of the BEM method (3.1) converge to a unique invariant measure* $\pi^{\Delta} \in \mathcal{P}(\mathbb{R}^n \times \mathbb{S})$ *with some exponential rate* $\xi_\triangle > 0$ *in the Wasserstein distance.*

Proof. For any initial data (x,i), by (3.4) and Chebyshev's inequality, we derive that $\{\delta_{(x,i)}\mathbf{P}^{\triangle}_{k\triangle}\}$ is tight, then one can extract a subsequence which converges weakly to an invariant measure denoted by $\pi^{\Delta} \in \mathcal{P}(\mathbb{R}^n\times\mathbb{S})$. It follows from (3.32) that

$$\mathbb{P}(r_k^i \ne r_k^j) = \mathbb{P}(\bar\tau > k) \le \mathrm{e}^{-\bar\gamma k\triangle} \tag{3.34}$$

for any $k>0$. Therefore, we derive from (3.21) and (3.34) that

$$\begin{aligned}W_p(\delta_{(x,i)}\mathbf{P}^{\triangle}_{k\triangle}, \delta_{(y,j)}\mathbf{P}^{\triangle}_{k\triangle}) \le &\mathbb{E}|X_k^{x,i} - X_k^{y,j}|^p + \mathbb{P}(r_k^i \ne r_k^j)\\ \le &C(1+|x|^p+|y|^p)\mathrm{e}^{-\xi_\triangle k\triangle}.\end{aligned} \tag{3.35}$$

where $\xi_\triangle := \varsigma \wedge \bar\gamma > 0$. Due to the Kolmogorov-Chapman equation and Lemma 3.3 one observes that for any $k,\ l>0$,

$$\begin{aligned}W_p(\delta_{(x,i)}\mathbf{P}^{\triangle}_{k\triangle}, \delta_{(x,i)}\mathbf{P}^{\triangle}_{(k+l)\triangle}) =&W_p(\delta_{(x,i)}\mathbf{P}^{\triangle}_{k\triangle}, \delta_{(x,i)}\mathbf{P}^{\triangle}_{l\triangle}\mathbf{P}^{\triangle}_{k\triangle})\\ \le&\int_{\mathbb{R}^n\times\mathbb{S}} W_p(\delta_{(x,i)}\mathbf{P}^{\triangle}_{k\triangle}, \delta_{(y,j)}\mathbf{P}^{\triangle}_{k\triangle})\mathbf{P}^{\triangle}_{l\triangle}(x,i;dy,j)\\ \le&\sum_{j\in\mathbb{S}}\int_{\mathbb{R}^n} C(1+|x|^p+|y|^p)\mathrm{e}^{-\xi_\triangle k\triangle}\mathbf{P}^{\triangle}_{l\triangle}(x,i;dy,j)\\ =&C(1+|x|^p+\mathbb{E}|X_l^{x,i}|^p)\mathrm{e}^{-\xi_\triangle k\triangle} \le C\mathrm{e}^{-\xi_\triangle k\triangle}.\end{aligned}$$

Thus, taking $l\to\infty$ implies

$$W_p(\delta_{(x,i)}\mathbf{P}^{\triangle}_{k\triangle}, \pi^{\Delta}) \le C\mathrm{e}^{-\xi_\triangle k\triangle} \to 0, \qquad k\to\infty, \tag{3.36}$$

namely, π^{Δ} is the unique invariant measure of $\{\delta_{(x,i)}\mathbf{P}^{\triangle}_{k\triangle}\}$. Assume $\nu_1^{\triangle}, \nu_2^{\triangle} \in \mathcal{P}(\mathbb{R}^n \times \mathbb{S})$ are the invariant measures of $(X_k^{x,i}, r_k^i)$ and $(X_k^{y,j}, r_k^j)$, respectively, we have

$$W_p(\nu_1^{\triangle},\nu_2^{\triangle}) \le \int_{(\mathbb{R}^n\times\mathbb{S})\times(\mathbb{R}^n\times\mathbb{S})} W_p(\delta_{(x,i)}\mathbf{P}^{\triangle}_{k\triangle}, \delta_{(y,j)}\mathbf{P}^{\triangle}_{k\triangle})\pi(dx\times di, dy\times dj),$$

where π is a coupling of $\nu_1^{\triangle}$ and $\nu_2^{\triangle}$. Therefore, the uniqueness of invariant measures follows from (3.35) immediately. □

The following theorem reveals that numerical invariant measure $\pi^{\triangle}$ converges in the Wassertein distance to the underlying one π.

THEOREM 3.6. *Under the conditions of Theorem 2.2,* $\lim_{\triangle\to 0} W_p(\pi,\pi^{\triangle}) = 0$. *Furthermore, if the drift term satisfies the polynomial growth condition, that is,*

$$|f(x,i)-f(y,i)|^2 \le c_i(1+|x|^q+|y|^q)|x-y|^2, \quad \forall x,y\in\mathbb{R}^n, i\in\mathbb{S},$$

then $W_p(\pi,\pi^{\triangle}) \le C\triangle^{\gamma}$ *for some* $\gamma\in(0,p/2)$*, where* c_i*,* q *are positive constants.*

Proof. Under Assumptions 2.1 and 2.2, by Theorem 2.2, Remark 2.3 and (3.36), for any $\triangle\in(0,\triangle^*)$ and any $\epsilon>0$, there is a $k>0$ sufficiently large such that

$$W_p(\delta_{(x,i)}\mathbf{P}_{k\triangle},\pi)+W_p(\delta_{(x,i)}\mathbf{P}^{\triangle}_{k\triangle},\pi^{\triangle}) \le C\mathrm{e}^{-\xi^* k\triangle} < \frac{\epsilon}{2} \tag{3.37}$$

where $\triangle^*$ is given by Theorem 3.5 and $\xi^* := \xi\wedge\xi_{\triangle}$. Moreover, for the fixed k by the convergence of finite time when $\triangle$ is sufficiently small, $W_p(\delta_{(x,i)}\mathbf{P}_{k\triangle},\delta_{(x,i)}\mathbf{P}^{\triangle}_{k\triangle}) < \frac{\epsilon}{2}$. Therefore the first desired assertion follows.

Furthermore, under the polynomial growth condition of f, by the similar way to [6], we can obtain that $W_p(\delta_{(x,i)}\mathbf{P}_{k\triangle},\delta_{(x,i)}\mathbf{P}^{\triangle}_{k\triangle}) \le Ce^{\nu k\triangle}\triangle^{p/2}$, for some positive constant ν. Let $\bar K$ be the integer part of constant $-p\ln\triangle/[2(\nu+\xi^*)\triangle]$, obviously, $\bar K\to\infty$ as $\triangle\to 0$. One observes that $e^{\nu\bar K\triangle}\triangle^{p/2} \le \triangle^{\frac{p\xi}{2(\nu+\xi^*)}}$, $\mathrm{e}^{-\xi\bar K\triangle} \le e^{\xi^*\triangle^*}\triangle^{\frac{p\xi}{2(\nu+\xi^*)}}$. Therefore, $W_p(\pi,\pi^{\triangle}) \le C\triangle^{\frac{p\xi}{2(\nu+\xi^*)}} =: C\triangle^{\gamma}$. □

REMARK 3.2. In Theorem 3.6 we not only give the convergence of invariant measures but also reveal the rate of the convergence is γ under the polynomial growth condition imposed on f. We also notice that Meyn and Tweedie's work [18] reveals the relationship of tightness, Harris recurrence and ergodicity for discrete-time Markov chains, they gave the generalization of Lyapunov-Foster criteria for the various ergodicity. However, these criteria are not applicable for (X_k,r_k) owing to the switching effects. Precisely, it is impossible from (3.14) to find a constant $0<\lambda\le 1$ such that $\mathbb{E}\big((1+|X_{k+1}|^2)^{\frac{p}{2}}|\mathcal{F}_{t_k}\big) \le \lambda(1+|X_k|^2)^{\frac{p}{2}}+C\triangle$ holds due to the changeable sign of $\beta_{r_k}+7\lambda/8$.

REMARK 3.3. By the virtue of Theorem 3.5, (X_k,r_k) is ergodic, moreover, the transition probability of (X_k,r_k) decays into its invariant measure exponentially under Wasserstein distance, see (3.36).

REMARK 3.4. Comparing with the convergence result of the EM scheme for SDE in [28], we release the restriction of the global Lipschitz continuity of the coefficients and deal with the convergence of invariant measures for nonlinear SDE with regime switching.

REMARK 3.5. Although many works pay attention to the approximation of invariant measures for SDEs, for example, [11, 16, 28], there are few works focusing on the approximation of invariant measures for switching diffusion processes, especially described by nonlinear systems. On the other hand, compared with the fast development of the finite-time numerical analysis for SPDEs, for examples, [9, 29], the results on long-time approximations for SPDEs are few. The methods developed in this paper provide ideas to deal with the invariant measure approximations for nonlinear SPDEs or SPDEs with regime switching. Owing to the importance this will be considered in our future work.

4. Examples. In this section, we consider two examples of nonlinear hybrid stochastic systems and provide simulations to illustrate the efficiency of the BEM method (3.1). We first consider a two-dimensional SDE with Markovian switching.

EXAMPLE 4.1. Consider (2.2) with $r(t)$ taking values in $\mathbb{S} = \{1, 2\}$ with generator $Q = \begin{pmatrix} -5 & 5 \\ 1 & -1 \end{pmatrix}$. The system is regarded as the Markovian switching between

$$\begin{cases} \mathrm{d}Y_1(t) = \big[2Y_1(t) - Y_1^3(t) - Y_1(t)Y_2^2(t)\big]\mathrm{d}t - 3\mathrm{d}B_1(t) + \mathrm{d}B_2(t), \\ \mathrm{d}Y_2(t) = \big[1 + Y_2(t) - Y_2^3(t) - Y_2(t)Y_1^2(t)\big]\mathrm{d}t + 4\mathrm{d}B_1(t), \end{cases} \tag{4.1}$$

and

$$\begin{cases} \mathrm{d}Y_1(t) & = \Big(Y_1(t) - 2Y_1(t)\sqrt{Y_1^2(t) + Y_2^2(t)} + 1\Big)\mathrm{d}t \\ & \quad + (2Y_1(t) - Y_2(t) + 2)\mathrm{d}B_1(t) + (Y_1(t) - Y_2(t))\mathrm{d}B_2(t), \\ \mathrm{d}Y_2(t) & = \Big(0.5Y_2(t) - 2Y_2(t)\sqrt{Y_1^2(t) + Y_2^2(t)} + 2\Big)\mathrm{d}t \\ & \quad + (Y_1(t) + 2Y_2(t))\mathrm{d}B_1(t) + (Y_1(t) + Y_2(t) - 4)\mathrm{d}B_2(t), \end{cases} \tag{4.2}$$

with the initial data $Y(0) = 1$, $r(0) = 1$, where $B(t) = (B_1(t), B_2(t))^T$ is a two-dimensional Brownian motion. Obviously, the diffusion coefficient g is global Lipschitz continuous with $h = 7$. Note that the drift coefficient f is neither the global Lipschitz continuous nor the linear growth, but we can derive that

$$(u - v)^T(f(u, 1) - f(v, 1)) \le 2|u - v|^2 - \frac{1}{4}(|u| - |v|)^4 \le 2|u - v|^2,$$

and

$$(u - v)^T(f(u, 2) - f(v, 2)) \le |u - v|^2 - 2(|u| + |v|)(|u| - |v|)^2 \le |u - v|^2,$$

i.e. Assumption 2.1 is satisfied with $\alpha_1 = 2$ and $\alpha_2 = 1$ for all $u, v \in \mathbb{R}^2$. We furthermore observe that

$$|u - v|^2|g(u, j) - g(v, j)|^2 - 2|(u - v)^T(g(u, j) - g(v, j))|^2 \le h_j|u - v|^4, \quad \forall j \in \mathbb{S},$$

holds with $h_1 = 0$ and $h_2 = -3$ for all $u, v \in \mathbb{R}^2$. Direct calculation leads to $\beta_1 = 2\alpha_1 + h_1 = 4$, $\beta_2 = 2\alpha_2 + h_2 = -1$. By solving the linear equation (2.1) we obtain the unique stationary distribution of $r(t)$, $\mu = (\mu_1, \mu_2) = (1/6, 5/6)$, then $\mu\beta = \mu_1\beta_1 + \mu_2\beta_2 = -1/6 < 0$. It follows from Theorem 2.2 that the exact solution $(Y(t), r(t))$ of (2.2) admits a unique invariant measure $\pi \in \mathcal{P}(\mathbb{R}^n \times \mathbb{S})$. By virtue of Theorems 3.5 and 3.6, for a given stepsize $\triangle$ the numerical solution of BEM scheme has a unique invariant measure $\pi^{\triangle} \in \mathcal{P}(\mathbb{R}^n \times \mathbb{S})$ approximating π in the Wasserstein metric. We apply the BEM scheme for numerical experiments. Since it is impossible to get the closed form of the solutions of the stochastic system with random switching between (4.1) and (4.2), we approximate the underlying solution by the numerical solution of BEM scheme (3.1). We regard the numerical solution with $\triangle = 2^{-17}$ as a more precise approximation comparing it with the numerical solution with stepsize $\triangle = 0.002$, see Figure 4.1.

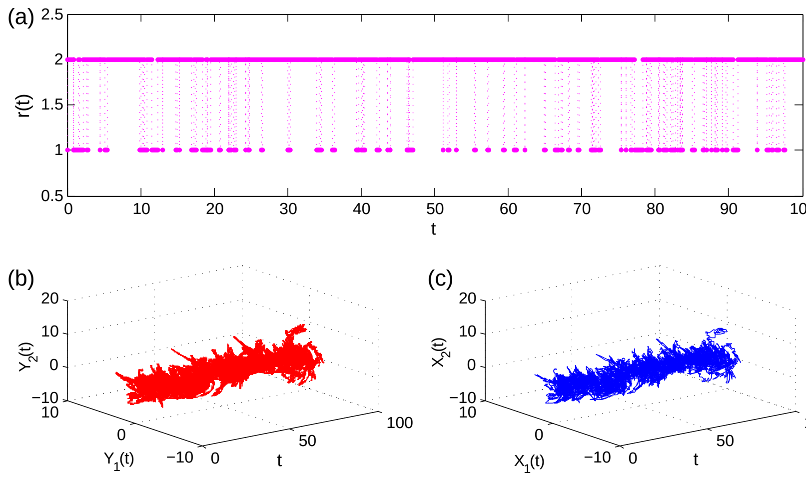

FIG. 4.1. **Example 4.1**. *(a) Computer simulation of a single path of Markov chain $r(t)$. (b) A sample path of exact solution $Y(t)$ in 3D settings. (c) A sample path of numerical solution $X(t)$ in 3D settings. The red line represents the exact solution (i.e. the BEM numerical solution with $\triangle = 2^{-17}$) while the blue line represents the BEM numerical solution with $\triangle = 0.002$.*

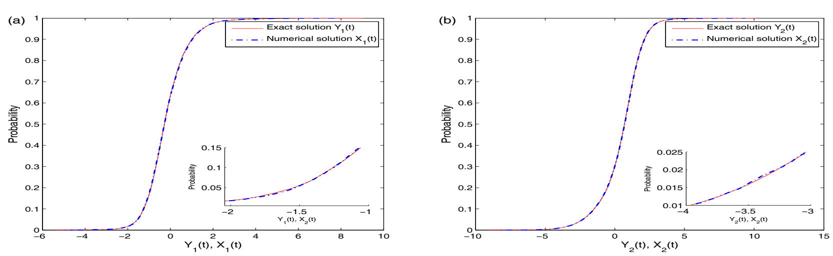

FIG. 4.2. **Example 4.1**. *(a) The ECDF for $Y_1(t)$. (b) The ECDF for $Y_2(t)$. The red solid line represents the exact solution of the switching system while the blue dashed line represents the numerical solution of the switching system.*

We simulate one path with 13107200 iterations and plot the empirical cumulative distribution function (ECDF) of numerical solution with $\triangle = 0.002$ in blue dashed line in Figure 4.2. The ECDF of exact solution is plotted on the same figure in a red solid line. The similarity of those two distributions is clearly seen, which indicates that the numerical stationary distribution is a good approximation to the theoretical one. To measure the similarity quantitatively, we use the Kolmogorov-Smirnov test [17] to test the alternative hypothesis that the numerical solution and exact solution are from different distributions against the null hypothesis that they are from the same distribution for both $Y_1(t)$ and $Y_2(t)$. With 3% significance level, the Kolmogorov-Smirnov test indicates that we cannot reject the null hypothesis. This example illustrates that numerical invariant measure converges to the underlying invariant measure.

In order to illustrate the validity, we consider the scalar hybrid cubic SDE (c.f. the stochastic Ginzburg-Laudau equation (4.52) in [10, p.125]) which drift coefficient isn't global Lipschitz continuous.

EXAMPLE 4.2. Let $r(t)$ be a Markov chain with the state space $\mathbb{S} = \{1, 2\}$ and the generator $Q = \begin{pmatrix} -q & q \\ 3 & -3 \end{pmatrix}$ for some $q > 0$. It is easy to see that its unique stationary distribution $\mu = (\mu_1, \mu_2) \in \mathbb{R}^{1\times 2}$ is given by $\mu_1 = \frac{3}{3+q}$, $\mu_2 = \frac{q}{3+q}$. Consider the scalar hybrid cubic SDE

$$\mathrm{d}Y(t) = (b(r(t))Y(t) + a(r(t))Y^3(t))\mathrm{d}t + \rho(r(t))Y(t)\mathrm{d}B(t), \tag{4.3}$$

with the initial data $Y(0) = 0.5$, $r(0) = 2$, where

$$b(1) = 1, \quad a(1) = -1, \quad \rho(1) = 2, \qquad b(2) = 2, \quad a(2) = -3, \quad \rho(2) = -1,$$

and $B(t)$ is a scalar Brownian motion. There exists a unique continuous solution $Y(t)$ to SDE (4.3) for any $Y(0) > 0$, which is global and represented by

$$Y(t) = \frac{0.5\exp\Big\{\int_0^t \Big[b(r(s)) - \frac{1}{2}\rho^2(r(s))\Big]\mathrm{d}s + \rho(r(s))\mathrm{d}B(s)\Big\}}{\sqrt{1 - 0.5\int_0^t a(r(s))\exp\Big\{\int_0^s \Big[2b(r(u)) - \rho^2(r(u))\Big]\mathrm{d}u + 2\rho(r(u))\mathrm{d}B(u)\Big\}\mathrm{d}s}}.$$

It is straightforward to see that $\alpha_1 = 1$, $\alpha_2 = 2$, $h_1 = -4$, and $h_2 = -1$. Direct calculation leads to $\beta_1 = -2$, $\beta_2 = 3$, then

$$\mu\beta = \mu_1\beta_1 + \mu_2\beta_2 < 0$$

holds with $q \in (0, 2)$. It follows from Theorem 2.2 that the exact solution $(Y(t), r(t))$ of (2.2) admits a unique invariant measure $\pi \in \mathcal{P}(\mathbb{R}^n \times \mathbb{S})$. By virtue of Theorems 3.5 and 3.6 the numerical solution of BEM scheme has a unique invariant measure $\pi^{\triangle} \in \mathcal{P}(\mathbb{R}^n \times \mathbb{S})$ approximating π in the Wasserstein metric.

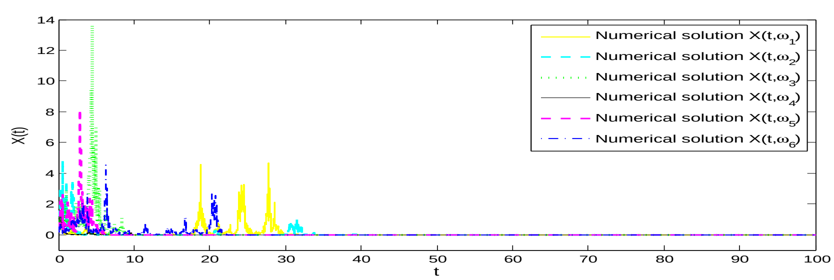

FIG. 4.3. **Example 4.2**. *Six trajectories of the BEM numerical solution with* 10^4 *iterations,* $Y(0) = 0.5$, $r(0) = 2$ *and stepsize* $\triangle = 0.001$.

We apply the BEM scheme to do numerical experiments. Choose $q = 1.5$ and stepsize $\triangle = 0.001$, we simulate 100 paths, each of which has 10^4 iterations. Figure 4.3 depicts six trajectories of the numerical solution of BEM scheme (3.1). Intuitively, some stationary behaviours display. Figure 4.4 (a) depicts the trajectory of the Markov chain. From this figure we find that the time the Markov chain staying on state 1 is more than on that of state 2. Figure 4.4(b) further depicts the trajectories of the exact solution $Y(t)$ and the corresponding BEM solution $X(t)$, and Figure 4.4(c) depicts the ECDFs of the exact solution and the BEM solution. The similarity of those two distributions is clear, which reveals that the numerical stationary distribution is a good approximation to the underlying one. Moreover, This example illustrates the existence of the stationary distribution as time goes to infinity. Thus instead of using numerous paths, we could just use few paths to picture the stationary distribution.

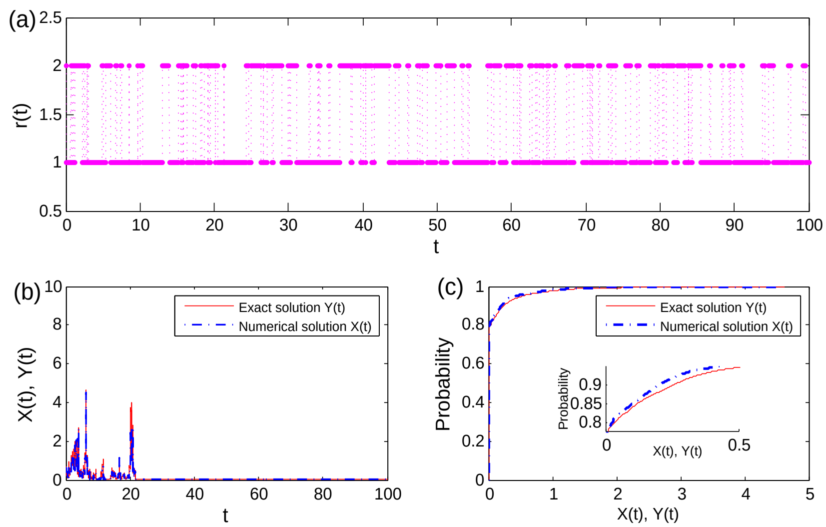

FIG. 4.4. **Example 4.2**. *(a) Computer simulation of a single path of Markov chain $r(t)$. (b) Sample paths of the exact solution and the BEM solution. (c) ECDFs for the exact solution and the BEM solution. The red solid line represents exact solution of the switching system while the blue dashed line represents the numerical solution.*

Acknowledgements. The authors would like to thank the associated editor and referee for their helpful comments and suggestions.

REFERENCES

[1] Anderson, W. J., *Continuous-Time Markov Chain*, Springer-Verlag, 1991.
[2] Athreya, K. B. and Pantula, S. G., *Mixing properties of Harris chains and autoregressive processes*, J. Appl. Prob., 23, 880-892, 1986.
[3] Bakhtin, Y. and Hurth, T., *Invariant densities for dynamical systems with random switching*, Nonlinearity, 25, 2937-2952, 2012.
[4] Bao, J., Shao, J. and Yuan, C., *Approximation of invariant measures for regime-switching diffusions*, Potential Anal., 44 (4), 707-727, 2016.
[5] Bardet, J. B., Guérin, H. and Malrieu, F., *Long time behavior of diffusion with Markov switching*, ALEA Lat. Am. J. Probab. Math. Stat., 7, 151-170, 2010.
[6] Higham D. J., Mao X. and Stuart A., *Strong convergence of Euler-type methods for nonlinear stochastic differential equations*, SIAM J. Numer. Anal. 40(3), 1041-1063, 2002.
[7] Higham, D. J., Mao, X. and Yuan, C., *Preserving exponential mean-square stability in the simulation of hybrid stochastic differential equations*, Numer. Math., 108, 295-325, 2007.
[8] Hutzenthaler M., Jentzen A. and Kloeden P., *Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 467, 1563-1576, 2011.
[9] Jentzen, A. and Kloeden, P. E., *Taylor approximations for stochastic partial differential equations*, CBMS-NSF Regional Conference Series in Applied Mathematics, 83. SIAM, Philadelphia, PA, 2011.
[10] Kloeden, P. E. and Platen, E., *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, New York, 1992.

[11] Liu, W. and Mao, X., *Numerical stationary distribution and its convergence for nonlinear stochastic differential equations*, J. Comput. Appl. Math., 276, 16-29, 2015.
[12] Mao, X., Yuan, C. and Yin, G., *Numerical method for stationary distribution of stochastic differential equations with Markovian switching*, J. Comput. Appl. Math., 174, 1-27, 2005.
[13] Mao, X. and Yuan, C., *Stochastic Differential Equations with Markovian Switching*, Imperial College Press, 2006.
[14] Mao, X., *Stochastic Differential Equations and Applications 2nd edition*, Horwood, Chichester, UK, 2007.
[15] Mao, X., Shen, Y. and Gray, A., *Almost sure exponential stability of backward Euler-Maruyama discretization for hybrid stochastic differential equation*, J. Comput. Appl. Math., 235, 1213-1226, 2011.
[16] Mattingly, J. C., Stuart, A. M., Higham, D. J., *Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise*, Stochastic Process. Appl., 101 (02), 185C232, 2002.
[17] Massey, Frank J., Jr., *The Kolmogorov-Smirnov test for goodness of fit*, J. Amer. Statist. Assoc., 46 (253), 68-78, 1951.
[18] Meyn, Sean P. and Tweedie, R. L., *Stability of Markovian processes. I. Criteria for discrete-time chains,* Adv. in Appl. Probab., 24(3), 542-574, 1992.
[19] Meyn, Sean P. and Tweedie, R. L., *Stability of Markovian processes. II. Continuous-time processes and sampled chains*, Adv. in Appl. Probab., 25(3), 487-517, 1993.
[20] Meyn, Sean P. and Tweedie, R. L., *Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes*, Adv. in Appl. Probab., 25(3), 518-548, 1993.
[21] Milstein G. N. and Tretyakov M. V., *Stochastic Numerics for Mathematical Physics*, Springer-Verlag, Berlin, 2004.
[22] Milstein G. N. and Tretyakov M. V., *Numerical integration of stochastic differential equations with nonglobally Lipschitz coefficients*, SIAM J. Numer. Anal. 43, 1139-1154, 2005.
[23] 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, 127-141, 2008.
[24] Pinsky, M. and Scheutzow, M., *Some remarks and examples concerning the transience and recurrence of random diffusions*, Ann. Inst. Henri. Poincaré, 28, 519-536, 1992.
[25] Rosenblatt, M. *Markov Processes, Structure and Asymptotic Behavior*, Springer-Verlag, New York, 1971.
[26] Shao, J. and Xi, F., *Strong ergodicity of the regime-switching diffusion processes*, Stoch. Proc. Appl., 123, 3903-3918, 2013.
[27] Shao, J., *Ergodicity of regime-switching diffusions in Wasserstein distances*, Stoch. Proc. Appl., 125, 739-758, 2015.
[28] Shardlow, T. and Stuart, A. M., *A perturbation theory for ergodic Markov chains and application to numerical approximations*, SIAM J. Numer. Anal., 37(4), 1120-1137, 2000.
[29] Shardlow, T.,*Stochastic perturbations of the Allen-Cahn equation*, Electron. J. Differential Equations, 47, 19 pp., 2000.
[30] Shiryaev, A. N., *Probability 2nd edition*, World Book Inc., 2004.
[31] Zhu, C. and Yin, G., *Asymptotic properties of hybrid diffusion systems*, SIAM J. Control Optim., 46, 1155-1179, 2007.
[32] Yin, G. and Zhu, C., *Hybrid Switching Diffusions: Properties and Applications*, Springer, 2010.
[33] Yuan, C. and Mao, X., *Stationary distributions of Euler-Maruyama-type stochastic difference equations with Markovian switching and their convergence*, J. Difference Equ. Appl., 11, 29-48, 2005.
[34] Zhou, S., *Strong convergence and stability of backward Euler Maruyama scheme for highly nonlinear hybrid stochastic differential delay equation*, Calcolo, 52, 445-473, 2015.
[35] Zong, X., Wu, F. and Huang, C., *The moment exponential stability criterion of nonlinear hybrid stochastic differential equations and its discrete approximations*, Proc. Roy. Soc. Edinburgh Sect. A, 146 (6), 1303-1328, 2016.