Available online at www.sciencedirect.com

ScienceDirect

Journal of Differential Equations 441 (2025) 113472

Journal of Differential Equations

www.elsevier.com/locate/jde

Ergodic semi-implicit approximations to periodic measures of stochastic differential equations with locally Lipschitz drifts—Error analysis in Wasserstein distance

Chunrong Feng [a,c], Yu Liu [b], Yujia Liu [b], Huaizhong Zhao [a,c,*]

[a] *Department of Mathematical Sciences, Durham University, DH1 3LE, UK*
[b] *School of Mathematics, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China*
[c] *Research Centre for Mathematics and Interdisciplinary Sciences, Shandong University, Qingdao, Shandong, 266237, China*

Received 26 November 2023; revised 25 March 2025; accepted 22 May 2025

Abstract

We study numerical approximations to the periodic measures of time-periodic stochastic differential equations. For those systems with locally Lipschitz coefficients, while the explicit Euler-Maruyama scheme does not work, we carry out semi-implicit Euler-Maruyama schemes to compute their periodic measures. We prove the local Doeblin condition for the numerical schemes uniformly with respect to discretization step size. This, together with a Lyapunov function argument due to the weakly dissipative condition, leads to the existence and uniqueness of periodic measures of numerical schemes, and geometric ergodicity with the convergence being independent of the step size in the discretization. The novelty of our approach is that without knowing any a priori information about periodic measure of the original problem, even the existence, we can prove its existence and ergodicity from that of periodic measure of the discretized numerical scheme.

MSC: 37H99; 60H10; 60H35

* Corresponding author.
E-mail addresses: chunrong.feng@durham.ac.uk (C. Feng), yu.liu@cumt.edu.cn (Y. Liu), yujia.liu@cumt.edu.cn (Y. Liu), huaizhong.zhao@durham.ac.uk (H. Zhao).

https://doi.org/10.1016/j.jde.2025.113472

Keywords: Periodic measure; Semi-implicit Euler-Maruyama scheme; Discretized semi-flows; Geometric ergodicity; Wasserstein distance

1. Introduction

Random periodicity attracts increasing attention since the works of [60] and [22]. It can describe many natural phenomena consisting of both periodicity and uncertainty. Examples include seasonal climate variations and the evolution of glacial periods, where alternations between the two climates that occurred approximately every 50,000 years were observed in the early last century; human activities including but not limited to agricultural and industrial productions, which are highly affected by periodic patterns. Some stochastic differential equations with time-periodic coefficients, such as the stochastic resonance model for climate change and a temperature model in the study of the weather derivative were proposed, respectively ([4], [5]). This work is devoted to developing their numerical analysis as one of the key tools in this analysis underpinning this kind of mathematical model.

The study of random periodicity is carried out in two different, but highly related indispensable ways, namely random periodic solutions and periodic measures, which give the mathematical definitions of random periodicity in the pathwise sense and in the sense of distributions, respectively ([19]). Periodicity in the sense of distribution was also studied in [12], [32]. The recent works ([21], [20]) discussed the existence of a periodic measure and its geometric ergodicity. In [19], the ergodic theory of random periodic processes started to emerge. The "equivalence" of the random periodic solutions and periodic measures and their characterization in terms of purely imaginary eigenvalues of the infinitesimal generator of the Markovian semigroup were obtained. The presence of pure imaginary eigenvalues distinguishes random periodic processes/periodic measures regime from stationary processes/mixing invariant measures. In the latter case, the Koopman-von Neumann Theorem says the infinitesimal generator has a unique eigenvalue 0 on the imaginary axis, and the eigenvalue is simple.

Our concept of random periodicity is the random counterpart of periodicity in the theory of dynamical systems. It has contributed to many important works in studying random periodicity in a variety of different topics, including bifurcations ([55]), random attractors ([3]), stochastic resonance ([10], [21], [20], [16]), random horseshoes ([30]), modelling El Niño phenomenon ([8]), stochastic oscillations ([13]), linear response and homogenizations ([6], [54]), large deviations principle ([24]), synchronization ([14]), random almost periodic solutions ([7], [46]), random periodic solutions of certain functional differential equations ([23]), certain stochastic differential equations and stochastic partial differential equations ([35], [49], [2]), invariant measures of quasi-periodic stochastic systems ([18], [35]).

However, it is difficult to give explicit formulae for periodic measures in many concrete models, so numerical approximation becomes critical in the study of stochastic dynamics. There are abundant important works on numerical analysis of SDEs on a finite horizon ([28], [31], [33], [41], [42]). A numerical analysis on approximation to invariant measures of SDEs by discretizing the pull-back flow was given in [38], [51], [52], [53], [57], [58]. Despite importance of random periodicity both at the theoretical and applied level, their numerical analysis has barely been investigated until very recently. Since its publication, the work [15] has prompted many recent works in this direction especially those in the last two years including [56], [43], [9], [59], [61], [62]. However, the current work is still the only work so far in the study on ergodic theory of numerical schemes and their approximation of periodic measures of SDEs with polynomial growth and weakly dissipative drift.

In this paper, we consider the following nonautonomous stochastic differential equations

$$dX(t) = b(t, X(t))dt + \sigma(t, X(t))dW_t, \quad t \geq s, \tag{1.1}$$

with initial condition $X(s) = x$, where $b : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$, W_t is a two-sided Wiener process in $\mathbb{R}^d$ on the Wiener probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We assume that b is τ-periodic in the time variable and locally Lipschitz weakly dissipative in the space variable, σ is τ-periodic in the time variable and Lipschitz in the space variable and nondegenerate. This equation has a unique solution, denoted by $X^{s,x}(t)$, $t \geq s$, throughout the paper.

Denote by $(\Omega, \mathcal{F}, \mathbb{P}, (\theta_s)_{s\in\mathbb{R}})$ the metric dynamical system associated with the canonical probability space $(\Omega, \mathcal{F}, \mathbb{P})$ for Brownian motion W in $\mathbb{R}^d$. Here $\theta_s : \Omega \to \Omega$ is defined by $(\theta_s\omega)(t) = W(t + s) - W(s)$ for all $s \in \mathbb{R}$ and it is a measurable map. It is measurably invertible, as $\theta^{-1}(s) = \theta(-s)$ exists for all $s \in \mathbb{R}$, so the inverse is also measurable. Denote $\mathcal{F}_t = \sigma\{W_s : -\infty < s \leq t\}$, $\Delta := \{(t, s) \in \mathbb{R}^2 : s \leq t\}$ and $u : \Delta \times \Omega \times \mathbb{R}^d \to \mathbb{R}^d$ as a periodic stochastic semi-flow of period τ if for all $(t, s) \in \Delta$ and $r \in [s, t]$

$$u(t, r, \omega) \circ u(r, s, \omega) = u(t, s, \omega), \tag{1.2}$$

and

$$u(t + \tau, s + \tau, \omega) = u(t, s, \theta_\tau\omega), \tag{1.3}$$

for almost all $\omega \in \Omega$. In the case of SDE (1.1), $u(t, s, \omega)x = X^{s,x}(t, \omega)$ defines a periodic semi-flow of period τ.

The work of Meyn and Tweedie [40] gave a criterion for the existence of invariant measures of stationary Markovian processes under a Lyapunov condition and the local Doeblin condition. More relevant interpretation for stochastic differential equations of this abstract and general ergodicity framework was provided in the work [38]. In [21], the authors gave a method to obtain the geometric ergodicity of periodic measures for periodic stochastic semi-flows. It was also proved that the periodic measure ρ_s exists and has a density function. In this paper, we study the numerical approximation of the periodic measure.

It is well known that the forward Euler-Maruyama scheme requires a global Lipschitz condition on the drift and diffusion coefficients ([38]). The analysis of periodic measures suffers the same drawback ([16]). However, an implicit Euler-type scheme supplies a tool to study many important SDE models with locally Lipschitz property on their drift terms. Semi-implicit Euler-Maruyama schemes on autonomous SDEs were studied in [28], [29], [33], [38]. Some more recent works ([1], [11], [36], [45], [47]) studied infinite-horizon problems on numerical approximations to invariant measures. In [48], the authors gave a perturbation theory for stationary Markov chains and invariant measures. Helped by this, they showed that the geometric ergodic property provides a way to obtain a uniform numerical error for the infinite-horizon problem.

Without a global Lipschitz assumption on the drift term, we apply some semi-implicit numerical schemes such as the backward Euler-Maruyama method and the split-step Euler-Maruyama method to approximate nonlinear random periodic flows. For any fixed initial time s and initial position x, we denote by $\{\hat{X}_n\}_{n=0,1,\ldots}$, the discrete approximation of the solution of (1.1) with step size $\Delta t = \tau/K$ and $\hat{X}_0 = x$, i.e. $\hat{X}_n := \hat{X}^{s,x}(s + n\Delta t)$. We prove that the discrete semi-flow has a periodic measure $\hat{\rho}_s^{\Delta t}$, $s \in \mathbb{R}$ and the transition probability converges to the periodic measure in the Wasserstein distance W_1 uniformly with respect to the step size Δt (Theorem 2.7). The uniformity is the main result of the first part of this paper. To prove this result, the local Doeblin condition and the Lyapunov function condition in one time-step Δt, in which the parameters depend on Δt, are not enough. Instead, we need to consider these two conditions for $\frac{\tau}{\Delta t}$ discrete steps corresponding to a time interval of length of period τ together and obtain estimates that are uniform in Δt ((2.13), (2.18)).

The other main result is Theorem 5.3 which proves that the cumulation of discretization errors in the Wasserstein distance W_1 is of the order of $\mathcal{O}(\Delta t)^{\frac{\eta}{2}}$ for the approximation of periodic measure for some $\eta \in (0, 1)$ i.e. for any $s \in [0, \tau)$,

$$W_1(\rho_s, \hat{\rho}_s^{\Delta t}) = \sup_{\phi \in Lip(1)} \left| \int_{\mathbb{R}^d} \phi(x)\rho_s(dx) - \int_{\mathbb{R}^d} \phi(x)\hat{\rho}_s^{\Delta t}(dx) \right| \leq C(\Delta t)^{\eta/2}, \tag{1.4}$$

where $C > 0$ is a constant independent of Δt. With the help of the perturbation theorem, the ergodicity of the approximating numerical scheme and the convergence to its periodic measure uniformly in the step size Δt., we can obtain (1.4) if the true periodic measure ρ_s exists (Proposition 5.1). The novelty of our approach is that without knowing any a priori information about $\rho_s, s \in \mathbb{R}$, even its existence, we can prove the existence and ergodicity of $\rho_s, s \in \mathbb{R}$ from that of periodic measure $\hat{\rho}_s^{\Delta t}, s \in \mathbb{R}$ of the discretized system (Theorem 5.3). In addition to the perturbation result, the explicit form of the d_β distance, the use of the Wasserstein distance and the estimate of transition probabilities of the true and approximating system (2.12) make it possible.

In the case where the numerical error does not accumulate with respect to the time duration of the random periodic flows, we have the coefficient $\eta = 1$, and thus the error is of order $1/2$. This is the case, for example, when the drift is strongly dissipative and Lipschitz. Under these stronger conditions, our result agrees with the previous result that the infinite horizon problem will possess the same order of numerical error as that of the finite time scheme ([15], [56]). Needless to say that the results in this paper go beyond these previous results without requiring global Lipschitz and strong dissipative conditions.

The results in this paper are applicable to many physically relevant SDEs, for instance, Benzi-Parisi-Sutera-Vulpiani's stochastic resonance model (BPSV model) for the ice-age transition in the climate change dynamics is given by SDE (1.1), with $b(t, x) = x - x^3 + A\cos(Bt)$ ([5]). This model was proved to have a unique periodic measure ([21]). This result implies the presence of transitions between the ice-age and the interglacial climates. A partial differential equation for expected transition time provides a method for studying the transition time in the stochastic resonance problem ([20]). The ergodic numerical analysis of the Euler scheme to periodic measures was given in [16] for SDEs with weakly dissipative and linear growth drifts, so that result was only applicable to a modified BPSV model. Compared with the result of [16], the implicit schemes release the global Lipschitz condition on the drift term and are applicable to the true BPSV model.

The rest of this paper will be organized as follows. In Section 2, we prove that the semi-implicit numerical schemes are ergodic and have discrete periodic measures which contains the main theorem of the first part of the paper (Theorem 2.7). In Section 3, we use a modified SDE (3.4) to analyze the approximation error for a finite horizon. In Section 4, we give the a priori estimates for the numerical schemes, original SDEs and an auxiliary SDE. These two sections provide the necessary preparations for the use of a perturbation theorem. In Section 5, to consider the infinite-horizon problem, we prove a perturbation theory on the periodic measures. This leads us to obtain the order of the numerical error of the approximate periodic measure to the true periodic measure and the main result of the second part of the paper Theorem 5.3. In Section 6, a numerical experiment on the BPSV model is carried out as a verification of the theoretical results proven in this paper.

2. Geometric ergodic periodic measures of SDEs and their semi-implicit discretizations

2.1. Assumptions and numerical schemes

A solution of stochastic differential equations (1.1) with coefficients being periodic in time with period τ, when it exists and is unique, generates a periodic semi-flow by setting $u(t,s)x = X^{s,x}(t)$. Moreover, for any $\Gamma \in \mathcal{B}(\mathbb{R}^d)$, $t, s \in \mathbb{R}$, $(t,s) \in \Delta$ denote the transition probability of u by $P(t,s,x,\Gamma) = \mathbb{P}(\{\omega : u(t,s,\omega)x \in \Gamma\})$. Then it satisfies the periodicity property

$$P(t+\tau, s+\tau, x, \cdot) = P(t,s,x,\cdot), \quad (t,s) \in \Delta, \tag{2.1}$$

and the measure preserving property of θ_τ. Define for $\phi \in \mathcal{B}_b(\mathbb{R}^d)$,

$$P_{t,s}\phi(x) := \mathbb{E}\phi(X^{s,x}(t)) = \int_{\mathbb{R}^d} \phi(y) P(t,s,x,dy),\ t \geq 0,$$

where $\mathcal{B}_b(\mathbb{R}^d)$ is the space of bounded and Borel measurable function from $\mathbb{R}^d$ to $\mathbb{R}$. Then it is well known that $P_{t,s} : \mathcal{B}_b(\mathbb{R}^d) \to \mathcal{B}_b(\mathbb{R}^d)$ defines a semigroup and satisfies the τ-periodic property:

$$P_{t+\tau,s+\tau} = P_{t,s},$$

and the semigroup property

$$P_{r,s} \circ P_{t,r} = P_{t,s},$$

for all $s \leq r \leq t$. The definition of periodic measure of the periodic Markovian semigroup is given below. Let $\mathcal{P}(\mathbb{R}^d)$ denote all probability measures on $\mathbb{R}^d$. The measure-valued function $\rho : \mathbb{R} \to \mathcal{P}(\mathbb{R}^d)$ is called a τ-periodic measure of the τ-periodic Markovian semigroup P if

$$P^*_{t,s}\rho_s(\cdot) := \int_{\mathbb{R}^d} P(t,s,x,\cdot)\rho_s(dx) = \rho_t(\cdot),\ \rho_{s+\tau} = \rho_s, \tag{2.2}$$

for all $s \in \mathbb{R}$, $t \in \mathbb{R}^+$. Such kind of ρ was also called $\tau-$periodic evolution system of measures ([12], [34], [37]). We use the name periodic measure for simplicity. In this paper, we use $|\cdot|$ to denote the Euclidean norm in $\mathbb{R}^d$ and use $\|\cdot\|$ to denote the matrix norm.

Condition (A1) *Functions b and σ are locally Lipschitz with respect to the spatial variables, $\frac{1}{2}$-Hölder continuous and τ-periodic with respect to the time variable, and σ is bounded.*

Condition (A2) *(Non-degeneracy) There exists a positive constant c such that for any $t \in \mathbb{R}$ and $x \in \mathbb{R}^d$, we have $\sum_{i,j} a_{ij}(t,x)x_i x_j \ge c|x|^2$, where $a = \sigma\sigma^*$.*

Condition (A3) *(Weak dissipativity) There exist constants $\beta > 0$ and $M > 0$ such that for any $t \in \mathbb{R}$ and $x \in \mathbb{R}^d$, $\langle x, b(t,x)\rangle \le -\beta|x|^2 + M$.*

If all (A1), (A2) and (A3) are satisfied, we call Condition (A) holds. Under this condition, it was proved in [21] that a periodic measure $\rho : (-\infty, +\infty) \to \mathcal{P}(\mathbb{R}^d)$ exists and is geometrically ergodic: there exist $C > 0$ and $\delta > 0$ such that

$$\|P(n\tau + s, s, x) - \rho_s\|_{TV} \le Ce^{-\delta n\tau}. \tag{2.3}$$

To approximate the periodic measure, we consider a semi-implicit scheme with step size $\Delta t = \tau/K > 0$, for some $K \in \mathbb{N}$, for SDE (1.1). In the following, to simplify notation especially in the numerical schemes, we denote $\hat{X}_i = \hat{X}^{s,x}(s + i\Delta t)$, $\hat{Z}_i := \hat{Z}^{s,x}(s + i\Delta t)$. Sometimes to simplify notation, we also set that $b^s(i\Delta t, x) := b(s + i\Delta t, x)$, $\sigma^s(i\Delta t, x) := \sigma(s + i\Delta t, x)$ and $\Delta^s W_i = W_{s+(i+1)\Delta t} - W_{s+i\Delta t}$. This reparameterization is very convenient to signify the starting time s and to make it as if from the time 0. We denote the exact solution as $X_t := X^{s,x}(s + t)$. The equation (1.1) is rewritten as

$$dX_t = b^s(t, X_t)dt + \sigma^s(t, X_t)d\tilde{W}_t, \quad t \ge 0, \tag{2.4}$$

where $\tilde{W}_t = W_{t+s} - W_s = (\theta_s\omega)(t)$.

We consider the split-step Euler-Maruyama scheme (SSEM) $\{\hat{X}_i\}_{i\in\mathbb{N}}$ given by

$$\begin{cases} \hat{X}_i^* = \hat{X}_i + b^s(i\Delta t, \hat{X}_i^*)\Delta t, \\ \hat{X}_{i+1} = \hat{X}_i^* + \sigma^s(i\Delta t, \hat{X}_i^*)\Delta^s W_i, \end{cases} \tag{2.5}$$

with initial condition $\hat{X}_0 = x$ and the backward Euler-Maruyama scheme (BEM) $\{\hat{Z}_i\}_{i\in\mathbb{N}}$ given by

$$\hat{Z}_{i+1} = \hat{Z}_i + b^s((i+1)\Delta t, \hat{Z}_{i+1})\Delta t + \sigma^s(i\Delta t, \hat{Z}_i)\Delta^s W_i, \tag{2.6}$$

with initial condition $\hat{Z}_0 = x$. The fully-implicit scheme involves an unbounded random variable, which requires a more stringent assumption on the model. Therefore, we consider schemes where the diffusion terms are in explicit form. This is why they are called semi-implicit schemes. In order to emphasize the dependence of $\hat{X}_i$ and $\hat{Z}_i$ on Δt, we may denote them by $\hat{X}_i^{\Delta t}$ and $\hat{Z}_i^{\Delta t}$. When no confusions occur, we still use the notation $\hat{X}_i$ and $\hat{Z}_i$ for simplicity. One can find more details for the semi-implicit schemes of autonomous cases in [33], [29], [50], [28].

Remark 2.1. The two schemes are equivalent. If $\hat{Z}_i$ satisfies (2.6), then $\hat{X}_i := \hat{Z}_i - b^s(i\Delta t, \hat{Z}_i)\Delta t$ solves (2.5) where $\hat{Z}_i$ coincides with $\hat{X}_i^*$, and vice versa.

There are many ways to generate the stochastic increment $\Delta^s W_i := W_{s+(i+1)\Delta t} - W_{s+i\Delta t}$. In this paper, we apply scaling Gaussian random variables in the approximation of numerical schemes i.e. $\Delta^s W_i = \sqrt{\Delta t}\mathcal{N}(0, 1)$. To apply numerical approximation, we give an additional assumption on the SDE.

Condition (B1) *The functions b and σ are C^1, and there exists $L_b > 0$ such that for any $t \in \mathbb{R}$ and $x, y \in \mathbb{R}^d$, we have $\langle x-y, b(t,x)-b(t,y)\rangle \le L_b |x-y|^2$.*

Function σ is globally Lipschitz with respect to the spatial variables, i.e. there exists a constant $L_\sigma > 0$ such that for any $t \in \mathbb{R}$ and $x, y \in \mathbb{R}^d$,

$$\|\sigma(t,x)-\sigma(t,y)\| \le L_\sigma |x-y| .$$

Here $\|\sigma\|$ is the matrix norm $\|\sigma\| = \sqrt{\mathrm{trace}(\sigma\sigma^*)}$ for a $\mathbb{R}^{d\times d}$ matrix σ.

Condition (B2) *The initial condition $X(s) = \xi$ is finite in any order of moment, i.e. for any integer $p \ge 2$, $\mathbb{E}|\xi|^p < \infty$.*

If both Conditions (B1) and (B2) are satisfied, we say Condition (B) holds. Sometimes we also need the following Condition.

Condition (C) *There exist $C > 0$ and positive integer q such that for all $x, y \in \mathbb{R}^d$ and $t \in \mathbb{R}$,*

$$|b(t,x)-b(t,y)|^2 \le C\left(1+|x|^{2q}+|y|^{2q}\right)|x-y|^2 .$$

Remark 2.2. For any fixed t, Δt, denote $G^{t,\Delta t}(x) = x - b(t,x)\Delta t + c$ with some given constant c. From Condition (B1), for any $t \in [0,\tau)$ and $x, y \in \mathbb{R}^d$, we have

$$\left\langle x-y, G^{t,\Delta t}(x)-G^{t,\Delta t}(y)\right\rangle = |x-y|^2 - \Delta t\,\langle x-y, b(t,x)-b(t,y)\rangle \ge (1-L_b\Delta t)\,|x-y|^2 .$$

We choose a step size $\Delta t < 1/L_b$ to satisfy the condition of the following theorem.

Theorem 2.3 *(Uniform Monotonicity Theorem [44], [50]). Suppose $G : \mathbb{R}^d \to \mathbb{R}^d$ is continuous and there exists $c > 0$ such that for any $x, y \in \mathbb{R}^d$,*

$$\langle x-y, G(x)-G(y)\rangle \ge c\,|x-y|^2 ,$$

then G is one-to-one and onto. Furthermore, G^{-1} is Lipschitz continuous and for any $x, y \in \mathbb{R}^d$,

$$\left|G^{-1}(x)-G^{-1}(y)\right| \le c^{-1}|x-y| .$$

Remark 2.4. Under Condition (B1), from Remark 2.2, we know that the map $G(x) = x - b(t,x)\Delta t$ is one-to-one when $\Delta t < 1/L_b$. Thus, G is invertible, and the first equation of (2.5) has a unique solution. As a direct consequence of Theorem 2.3 and Remark 2.1, the numerical schemes (2.5) and (2.6) are well-defined.

2.2. Discrete periodic measures

Take $\Delta t = \tau/K$ with an integer $K > 0$ and consider the discrete counterpart of periodic stochastic semi-flow with period K. Denote the discrete Wiener shift $\hat\theta : \Omega \to \Omega$ by $\hat\theta = \theta_{\Delta t}$, then $\hat\theta^K = \theta_\tau$. Denote $\hat\Delta := \{(i,j) \in \mathbb{Z}^2 : j \le i\}$. For a fixed parameter $s \in [0,\tau)$, define $\hat u^s : \hat\Delta \times \Omega \times \mathbb{R}^d \to \mathbb{R}^d$ as a stochastic semi-flow satisfying for all $(i,j) \in \hat\Delta$ and $j \le l \le i$,

$$\hat u^s_{i,l} \circ \hat u^s_{l,j} = \hat u^s_{i,j}, \quad a.s., \tag{2.7}$$

and the random periodicity condition

$$\hat{u}^s_{i+K,j+K}(\omega)=\hat{u}^s_{i,j}(\hat{\theta}^K\omega)=\hat{u}^s_{i,j}(\theta_\tau\omega),\quad \text{for almost all } \omega\in\Omega. \tag{2.8}$$

The numerical scheme (2.5) possesses a discrete semi-flow defined as

$$\hat{u}^s_{i,j}(\omega)x=\hat{X}_{i-j}=\hat{X}^{s+j\Delta t,x}(s+i\Delta t)(\omega),$$

where $(i,j)\in\hat{\Delta}$. Here $\hat{X}_0=\hat{X}^{s+j\Delta t,x}(s+j\Delta t)=x$ as initial condition of the scheme at time $s+j\Delta t$. The transition probability is

$$\hat{P}^s_{i,j}(x,\Gamma)=\hat{P}^s_{i,j}I_\Gamma(x):=\mathbb{P}\{\hat{u}^s_{i,j}(\omega)x\in\Gamma\},\quad \Gamma\in\mathcal{B}(\mathbb{R}^d),$$

the corresponding semigroup $\hat{P}^s_{\cdot,\cdot}$ is given by

$$\hat{P}^s_{i,j}\phi(x)=\int_{\mathbb{R}^d}\hat{P}^s_{i,j}(x,dy)\phi(y),\quad (i,j)\in\hat{\Delta},$$

for any $\phi\in\mathcal{B}_b(\mathbb{R}^d)$. From the Markovian property and (2.7), it is easy to see the semigroup property of $\hat{P}^s$, i.e. $\hat{P}^s_{l,j}\circ\hat{P}^s_{i,l}=\hat{P}^s_{i,j}$ for all $(i,j)\in\hat{\Delta}$ and $j\le l\le i$, and from (2.8) and the measure preserving property of θ_τ, it follows that $\hat{P}^s_{i+K,j+K}=\hat{P}^s_{i,j}$, for all $(i,j)\in\hat{\Delta}$. We call a measure-valued function $\hat{\rho}^s_\cdot:\mathbb{Z}\to\mathcal{P}(\mathbb{R}^d)$ a periodic measure of the semigroup $\hat{P}^s_{i,j}$ with period K if

$$\int_{\mathbb{R}^d}\hat{P}^s_{i,j}(x,\Gamma)\hat{\rho}^s_j(dx)=\hat{\rho}^s_i(\Gamma),\quad (i,j)\in\hat{\Delta},$$

and

$$\hat{\rho}^s_{i+K}=\hat{\rho}^s_i,\quad i\in\mathbb{Z}.$$

To consider the existence of discrete periodic measures under a Wasserstein distance, we apply the Harris small set and Lyapunov function method ([27], [40]). This method was refined in Theorem 1.3 in [26] where the authors gave the measure contraction result under a weighted total variation distance on $\mathcal{P}(\mathbb{R}^d)$,

$$d_\beta(\mu_1,\mu_2)=\int_{\mathbb{R}^d}(1+\beta V(x))\,|\mu_1-\mu_2|\,(dx),$$

where $\beta>0$. It was known that d_β is complete for the space of probability measures $\mathcal{P}(\mathbb{R}^d)$. In the following, let $P:\mathcal{B}_b(\mathbb{R}^d)\to\mathcal{B}_b(\mathbb{R}^d)$ be a linear map and $P^*:\mathcal{P}(\mathbb{R}^d)\to\mathcal{P}(\mathbb{R}^d)$ be its conjugate. Denote $P(x,\cdot):=PI_\cdot(x)=P^*\delta_x(\cdot)$.

Lemma 2.5. *[26] Assume there exist a measurable function $V : \mathbb{R}^d \to [0, \infty)$ and constants $C \geq 0$ and $\gamma \in (0, 1)$ such that*

$$(PV)(x) \leq \gamma V(x) + C, \quad x \in \mathbb{R}^d, \tag{2.9}$$

and there exist a constant $\zeta \in (0, 1)$ and $\nu \in \mathcal{P}(\mathbb{R}^d)$ such that

$$\inf_{x \in \mathcal{C}} P(x, \cdot) \geq \zeta \nu(\cdot), \tag{2.10}$$

with $\mathcal{C} = \{x \in \mathbb{R}^d : V(x) \leq R\}$ for some $R > \frac{2C}{1-\gamma}$. Then for any $\mu_1, \mu_2 \in \mathcal{P}(\mathbb{R}^d)$

$$d_\beta(P^*\mu_1, P^*\mu_2) \leq \alpha d_\beta(\mu_1, \mu_2),$$

where $\alpha = \max\{1 - \frac{\zeta}{2}, \frac{2+\beta(R\gamma+2C)}{2+\beta R}\} < 1$ and $\beta = \frac{\zeta}{2C}$.

Remark 2.6. (i) When $\beta = \frac{\zeta}{2C}$, $R > \frac{2C}{1-\gamma}$,

$$\frac{2 + \beta(R\gamma + 2C)}{2 + \beta R} = 1 - \frac{\beta(R(1-\gamma) - 2C)}{2 + \beta R} < 1.$$

(ii) A function V satisfying (2.9) is called a Lyapunov function, and the second condition of Lemma 2.5, i.e. (2.10), is called the local Doeblin condition.

In the following two theorems, we give the existence of discrete periodic measures and the geometric contraction of numerical approximations (SSEM and BEM schemes). Note the convergence of $\hat{P}^s_{i,-kK+i}$ to $\hat{\rho}^s_i$ as $k \to \infty$ is proved to be uniform in Δt for $\Delta t < \frac{1}{L_b}$.

Theorem 2.7. *Assume Conditions (A), (B) and (C). If $\Delta t < 1/L_b$, then the split-step Euler-Maruyama scheme (2.5) is geometrically ergodic, i.e. for any fixed $s \in \mathbb{R}$, there exists a periodic measure $\hat{\rho}^s_\cdot : \mathbb{Z} \to \mathcal{P}(\mathbb{R}^d)$ such that*

$$d_\beta\left(\hat{P}^s_{i,-kK+i}(x, \cdot), \hat{\rho}^s_i\right) \leq C(1 + |x|^2)e^{-\delta(k\tau + i\Delta t)}, \quad k \in \mathbb{N}, \tag{2.11}$$

for some constants $C, \delta > 0$ and any $i \in \mathbb{Z}$.

Proof. Firstly, we will verify the local Doeblin condition hold for the discrete transition semigroup generated by the numerical approximation by scheme SSEM or BEM. We will mainly use Theorem 6.2 in [38]. It is not difficult to generalize this theorem to the semigroup $P_{t,s}$ as the Chapman-Kolmogorov equality still holds. Here, we have $K\Delta t = \tau$ is fixed for all Δt sufficiently small. For any open set $\mathcal{O}$ and compact set $\mathcal{C}$, define for any $\delta > 0$,

$$\mathcal{O}^c_\delta := \{x : \mathrm{dist}(x, \mathcal{O}^c) < \delta\}.$$

Then for any $x \in \mathcal{C}$,

$$\begin{aligned}\mathbb{P}(X^{s,x}(s+\tau)\in\mathcal{O}) =&\mathbb{P}(X_\tau\in\mathcal{O})\\ =&\mathbb{P}(X_\tau-\hat{X}_K+\hat{X}_K\in\mathcal{O})\\ \geq&\mathbb{P}(\hat{X}_K\in\mathcal{O}_\delta, |X_\tau-\hat{X}_K|<\delta)\\ \geq&\mathbb{P}(\hat{X}_K\in\mathcal{O}_\delta)-\mathbb{P}(|X_\tau-\hat{X}_K|\geq\delta).\end{aligned}$$

So, by Theorem 4.5, (note: the proof of Theorem 4.5 does not use any argument of this theorem)

$$\begin{aligned}\mathbb{P}(\hat{X}_K\in\mathcal{O})-\mathbb{P}(X_\tau\in\mathcal{O}) \leq&\mathbb{P}(\hat{X}_K\in\mathcal{O}\backslash\mathcal{O}_\delta)+\mathbb{P}(|X_\tau-\hat{X}_K|\geq\delta)\\ \leq&\mathbb{P}(\hat{X}_K\in\mathcal{O}\backslash\mathcal{O}_\delta)+\frac{1}{\delta^2}\mathbb{E}[|X_\tau-\hat{X}_K|^2]\\ \leq&\mathbb{P}(\hat{X}_K\in\mathcal{O}\backslash\mathcal{O}_\delta)+\frac{1}{\delta^2}Ce^{\lambda\tau}(\Delta t)^2\\ \to&0,\end{aligned}$$

as $\Delta t\to 0$ first and then $\delta\to 0$. This verifies the first condition in Condition 6.1 of [38], i.e.

$$\sup_{x\in\mathcal{C}}\left|\hat{P}^s_{K,0}(x,\mathcal{O})-P(s+\tau,s,x,\mathcal{O})\right|\xrightarrow{\Delta t\to 0}0. \tag{2.12}$$

On the other hand, as the noise term is assumed to be non-degenerate, so it is easy to verify the density $p(t,s,x,y)$ and $\hat{p}(s+\Delta t,s,x,y)$ exist, and $p(t,s,x,y)$ is jointly continuous in $(t,s,x,y)\in\Delta\times\mathcal{C}\times\mathcal{C}$. Denote k_0 be an integer such that $k_0\Delta t\geq\frac{\tau}{3}$. By Theorem 1.2 in [39], we can get for $k_1,k_2\geq k_0$ and $k_1+k_2=K$, the density $p(s+k_1\Delta t,s,x,y)$ is differentiable in (x,y) with derivative bounded independently of Δt sufficiently small, for $k_i\Delta t\geq\frac{\tau}{3}$ fixed and for some $y^*\in\mathrm{Int}(\mathcal{C})$,

$$P(s+k_i\Delta t,s,x,\mathcal{B}_{\delta_1}(y^*))>0,\ i=1,2,\quad\forall x\in\mathcal{C}.$$

Because of (2.12), we can see

$$\left|p(s+k_i\Delta t,s,x,y)-\hat{p}(s+k_i\Delta t,s,x,y)\right|<\varepsilon,$$

for sufficiently small Δt. Therefore the second condition in Condition 6.1 of [38] is also verified. Then by Theorem 6.2 of [38], we have

$$\inf_{x\in\mathcal{C}}\hat{P}^s_{K,0}(x,\Gamma)>\zeta\nu(\Gamma), \tag{2.13}$$

where ζ is independent of Δt. Thus we obtained the local Doeblin condition (2.10) as $\mathcal{C}=\{x\in\mathbb{R}^d: V(x)\leq R\}$ is a compact set, where $V(x)=1+|x|^2$.

Next, we prove the Lyapunov condition (2.9) with $V(x)$. For the numerical scheme SSEM (2.5), from the first equation and Condition (A3), we have

$$\begin{aligned}
|\hat{X}_i|^2 =& \left|\hat{X}_i^* - (\Delta t) b^s(i\Delta t, \hat{X}_i^*)\right|^2 \\
=& |\hat{X}_i^*|^2 - 2\left\langle \hat{X}_i^*, (\Delta t) b^s(i\Delta t, \hat{X}_i^*)\right\rangle + \left|(\Delta t) b^s(i\Delta t, \hat{X}_i^*)\right|^2 \\
\geq& (1+2\beta\Delta t)|\hat{X}_i^*|^2 - 2M\Delta t,
\end{aligned}$$

where M is the constant in Condition (A3). So,

$$|\hat{X}_i^*|^2 \leq (1+2\beta\Delta t)^{-1}\left(|\hat{X}_i|^2 + 2M\Delta t\right). \tag{2.14}$$

Define $\hat{\mathcal{F}}_i = \mathcal{F}_{s+i\Delta t}$. As $\hat{X}_i^*$, $\hat{X}_i$ are measurable with respect to $\hat{\mathcal{F}}_i$ and independent of $\Delta^s W_i$, and $\mathbb{E}\left[\Delta^s W_i \middle| \hat{\mathcal{F}}_i\right] = 0$, it follows that

$$\begin{aligned}
&\mathbb{E}\left[|\hat{X}_{i+1}|^2 \middle| \hat{\mathcal{F}}_i\right] \qquad (2.15)\\
=& \mathbb{E}\left[|\hat{X}_i^*|^2 \middle| \hat{\mathcal{F}}_i\right] + 2\mathbb{E}\left[\left\langle \hat{X}_i^*, \sigma^s(i\Delta t, \hat{X}_i^*)\Delta^s W_i\right\rangle \middle| \hat{\mathcal{F}}_i\right] + \mathbb{E}\left[|\sigma^s(i\Delta t, \hat{X}_i^*)\Delta^s W_i|^2 \middle| \hat{\mathcal{F}}_i\right] \\
\leq& |\hat{X}_i^*|^2 + \|\sigma^s(i\Delta t, \hat{X}_i^*)\|^2 \mathbb{E}\left[\left|\Delta^s W_i\right|^2 \middle| \hat{\mathcal{F}}_i\right] \\
\leq& (1+2\beta\Delta t)^{-1}|\hat{X}_i|^2 + C'\Delta t.
\end{aligned}$$

Here we can find such a constant C' from the uniform boundedness assumption of σ and $\mathbb{E}\left[|\Delta^s W_i|^2 \middle| \hat{\mathcal{F}}_i\right] = \Delta t$. As $(1+2\beta\Delta t)^{-1} < 1$ with strictly positive β and Δt, so there exist constants $r = \frac{1}{1+2\beta\Delta t} \in (0,1)$ and $C_1 \geq 0$ such that

$$\begin{aligned}
\mathbb{E}\left[V(\hat{X}_{i+1}) \middle| \hat{\mathcal{F}}_i\right] =& 1 + \mathbb{E}\left[|\hat{X}_{i+1}|^2 \middle| \hat{\mathcal{F}}_i\right] \\
=& 1 + r|\hat{X}_i|^2 + C'\Delta t \\
=& rV(\hat{X}_i) + 1 - r + C'\Delta t \\
\leq& rV(\hat{X}_i) + 2\beta\Delta t + C'\Delta t \\
\leq& rV(\hat{X}_i) + C_1\Delta t.
\end{aligned} \tag{2.16}$$

This implies that

$$\hat{P}^s_{i+1,i} V(x) \leq rV(x) + C_1\Delta t, \quad i = 0, 1, \ldots. \tag{2.17}$$

By the semigroup property of $\hat{P}^s_{\cdot,\cdot}$, it is easy to obtain that

$$\begin{aligned}
\hat{P}^s_{K,0} V(x) \leq& r^K V(x) + \sum_{i=0}^{K-1} r^i C_1 \Delta t \\
=& (1+2\beta\Delta t)^{\frac{\tau}{\Delta t}} V(x) + C_1\Delta t \frac{1-r^K}{1-r}
\end{aligned}$$

$$
\begin{aligned}
&\le e^{-\frac{1}{2}\beta\tau}V(x)+C_1\Delta t\frac{1+2\beta\Delta t}{2\beta\Delta t}\\
&=e^{-\frac{1}{2}\beta\tau}V(x)+C_2,
\end{aligned}
\tag{2.18}
$$

where C_2 is independent of Δt.

Set $P=\hat{P}^s_{K,0}$. From $\hat{P}^s_{2K,K}=\hat{P}^s_{K,0}$, then $\hat{P}^s_{kK,0}=(\hat{P}^s_{K,0})^k=P^k$, and

$$
d_\beta(P^k(x,\cdot),P^{k-1}(x,\cdot))=d_\beta((P^{k-1})^*P(x,\cdot),(P^{k-1})^*\delta_x).
$$

By Lemma 2.5, we have

$$
\begin{aligned}
d_\beta(P^k(x,\cdot),P^{k-1}(x,\cdot)) &=d_\beta(\hat{P}^s_{kK,0}(x,\cdot),\hat{P}^s_{(k-1)K,0}(x,\cdot))\\
&\le\alpha d_\beta(\hat{P}^s_{(k-1)K,0}(x,\cdot),\hat{P}^s_{(k-2)K,0}(x,\cdot))\\
&\le\alpha^{k-1}d_\beta(\hat{P}^s_{K,0}(x,\cdot),\delta_x)\\
&=\alpha^{k-1}\int_{\mathbb{R}^d}(1+\beta V(y))\left|\hat{P}^s_{K,0}-\delta_x\right|(dy)\\
&\le\alpha^{k-1}\left[\int_{\mathbb{R}^d}(1+\beta V(y))\hat{P}^s_{K,0}(dy)+\int_{\mathbb{R}^d}(1+\beta V(y))\delta_x(dy)\right]\\
&\le\alpha^{k-1}\left(2+\beta\mathbb{E}\left(V(\hat{X}_K)\right)+\beta V(x)\right)\\
&\le\alpha^{k-1}\left(2+\beta(e^{-\frac{1}{2}\beta\tau}V(x)+C_2)+\beta V(x)\right)\\
&=\alpha^{k-1}\left(2+\beta(e^{-\frac{1}{2}\beta\tau}+1)(1+|x|^2)+\beta C_2\right)\\
&\le C\alpha^{k-1}(1+|x|^2),
\end{aligned}
$$

where $C>0$ is a constant and $0<\alpha<1$ is a constant. It follows that for any $l\ge k$,

$$
\begin{aligned}
d_\beta(\hat{P}^l(x,\cdot),\hat{P}^k(x,\cdot)) &=d_\beta((\hat{P}^k)^*\hat{P}^{l-k}(x,\cdot),(\hat{P}^k)^*\delta_x)\\
&\le C\alpha^k d_\beta(\hat{P}^s_{(l-k)K,0}(x,\cdot),\delta_x)\\
&\le C\alpha^k\left(2+\beta(e^{-\frac{1}{2}\beta\tau(l-k)}V(x)+C_2)+\beta V(x)\right)\\
&\le C\alpha^k(1+|x|^2),
\end{aligned}
$$

so we know $\{\hat{P}^k\}_k$ is a Cauchy sequence. Then we have the existence of its limit which is an invariant measure of $\hat{P}^k$, denoted by $\hat{\rho}^s_0$. Let $l\to\infty$, we have

$$
d_\beta(\hat{P}^s_{kK,0}(x,\cdot),\hat{\rho}^s_0)=d_\beta(\hat{P}^k(x,\cdot),\hat{\rho}^s_0)\le C\alpha^k(1+|x|^2)=C(\alpha^{1/K})^{kK}(1+|x|^2). \tag{2.19}
$$

Let $\delta>0$ be chosen such that $e^{-\delta\Delta t}=\alpha^{\frac{1}{K}}$, then

$$d_\beta(\hat{P}^s_{0,-kK}(x,\cdot),\hat{\rho}^s_0) \le Ce^{-\delta k\tau}(1+|x|^2). \tag{2.20}$$

Note from (2.17) and Lemma 2.5 again, we also have for any $\mu_1,\mu_2 \in \mathcal{P}(\mathbb{R}^d)$, and all $i \in \mathbb{Z}$,

$$d_\beta((\hat{P}^s_{i+1,i})^*\mu_1,(\hat{P}^s_{i+1,i})^*\mu_2) \le Ce^{-\delta\Delta t}d_\beta(\mu_1,\mu_2). \tag{2.21}$$

Define

$$\hat{\rho}^s_i = (\hat{P}^s_{i,0})^*\hat{\rho}^s_0, \quad i=0,1,\dots. \tag{2.22}$$

Then it is easy to see from the semigroup property, (2.20) and (2.21) that for $i=0,1,\dots, k\in\mathbb{N}$,

$$\begin{aligned} d_\beta(\hat{P}^s_{i,-kK}(x,\cdot),\hat{\rho}^s_i) =& d_\beta((\hat{P}^s_{i,0})^*\hat{P}^s_{0,-kK}(x,\cdot),(\hat{P}^s_{i,0})^*\hat{\rho}^s_0) \\ \le& e^{-\delta i\Delta t}d_\beta(\hat{P}^s_{0,-kK}(x,\cdot),\hat{\rho}^s_0) \\ \le& Ce^{-\delta(k\tau+i\Delta t)}(1+|x|^2). \end{aligned} \tag{2.23}$$

In the following we want to prove the above also holds for $i\in\mathbb{Z}\backslash\mathbb{N}$. From (2.20) and periodicity of $\hat{P}^s_{\cdot,\cdot}$, we know that

$$d_\beta(\hat{P}^s_{-K,-kK}(x,\cdot),\hat{\rho}^s_0) \le e^{-\delta(k-1)\tau}(1+|x|^2).$$

Define $\hat{\rho}^s_{-K}=\hat{\rho}^s_0$, then

$$d_\beta(\hat{P}^s_{-K,-kK}(x,\cdot),\hat{\rho}^s_{-K}) \le e^{-\delta(k-1)\tau}(1+|x|^2).$$

Define

$$\hat{\rho}^s_i = (\hat{P}^s_{i,-K})^*\hat{\rho}^s_{-K}, \quad i=-K+1,-K+2,\dots, \tag{2.24}$$

then it is obvious that $\hat{\rho}^s_i$ defined in (2.24) is consistent with that defined in (2.22) for all $i\in\mathbb{N}$. On the other hand, (2.21) holds for all $i\in\mathbb{Z}$. Thus it follows that for $i\ge -K, k\ge 1$, similarly as (2.23),

$$\begin{aligned} d_\beta(\hat{P}^s_{i,-kK}(x,\cdot),\hat{\rho}^s_i) =& d_\beta((\hat{P}^s_{i,-K})^*\hat{P}^s_{-K,-kK}(x,\cdot),(\hat{P}^s_{i,-K})^*\hat{\rho}^s_{-K}) \\ \le& e^{-\delta\Delta t(i+K)}Ce^{-\delta(k-1)\tau}(1+|x|^2) \\ =& Ce^{-\delta(k\tau+i\Delta t)}(1+|x|^2). \end{aligned}$$

The result can extend similarly to all $i<-K$ as well. The result of this theorem is proved. □

Remark 2.8. For the backward Euler-Maruyama scheme (2.6), define, for $(i,j)\in\hat{\Delta}$,

$$\tilde{P}^s_{i,j}(x,\Gamma) = \mathbb{P}\{\hat{Z}_i\in\Gamma \big| \hat{Z}_j = x\}.$$

The map $\hat{X}_i = \hat{Z}_i - b^s(i\Delta t, \hat{Z}_i)\Delta t = G_i^{s,\Delta t}(\hat{Z}_i)$ between the SSEM scheme $\hat{X}_i$ and the BEM scheme $\hat{Z}_i$ is one-to-one with step size $\Delta t < 1/L_b$. By Remark 2.1, we have that for any set Γ, $(i, j) \in \hat{\Delta}$,

$$
\begin{aligned}
\tilde{P}^s_{i,j}(x,\Gamma) =&\mathbb{P}\{\hat{Z}_{i-j} \in \Gamma\}\\
=&\mathbb{P}\{(G_i^{s,\Delta t})^{-1}\hat{X}^{s+j\Delta t, G_j^{s,\Delta t}(x)}(s+i\Delta t) \in \Gamma\}\\
=&\mathbb{P}\{\hat{X}^{s+j\Delta t, G_j^{s,\Delta t}(x)}(s+i\Delta t) \in G_i^{s,\Delta t}(\Gamma)\}\\
=&\hat{P}^s_{i,j}(G_j^{s,\Delta t}(x), G_i^{s,\Delta t}(\Gamma))\\
=&(G_i^{s,\Delta t})^{-1}\hat{P}^s_{i,j}(G_j^{s,\Delta t}(x), \Gamma).
\end{aligned}
\tag{2.25}
$$

In particular,

$$
\begin{aligned}
\tilde{P}^s_{i,-kK}(x,\Gamma) =&\hat{P}^s_{i,-kK}(G_{-kK}^{s,\Delta t}(x), G_i^{s,\Delta t}(\Gamma))\\
=&\hat{P}^s_{i,-kK}(G_{-K}^{s,\Delta t}(x), G_i^{s,\Delta t}(\Gamma)).
\end{aligned}
\tag{2.26}
$$

Now define

$$
\tilde{\rho}^s_i(\Gamma) = \hat{\rho}^s_i(G_i^{s,\Delta t}(\Gamma)) = ((G_i^{s,\Delta t})^{-1}\hat{\rho}^s_i)(\Gamma). \tag{2.27}
$$

From the periodicity of $\hat{\rho}^s_\cdot$ and $G_\cdot^{s,\Delta t}$, we see that $i \mapsto \tilde{\rho}^s_i$ is also K-periodic. Moreover for $(i, j) \in \hat{\Delta}$, $\Gamma \in \mathcal{B}(\mathbb{R}^d)$,

$$
\begin{aligned}
(\tilde{P}^s_{i,j})^*\tilde{\rho}^s_j(\Gamma) =& \int_{\mathbb{R}^d} \hat{P}^s_{i,j}(G_j^{s,\Delta t}(y), G_i^{s,\Delta t}(\Gamma))\hat{\rho}^s_j(G_j^{s,\Delta t}(dy))\\
=&\int_{\mathbb{R}^d} \hat{P}^s_{i,j}(x, G_i^{s,\Delta t}(\Gamma))\hat{\rho}^s_j(dx)\\
=&\hat{\rho}^s_i(G_i^{s,\Delta t}(\Gamma)) = \tilde{\rho}^s_i(\Gamma).
\end{aligned}
$$

Thus $\tilde{\rho}^s_\cdot$ is a periodic measure of $\tilde{P}^s_{\cdot,\cdot}$.

Similar to the proof of Theorem 2.7, we have the geometric ergodicity of BEM.

Theorem 2.9. *Assume Conditions (A), (B) and (C). If $\Delta t < 1/L_b$, then the backward Euler-Maruyama scheme (2.6) is geometrically ergodic i.e. for any fixed s, there exists a periodic measure $\tilde{\rho}^s : \mathbb{Z} \to \mathcal{P}(\mathbb{R}^d)$ of period K such that*

$$
d_\beta\left(\tilde{P}^s_{i,-kK}(x,\cdot), \tilde{\rho}^s_i\right) \le C(1+|x|^2)e^{-\delta(k\tau+i\Delta t)}, \quad k \in \mathbb{N}, \tag{2.28}
$$

for some constants $C, \delta > 0$ and $i = 0, \dots, K-1$.

3. A priori estimates for SDEs and their semi-implicit numerical scheme

The main objective of the rest of this paper is to estimate the error of the periodic measure and its numerical approximation. For this purpose we need a priori estimates for both the solution of SDE (1.1) and its numerical approximation.

3.1. A priori estimates of the semi-implicit numerical schemes

First we show the numerical approximations under Conditions (A) and (B) possess bounded $2p$-th moments.

Proposition 3.1. *Assume Conditions (A) and (B), $0 < \Delta t < 1/L_b$, then for any integer p, there exist constants $C_p, \lambda > 0$ such that for any $x \in \mathbb{R}^d$ and $n \in \mathbb{N}$, $s \in \mathbb{R}$,*

$$\mathbb{E}|\hat{X}^{s,x}(s+n\Delta t)|^{2p} \le C_p \left(1 + |x|^{2p} \exp(-\lambda p n \Delta t)\right),$$

Proof. For simplicity in this proof, we denote $\hat{X}_n = \hat{X}^{s,x}(s+n\Delta t)$. Firstly by applying Young's inequality, we have for any $a, b > 0$,

$$(a+b)^p \le \left[1 + \sum_{l=1}^{p-1} \binom{p}{l} \frac{l}{p} \delta_l^{\frac{p}{l}}\right] a^p + \left[\sum_{l=1}^{p-1} \binom{p}{l} \frac{p-l}{p} \delta_l^{-\frac{p}{p-l}}\right] b^p.$$

Here $\delta_l > 0$ can be any fixed constants. For any given integer p and Δt, we choose each δ_l such that $\frac{l}{p}\delta_l^{\frac{p}{l}} = (\varepsilon\Delta t)^l$. Then,

$$(a+b)^p \le \left[1 + \sum_{l=1}^{p-1} \binom{p}{l} (\varepsilon\Delta t)^l\right] a^p + C_{p,\Delta t,\varepsilon} b^p \le (1+\varepsilon\Delta t)^p a^p + C_{p,\Delta t,\varepsilon} b^p. \tag{3.1}$$

Here $C_{p,\Delta t,\varepsilon}$ is a constant depending on p, Δt, ε. Using the estimations (2.14) and (2.15), applying the inequality (3.1), we have for each i,

$$\begin{aligned}
|\hat{X}_{i+1}|^{2p} &= \left(|\hat{X}_i^* + \sigma^s(i\Delta t, \hat{X}_i^*)\Delta^s W_i|^2\right)^p \\
&= \left(|\hat{X}_i^*|^2 + 2\left\langle \hat{X}_i^*, \sigma^s(i\Delta t, \hat{X}_i^*)\Delta^s W_i\right\rangle + |\sigma^s(i\Delta t, \hat{X}_i^*)\Delta^s W_i|^2\right)^p \\
&\le \left((1+\varepsilon_1\Delta t)|\hat{X}_i^*|^2 + C_{\Delta t,\varepsilon_1}\left|\sigma^s(i\Delta t, \hat{X}_i^*)\Delta^s W_i\right|^2\right)^p \\
&\le \left(\frac{1+\varepsilon_1\Delta t}{1+2\beta\Delta t}\left(|\hat{X}_i|^2 + 2C\Delta t\right) + C_{\Delta t,\varepsilon_1}\left|\sigma^s(i\Delta t, \hat{X}_i^*)\Delta^s W_i\right|^2\right)^p \\
&\le \left(\frac{(1+\varepsilon_1\Delta t)(1+\varepsilon_2\Delta t)}{1+2\beta\Delta t}\left(|\hat{X}_i|^2 + 2C\Delta t\right)\right)^p + C_{p,\Delta t,\varepsilon_1,\varepsilon_2}\left|\sigma^s(i\Delta t, \hat{X}_i^*)\Delta^s W_i\right|^{2p} \\
&\le \left(\frac{(1+\varepsilon_1\Delta t)(1+\varepsilon_2\Delta t)(1+\varepsilon_3\Delta t)}{1+2\beta\Delta t}|\hat{X}_i|^2\right)^p
\end{aligned}$$

$$+ C_{p,\Delta t,\varepsilon_3} (2C\Delta t)^p + C_{p,\Delta t,\varepsilon_1,\varepsilon_2} \left|\sigma^s(i\Delta t, \hat{X}_i^*)\Delta^s W_i\right|^{2p}.$$

It is worth noting that if we take $\varepsilon = \frac{\beta}{2+\beta\Delta t}$, then

$$\frac{1+\varepsilon\Delta t}{1+\beta\Delta t} = \frac{1+\frac{\beta}{2+\beta\Delta t}\Delta t}{1+\beta\Delta t} = \frac{2+\beta\Delta t+\beta\Delta t}{(1+\beta\Delta t)(2+\beta\Delta t)} = \frac{1}{1+\frac{\beta}{2}\Delta t}.$$

So we choose $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, $\varepsilon_3 > 0$ in succession to satisfy that

$$\frac{(1+\varepsilon_1\Delta t)(1+\varepsilon_2\Delta t)(1+\varepsilon_3\Delta t)}{1+2\beta\Delta t} = \frac{1}{1+\frac{\beta}{4}\Delta t}.$$

Then we take $\lambda = \frac{\ln\left(1+\frac{\beta}{4}\Delta t\right)}{\Delta t} > 0$ to have

$$\frac{(1+\varepsilon_1\Delta t)(1+\varepsilon_2\Delta t)(1+\varepsilon_3\Delta t)}{1+2\beta\Delta t} = e^{-\lambda\Delta t} < 1.$$

By the boundedness of $\sigma^s(i\Delta t, \hat{Y}_i)$ and its independence with $\Delta^s W_i$, we have that

$$\mathbb{E}\left(|\hat{X}_{i+1}|^{2p}\big|\hat{\mathcal{F}}_i\right) \le e^{-\lambda p\Delta t}|\hat{X}_i|^{2p} + C_{p,\sigma}. \tag{3.2}$$

Here λ, $C_{p,\sigma}$ could depend on Δt, but independent of i. Iterating (3.2), we obtain that

$$\mathbb{E}|\hat{X}_n|^{2p} \le e^{-\lambda pn\Delta t}|x|^{2p} + C_{p,\sigma}\sum_{i=0}^{n} e^{-\lambda pi\Delta t}.$$

The result of this proposition follows from the fact that $\sum_{i=0}^{n} e^{-\lambda pi\Delta t} < \frac{1}{1-e^{-\lambda p\Delta t}} < \infty$. $\square$

Denote $G_t^{s,\Delta t}(x) = x - b^s(t,x)\Delta t = x - b(t+s,x)\Delta t$ and take $\Delta t < 1/L_b$ with one-sided Lipschitz coefficient L_b of b. It is easy to see there is a unique solution to $x = y + b^s(t,x)\Delta t$, for any give $y \in \mathbb{R}^d$, denoted by $x = (G_t^{s,\Delta t})^{-1}(y)$. In [28], the authors presented an idea to analyze the semi-implicit scheme by an explicit numerical scheme in a finite horizon for autonomous SDEs. In the following, we will extend the analysis to the non-autonomous SDEs. This is an important step to analyze the semi-implicit scheme in infinite horizon. For this, set

$$\hat{b}^s(t,x) = b^s(t,(G_t^{s,\Delta t})^{-1}(x)), \quad \hat{\sigma}^s(t,x) = \sigma^s(t,(G_t^{s,\Delta t})^{-1}(x)) \tag{3.3}$$

and consider

$$d\hat{S}_t = \hat{b}^s(t,\hat{S}_t)dt + \hat{\sigma}^s(t,\hat{S}_t)d\tilde{W}_t, \tag{3.4}$$

with $\hat{S}_0 = x$, where $\tilde{W}_t = W_{t+s} - W_s = (\theta_s\omega)(t)$. Note that by rewriting the SSEM (2.5), we obtain

$$\hat{X}_{i+1} = \hat{X}_i + \hat{b}^s(i\Delta t, \hat{X}_i)\Delta t + \hat{\sigma}^s(i\Delta t, \hat{X}_i)\Delta^s W_i.$$

It is clear that the above is an explicit Euler-Maruyama scheme of modified SDE (3.4) which will help to carry out the desired analysis. We first give the following estimates in Lemma 3.2.

Lemma 3.2. *Assume Condition (A3), (B1) and take $\Delta t < \frac{1}{2L_b}$, then for any $x, y \in \mathbb{R}^d$ and $t \in \mathbb{R}$*

$$|\hat{b}^s(t,x)| \leq \frac{1}{1 - L_b\Delta t} \left|b^s(t,x)\right|, \tag{3.5}$$

$$\left|(G_t^{s,\Delta t})^{-1}(x) - (G_t^{s,\Delta t})^{-1}(y)\right|^2 \leq \frac{1}{1 - 2L_b\Delta t} |x-y|^2, \tag{3.6}$$

$$\left\langle x - y, \hat{b}^s(t,x) - \hat{b}^s(t,y)\right\rangle \leq \frac{L_b}{1 - 2L_b\Delta t} |x-y|^2, \tag{3.7}$$

$$\left\|\hat{\sigma}^s(t,x) - \hat{\sigma}^s(t,y)\right\| \leq \frac{L_\sigma}{1 - 2L_b\Delta t} |x-y|^2, \tag{3.8}$$

$$\left\langle x, \hat{b}^s(t,x)\right\rangle \leq M - \frac{\beta}{1 + 2\beta\Delta t} |x|^2. \tag{3.9}$$

The proof of (3.5) follows a homotopy argument as in [25] and the rest parts of Lemma 3.2 follows the proof in [28] (Lemma 3.4). The only difference is that the functions $\hat{b}$, $\hat{\sigma}$ depend on a time variable t here. But this does not create any difficulty to the proof, so we omit the proof in this paper.

Now we consider the solution of (3.4), we have the following estimate.

Lemma 3.3. *Assume Conditions (A) and (B) and take $\Delta t < \frac{1}{2L_b}$. Then for any integer $p \geq 2$, there exists $C_p, L_p > 0$, such that for any $t > 0$,*

$$\mathbb{E}|\hat{S}_t|^p \leq C_p(1 + |\hat{S}_0|^p \exp(-L_p t)). \tag{3.10}$$

Proof. From (3.9), we have a corresponding weakly dissipative condition for the modified SDE (3.4). Following the proof of Proposition 3.5, we conclude the result with some new constant C_p and L_p. $\square$

Now we consider

$$\bar{X}_t = \hat{X}_i + (t - i\Delta t)\hat{b}^s(i\Delta t, \hat{X}_i) + \hat{\sigma}^s(i\Delta t, \hat{X}_i)\Delta^s W_i(t), \tag{3.11}$$

where $\tilde{W}_r = W_{r+s} - W_s = (\theta_s\omega)(r)$, $\hat{X}_r := \hat{X}_i^{\Delta t}$, $\hat{r} = i\Delta t$ for $r \in [s + i\Delta t, s + (i+1)\Delta t)$ and $\bar{X}_0 = \bar{\xi}$. Here $\hat{X}_i$ is given in SSEM (2.5). Note $\bar{X}_{i\Delta t} = \hat{X}_i$, so (3.11) is a continuous version of the SSEM Scheme agrees with $\hat{X}_i$ for $i = i\Delta t$. This can be rewritten as (3.11) below.

Lemma 3.4. *Assume Conditions (A) and (B), take $\Delta t < \frac{1}{2L_b}$. Then for any integer $p \geq 2$, there exists a constant $C = C(p) > 0$, such that for any $t > 0$, the process $\bar{X}$ defined in (3.11) satisfies*

$$\mathbb{E}\left|\bar{X}_t\right|^{2p} \leq C_p. \tag{3.12}$$

Proof. Note from (2.5),

$$\hat{X}_i^* = \hat{X}_i + b^s(i\Delta t, \hat{X}_i^*)\Delta t = \hat{X}_i + \hat{b}^s(i\Delta t, \hat{X}_i)\Delta t. \tag{3.13}$$

Then it can be proved as in (2.14) that $|\hat{X}_i^*|^2 \le (1+2\beta\Delta t)^{-1}\left(|\hat{X}_i|^2 + 2C\Delta t\right)$. Denote $a = \frac{t-i\Delta t}{\Delta t}$. Hence from (3.11) and (3.13), we have that for any $t \in [i\Delta t, (i+1)\Delta t)$,

$$\begin{aligned}\left|\bar{X}_t\right|^{2p} =& \left|(1-a)\hat{X}_i + a\hat{X}_i^* + \hat{\sigma}^s(i\Delta t, \hat{X}_i)\Delta^s W_i(t)\right|^{2p}\\ \le& 2^{2p}\left((1-a)|\hat{X}_i|^{2p} + a|\hat{X}_i^*|^{2p} + \left\|\hat{\sigma}^s(i\Delta t, \hat{X}_i)\Delta^s W_i(t)\right\|^{2p}\right).\end{aligned}$$

From (2.14), boundedness of σ and Proposition 3.1, we obtain the result of this lemma. □

3.2. A priori estimates for the solution of equation (1.1)

Denote by $X_t := X^{s,x}(s+t)$ for simplicity. We also need p-th moment estimates for the solution of equation (1.1) as given in the following proposition.

Proposition 3.5. *Assume Conditions (A1) and (A3). Then for any integer $p \ge 2$, there exist $C_p, L_p > 0$, such that for any $t > 0$ and $x \in \mathbb{R}^d$,*

$$\mathbb{E}\,|X_t|^p \le C_p(1 + |x|^p \exp(-L_p t)). \tag{3.14}$$

Proof. The Brownian motion is also shifted, $\tilde{W}_t = W_{t+s} - W_s = (\theta_s\omega)(t)$. Applying Itô's formula and Conditions (A1), (A3), we have

$$\begin{aligned}d\left(e^{\delta t}|X_t|^p\right) \le& (\delta - p\beta)\, e^{\delta t}|X_t|^p\, dt + p\sigma(t, X_t)e^{\delta t}|X_t|^{p-1}\, d\tilde{W}_t\\ &+ \left(pM + \frac{p(p-1)}{2}C_\sigma^2\right)e^{\delta t}|X_t|^{p-2}\, dt,\end{aligned} \tag{3.15}$$

where C_σ is the bound of function σ, M and β are as given in the weak dissipativity Condition (A3). Denote $C_{p,\sigma} = pM + \frac{p(p-1)}{2}C_\sigma^2$. Let τ_N be the first exit time of the process X_t from the ball of radius N centred at 0, then $\mathbb{E}\int_s^{T\wedge\tau_N} \sigma(t, X_t)|X_t|^p\, dW_t = 0$ for arbitrary p. Now take expectation on both sides of (3.15) after integrating from s to $T \wedge \tau_N$, together with Young's inequality, we have

$$\begin{aligned}&\mathbb{E}e^{\delta(T\wedge\tau_N - s)}\left|X_{T\wedge\tau_N}\right|^p\\ &\le |x|^p + (\delta - p\beta)\,\mathbb{E}\int_s^{T\wedge\tau_N} e^{\delta(t-s)}|X_t|^p\, dt + C_{p,\sigma}\mathbb{E}\int_s^{T\wedge\tau_N} e^{\delta(t-s)}|X_t|^{p-2}\, dt \qquad (3.16)\\ &\le |x|^p + \frac{2C_{p,\sigma}}{p\delta\varepsilon^{\frac{p}{2}}}\mathbb{E}(e^{\delta(T\wedge\tau_N - s)} - 1) + K_0\mathbb{E}\int_0^{T\wedge\tau_N} e^{\delta t}|X_t|^p\, dt,\end{aligned}$$

where $K_0 = \delta - p\beta + \frac{(p-2)C_{p,\sigma}}{p}\varepsilon^{\frac{p}{p-2}}$ and we applied Young's inequality

$$|X_t|^{p-2} \le \frac{(|X_t|^{p-2}\varepsilon)^{\frac{p}{p-2}}}{\frac{p}{p-2}} + \frac{\left(\frac{1}{\varepsilon}\right)^{\frac{p}{2}}}{\frac{p}{2}} = \frac{p-2}{p}\varepsilon^{\frac{p}{p-2}}|X_t|^p + \frac{2}{p\varepsilon^{\frac{p}{2}}}$$

with $\varepsilon < \left(\frac{p^2\beta}{(p-2)C_{p,\sigma}}\right)^{\frac{p-2}{p}}$.

Now we consider $\delta = 0$. In this situation, $K_0 < 0$ and we have $\mathbb{E}\left|X_{T\wedge\tau_N}\right|^p \le |x|^p < \infty$. Let

$$\Omega_0 = \{\omega : \lim_{N\to\infty} \tau_N \le T\},$$

then note $|x|^p \ge \mathbb{E}\left|X_{T\wedge\tau_N}\right|^p \ge \mathbb{E}\mathbb{1}_{\Omega_0}\left|X_{T\wedge\tau_N}\right|^p \ge N^p\mathbb{P}(\Omega_0)$. Thus $\mathbb{P}(\Omega_0) \le \frac{1}{N^p}|x|^p \to 0$ as $N \to \infty$. So $\mathbb{P}(\Omega_0) = 0$. Note τ_N is non-decreasing in N a.s. So $\mathbb{P}(\Omega_0) = 0$ suggests $\lim_{N\to\infty}\tau_N \ge T$ a.s. But T is arbitrary. Hence, $\tau_N \to \infty$ as $N \to \infty$. Next we let N go to ∞ in (3.16), by Fatou's lemma

$$\begin{aligned}
e^{\delta(T-s)}\mathbb{E}|X_T|^p \le& \mathbb{E}\lim_{N\to\infty} e^{\delta(T\wedge\tau_N-s)}|X_{T\wedge\tau_N}|^p \\
\le& \liminf_{N\to\infty}\mathbb{E}e^{\delta(T\wedge\tau_N-s)}|X_{T\wedge\tau_N}|^p \\
\le& \liminf_{N\to\infty}\left[|x|^p + \frac{2C_{p,\sigma}}{p\delta\varepsilon^{\frac{p}{2}}}\mathbb{E}(e^{\delta(T\wedge\tau_N-s)}-1) + K_0\int_s^{T\wedge\tau_N-s} e^{\delta(t-s)}\mathbb{E}|X_t|^p\,dt\right] \\
\le& |x|^p + \frac{2C_{p,\sigma}}{p\delta\varepsilon^{\frac{p}{2}}}(e^{\delta(T-s)}-1) + K_0\int_s^T e^{\delta(t-s)}\mathbb{E}|X_t|^p dt.
\end{aligned}\tag{3.17}$$

Now we choose the constant δ to guarantee $K_0 > 0$ and apply the Gronwall inequality on (3.17),

$$e^{\delta T}\mathbb{E}|X_T|^p \le \frac{2C_{p,\sigma}}{p\delta\varepsilon^{\frac{p}{2}}}e^{\delta T} + e^{K_0T}\left(|x|^p - \frac{2C_{p,\sigma}}{p\delta\varepsilon^{\frac{p}{2}}}\right) + \frac{2K_0C_{p,\sigma}}{p(\delta-K_0)\delta\varepsilon^{\frac{p}{2}}}(e^{\delta T}-e^{K_0T}). \tag{3.18}$$

Then (3.14) follows as $K_0 < \delta$. $\square$

4. Convergence of numerical schemes in finite horizon

In the previous sections, we have given the numerical schemes to stochastic periodic systems and discussed the existence of their periodic measures. In order to analyze the error of the approximated periodic measures to the true periodic measures, we first need to study the numerical approximation to the true solution of the SDE (1.1) in finite horizon first.

As $b(t,0)$ is continuous and periodic in t, we know that it is uniformly bounded for any $t \in \mathbb{R}$. From Condition (C), we have a polynomial growth of b in the spacial variable as

$$|b(t,x)|^2 \le 2|b(t,0)|^2 + 2C\left(1+|x|^{2q}\right)|x|^2.$$

Applying Young's inequality and rearranging the constants, we have that

$$|b(t,x)|^2 \le C_q(1+|x|^{2q+2}). \tag{4.1}$$

Assuming Conditions (A) and (B) and applying Proposition 3.5 and Proposition 3.1, we have that

$$\mathbb{E}\,|b(t,X_t)|^2 \le C_q\left(1+\mathbb{E}\,|X_t|^{2q+2}\right) \le C_q(1+\mathbb{E}\,|\xi|^{2q+2}\exp(-nL_{2q+2}\Delta t)) \tag{4.2}$$

and

$$\mathbb{E}|b(t,\hat{X}_n)|^2 \le C'_q\left(1+\mathbb{E}|\hat{X}_n|^{2q+2}\right) \le C'_q(1+\mathbb{E}\,|\xi|^{2q+2}\exp(-nL_{2q+2}\Delta t)). \tag{4.3}$$

Under Condition (B2), we have the boundedness of b under the expectation following (4.2) and (4.3).

Lemma 4.1. *Assume Conditions (A), (B) and (C) and take $\Delta t < \frac{1}{2L_b}$. Then there exist $q \in \mathbb{Z}^+$ and constants $C_i = C_i(q)$, $i = 1,2,3$ such that for any $x, y \in \mathbb{R}^d$ and $t \in \mathbb{R}$, $\hat{b}^s$, $\hat{\sigma}^s$ defined by (3.3) satisfy*

$$|\hat{b}^s(t,x)-\hat{b}^s(t,y)|^2 \le C_1(1+|x|^{2q}+|y|^{2q})\,|x-y|^2\,, \tag{4.4}$$

$$|b^s(t,x)-\hat{b}^s(t,x)|^2 \le C_2(1+|x|^{4q+2})(\Delta t)^2, \tag{4.5}$$

$$\left\|\sigma^s(t,x)-\hat{\sigma}^s(t,x)\right\|^2 \le C_3(1+|x|^{2q+2})(\Delta t)^2. \tag{4.6}$$

Proof. Similar to the argument in (2.14), we have $|(G_t^{s,\Delta t})^{-1}(x)|^2 \le \frac{1}{1+2\beta\Delta t}(|x|^2+2C\Delta t)$. It implies that $(G_t^{s,\Delta t})^{-1}(x)$ has the same growth order as x, i.e.

$$|(G_t^{s,\Delta t})^{-1}(x)|^p \le C_p(1+|x|^p) \tag{4.7}$$

for any fixed integer $p \ge 2$. From Condition (C) and (3.6)

$$\begin{aligned}
&|\hat{b}^s(t,x)-\hat{b}^s(t,y)|^2\\
\le& C\left(1+|(G_t^{s,\Delta t})^{-1}(x)|^{2q}+|(G_t^{s,\Delta t})^{-1}(y)|^{2q}\right)\left|(G_t^{s,\Delta t})^{-1}(x)-(G_t^{s,\Delta t})^{-1}(y)\right|^2\\
\le& C\left(1+|x|^{2q}+|y|^{2q}\right)|x-y|^2\,.
\end{aligned}$$

Here C is a constant. Next from the fact that

$$\hat{b}^s(t,x)=b^s(t,(G_t^{s,\Delta t})^{-1}(x))=b^s\left(t,x+b^s\left(t,(G_t^{s,\Delta t})^{-1}(x)\right)\right)=b^s(t,x+\hat{b}^s(t,x)\Delta t),$$

it follows that

$$
\begin{aligned}
|b^s(t,x)-\hat{b}^s(t,x)|^2 \leq & C\left(1+|x|^{2q}+|(G_t^{s,\Delta t})^{-1}(x)|^{2q}\right)|\hat{b}^s(t,x)|^2(\Delta t)^2 \\
\leq & C_2\left(1+|x|^{2q}\right)|b^s(t,(G_t^{s,\Delta t})^{-1}(x))|^2(\Delta t)^2 \\
\leq & C_2\left(1+|x|^{2q}\right)\left(1+|(G_t^{s,\Delta t})^{-1}(x)|^{2q+2}\right)(\Delta t)^2 \\
\leq & C_2(1+|x|^{4q+2})(\Delta t)^2,
\end{aligned}
$$

where (4.1), (4.7) were used and constant C_2 can differ from time to time. Finally as σ is globally Lipschitz, $\hat{\sigma}^s(t,x)=\sigma^s(t,x+\hat{b}^s(t,x)\Delta t)$, (4.7), with a similar argument, it follows that

$$
\begin{aligned}
\left\|\sigma^s(t,x)-\hat{\sigma}^s(t,x)\right\|^2 \leq & L_\sigma|\hat{b}^s(t,x)|^2(\Delta t)^2 \\
= & L_\sigma\left|b^s(t,(G_t^{s,\Delta t})^{-1}(x))\right|^2(\Delta t)^2 \\
\leq & C_3\left(1+|(G_t^{s,\Delta t})^{-1}(x)|^{2q+2}\right)(\Delta t)^2 \\
\leq & C_3(1+|x|^{2q+2})(\Delta t)^2.
\end{aligned}
$$

Here C_3 is a constant that may differ from time to time. □

Recall the notations in the proof of Proposition 3.5 for simplicity: $X_t=X^{s,x}(s+t)$, $t\geq 0$, is the solution of (1.1) with $X_0=x$. $\hat{X}_i=\hat{X}^{s,x}(s+i\Delta t)$ is the SSEM scheme approximation with $\hat{X}_0=x$ and $t_i=i\Delta t$.

Lemma 4.2. *Assume Conditions (A), (B) and (C). Then for $0\leq t\leq T$ with a given $T>0$, the solutions of original SDE (1.1) and modified SDE (3.4) satisfy*

$$
\sup_{0\leq t\leq T}\mathbb{E}|X_t-\hat{S}_t|^2\leq Ce^{K_1T}(\Delta t)^2,
$$

for some constants K_1 and $C>0$.

Proof. Define $e(t)=X_t-\hat{S}_t$, so $e(0)=0$. We apply Itô's formula to have

$$
|e(t)|^2=2\int_0^t\left\langle b^s(r,X_r)-\hat{b}^s(r,\hat{S}_r),e(r)\right\rangle dr+\int_0^t\|\sigma^s(r,X_r)-\hat{\sigma}^s(r,\hat{S}_r)\|^2dr+M(t),
$$

where

$$
M(t)=2\int_0^t\left\langle e(r),(\sigma^s(r,X_r)-\hat{\sigma}^s(r,\hat{S}_r)d\tilde{W}_r\right\rangle
$$

and $\tilde{W}_t=W_{t+s}-W_s=(\theta_s\omega)(t)$. Note

$$
\begin{aligned}
&2\int_0^t \left\langle b^s(r, X_r) - \hat{b}^s(r, \hat{S}_r), e(r) \right\rangle dr \\
=&2\int_0^t \left\langle b^s(r, X_r) - b^s(r, \hat{S}_r), e(r) \right\rangle dr + 2\int_0^t \left\langle b^s(r, \hat{S}_r - \hat{b}^s(r, \hat{S}_r), e(r) \right\rangle dr \\
\le&2L_b \int_0^t |e(r)|^2\, dr + 2\int_0^t |b^s(r, \hat{S}_r - \hat{b}^s(r, \hat{S}_r)|^2 dr + 2\int_0^t |e(r)|^2\, dr \\
\le&2(L_b+1)\int_0^t |e(r)|^2\, dr + 2C_2(\Delta t)^2 \int_0^t \left[1 + |\hat{S}_r|^{4q+2}\right] dr
\end{aligned}
$$

and

$$
\begin{aligned}
&\int_0^t \|\sigma^s(r, X_r) - \hat{\sigma}^s(r, \hat{S}_r)\|^2 dr \\
\le&2\int_0^t \|\sigma^s(r, X_r) - \sigma(r, \hat{S}_r)\|^2 dr + 2\int_0^t \|\sigma^s(r, \hat{S}_r) - \hat{\sigma}^s(r, \hat{S}_r)\|^2 dr \\
\le&L_\sigma \int_0^t |e(r)|^2\, dr + 2C_3(\Delta t)^2 \int_0^t \left[1 + |\hat{S}_r|^{2q+2}\right] dr.
\end{aligned}
$$

So it follows from above and Lemma 3.3 that

$$
\begin{aligned}
\mathbb{E}\,|e(t)|^2 \le&(2(L_b+1)+L_\sigma)\int_0^t \mathbb{E}\,|e(r)|^2\, dr + 2(C_2+C_3)(\Delta t)^2 \int_0^t \left[1 + \mathbb{E}|\hat{S}_r|^{4q+2}\right] dr \\
\le&K_1 \int_0^t \mathbb{E}\,|e(r)|^2\, dr + K_2(\Delta t)^2 t,
\end{aligned}
$$

where K_1, K_2 are independent of Δt. From the Gronwall inequality,

$$
\mathbb{E}\,|e(t)|^2 \le K_2(\Delta t)^2 \int_0^t e^{K_1(t-r)} dr \le \frac{K_2}{K_1} e^{K_1 t} (\Delta t)^2.
$$

The desired result follows. □

Lemma 4.3. *Assume Conditions (A), (B) and (C), then the local error of (3.11) satisfies for any $p \ge 1$,*

$$\sup_{t\geq 0}\mathbb{E}|\bar{X}_t-\hat{X}_t|^{2p}\leq C_p(\Delta t)^p.$$

Proof. Recall the definitions of $\bar{X}$ and $\hat{X}_i$ in (3.11),

$$\bar{X}_t-\hat{X}_i=\hat{b}^s(i\Delta t,\hat{X}_i)(t-i\Delta t)+\hat{\sigma}^s(i\Delta t,\hat{X}_i)\Delta^s W_i(t),$$

and the coincidence of this two processes on grip points, i.e. $\hat{X}_r:=\hat{X}_i$, $\hat{r}=i\Delta t$ for $r\in[s+i\Delta t, s+(i+1)\Delta t)$. We apply (3.5), (4.1) and the boundedness of σ to deduce that

$$\begin{aligned}\mathbb{E}|\bar{X}_t-\hat{X}_i|^2\leq&2(\Delta t)^2\mathbb{E}\left|\hat{b}^s(i\Delta t,\hat{X}_i)\right|^2+2\mathbb{E}\left|\hat{\sigma}^s(i\Delta t,\hat{X}_i)\Delta^s W_i(t)\right|^2\\ \leq&\frac{2(\Delta t)^2}{(1-L_b\Delta t)^2}\mathbb{E}\left|b^s(t,\hat{X}_i))\right|^2+2C_\sigma^2\Delta t\\ \leq&\frac{2C_q(\Delta t)^2}{(1-L_b\Delta t)^2}\left(1+\mathbb{E}|\hat{X}_i|^{2q+2}\right)+2C_\sigma^2\Delta t.\end{aligned}$$

The desired result when $p=1$ follows from the a priori estimate in Proposition 3.1 and the notation of $\hat{X}_t$ and $\hat{X}_i$. Moreover, the above inequality can be extended to any power $2p$ $(p\geq 1)$ and the result of this lemma follows immediately from Proposition 3.1 again. □

Lemma 4.4. *Assume Conditions (A), (B) and (C), then the solutions of modified SDE (3.4) and continuous-time extension (3.11) satisfy*

$$\sup_{0\leq t\leq T}\mathbb{E}|\hat{S}_t-\bar{X}_t|^2\leq Ce^{K_3T}(\Delta t),$$

for some constants $K_3, C>0$.

Proof. Define $e(t)=\hat{S}_t-\bar{X}_t$. First $e(0)=0$. Again Itô's formula leads to

$$\begin{aligned}|e(t)|^2=&2\int_0^t\left\langle\hat{b}^s(r,\hat{S}_r)-\hat{b}^s(\hat{r},\hat{X}_r),e(r)\right\rangle dr\\ &+\int_0^t\|\hat{\sigma}^s(r,\hat{S}_r)-\hat{\sigma}^s(\hat{r},\hat{X}_r)\|^2dr+M(t),\end{aligned}\tag{4.8}$$

where

$$M(t)=2\int_0^t\left\langle e(r),\left(\hat{\sigma}^s(r,\hat{S}_r)-\hat{\sigma}^s(\hat{r},\hat{X}_r)\right)d\tilde{W}_r\right\rangle.$$

Note first from Lemma 3.2, (4.4) and Condition (A1)

$$
\begin{aligned}
&2\int_0^t \left\langle \hat{b}^s(r,\hat{S}_r)-\hat{b}^s(\hat{r},\hat{X}_r), e(r)\right\rangle dr \\
=&2\int_0^t \left\langle \hat{b}^s(r,\hat{S}_r)-\hat{b}^s(r,\bar{X}_r), e(r)\right\rangle dr \\
&+2\int_0^t \left\langle \hat{b}^s(r,\bar{X}_r)-\hat{b}^s(r,\hat{X}_r), e(r)\right\rangle dr+2\int_0^t \left\langle \hat{b}^s(r,\hat{X}_r)-\hat{b}^s(\hat{r},\hat{X}_r), e(r)\right\rangle dr \\
\le&\frac{2L_b}{1-2L_b\Delta t}\int_0^t |e(r)|^2\,dr+2\int_0^t |e(r)|^2\,dr \\
&+\int_0^t \left|\hat{b}^s(r,\bar{X}_r)-\hat{b}^s(r,\hat{X}_r)\right|^2 dr+\int_0^t \left|\hat{b}^s(r,\hat{X}_r)-\hat{b}^s(\hat{r},\hat{X}_r)\right|^2 dr \\
\le&\left(2+\frac{2L_b}{1-2L_b\Delta t}\right)\int_0^t |e(r)|^2\,ds \\
&+C_1\int_0^t \left(1+|\bar{X}_r|^q+|\hat{X}_r|^q\right)|\bar{X}_r-\hat{X}_r|^2dr+L_b(\Delta t)t.
\end{aligned}
$$

Then take expectation and apply Cauchy-Schwarz inequality to have

$$
\begin{aligned}
&2\mathbb{E}\int_0^t \left\langle \hat{b}^s(r,\hat{S}_r)-\hat{b}^s(\hat{r},\hat{X}_r), e(r)\right\rangle dr \\
\le&\left(2+\frac{2L_b}{1-2L_b\Delta t}\right)\int_0^t \mathbb{E}\,|e(r)|^2\,dr \\
&+C_q\left(\int_0^t \mathbb{E}\left(1+|\bar{X}_r|^q+|\hat{X}_r|^q\right)^2dr\cdot\int_0^t \mathbb{E}|\bar{X}_r-\hat{X}_r|^4dr\right)^{\frac{1}{2}}+tL_b(\Delta t).
\end{aligned}
\tag{4.9}
$$

Moreover,

$$
\mathbb{E}\int_0^t \|\hat{\sigma}^s(r,\hat{S}_r)-\hat{\sigma}^s(\hat{r},\hat{X}_r)\|^2dr
$$

$$\leq 3\int_0^t \mathbb{E}\|\hat{\sigma}^s(r,\hat{S}_r)-\hat{\sigma}^s(r,\bar{X}_r)\|^2dr+3\int_0^t \mathbb{E}\|\hat{\sigma}^s(r,\bar{X}_r)-\hat{\sigma}^s(r,\hat{X}_r)\|^2dr$$

$$+3\int_0^t \mathbb{E}\|\hat{\sigma}^s(r,\hat{X}_r)-\hat{\sigma}^s(\hat{r},\hat{X}_r)\|^2dr$$

$$\leq 3L_\sigma\int_0^t \mathbb{E}\,|e(r)|^2\,dr+3L_\sigma\int_0^t \mathbb{E}|\bar{X}_r-\hat{X}_r|^2dr+3tL_\sigma(\Delta t). \tag{4.10}$$

Putting (4.8), (4.9) and (4.10) together applying Lemma 4.3 and a priori estimates for $\bar{X}$ (Lemma 3.4) and $\hat{X}$ (Proposition 3.1), we have some constants K_3 and K_4 independent of Δt or t, such that

$$\mathbb{E}\,|e(t)|^2\leq K_3\int_0^t \mathbb{E}\,|e(r)|^2\,dr+K_4(\Delta t)t.$$

Then the desired result follows from the Gronwall inequality. □

Now we simply apply the triangle inequality with the results of Lemma 4.2 and Lemma 4.4 to derive the following result.

Theorem 4.5. *Assume Conditions (A), (B) and (C), and take $\Delta t=\frac{\tau}{K}<\frac{1}{2L_b}$ with some integer K. Then the solution $X^{s,x}(s+i\Delta t)$ of original SDE (1.1) and its numerical approximation $\hat{X}_i^{\Delta t}$ given by the SSEM scheme (2.5) satisfy that for any fixed integer n,*

$$\sup_{0\leq i\leq n}\mathbb{E}\left|X^{s,x}(s+i\Delta t)-\hat{X}_i^{\Delta t}\right|^2\leq Ce^{\lambda n\Delta t}(\Delta t),$$

where $\lambda=max\{K_1,K_3\}$ with K_1, K_3 given in Lemma 4.2 and 4.4 respectively.

Proof. Note first as demonstrated in Section 3.1, the SSEM scheme is equivalent to the forward Euler-Maruyama scheme of the modified SDE (3.4) i.e. $\hat{X}_i^{\Delta t}=\hat{X}_i=\bar{X}(i\Delta t)$. Thus,

$$\sup_{0\leq i\leq n}\mathbb{E}\left|X^{s,x}(s+i\Delta t)-\hat{X}_i^{\Delta t}\right|^2$$

$$\leq 2\sup_{0\leq i\leq n}\left(\mathbb{E}\left|X^{s,x}(s+i\Delta t)-\hat{S}_{i\Delta t}\right|^2+\mathbb{E}\left|\hat{S}_{i\Delta t}-\bar{X}_{i\Delta t}\right|^2\right)\leq Ce^{\lambda n\Delta t}(\Delta t),$$

and then the desired result follows from Lemma 4.2 and Lemma 4.4. □

Remark 4.6. (i) From the equivalence between BEM and SSEM as in Remark 2.1, we have that $\hat{Z}_i=\hat{X}_i+b^s(i\Delta t,\hat{Z}_i)\Delta t=\hat{X}_i+b^s(i\Delta t,(G^{s,\Delta t}_{i\Delta t})^{-1}(\hat{X}_i))\Delta t=\hat{X}_i+\hat{b}^s(i\Delta t,\hat{X}_i)\Delta t$. Thus,

$$\mathbb{E}|\hat{X}_i-\hat{Z}_i|^2=\mathbb{E}\left|\hat{X}_i-(\hat{X}_i+\hat{b}^s(i\Delta t,\hat{X}_i)\Delta t)\right|^2\leq C_q(1+\mathbb{E}|\hat{X}_i|^{2q+2})(\Delta t)^2,$$

where the inequality follows from (4.4). Then we combine the above inequality, Proposition 3.1 and Theorem 4.5 to conclude that

$$\sup_{0\leq i\leq n}\mathbb{E}\left|X_{i\Delta t}-\hat{Z}_i^{\Delta t}\right|^2\leq Ce^{\lambda n\Delta t}(\Delta t).$$

This gives the numerical error of the BEM scheme.

(ii) Theorem 4.5 which is about the error estimates of a finite time horizon problem will play a key role in the infinite horizon error analysis of periodic measures in the Wasserstein distance in the next section.

5. Discretization and pathway to periodic measures and error estimates

Denote by $\mathcal{L}(\hat{X})$ the law of process $\hat{X}$. Consider a Markov chain $\{\hat{X}_n^h\}_{n=1,2,\ldots}$ with an arbitrary time-increment h, generated by the SSEM numerical approximation $\hat{X}^{s,x}(s+nh)$, to the periodic Markov process $X^{s,x}(s+nh)$, with its transition probability as

$$\mathcal{L}(\hat{X}^{s,x}(s+nh))(B)=\hat{P}(s+nh,s,x,B)=\mathbb{P}\{\omega\in\Omega:\hat{X}^{s,x}(s+nh)\in B\},\quad B\in\mathcal{B}(\mathbb{R}^d).$$

We also consider the Markov chain $\{X^{s,x}(s+nh)\}_{n=1,2,\ldots}$ generated by the exact solution of (1.1). The corresponding transition probability is

$$\mathcal{L}(X^{s,x}(s+nh))(B)=P(s+nh,s,x,B)=\mathbb{P}\{\omega\in\Omega:X^{s,x}(s+nh)\in B\},\quad B\in\mathcal{B}(\mathbb{R}^d).$$

For a given $p\geq 1$, we consider the following subspace of $\mathcal{P}(\mathbb{R}^d)$,

$$\mathcal{P}_p(\mathbb{R}^d)=\left\{\mu\in\mathcal{P}(\mathbb{R}^d)\middle|\,\|\mu\|_p=\left(\int_{\mathbb{R}^d}|x|^p\,\mu(dx)\right)^{\frac{1}{p}}<\infty\right\}.$$

This space is a Polish space under the following p-Wasserstein distance

$$W_p(\mu_1,\mu_2)=\inf_{\nu\in\mathcal{C}(\mu_1,\mu_2)}\left(\int_{\mathbb{R}^d\times\mathbb{R}^d}|x-y|^p\,\nu(dx,dy)\right)^{\frac{1}{p}},\quad \mu_1,\mu_2\in\mathcal{P}_p(\mathbb{R}^d),$$

where $\mathcal{C}(\mu_1,\mu_2)$ is the set of all couplings of μ_1 and μ_2 containing all probability measures on $\mathbb{R}^d\times\mathbb{R}^d$ with marginal distributions μ_1 and μ_2. For $\mu\in\mathcal{P}_p(\mathbb{R}^d)$ and $\phi:\mathbb{R}^d\to\mathbb{R}$ measurable with $\sup_{x\in\mathbb{R}^d}\frac{\phi(x)}{1+|x|^p}<\infty$, denote $\mu(\phi)=\int_{\mathbb{R}^d}\phi d\mu$. Furthermore, let $Lip(1)$ be the collection of all functions $f:\mathbb{R}^d\to\mathbb{R}$ that are Lipschitz continuous with Lipschitz constant $|f|_{Lip}\leq 1$.

Consider W_1 and its dual representation gives

$$
\begin{aligned}
W_1(\mu_1,\mu_2) &= \sup_{f\in Lip(1)} \left| \int_{\mathbb{R}^d} f(x)\mu_1(dx) - \int_{\mathbb{R}^d} f(x)\mu_2(dx) \right| \\
&= \sup_{f\in Lip(1)} \left| \int_{\mathbb{R}^d} (f(x)-f(0))(\mu_1-\mu_2)(dx) \right| \\
&\leq \sup_{f\in Lip(1)} \left| \int_{\mathbb{R}^d} |f(x)-f(0)|\,|\mu_1-\mu_2|\,(dx) \right| \\
&\leq \int_{\mathbb{R}^d} |x|\,|\mu_1-\mu_2|\,(dx).
\end{aligned}
\tag{5.1}
$$

For notational convenience, denote $\mu=\mu_1-\mu_2$. If we assume $V(x)\geq |x|$ and

$$
\mathcal{G} = \{\text{measurable } \phi : \mathbb{R}^d \to \mathbb{R}, |\phi(x)| \leq V(x)\}.
$$

We introduce the notation $\mu^+=\max\{\mu,0\}$ and $\mu^-=\max\{-\mu,0\}$. It is easy to see from (5.1) that

$$
\begin{aligned}
W_1(\mu_1,\mu_2) &\leq \int_{\mathbb{R}^d} V(x)\mu^+(dx) + \int_{\mathbb{R}^d} V(x)\mu^-(dx) \\
&= \int_{\mathbb{R}^d} (V(x)\mathbb{I}_{supp(\mu^+)})\mu^+(dx) - \int_{\mathbb{R}^d} (-V(x)\mathbb{I}_{supp(\mu^-)})\mu^-(dx) \\
&\leq \sup_{\phi\in\mathcal{G}} \left[\int_{\mathbb{R}^d} \phi(x)\mu^+(dx) - \int_{\mathbb{R}^d} \phi(x)\mu^-(dx) \right] \\
&= \sup_{\phi\in\mathcal{G}} \left[\int_{\mathbb{R}^d} \phi(x)\mu_1(dx) - \int_{\mathbb{R}^d} \phi(x)\mu_2(dx) \right].
\end{aligned}
\tag{5.2}
$$

Take $V_p(x) = 1+|x|^{2p}$ and let

$$
\begin{aligned}
\mathcal{G}_p = \Big\{ \phi : \mathbb{R}^d \to \mathbb{R} \text{ such that } \ |\phi(x)| \leq V_p(x), \\
|\phi(x)-\phi(y)| \leq C(1+|x|^{2p-1}+|y|^{2p-1})|x-y| \text{ for a constant } C \Big\}.
\end{aligned}
$$

It is easy to see that

$$W_1(\mu_1,\mu_2) \le \sup_{\phi\in\mathcal{G}_p}\left[\int_{\mathbb{R}^d}\phi(x)\mu_1(dx) - \int_{\mathbb{R}^d}\phi(x)\mu_2(dx)\right]. \tag{5.3}$$

We give the following two hypotheses needed for a perturbation lemma. The idea of the perturbation theory was introduced in [48] to analyze the long time behavior of an approximation scheme to invariant measures. However, we substitute the geometric ergodicity condition of the original continuous semigroup with that of the discretized approximating process. This approach offers a significant advantage that it eliminates the necessity to directly establish the ergodicity of the original continuous-time process as a priori information. In fact, we will demonstrate that the ergodicity of the discretized approximating process implies the ergodicity of the original continuous process.

The theory is presented in an abstract form, applicable to the solutions of SDEs and their discretization schemes. Although the theory is general, it can be verified and applied to some specific SDEs discussed in the previous sections.

Hypothesis (I) (Geometric ergodicity) For some constant $C_1 > 0$, $0 < r < 1$, we have for all $s \in [0,\tau), n \ge 0$,

$$W_1\left(\mathcal{L}(\hat{X}^{s,x}(s+nh)), \hat{\rho}_{s+nh}\right) \le \sup_{\phi\in\mathcal{G}}\left|\mathbb{E}\phi(\hat{X}^{s,x}(s+nh)) - \hat{\rho}_{s+nh}(\phi)\right| \le C_1 r^{nh} V(x),$$

where $\hat{\rho}_\cdot$ is the periodic measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ of discrete time Markov process $\{\hat{X}_n\}_{n=1,2,\dots}$.

Hypothesis (II) There exist some constant $C_2 > 0$, $R \ge 1$, such that for all $s \in [0,\tau), n \ge 0$,

$$\begin{aligned} W_1\left(\mathcal{L}(X^{s,x}(s+nh)), \mathcal{L}(\hat{X}^{s,x}(s+nh))\right) \le& \sup_{\phi\in\mathcal{G}}\left|\mathbb{E}\phi(X^{s,x}(s+nh)) - \mathbb{E}\phi(\hat{X}^{s,x}(s+nh))\right| \\ \le& C_2 R^{nh} V(x)\varepsilon, \end{aligned}$$

where ε is a parameter related to the process $X^{s,x}(\cdot)$ and the process $\hat{X}^{s,x}(\cdot)$.

In the above hypotheses, all constants C_1, C_2, r, R are independent of h. Hypothesis (I) is about geometric ergodicity that represents the long time behavior of the approximating process. Hypothesis (II) is on an error accumulation in a finite time horizon. Hypotheses (I) and (II) together enable us to obtain a uniform error estimate over an arbitrary long time horizon. In applications to numerical analysis, the perturbation ε is essentially a numerical error of form $(\Delta t)^p$ with some order p. In the following lemma, for a given V, we denote $\bar{\rho}(V) := \sup_{s\in[0,\tau)}\rho_s(V)$ and $\bar{\hat{\rho}}(V) := \sup_{s\in[0,\tau)}\hat{\rho}_s(V)$ respectively.

Proposition 5.1. *Under Hypotheses (I) and (II), there exist constants $C > 0$, $\varepsilon_0 > 0$ such that for any $0 < \varepsilon < \varepsilon_0$, there is a constant $N = \frac{\log\varepsilon}{h\log(r/R^2)}$, such that*

$$W_1\left(\mathcal{L}(X^{s,x}(s+nh)), \hat{\rho}_{s+nh}\right) \le 2C\max\{\bar{\hat{\rho}}(V), V(x)\}\varepsilon^\eta, \quad \text{for any } n \ge N, \tag{5.4}$$

where $\eta = \frac{\log r}{\log r - 2\log R} \in (0,1]$ and $\varepsilon_0 = \min\left(\left(\frac{r}{R^2}\right)^{2h}, \left(2C_1 r^{-h} + 2C_2\right)^{-\frac{1}{\eta}}\right)$. Furthermore, if the Markov chain $\{X^{s,x}(s+nh)\}_{n=1,2,\dots}$ has a periodic measure $\rho_\cdot$, then we have for any $s \in [0,\tau)$,

$$W_1\left(\rho_s, \hat{\rho}_s\right) \leq 2C\{\bar{\rho}(V) + \bar{\bar{\rho}}(V)\}\varepsilon^{\eta}. \tag{5.5}$$

Proof. For the given constants $0 < r < 1$ and $R > 1$ and fixed h as in Hypotheses (I) and (II), we have that for any $0 < \varepsilon \leq \varepsilon_0 \leq \left(\frac{r}{R^2}\right)^{2h}$, the constant $N(\varepsilon) = \frac{\log \varepsilon}{h \log\left(r/R^2\right)} \geq 2$ satisfying

$$\varepsilon R^{2Nh} = r^{Nh} = \varepsilon^{\eta}.$$

From Hypothesis (I), we have for any $s \in [0, \tau)$, $L \geq N - 1$,

$$\begin{aligned} W_1\left(\mathcal{L}(\hat{X}_L), \hat{\rho}_{s+Lh}\right) \leq & \sup_{\phi \in \mathcal{G}} \left|\mathbb{E}\phi(\hat{X}_L) - \hat{\rho}_{s+Lh}(\phi)\right| \\ \leq & C_1 r^{Lh} V(x) \leq C_1 r^{-h} r^{Nh} V(x) = C_1 r^{-h} V(x) \varepsilon^{\eta}. \end{aligned}$$

Also from Hypothesis (II), we have for any $i = 0, 1, \ldots, K-1$, $L \leq 2N$,

$$\begin{aligned} W_1\left(\mathcal{L}(X^{s,x}(s+Lh)), \mathcal{L}(\hat{X}_L)\right) \leq & \sup_{\phi \in \mathcal{G}} \left|\mathbb{E}\phi(X^{s,x}(s+Lh)) - \mathbb{E}\phi(\hat{X}_L)\right| \\ \leq & C_2 R^{Lh} V(x)\varepsilon \leq C_2 R^{2Nh} V(x)\varepsilon = C_2 V(x)\varepsilon^{\eta}. \end{aligned}$$

Combining the above two results, we have for any $L \in [N-1, 2N]$,

$$W_1\left(\mathcal{L}(X^{s,x}(s+Lh)), \hat{\rho}_{s+Lh}\right) \leq \sup_{\phi \in \mathcal{G}} \left|\mathbb{E}\phi(X^{s,x}(s+Lh)) - \hat{\rho}_{s+Lh}(\phi)\right| \leq CV(x)\varepsilon^{\eta}, \tag{5.6}$$

where $C = C_1 r^{-h} + C_2$. Thus it follows from (5.6) that

$$\begin{aligned} \mathcal{L}(X^{s,x}(s+Lh))(V) = & \mathbb{E}V(X^{s,x}(s+Lh)) \leq \hat{\rho}_{s+Lh}(V) + \left|\mathbb{E}V(X^{s,x}(s+Lh)) - \hat{\rho}_{s+Lh}(V)\right| \\ \leq & \hat{\rho}_{s+Lh}(V) + CV(x)\varepsilon^{\eta} \leq \bar{\bar{\rho}}(V) + CV(x)\varepsilon^{\eta}, \end{aligned} \tag{5.7}$$

where $L \in [N-1, 2N]$.

Now we prove by induction that for any $s \in [0, \tau)$, $M \in \mathbb{N}$ and any $n \in [MN, (M+1)N]$

$$W_1\left(\mathcal{L}(X^{s,x}(s+nh)), \hat{\rho}_{s+nh}\right) \leq \bar{\bar{\rho}}(V) \sum_{j=1}^{M-1} (C\varepsilon^{\eta})^j + V(x)(C\varepsilon^{\eta})^M. \tag{5.8}$$

First note that when $M = 1$, the first summation of (5.8) is 0, so the inequality holds due to (5.6). Next we assume that (5.8) holds for a fixed $M \in \mathbb{N}$. Consider any integer $m \in [(M+1)N, (M+2)N]$. Let n be the closest integer to $m - N$ subject to the constraint that $n \in [MN, (M+1)N]$. So $n \in [m-N-1, m-N+1]$ and hence the integer $L := m - n \in [N-1, N+1] \subset [N-1, 2N]$ as $N \geq 2$. From the semi-flow property, $X^{s,x}(s+mh) = X^{s+Lh,y}(s+mh)\big|_{y=X^{s,x}_{s+Lh}}$, the Markov property, the induction assumption (5.8) for $n \in [MN, (M+1)N]$, (5.7), we have that

$$
\begin{aligned}
&W_1\left(\mathcal{L}(X^{s,x}(s+mh)),\hat{\rho}_{s+mh}\right)\\
=&\sup_{f\in Lip(1)}\left|\int f(y)\mathcal{L}(X^{s,x}(s+mh))(dy)-\int f(y)\hat{\rho}_{s+mh}(dy)\right|\\
=&\sup_{f\in Lip(1)}\left|\iint f(y)\mathcal{L}(X^{s+Lh,z}(s+mh))(dy)\mathcal{L}(X^{s,x}(s+Lh))(dz)\right.\\
&\qquad\qquad\left.-\iint f(y)\hat{\rho}_{s+mh}(dy)\mathcal{L}(X^{s,x}(s+Lh))(dz)\right|\\
\le&\int\sup_{f\in Lip(1)}\left|\int f(y)\mathcal{L}(X^{s+Lh,z}(s+mh))(dy)-\int f(y)\hat{\rho}_{s+mh}(dy)\right|\mathcal{L}(X^{s,x}(s+Lh))(dz)\\
=&\int W_1\left(\mathcal{L}(X^{s+Lh,z}(s+Lh+nh)),\hat{\rho}_{s+Lh+nh}\right)\mathcal{L}(X^{s,x}(s+Lh))(dz)\\
\le&\int\left(\bar{\hat{\rho}}(V)\sum_{j=1}^{M-1}(C\varepsilon^\eta)^j+V(z)(C\varepsilon^\eta)^M\right)\mathcal{L}(X^{s,x}(s+Lh))(dz)\\
=&\bar{\hat{\rho}}(V)\sum_{j=1}^{M-1}(C\varepsilon^\eta)^j+\mathcal{L}(X^{s,x}(s+Lh))(V)(C\varepsilon^\eta)^M\\
\le&\bar{\hat{\rho}}(V)\sum_{j=1}^{M-1}(C\varepsilon^\eta)^j+\left(\bar{\hat{\rho}}(V)+CV(x)\varepsilon^\eta\right)(C\varepsilon^\eta)^M\\
=&\bar{\hat{\rho}}(V)\sum_{j=1}^{M}(C\varepsilon^\eta)^j+V(x)(C\varepsilon^\eta)^{M+1}.
\end{aligned}
$$

By induction principle, (5.8) holds for arbitrary $M>0$. As $\varepsilon\le\varepsilon_0\le\left(2C_1r^{-h}+2C_2\right)^{-\frac{1}{\eta}}$, we have $C\varepsilon^\eta\le 1/2$. Then (5.4) is proved as

$$
W_1\left(\mathcal{L}(X^{s,x}(s+nh)),\hat{\rho}_{s+nh}\right)\le\max\{\bar{\hat{\rho}}(V),V(x)\}\sum_{j=1}^{M+1}(C\varepsilon^\eta)^j\le 2\max\{\bar{\hat{\rho}}(V),V(x)\}C\varepsilon^\eta.
$$

At last, (5.5) follows from the dual representation of Wasserstein distance, definition of periodic measures and averaging of Wasserstein distance of (5.4) with respect to ρ_s:

$$
\begin{aligned}
W_1\left(\rho_s,\hat{\rho}_s\right)=&\sup_{f\in Lip(1)}\left|\int f(y)\rho_s(dy)-\int f(y)\hat{\rho}_s(dy)\right|\\
=&\sup_{f\in Lip(1)}\left|\iint f(y)P(s+m\tau,s,x,dy)\rho_s(dx)-\iint f(y)\hat{\rho}_s(dy)\rho_s(dx)\right|\\
\le&\int\sup_{f\in Lip(1)}\left|\int f(y)P(s+m\tau,s,x,dy)-\int f(y)\hat{\rho}_s(dy)\right|\rho_s(dx)
\end{aligned}
$$

$$
\begin{aligned}
&=\int W_1\left(\mathcal{L}(X^{s,x}(s+m\tau)),\hat{\rho}_{s+m\tau}\right)\rho_s(dx)\\
&\leq 2\{\bar{\hat{\rho}}(V)+\bar{\rho}(V)\}C\varepsilon^{\eta}. \quad \square
\end{aligned}
$$

Remark 5.2. In Proposition 5.1, the choice of increment h of Markov chain is arbitrary. However, it is worth noting that to impose fewer restrictions on

$$
\varepsilon_0=\min\left(\left(\frac{r}{R^2}\right)^{2h},\left(2C_1r^{-h}+2C_2\right)^{-\frac{1}{\eta}}\right),
$$

a smaller value of h would be preferable and more practical. In theoretical arguments, one could always fix h in advance, such as setting $h=\tau$, without introducing any rigorous issues in the proof.

In the following theorem, we assume the test function $\phi:\mathbb{R}^d\to\mathbb{R}$ being measurable and satisfying that there exist some constant $C>0$ and integer p such that for any $x,y\in\mathbb{R}^d$, we have

$$
|\phi(x)|\leq C(1+|x|^{2p})=CV_p(x) \tag{5.9}
$$

and

$$
|\phi(x)-\phi(y)|\leq C\left(1+|x|^{2p-1}+|y|^{2p-1}\right)|x-y|. \tag{5.10}
$$

The test functions $\phi\in\mathcal{G}_p$. Consider the SSEM scheme (2.5) generates a discretized periodic measure $\hat{\rho}_s^{\Delta t}$ for $s\in[0,\tau)$ with a step size $\Delta t=\frac{\tau}{K}$. To consider the numerical error and convergence rate of these two measures, we apply the law of large numbers following the ergodicity (see [17]) to approximate the measures as

$$
\hat{\rho}_s^{\Delta t}(\phi)=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\phi\left(\hat{X}_n\right). \tag{5.11}
$$

Similarly the approximation of the BEM scheme (2.6) generates a discretized periodic measure $\tilde{\rho}_s^{\Delta t}$ by the law of large numbers:

$$
\tilde{\rho}_s^{\Delta t}(\phi)=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\phi\left(\hat{Z}_n\right). \tag{5.12}
$$

Note that (5.11) and (5.12) are not used immediately in the following error analysis of the periodic measures. But they are extremely useful in simulation of periodic measures in the next section. In the next theorem, we obtain error estimates of the approximation to the periodic measure $\rho_\cdot$ by the discretized periodic measures $\hat{\rho}_\cdot^{\Delta t}$ and $\tilde{\rho}_\cdot^{\Delta t}$ in the Wasserstein distance in terms of the step size Δt.

Theorem 5.3. *Assume Conditions (A), (B) and (C). Then the Markovian process defined by SDE (1.1) has a periodic measure ρ_s, $s \in \mathbb{R}$ and there exist $C > 0$ and $\Delta t_0 > 0$ such that for any $\Delta t \in (0, \Delta t_0]$ and $s \in [0, \tau)$, we have*

$$W_1\left(\hat{\rho}_s^{\Delta t}, \rho_s\right) \leq C(\Delta t)^{\eta/2}, \quad W_1\left(\tilde{\rho}_s^{\Delta t}, \rho_s\right) \leq C(\Delta t)^{\eta/2}, \tag{5.13}$$

where $\eta = \frac{\delta}{\delta+\lambda}$ with δ given in Theorem 2.7 and λ given in Theorem 4.5.

Proof. By (5.3), Proposition 3.1, Proposition 3.5 and Theorem 4.5, we have that

$$\begin{aligned}
& W_1\left(\mathcal{L}(X_{nh}), \mathcal{L}(\hat{X}_n)\right) \\
\leq & \sup_{\phi \in \mathcal{G}_p} \left|\mathbb{E}\phi(X_{nh}) - \mathbb{E}\phi(\hat{X}_n)\right| \\
\leq & \sup_{\phi \in \mathcal{G}_p} \mathbb{E}\left|\phi(X_{nh}) - \phi(\hat{X}_n)\right| \\
\leq & \sqrt{3}\mathbb{E}\left[\left(1 + |X_{nh}|^{2p-1} + \left|\hat{X}_n\right|^{2p-1}\right)\left(\left|X_{nh} - \hat{X}_n\right|\right)\right] \\
\leq & C\left[\mathbb{E}\left(1 + |X_{nh}|^{4p-2} + \left|\hat{X}_n\right|^{4p-2}\right)\right]^{1/2}\left[\mathbb{E}\left(\left|X_{nh} - \hat{X}_n\right|^2\right)\right]^{1/2} \\
\leq & C\left(1 + |x|^{2p-1}\right)e^{\frac{\lambda}{2}nh}(\Delta t)^{1/2},
\end{aligned}$$

where $h = l\Delta t$ for some integer l and λ is obtained from Theorem 4.5. We take $R = e^{\frac{\lambda}{2}} \geq 1$ and $V_p = 1 + |x|^{2p}$ to achieve

$$\left|\mathbb{E}\phi(X_{nh}) - \mathbb{E}\phi(\hat{X}_n)\right| \leq CR^{nh}V_p(x)(\Delta t)^{1/2}. \tag{5.14}$$

Thus Hypothesis (II) is verified with Lyapunov function $V_p(x)$ and $\varepsilon = (\Delta t)^{1/2}$. Moreover, recall Theorem 2.7, in which the existence of discrete periodic measure $\hat{\rho}_{\cdot}^{\Delta t}$ was proved and Hypothesis (I) was verified for $\Delta t < \frac{1}{L_b}$, where L_b is the one-sided Lipschitz constant of b. We take $r = e^{-\delta} < 1$ and $V_p = 1 + |x|^{2p}$ to have

$$\sup_{\phi \in \mathcal{G}} \left|\mathbb{E}\phi(\hat{X}^{s,x}(s+nh)) - \hat{\rho}_{s+nh}(\phi)\right| \leq CV_p(x)e^{-\delta nh},$$

where δ is given in Theorem 2.7. It follows from Proposition 5.1 that

$$W_1(P(s+n\Delta t, s, x, \cdot), \hat{\rho}_{s+n\Delta t}^{\Delta t}) \leq 2C\max\{\bar{\hat{\rho}}(V), V(x)\}(\Delta t)^{1/2},$$

where $P(s+n\Delta t, s, x, \cdot) = \mathcal{L}(X^{s,x}(s+n\Delta t))$. Then triangle inequality implies that $\{\hat{\rho}_s^{\Delta t} : s \in \mathbb{R}\}$ forms a convergent family of measure valued functions with period τ as $\Delta t \to 0$. Denote this limit by ρ_s, $s \in \mathbb{R}$ and it is easy to see that $s \mapsto \rho_s$ is also τ-periodic. Note when $t = n\Delta t$ fixed, for any open set $\Gamma \in \mathbb{R}^d$,

$$\begin{aligned}\int_{\mathbb{R}^d} P(t+s,s,x,\Gamma)d\rho_s =& \int_{\mathbb{R}^d} P(t+s,s,x,\Gamma)d(\rho_s - \hat{\rho}_s^{\Delta t})(x) \\ &+ \int_{\mathbb{R}^d} \left(P(t+s,s,x,\Gamma) - \hat{P}_{n,0}^{\Delta t}(x,\Gamma)\right) d\hat{\rho}_s^{\Delta t}(x) \\ &+ \int_{\mathbb{R}^d} \hat{P}_{n,0}^{\Delta t}(x,\Gamma) d\hat{\rho}_s^{\Delta t}(x).\end{aligned} \tag{5.15}$$

For any fixed ε, the first term of the RHS is less than ε when Δt is sufficiently small as $\hat{\rho}_s^{\Delta t}$ tends to ρ_s in the Wasserstein W_1 metric and $P(t+s,s,x,\Gamma)$ is Lipschitz in x with bounded Lipschitz constant. For the second term, we use first

$$\begin{aligned}&\int_{\mathbb{R}^d} \left(P(t+s,s,x,\Gamma) - \hat{P}_{n,0}^{s}(x,\Gamma)\right) d\hat{\rho}_s^{\Delta t}(x) \\ =& \int_{V(x)\le M} \left(P(t+s,s,x,\Gamma) - \hat{P}_{n,0}^{s}(x,\Gamma)\right) d\hat{\rho}_s^{\Delta t}(x) \\ &+ \int_{V(x)> M} \left(P(t+s,s,x,\Gamma) - \hat{P}_{n,0}^{s}(x,\Gamma)\right) d\hat{\rho}_s^{\Delta t}(x).\end{aligned} \tag{5.16}$$

But from (2.18), we derive that

$$\hat{P}_{kK,0}^{s} V(x) \le (e^{-\frac{1}{2}\beta\tau})^k V(x) + \frac{C_2}{1-e^{-\frac{1}{2}\beta\tau}},$$

and from (2.19) it is easy to see that

$$\int_{\mathbb{R}^d} V(x) d\hat{\rho}_s^{\Delta t}(x) \le \frac{C_2}{1-e^{-\frac{1}{2}\beta\tau}} =: C.$$

Thus by Chebyshev's inequality we have

$$\hat{\rho}_s^{\Delta t}(V(x)\ge M) \le \frac{C}{M}.$$

So for any fixed ε, there exists $M>0$ such that $\hat{\rho}_s^{\Delta t}(V(x)\ge M) \le \frac{C}{M} < \frac{1}{2}\varepsilon$. It turns out that

$$\left| \int_{V(x)>M} \left(P(t+s,s,x,\Gamma) - \hat{P}_{n,0}^{s}(x,\Gamma)\right) d\hat{\rho}_s^{\Delta t}(x) \right| < \varepsilon. \tag{5.17}$$

For this particular chosen $M > 0$, on $x \in \{V(x) \leq M\}$, from the same convergence as that of (2.12), we have that $\left| P(t+s, s, x, \Gamma) - \hat{P}_{n,0}(x, \Gamma) \right| \to 0$ as $\Delta t \to 0$, thus when Δt is sufficiently small,

$$\left| \int\limits_{V(x)\leq M} \left(P(t+s, s, x, \Gamma) - \hat{P}^{s}_{n,0}(x, \Gamma) \right) d\hat{\rho}_s^{\Delta t}(x) \right| < \varepsilon. \tag{5.18}$$

This together with (5.17), gives the desired estimate the second term of (5.15). The last term of (5.15) gives that

$$\int\limits_{\mathbb{R}^d} \hat{P}^{\Delta t}_{n,0}(x, \Gamma) d\hat{\rho}_s^{\Delta t}(x) = \hat{\rho}^{\Delta t}_{n\Delta t+s}(\Gamma) = \hat{\rho}^{\Delta t}_{t+s}(\Gamma) \xrightarrow{\Delta t \to 0} \rho_{t+s}(\Gamma). \tag{5.19}$$

It then follows from (5.15), (5.16), (5.17), (5.18) and (5.19) that

$$\int\limits_{\mathbb{R}^d} P(t+s, s, x, \Gamma) d\rho_s = \rho_{t+s}(\Gamma), \tag{5.20}$$

for any $t \geq 0$, $s \in \mathbb{R}$ and Γ being an open set. Then by a standard π-system argument, we have that (5.20) holds for any Borel set $\Gamma \in \mathbb{R}^d$. Thus, ρ_s, $s \in \mathbb{R}$ is a periodic measure of the transition semigroup $P_{t,s}$.

It turns out now that the first claim of (5.13) follows from Lemma 5.1 where it is trivial to see that $\eta = \frac{\log r}{\log r - 2\log R} = \frac{\delta}{\delta+\lambda}$. Similarly, by Theorem 2.9 and Remark 4.6, one can verify Hypotheses (I) and (II) with respect to $\tilde{\rho}^{\Delta t}_{\cdot}$. The second conclusion of (5.13) also follows from Lemma 5.1 immediately. $\square$

Remark 5.4. (i) It is worth noting from the proof of Theorem 2.9 that we do not have to know the form of the periodic measure or even the existence of the periodic measure of the original Markovian system as a priori information for this analysis. Our result establishes the existence of a periodic measure following that of the discretized approximating system. This novel approach was made possible by ensuring that the convergence of the transition probability of the discretized approximating system to its periodic measure is uniform in the discretization step size Δt, as the pull-back time approaches infinity.

(ii) We can also use the ergodicity of the discrete numerical approximation to deduce the ergodicity of the original system. This can be seen from the following brief argument that for the SSEM scheme, we can use Theorem 2.7 and 5.3 to get:

$$\begin{aligned} W_1\left(P(t, s, x, \cdot), \rho_t\right) =& W_1\left(\mathcal{L}(X^{s,x}(t)), \rho_t\right) \\ \leq& W_1\left(\mathcal{L}(X^{s,x}(t)), \hat{\mathcal{L}}(X^{s,x}(t))\right) + W_1\left(\hat{\mathcal{L}}(X^{s,x}(t)), \hat{\rho}_t\right) + W_1\left(\hat{\rho}_t, \rho_t\right) \\ \leq& C R^{t-s} V_p(x)(\Delta t)^{\frac{1}{2}} + C(1+|x|^2)e^{-\delta(t-s)} + C|\Delta t|^{\frac{\eta}{2}}, \end{aligned}$$

and then let $\Delta t \to 0$ to conclude the result. For the BEM scheme, we can use Theorem 2.9 and Theorem 5.3 to get the result.

(iii) It is worth noting that in a special case with a further strong dissipative condition, which required that

$$2\langle x-y, b(t,x)-b(t,y)\rangle + \|\sigma(t,x)-\sigma(t,y)\|^2 \le -C\|x-y\|^2,$$

for some $C>0$, the numerical error does not aggregate with respect to time. One can follow the proof as in [15] to obtain a convergence result of paths with different initial values.

With the strong dissipative condition, the numerical error analysis in Lemma 4.2 and 4.4 gives estimates with negative constants K_1 and K_3. Thus, the conclusion in Theorem 4.5 becomes

$$\sup_{0\le i\le n} \mathbb{E}\left|X^{s,x}(s+i\Delta t)-\hat{X}_i^{\Delta t}\right|^2 \le Ce^{-\lambda' n\Delta t}(\Delta t),$$

for some $\lambda'>0$. Consequently, the numerical error does not aggregate with time-length of simulation, which ensures the Hypothesis (II) with $R=1$ in this case. Therefore the infinite horizon problem possesses the same order of numerical error as finite ones by $\eta=\frac{\log r}{\log r-2\log 1}=1$. In this special case, the result agrees with the result in [56] and our earlier result [15]. Needless to say that the results in this paper go much beyond the results requiring a strong dissipativity assumption e.g. in [15] and [56].

6. Numerical experiments

In this section, we simulate the Benzi-Parisi-Sutera-Vulpiani stochastic resonance model of the ice age transition in the dynamics of climate change with $b(t,x)=x-x^3+0.12\cos(0.001t)$ and $\sigma(x)=0.285(2+\cos(x))$. To rescale the period as integer $\tau=5000$, we take $b(t,x)=0.4\pi(x-x^3+0.12\cos(0.0004\pi t))$ and $\sigma(x)=0.285\times\sqrt{0.4\pi}(2+\cos(x))$.

In such a model, the computational difficulty in giving an approximation of the periodic measure comes from the fact that the period is very large. One way to conquer the difficulty would be to rescale the model to keep the resonance property by choosing a different scaling coefficient C as $b(t,x)=C(x-x^3+0.12\cos(0.001Ct))$ and $\sigma(x)=0.285\times\sqrt{C}$. But if we change the period, for example, to $\tau=5$, then the coefficients of the SDE will also change accordingly. As a result, $C=400\pi$ and so the approximation requires a very small step size to satisfy $\Delta t<1/2L_b=1/2C=1/800\pi\approx 0.0008$. With such a small step size, the computing time is essentially the same as the problem with the original large period. This does not seem to be an efficient way to address the problem.

From the ergodicity, we can compute the discrete periodic measure $\hat{\rho}$ from one sample path by the law of large numbers due to the ergodicity [19], see (5.11) and (5.12). But, to deal with such a large period problem, one can split the approximations into a collection of several sample paths with independent Brownian motions. We apply multiprocess computing to reduce the total cost of computation time. We may lose some accuracy from convergence at the beginning of the computation for each sample path. However, our method, unlike the Monte Carlo method, uses a small number of sample paths. Thus, this kind of error is, in fact, negligible by the geometric convergence for each sample path.

From Lemma 4.2, Lemma 4.4, Theorem 4.5 and their proofs, we see that the coefficient $\lambda=\max\{K_1,K_3\}=\max\{2(L_b+1)+L_\sigma, 2+\frac{L_b}{1-2L_b\Delta t}+3L_\sigma\}$. It is worth noting that

$$\lim_{\Delta t\to \frac{1}{2L_b}} \lambda = \lim_{\Delta t\to \frac{1}{2L_b}} K_3 = \lim_{\Delta t\to \frac{1}{2L_b}} \frac{L_b}{1-2L_b\Delta t} = \infty.$$

Thus, though it is allowed to take $\lambda < \frac{1}{2L_b}$ to be close to $\frac{1}{2L_b}$, it does not make very good sense to take λ to be close to $\frac{1}{2L_b}$. This results in a large value of λ, causing $\eta = \frac{\delta}{\lambda+\delta}$ to be small, and ultimately leading to a suboptimal error rate order. However, in the case of $L_\sigma < L_b$, if one chooses $\Delta t \le \frac{3L_b-2L_\sigma}{4L_b(L_b-L_\sigma)}$, then $\lambda = K_1$ and η is independent of Δt, and the error rate is always $(\Delta t)^{\eta/2}$.

In the BPSV model, $L_b = 0.4\pi \approx 1.2566$, $L_\sigma = 0.285\sqrt{0.4\pi} \approx 0.3195$. Thus $K_1 \approx 4.8327$ and an efficient choice of Δt makes $K_3 \approx K_1$, which gives $\Delta t \le \frac{1}{2L_b} - \frac{1}{K_1-2-3L_\sigma} \approx 0.1311$. So we choose $\Delta t = \frac{1}{80}, \frac{1}{64}, \frac{1}{50}, \frac{1}{20}, \frac{1}{16}, \frac{1}{10}$ in the numerical experiments. As the problem has no explicit solution, we may regard the outcome when we take $\Delta t = 1/400$ as the true solution. To achieve the best possible accuracy, we have performed the computations for the time up to 185000τ=185000×5000 for each Δt. With the help of multiprocess computing, we are able to generate $185000\times \frac{5000}{\Delta t}$ data for each Δt.

To compare the numerical error in the Wasserstein distance W_1, we generate approximations of $\hat{X}^{s,x}(s+i\Delta t)$ with different step size $\Delta t = \tau/K$ and collect the data on period points $\{\hat{X}^{s,x}(s+nK\Delta t)\}_{n=1,2,\dots,N}$. Then we sort the data in the ascending order for each Δt, denoted by $\{\hat{X}_n^{\Delta t}\}_{n=1,2,\dots,N}$. Finally, the numerical error in the Wasserstein distance W_1 is estimated by

$$W_1(\rho_s, \hat{\rho}_s^{\Delta t}) = \frac{1}{N}\sum_{n=1}^{N}\left|\hat{X}_n^{1/400} - \hat{X}_n^{\Delta t}\right|,$$

which agrees the result by applying the “wasserstein1d” function in the R language as mentioned in [17]. Here, due to our large amount of computations mentioned above, we can take $N = 185000$.

The numerical errors for $\Delta t = \frac{1}{80}, \frac{1}{64}, \frac{1}{50}, \frac{1}{20}, \frac{1}{16}, \frac{1}{10}$ are presented in Fig. 1 in the log-log scale. They align with the 0.5 order line very well.

We also simulate paths and take the values of $\hat{X}^{s,x}(s+n\tau)$ for $\{s_j\}_{j=0,1,\dots,7}$ where $s_j = \frac{j\tau}{8}$, can then compute their density for each s_j by the law of large numbers (5.11). The histogram presented as in Fig. 2 for each s_j illustrates the approximations to the density of the periodic measure ρ_{s_i}. It is noted that the density patterns change according to the periodic term $0.12\cos(0.0004\pi t)$ in the drift $b(t,x)$.

Finally, we summarize some important points from the numerical experiment as follows. The semi-implicit Euler schemes provide a numerical tool to estimate periodic measures of locally Lipschitz drift systems. The cost of the implicit schemes is higher than that of the explicit scheme discussed in [16], but they allow for larger step sizes in practical experiments and work for problems with polynomial growth coefficients. In addition, the Wasserstein distance has been studied to analyze the errors of the approximate periodic measure ρ_s for each $s \in [0,\tau)$ of the semi-implicit schemes to the true periodic measure. A rigorous proof of the error analysis is given, and the numerical experiment is carried out as a verification of the theoretical results. It is noted that the results of the numerics align with the theoretical results proved in this paper.

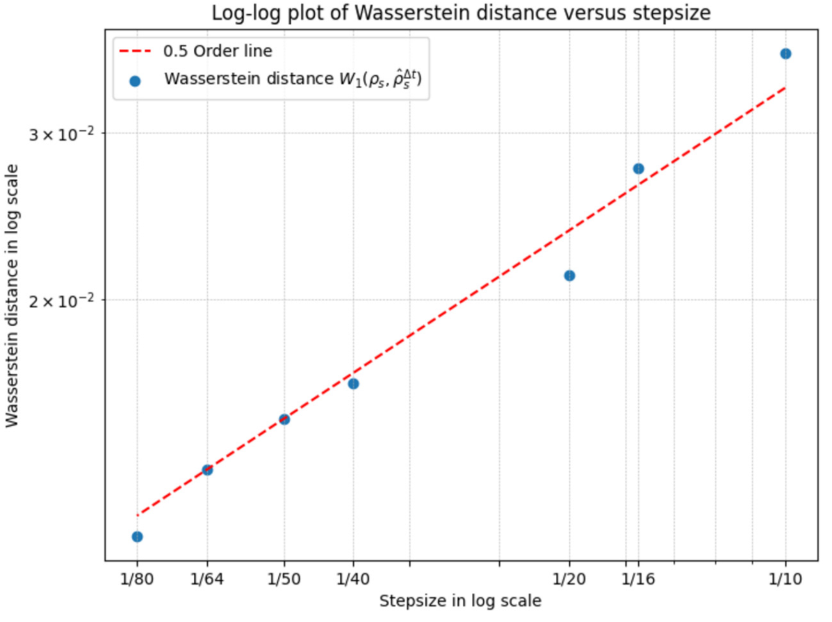


Fig. 1. Error of approximation to periodic measure versus step size in log-log graph.

Acknowledgments

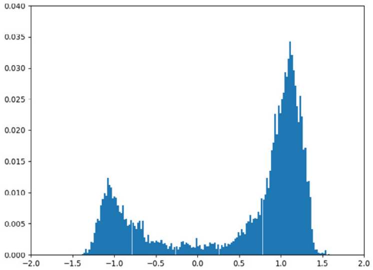
We would like to express our gratitude to the anonymous referee for their constructive comments which have helped us to improve the paper significantly. The project was supported by an EPSRC grant (ref. EP/S005293/2) and the Royal Society Newton Fund (ref. NIF\R1\221003).

Data availability

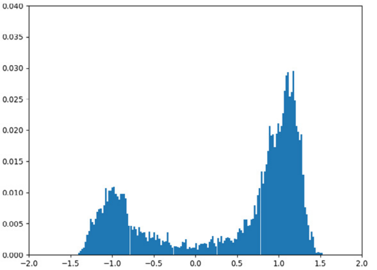
Data will be made available on request.

References

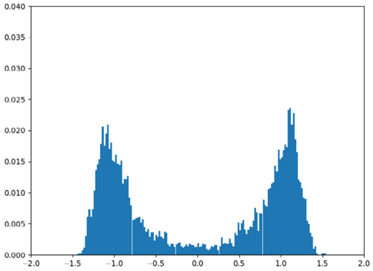
[1] L. Angeli, D. Crisan, M. Ottobre, Uniform in time convergence of numerical schemes for stochastic differential equations via strong exponential stability: Euler methods, split-step and tamed schemes, arXiv:2303.15463, 2023.
[2] J. Bao, Y. Wu, Random periodic solutions for stochastic differential equations with non-uniform dissipativity, arXiv:2202.09771, 2022.
[3] P.W. Bates, K. Lu, B. Wang, Attractors of non-autonomous stochastic lattice systems in weighted spaces, Phys. D 289 (2014) 32–50.
[4] F.E. Benth, J. Šaltytė Benth, The volatility of temperature and pricing of weather derivatives, Quant. Finance 7 (2007) 553–561.
[5] R. Benzi, G. Parisi, A. Sutera, A. Vulpiani, A theory of stochastic resonance in climatic change, SIAM J. Appl. Math. 43 (1983) 565–578.
[6] M. Branicki, K. Uda, Time-periodic measures, random periodic orbits, and the linear response for dissipative non-autonomous stochastic differential equations, Res. Math. Sci. 8 (2021) 42.
[7] D. Cheban, Z. Liu, Periodic, quasi-periodic, almost periodic, almost automorphic, Birkhoff recurrent and Poisson stable solutions for stochastic differential equations, J. Differ. Equ. 269 (2020) 3652–3685.
[8] M.D. Chekroun, E. Simonnet, M. Ghil, Stochastic climate dynamics: random attractors and time-dependent invariant measures, Phys. D 240 (2011) 1685–1700.



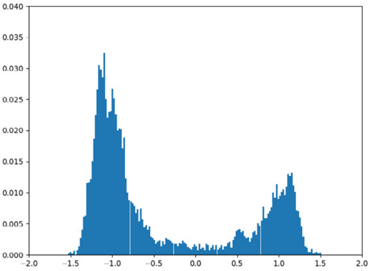
(a) Distribution of $\hat{\rho}_0$



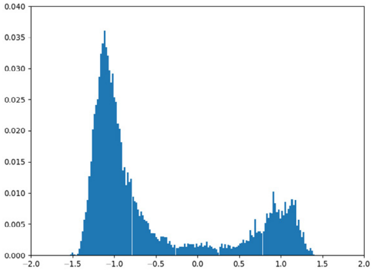
(b) Distribution of $\hat{\rho}_{625}$



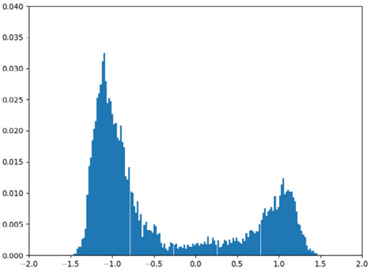
(c) Distribution of $\hat{\rho}_{1250}$



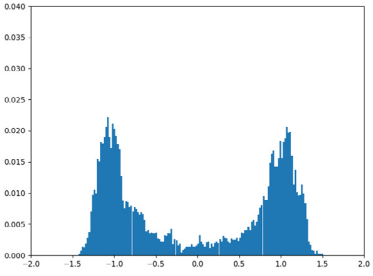
(d) Distribution of $\hat{\rho}_{1875}$



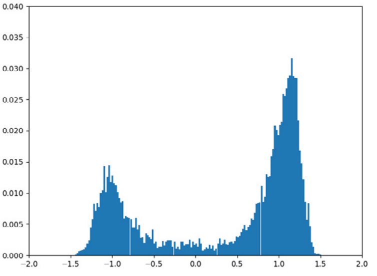
(e) Distribution of $\hat{\rho}_{2500}$



(f) Distribution of $\hat{\rho}_{3125}$



(g) Distribution of $\hat{\rho}_{3750}$



(h) Distribution of $\hat{\rho}_{4375}$

Fig. 2. Approximations of periodic measure with $\Delta t = 0.01$ and 10000 periods.

[9] Z. Chen, L. Cao, L. Chen, Stochastic theta methods for random periodic solution of stochastic differential equations under non-globally Lipschitz conditions, Numer. Algorithms (2024) 1–31.
[10] A.M. Cherubini, J.S.W. Lamb, M. Rasmussen, Y. Sato, A random dynamical systems perspective on stochastic resonance, Nonlinearity 30 (2017) 2835–2853.
[11] D. Crisan, P. Dobson, M. Ottobre, Uniform in time estimates for the weak error of the Euler method for SDEs and a pathwise approach to derivative estimates for diffusion semigroups, Trans. Am. Math. Soc. 374 (2021) 3289–3330.
[12] G. Da Prato, A. Lunardi, Ornstein-Uhlenbeck operators with time periodic coefficients, J. Evol. Equ. 7 (2007) 587–614.
[13] M. Engel, C. Kuehn, A random dynamical systems perspective on isochronicity for stochastic oscillations, Commun. Math. Phys. 386 (2021) 1603–1641.
[14] M. Engel, G. Olicón-Méndez, N. Wehlitz, S. Winkelmann, Synchronization and random attractors in reaction jump processes, J. Dyn. Differ. Equ. (2024).
[15] C. Feng, Y. Liu, H. Zhao, Numerical approximation of random periodic solutions of stochastic differential equations, Z. Angew. Math. Phys. 68 (2017) 119.
[16] C. Feng, Y. Liu, H. Zhao, Ergodic numerical approximation to periodic measures of stochastic differential equations, J. Comput. Appl. Math. 398 (2021) 113701.
[17] C. Feng, Y. Liu, H. Zhao, Periodic measures and Wasserstein distance for analysing periodicity of time series datasets, Commun. Nonlinear Sci. Numer. Simul. 120 (2023) 107166.
[18] C. Feng, B. Qu, H. Zhao, Random quasi-periodic paths and quasi-periodic measures of stochastic differential equations, J. Differ. Equ. 286 (2021) 119–163.
[19] C. Feng, H. Zhao, Random periodic processes, periodic measures and ergodicity, J. Differ. Equ. 269 (2020) 7382–7428.
[20] C. Feng, H. Zhao, J. Zhong, Expected exit time for time-periodic stochastic differential equations and applications to stochastic resonance, Phys. D 417 (2021) 132815.
[21] C. Feng, H. Zhao, J. Zhong, Existence of geometric ergodic periodic measures of stochastic differential equations, J. Differ. Equ. 359 (2023) 67–106.
[22] C. Feng, H. Zhao, B. Zhou, Pathwise random periodic solutions of stochastic differential equations, J. Differ. Equ. 251 (2011) 119–149.
[23] L. Gao, L. Yan, On random periodic solution to a neutral stochastic functional differential equation, Math. Probl. Eng. (2018) 8353065.
[24] P. Gao, Y. Liu, Y. Sun, Z. Zheng, Large deviations principle for stationary solutions of stochastic differential equations with multiplicative noise, arXiv:2206.02356, 2022.
[25] E. Hairer, G. Wanner, Solving Ordinary Differential Equations. II, second ed., Springer Series in Computational Mathematics, vol. 14, Springer-Verlag, Berlin, 1996. Stiff and differential-algebraic problems.
[26] M. Hairer, J.C. Mattingly, Yet another look at Harris' ergodic theorem for Markov chains, in: Seminar on Stochastic Analysis, Random Fields and Applications VI, in: Progr. Probab., vol. 63, Birkhäuser/Springer, Basel AG, Basel, 2011, pp. 109–117.
[27] T.E. Harris, The Existence of Stationary Measures for Certain Markov Processes, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. II, University of California Press, Berkeley-Los Angeles, Calif., 1956, pp. 113–124.
[28] D.J. Higham, X. Mao, A.M. Stuart, Strong convergence of Euler-type methods for nonlinear stochastic differential equations, SIAM J. Numer. Anal. 40 (2002) 1041–1063.
[29] Y. Hu, Semi-implicit Euler-Maruyama scheme for stiff stochastic equations, in: Stochastic Analysis and Related Topics, V, Silivri, 1994, in: Progr. Probab., vol. 38, Birkhäuser Boston, Boston, MA, 1996, pp. 183–202.
[30] W. Huang, Z. Lian, K. Lu, Dynamical complexity of Anosov systems driven by a quasi-periodic force, Sci. China, Math. 68 (2025) 89–136.
[31] A. Jentzen, P. Kloeden, Taylor expansions of solutions of stochastic partial differential equations with additive noise, Ann. Probab. 38 (2010) 532–569.
[32] R. Khasminskii, Stochastic Stability of Differential Equations, second ed., Stochastic Modelling and Applied Probability, vol. 66, Springer, Heidelberg, 2012. With contributions by G.N. Milstein and M.B. Nevelson.
[33] P.E. Kloeden, E. Platen, Numerical Solution of Stochastic Differential Equations, Applications of Mathematics (New York), vol. 23, Springer-Verlag, Berlin, 1992.
[34] M. Kunze, L. Lorenzi, A. Lunardi, Nonautonomous Kolmogorov parabolic equations with unbounded coefficients, Trans. Am. Math. Soc. 362 (2010) 169–198.
[35] R. Liu, K. Lu, Exponential mixing and limit theorems of quasi-periodically forced 2D stochastic Navier–Stokes equations in the hypoelliptic setting, Commun. Math. Phys. 406 (2025) 55.

[36] W. Liu, X. Mao, Y. Wu, The backward Euler-Maruyama method for invariant measures of stochastic differential equations with super-linear coefficients, Appl. Numer. Math. 184 (2023) 137–150.
[37] L. Lorenzi, A. Lunardi, A. Zamboni, Asymptotic behavior in time periodic parabolic problems with unbounded coefficients, J. Differ. Equ. 249 (2010) 3377–3418.
[38] J.C. Mattingly, A.M. Stuart, D.J. Higham, Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise, Stoch. Process. Appl. 101 (2002) 185–232.
[39] S. Menozzi, A. Pesce, X. Zhang, Density and gradient estimates for non degenerate Brownian SDEs with unbounded measurable drift, J. Differ. Equ. 272 (2021) 330–369.
[40] S.P. Meyn, R.L. Tweedie, Markov Chains and Stochastic Stability, Communications and Control Engineering Series, Springer-Verlag, London, Ltd., London, 1993.
[41] G.N. Milstein, Numerical Integration of Stochastic Differential Equations, Mathematics and Its Applications, vol. 313, Kluwer Academic Publishers Group, Dordrecht, 1995. Translated and revised from the 1988 Russian original.
[42] G.N. Milstein, M.V. Tretyakov, Stochastic Numerics for Mathematical Physics, Scientific Computation, Springer-Verlag, Berlin, 2004.
[43] A. Moradi, R. D'Ambrosio, Random periodic solutions of SDEs: existence, uniqueness and numerical issues, Commun. Nonlinear Sci. Numer. Simul. 128 (2024) 107586.
[44] J.M. Ortega, W.C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Classics in Applied Mathematics, vol. 30, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000. Reprint of the 1970 original.
[45] G. Pagès, F. Panloup, Unadjusted Langevin algorithm with multiplicative noise: total variation and Wasserstein bounds, Ann. Appl. Probab. 33 (2023) 726–779.
[46] P. Raynaud de Fitte, Almost periodicity and periodicity for nonautonomous random dynamical systems, Stoch. Dyn. 21 (2020) 2150034.
[47] J.M. Sanz-Serna, K.C. Zygalakis, Wasserstein distance estimates for the distributions of numerical approximations to ergodic stochastic differential equations, J. Mach. Learn. Res. 22 (2021) 242.
[48] T. Shardlow, A.M. Stuart, A perturbation theory for ergodic Markov chains and application to numerical approximations, SIAM J. Numer. Anal. 37 (2000) 1120–1137.
[49] J. Song, X. Song, Q. Zhang, Nonlinear Feynman-Kac formulas for stochastic partial differential equations with space-time noise, SIAM J. Math. Anal. 51 (2019) 955–990.
[50] A.M. Stuart, A.R. Humphries, Dynamical Systems and Numerical Analysis, Cambridge Monographs on Applied and Computational Mathematics, vol. 2, Cambridge University Press, Cambridge, 1996.
[51] D. Talay, Second-order discretization schemes of stochastic differential systems for the computation of the invariant law, Stoch. Stoch. Rep. 29 (1990) 13–36.
[52] D. Talay, L. Tubaro, Expansion of the global error for numerical schemes solving stochastic differential equations, Stoch. Anal. Appl. 8 (1990) 483–509, (1991).
[53] A. Tocino, R. Ardanuy, Runge-Kutta methods for numerical solution of stochastic differential equations, J. Comput. Appl. Math. 138 (2002) 219–241.
[54] K. Uda, Averaging principle for stochastic differential equations in the random periodic regime, Stoch. Process. Appl. 139 (2021) 1–36.
[55] B. Wang, Existence, stability and bifurcation of random complete and periodic solutions of stochastic parabolic equations, Nonlinear Anal. 103 (2014) 9–25.
[56] Y. Wu, C. Yuan, The Galerkin analysis for the random periodic solution of semilinear stochastic evolution equations, J. Theor. Probab. 37 (2024) 133–159.
[57] A. Yevik, H. Zhao, Numerical approximations to the stationary solutions of stochastic differential equations, SIAM J. Numer. Anal. 49 (2011) 1397–1416.
[58] C. Yuan, X. Mao, Stability in distribution of numerical solutions for stochastic differential equations, Stoch. Anal. Appl. 22 (2004) 1133–1150.
[59] T. Zhang, Y. Liu, H. Qu, Global mean-square exponential stability and random periodicity of discrete-time stochastic inertial neural networks with discrete spatial diffusions and Dirichlet boundary condition, Comput. Math. Appl. 141 (2023) 116–128.
[60] H. Zhao, Z.-H. Zheng, Random periodic solutions of random dynamical systems, J. Differ. Equ. 246 (2009) 2020–2038.
[61] M. Zhu, Y. Hu, Approximations of random periodic solutions for path-dependent stochastic differential equations with finite/infinite delay, Math. Methods Appl. Sci. (2024).
[62] M. Zhu, M. Tang, Approximation of random periodic solutions for neutral type sdes with non-uniform dissipativity, Front. Math. (2025).