Ergodic Numerical Approximation to Periodic Measures of Stochastic Differential Equations

Chunrong Feng[1], Yu Liu[1], and Huaizhong Zhao[1,2]

[1] *Department of Mathematical Sciences, Durham University, DH1 3LE, UK*
[2] *Research Centre for Mathematics and Interdisciplinary Sciences, Shandong University, Qingdao, Shandong, 266237, China*
chunrong.feng@durham.ac.uk, yu.liu3@durham.ac.uk, huaizhong.zhao@durham.ac.uk

Abstract

In this paper, we consider numerical approximation to periodic measure of a time periodic stochastic differential equations (SDEs) under weakly dissipative condition. For this we first study the existence of the periodic measure ρ_t and the large time behaviour of $\mathcal{U}(t+s,s,x) := \mathbb{E}\phi(X_t^{s,x}) - \int \phi d\rho_t$, where $X_t^{s,x}$ is the solution of the SDEs and ϕ is a test function being smooth and of polynomial growth at infinity. We prove $\mathcal{U}$ and all its spatial derivatives decay to 0 with exponential rate on time t in the sense of average on initial time s. We also prove the existence and the geometric ergodicity of the periodic measure of the discretized semi-flow from the Euler-Maruyama scheme and moment estimate of any order when the time step is sufficiently small (uniform for all orders). We thereafter obtain that the weak error for the numerical scheme of infinite horizon is of the order 1 in terms of the time step. We prove that the choice of step size can be uniform for all test functions ϕ. Subsequently we are able to estimate the average periodic measure with ergodic numerical schemes.

Keywords: Periodic measure; Fokker-Planck equation; discretized semi-flows; geometrical ergodicity; weak approximation.

Mathematics Subject Classifications (2010): 37H99, 60H10, 60H35.

1 Introduction

Random periodicity is ubiquitous in the real world from daily temperature process to economic cycles. The concepts of random periodic paths and periodic measures were introduced and their ergodicity was obtained recently ([10],[11],[12],[15],[37]). They are two different indispensable ways in the pathwise sense and in distributions to describe random periodicity. The "equivalence" of the random periodic solutions and periodic measures and their characterisation in terms of purely imaginary eigenvalues of the infinitesimal generator of the Markovian semigroup were obtained in [12]. The presence of pure imaginary eigenvalues distinguishes the random periodic processes/periodic measures regime from that of the 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.

As in the case of deterministic dynamical systems where periodic motion has been in the central stage of its study, the relevance of random periodic paths and periodic measures to theoretical and applied problems arising in stochastic dynamical systems has begun to be realised. In particular, there has been progress in the study of some topics in stochastic dynamics e.g. bifurcations (Wang [34]), random attractors (Bates, Lu and Wang [2]), stochastic resonance

(Cherubini, Lamb, Rasmussen and Sato [6], Feng, Zhao and Zhong [13],[14]), random horseshoes (Huang, Lian and Lu [19]), modelling the El Nîno phenomenon (Chekroun, Simonnet and Ghil [5]), isochronicity of stochastic oscillations (Engel and Kuehn [7]), and invariant measures of quasi-periodic stochastic systems (Feng, Qu and Zhao [9]).

However, it is difficult to construct random periodic solutions explicitly for many problems. So numerical approximation is critical in the study of stochastic dynamics in addition to the study of random periodic dynamics theory. There are numerous works on numerical analysis of SDEs on a finite horizon ([21],[25],[20],[26]). A numerical analysis of approximation to invariant measures of SDEs through discretizing the pull-back, was given in [23],[28],[30],[31],[35],[36]. Numerical approximations to stable zero solutions of SDEs were given in [17],[21]. Despite the importance both on the theoretical and applied aspects of the random periodic regime, its numerical analysis has barely been developed. The only result we know is the pathwise approximations of the random periodic solutions of SDEs discussed in [8]. In this paper we study the weak approximation to periodic measures.

We consider the following non-autonomous stochastic differential equations on $\mathbb{R}^d$

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

with initial condition $X_s = x$ where $b : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d, \sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$, 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 weakly dissipative in the space variable. For a technical reason, here we only consider the case when σ is time independent as we use the results in [13],[18]. Denote by $X_t^{s,x}$ the solution of (1.1) throughout the paper.

The existence of the periodic measure was studied in [13]. Under the assumption that the drift term is weakly dissipative and the diffusion term is non-degenerate, it was proved that the periodic measure ρ_s exists and has a density function, denoted by $q(s, x)$. To obtain $q(s, x)$, one could solve the corresponding infinite horizon Fokker-Planck equation $\frac{\partial}{\partial s}q(s, x) = \tilde{\mathcal{L}}^*(s)q(s, x)$ with additional condition $q(s + \tau, x) = q(s, x)$, where $\tilde{\mathcal{L}}^*(s)$ is the adjoint of the infinitesimal generator of the process X_t. This partial differential equation is generally difficult to solve explicitly. But we will show that theoretically it plays an essential role in establishing the theory of numerical schemes of weak approximations.

We apply numerical schemes such as Euler-Maruyama method to estimate the periodic measure. For any fixed $i \in \mathbb{Z}$, denoted by $\{\widehat{X}^{i\Delta t,x}_{i\Delta t+n\Delta t}\}_{n=0,1,\cdots}$ the discrete approximation of the solution of (1.1) with step size $\Delta t = \frac{\tau}{N}$ and $\widehat{X}^{i\Delta t,x}_{i\Delta t} = x$. We prove that the discrete semi-flow is geometrically ergodic and has a periodic measure $\hat{\rho}^{\Delta t}_i(\cdot)$, $i \in \mathbb{Z}$.

In this paper, we use the idea of lifting the flow and periodic measure to the cylinder $[0, \tau)\times\mathbb{R}^d$ proposed in [12]. With the help of this tool, our main result is to prove that the cumulation of discretization errors is of the order of $\mathcal{O}(\Delta t)$ for the approximation of the average of periodic measure, i.e. for any $\phi \in C^\infty_p\left(\mathbb{R}^d\right)$

$$\left|\int_{\mathbb{R}^d} \phi(x)\bar{\rho}(dx) - \int_{\mathbb{R}^d} \phi(x)\bar{\hat{\rho}}^{\Delta t}(dx)\right| \leq C\Delta t, \tag{1.2}$$

where $\bar{\rho} := \frac{1}{\tau}\int_0^\tau \rho_s ds$ ([12]), $\bar{\hat{\rho}}^{\Delta t} := \frac{1}{N}\sum_{i=0}^{N-1} \hat{\rho}^{\Delta t}_i$, $C^\infty_p\left(\mathbb{R}^d\right)$ is the space of smooth functions with the property that themselves and all their derivatives have at most polynomial growth at infinity. In fact, (1.2) only holds for Δt being small enough and the choice of step size can be uniform for all $\phi \in C^\infty_p$. For this, the uniformity of the step size working for all moment estimates of $\widehat{X}^{s,x}_{s+n\Delta t}$ is derived in Proposition 4.3. The error estimate (1.2) can also be numerically verified.

The results in this paper are applicable for many physically relevant SDEs, for instance, Benzi-Parisi-Sutera-Vulpiani's stochastic resonance model for the ice-age transition in climate change dynamics is SDE (1.1), with $b(t,x) = x - x^3 + A\cos(Bt)$ and $\sigma(x) = \sigma$ being constant ([4]). It was proved that this model has a unique periodic measure ([13]). This result implies the transition between ice-age and interglacial climates. A partial differential equation for expected transition time was given as well ([14]). This paper gives the weak approximation of numerical scheme for the SDE (1.1) with a modified drift which is nearly the same as the above b when $x \in [-4,4]$ and linear when x is far from this interval. This modified model provides the same climate dynamics as that of the original one of Benzi-Parisi-Sutera-Vulpiani since the global earth temperature cannot be outside of $[265, 305]$ in Kelvin scale.

We first study the lifts of semi-flows and corresponding Fokker-Planck equation for the density of the periodic measure. The infinitesimal generator $\tilde{\mathcal{L}}$ does not satisfy the non-degeneracy property with respect to initial time s. Under the weakly dissipative condition, we then obtain the exponential contraction of initial distribution to the periodic measure and all its spatial derivatives in the average with respect to initial time s. Finally, the numerical analysis on the cumulation of discretization errors is derived from these estimates and numerical experiments of error analysis are carried out for some specific SDEs arising in climate dynamics.

2 Preliminary results and notation

2.1 Lifts of semi-flows, random periodic paths and periodic measures

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 $\mathbb{R}^d$, where $\theta_s : \Omega \to \Omega$ defined by $(\theta_s\omega)(t) = W(t+s) - W(s)$, is measurably invertible for all $s \in \mathbb{R}$. Denote $\Delta := \{(t,s) \in \mathbb{R}^2, s \le t\}$ and let $u : \Delta \times \Omega \times \mathbb{R}^d \to \mathbb{R}^d$ be a periodic stochastic semi-flow of period τ satisfying for all $(t,s) \in \Delta$ and $r \in [s,t]$

$$u(t,r,\omega) \circ u(r,s,\omega) = u(t,s,\omega), \quad \text{for almost all } \omega \in \Omega,$$

and

$$u(t+\tau, s+\tau, \omega) = u(t,s,\theta_\tau\omega), \quad \text{for almost all } \omega \in \Omega.$$

Here $\tau > 0$ is a deterministic real number. Solutions of stochastic differential equations (1.1) with coefficients being periodic in time with period τ, when they exist and unique, generates a periodic semi-flow $u(t,s)x = X_t^{s,x}$, which satisfies the above two properties. As we consider periodic measures in this paper, so perfection is not needed here.

Consider the case when $u(t+s,s,\cdot)$ is a Markovian semi-flow on a filtered dynamical system $(\Omega, \mathcal{F}, \mathbb{P}, (\theta_t)_{t\in\mathbb{R}}, (\mathcal{F}_s^t)_{s\le t})$, i.e. for any $s,t,r \in \mathbb{R},\ s \le t$, we have $\theta_r^{-1}\mathcal{F}_s^t = \mathcal{F}_{s+r}^{t+r}$ and $u(t+s,s,\cdot)$ is independent of $\mathcal{F}_{-\infty}^s$, where $\mathcal{F}_{-\infty}^s := \bigvee_{r\le s}\mathcal{F}_r^s$. For any $\Gamma \in \mathcal{B}(\mathbb{R}^d)$, $t \in \mathbb{R}^+$, $s \in \mathbb{R}$, denote the transition probability of u by $P(t+s,s,x,\Gamma) = \mathbb{P}(\{\omega : u(t+s,s,\omega)x \in \Gamma\})$. From the periodicity of semi-flow u and the measure preserving property of θ_τ, the transition probability $P(t+s,s,x,\cdot)$ satisfies the periodic relation

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

Define for $\phi \in \mathcal{B}_b(\mathbb{R}^d)$, the space of bounded and Borel measurable function from $\mathbb{R}^d$ to $\mathbb{R}$,

$$\mathcal{T}(t+s,s)\phi(x) := \mathbb{E}\phi(X_{t+s}^{s,x}) = \int_{\mathbb{R}^d} \phi(y)P(t+s,s,x,dy),\ t \ge 0.$$

Then it is well-known that $\mathcal{T}(t+s,s) : \mathcal{B}_b(\mathbb{R}^d) \to \mathcal{B}_b(\mathbb{R}^d)$ defines a semigroup and satisfies the τ-periodic property:

$$\mathcal{T}(t+s+\tau, s+\tau) = \mathcal{T}(t+s,s).$$

This follows from (2.1) and the definition of $\mathcal{T}(t+s,s)$ easily. Moreover for any probability measure $\rho \in \mathcal{P}(\mathbb{R}^d)$, the space of probability measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, define

$$(\mathcal{T}^*(t+s,s)\rho)(\Gamma) = \int_{\mathbb{R}^d} P(t+s,s,x,\Gamma)\rho(dx).$$

The definition of periodic measure of the periodic Markovian semi-group is given as follows. The existence of the periodic measure was proved in [13] for a wide class of SDEs.

Definition 2.1. *([12]) The measure valued function $\rho : \mathbb{R} \to \mathcal{P}(\mathbb{R}^d)$ is called a τ-periodic measure of the τ-periodic Markovian semi-group $\mathcal{T}$ if*

$$\mathcal{T}^*(t+s,s)\rho_s = \rho_{t+s},\ \rho_{s+\tau} = \rho_s, \quad \forall s \in \mathbb{R},\ t \in \mathbb{R}^+. \tag{2.2}$$

The idea of lifting a stochastic periodic semi-flow to a cocycle on a cylinder in [12] plays an important role in this paper. As for this paper, the relevant part is briefly discussed below. Let $\mathbb{S} = [0,\tau) \times \mathbb{R}^d$, the lifted cocycle arising from SDE (1.1) with coordinates $\tilde{X}_s = (s, X_s) \in \mathbb{S}$ is given by

$$d\tilde{X}_t = \tilde{b}(\tilde{X}_t)dt + \tilde{\sigma}(\tilde{X}_t)d\tilde{W}(t),$$

where $\tilde{X}_0^{0,\tilde{x}} = \tilde{x} = (s,x)$, $\tilde{W} = (\tilde{W}_0, W)$, $\tilde{W}_0$ is a one-dimensional Brownian motion which is independent of W, $\tilde{b}(\tilde{X}_t) = \begin{pmatrix} 1 \\ b(t,X_t) \end{pmatrix}$, $\tilde{\sigma}(\tilde{X}_t) = \begin{pmatrix} 0 & 0 \\ 0 & \sigma(X_t) \end{pmatrix}$. One can enlarge the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, still denoted by $(\Omega, \mathcal{F}, \mathbb{P})$, as the canonical probability space for $\mathbb{R}^{d+1}$ Brownian motion $\tilde{W}$.

It is easy to see that the infinitesimal generator of the process $\tilde{X}$ is given by

$$\tilde{\mathcal{L}} = \sum_{i=1}^d b_i(s,x)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^d a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \frac{\partial}{\partial s} =: \mathcal{L}(s) + \frac{\partial}{\partial s}, \tag{2.3}$$

where $a = (a_{ij}) = \sigma\sigma^T$, and $\mathcal{U}(t+s,s,x) := \mathcal{T}(t+s,s)\phi(x)$ satisfies

$$\frac{\partial}{\partial t}\mathcal{U}(t+s,s,x) = \tilde{\mathcal{L}}\mathcal{U}(t+s,s,x), \quad \mathcal{U}(s,s,x) = \phi(x), \tag{2.4}$$

provided $\mathcal{U}$ is sufficiently smooth. Meanwhile, the transition probability and the periodic measure are lifted to

$$\begin{aligned}\tilde{\rho}_s(C \times \Gamma) =& \delta_{(s \ mod \ \tau)}(C)\rho_s(\Gamma),\\ \tilde{P}(t,\tilde{x}, C\times\Gamma) =& \delta_{(t+s \ mod \ \tau)}(C)P(t+s,s,x,\Gamma) = \mathbb{P}(\{\omega : X_{t+s}^{s,x} \in \Gamma\}),\end{aligned}$$

where $C \in \mathcal{B}([0,\tau))$ and $\Gamma \in \mathcal{B}(\mathbb{R}^d)$. It was shown in [12] that $\tilde{P}$ generates a homogeneous semigroup $\tilde{\mathcal{T}}$ defined by $(\tilde{\mathcal{T}}\tilde{\phi})(\tilde{x}) = \int_{\mathbb{S}} \tilde{\phi}(\tilde{y})\tilde{P}(t,\tilde{x},d\tilde{y})$ and $\tilde{\rho}_s$ is a periodic measure of the lifted semigroup $\tilde{\mathcal{T}}$. It was also noticed that $\bar{\tilde{\rho}} = \frac{1}{\tau}\int_0^\tau \tilde{\rho}_s ds$ is an invariant measure of $\tilde{P}$ following a standard procedure of Fubini theorem. It is easy to see that for a measurable function $\phi : \mathbb{R}^d \to \mathbb{R}$,

$$\begin{aligned}\int_0^\tau \int_{\mathbb{R}^d} \phi(x)\bar{\tilde{\rho}}(dt,dx) =& \frac{1}{\tau}\int_0^\tau \int_0^\tau \int_{\mathbb{R}^d} \phi(x)\delta_{(s \ mod \ \tau)}(dt)\rho_s(dx)ds \\ =& \frac{1}{\tau}\int_0^\tau \int_{\mathbb{R}^d} \phi(x)\rho_s(dx)ds \\ =& \int_{\mathbb{R}^d} \phi(x)\bar{\rho}(dx),\end{aligned} \tag{2.5}$$

where $\bar{\rho} = \frac{1}{\tau}\int_0^\tau \rho_s ds$. This will be used in later part of this paper.

2.2 Assumptions and some preliminary estimates

Assume

Condition (1) *The functions b, σ are of class C^∞ with σ being bounded, b and σ having bounded derivatives of any order and b being τ-periodic with respect to time.*

Condition (2) *(Uniform ellipticity) There exists a positive constant α such that for any $x, y \in \mathbb{R}^d$, we have $\sum_{i,j} a_{ij}(y)x_i x_j \geq \alpha |x|^2$.*

Condition (3) *(Weak dissipativity) There exist constants $\beta > 0$ and $C > 0$ such that for any $t \in \mathbb{R}^+$ and any $x \in \mathbb{R}^d$, $x \cdot b(t,x) \leq -\beta |x|^2 + C$.*

Under conditions (1)-(3), it was proved in [13] that the periodic measure $\rho : (-\infty, +\infty) \to \mathcal{P}(\mathbb{R}^d)$ exists and is geometrically ergodic:

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

We now discuss the existence of the density function $q(s,x)$ of the periodic measure ρ_s. Set the Fokker-Planck operator as follows

$$\tilde{\mathcal{L}}^*(s)\cdot = -\sum_{i=1}^{d} \frac{\partial}{\partial x_i}\left(b_i(s,x)\cdot\right) + \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2}{\partial x_i \partial x_j}\left(a_{ij}(x)\cdot\right),$$

and

$$\tilde{\mathcal{L}}^* = \tilde{\mathcal{L}}^*(s) - \frac{\partial}{\partial s}.$$

Proposition 2.2. *Assume Conditions (1), (2) and (3). Then the periodic measure ρ_s has a density $q(s,x)$ with respect to the Lebesgue measure in $\mathbb{R}^d$, and the density is the unique bounded solution of the Fokker-Planck equation*

$$\frac{\partial}{\partial s} q(s,x) = \tilde{\mathcal{L}}^*(s) q(s,x), \tag{2.6}$$

satisfying that for any $s \in [0,\tau)$, $q(s+\tau,x) = q(s,x)$ and $q(s,x) \to 0$ as $|x| \to \infty$.

Proof. Under the assumption of this proposition, the τ-periodic two-parameter Markov transition probability $P(t,s,x,\Gamma)$ has a density $p(t,s,x,y)$. Thus we have the representation of periodic measure ρ_s as follows, for any $\Gamma \in \mathcal{B}(\mathbb{R}^d)$,

$$\rho_s(\Gamma) = \int_{\mathbb{R}^d}\int_{\Gamma} p(s+\tau,s,x,y)dy\rho_s(dx) = \int_{\Gamma}\int_{\mathbb{R}^d} p(s+\tau,s,x,y)\rho_s(dx)dy,$$

where we applied Fubini's theorem. Hence we get the formula of the density of ρ_s as

$$q(s,y) = \int_{\mathbb{R}^d} p(s+\tau,s,x,y)\rho_s(dx). \tag{2.7}$$

It is easy to prove the periodicity of the density $q(s,y)$ by the periodic property of both $p(t,s,x,y)$ and ρ_s. Moreover, we have that for any $\Gamma \in \mathcal{B}(\mathbb{R}^d)$,

$$\begin{aligned}\rho_{t+\tau}(\Gamma) &= \int_{\Gamma}\int_{\mathbb{R}^d} p(t+\tau,s+\tau,z,y)\int_{\mathbb{R}^d} p(s+\tau,s,x,z)\rho_s(dx)dzdy \\ &= \int_{\Gamma}\int_{\mathbb{R}^d} p(t,s,z,y)q(s,z)dzdy.\end{aligned}$$

As the periodic measure satisfies $\rho_t(\Gamma) = \rho_{t+\tau}(\Gamma)$, the above implies

$$q(t,y) = \int_{\mathbb{R}^d} p(t,s,z,y)q(s,z)dz. \tag{2.8}$$

It is well known that $p(t,s,y,x)$ satisfies the Fokker-Planck equation $\partial_t p(t,s,y,x) = \tilde{\mathcal{L}}^*(t)p(t,s,y,x)$. Therefore,

$$\begin{aligned}
\partial_t q(t,x) &= \int_{\mathbb{R}^d} \tilde{\mathcal{L}}^*(t)p(t,s,y,x)q(s,y)dy \\
&= -\sum_{i=1}^d \frac{\partial}{\partial x_i}\left(b_i(t,x)\int_{\mathbb{R}^d} p(t,s,y,x)q(s,y)dy\right) \\
&\quad +\frac{1}{2}\sum_{i,j=1}^d \frac{\partial^2}{\partial x_i \partial x_j}\left(a_{ij}(x)\int_{\mathbb{R}^d} p(t,s,y,x)q(s,y)dy\right) \\
&= \tilde{\mathcal{L}}^*(t)\int_{\mathbb{R}^d} p(t,s,y,x)q(s,y)dy = \tilde{\mathcal{L}}^*(t)q(t,x),
\end{aligned}$$

which implies the density $q(s,y)$ satisfies the equation (2.6). The claim that $q(t,y) \to 0$ as $y \to \infty$ follows from (2.8) and the fact that when $|y| \to \infty$, we have $p(t,s,z,y) \to 0$. □

Corollary 2.3. *If the density function $q(s,x)$ of periodic measure satisfies the equation (2.6), then for any τ-periodic function $f \in C_p^\infty$, we have*

$$\int_0^\tau \int_{\mathbb{R}^d} \tilde{\mathcal{L}}f(s,x)q(s,x)dxds = 0.$$

Proof. The main ingredient of proof is to apply integration by parts. Note first

$$\begin{aligned}
\int_0^\tau \int_{\mathbb{R}^d} b_i(s,x)\frac{\partial}{\partial x_i}f(s,x)q(s,x)dxds &= -\int_0^\tau \int_{\mathbb{R}^d} f(s,x)\frac{\partial}{\partial x_i}\left(b_i(s,x)q(s,x)\right)dxds, \\
\int_0^\tau \int_{\mathbb{R}^d} a_{ij}(x)\frac{\partial^2}{\partial x_i x_j}f(s,x)q(s,x)dxds &= \int_0^\tau \int_{\mathbb{R}^d} f(s,x)\frac{\partial^2}{\partial x_i x_j}\left(a_{ij}(x)q(s,x)\right)dxds.
\end{aligned}$$

Here we used the property that $q(s,x)$ vanishes as x goes to ∞ when we performed the integration by parts. Applying the periodicity with respect to time s of function f and density function q in the third part, we have

$$\int_0^\tau \int_{\mathbb{R}^d} \frac{\partial}{\partial s}f(s,x)q(s,x)dxds = -\int_0^\tau \int_{\mathbb{R}^d} f(s,x)\frac{\partial}{\partial s}q(s,x)dxds.$$

Therefore, by the Fokker-Planck equation on the density function $q(s,x)$, we have

$$\int_0^\tau \int_{\mathbb{R}^d} \tilde{\mathcal{L}}f(s,x)q(s,x)dxds = \int_0^\tau \int_{\mathbb{R}^d} f(s,x)\tilde{\mathcal{L}}^*q(s,x)dxds = 0.$$

□

Proposition 2.4. *Assume Conditions (1) and (3). Then for any $p \in \mathbb{N}$, there exist strictly positive constants C_p and γ_p, such that for any $t > 0$ and $x \in \mathbb{R}^d$,*

$$\mathbb{E}\left|X_{t+s}^{s,x}\right|^p \le C_p(1 + |x|^p \exp(-\gamma_p t)). \tag{2.9}$$

Proof. Denote by $X_t := X^{s,x}_{t+s}$ for simplicity. Applying Itô's formula and Conditions (1), (3), we have the estimate

$$
\begin{aligned}
d\left(e^{\delta t}\left|X_t\right|^p\right) \leq& (\delta - p\beta)\, e^{\delta t}\left|X_t\right|^p dt + p\sigma(X_t)e^{\delta t}\left|X_t\right|^{p-1} dW_t \\
&+ \left(pC\sigma + \binom{p}{2}C_\sigma^2\right) e^{\delta t}\left|X_t\right|^{p-2} dt,
\end{aligned}
$$

where C_σ is the bound of function σ, C and β are the constants in the weakly dissipative condition. For convenience, here we denote $C_{p,\sigma} = pC + \binom{p}{2}C_\sigma^2$. Let τ_N be the first exit time of the process X_t from the ball of radius N. Consider the expectation of the integral $\mathbb{E}\int_0^{T\wedge\tau_N} \left|X_t\right|^p dW_t = 0$ for arbitrary p. Now take expectation on both sides after integrating from 0 to $T\wedge\tau_N$, together with Young's inequality, we have

$$
\begin{aligned}
&\mathbb{E}e^{\delta(T\wedge\tau_N)}\left|X_{T\wedge\tau_N}\right|^p \\
&\leq |x|^p + (\delta - p\beta)\,\mathbb{E}\int_0^{T\wedge\tau_N} e^{\delta t}\left|X_t\right|^p dt + C_{p,\sigma}\mathbb{E}\int_0^{T\wedge\tau_N} e^{\delta t}\left|X_t\right|^{p-2} dt \\
&\leq |x|^p + \frac{2C_{p,\sigma}}{p\delta\varepsilon^{\frac{p}{2}}}\mathbb{E}(e^{\delta(T\wedge\tau_N)} - 1) + K_1\mathbb{E}\int_0^{T\wedge\tau_N} e^{\delta t}\left|X_t\right|^p dt,
\end{aligned}
$$

where $K_1 = \delta - p\beta + \frac{(p-2)C_{p,\sigma}}{p}\varepsilon^{\frac{p}{p-2}}$, $\varepsilon < \left(\frac{p^2\beta}{(p-2)C_{p,\sigma}}\right)^{\frac{p-2}{p}}$ is chosen such that $K_1 - \delta < 0$. The choice of the constant δ guarantees $K_1 > 0$.

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

Then we let N go to ∞ to obtain

$$
e^{\delta T}\mathbb{E}\left|X_T\right|^p \leq |x|^p + \frac{2C_{p,\sigma}}{p\delta\varepsilon^{\frac{p}{2}}}(e^{\delta T} - 1) + K_1\int_0^T e^{\delta t}\mathbb{E}\left|X_t\right|^p dt. \tag{2.10}
$$

Apply Gronwall's inequality on (2.10),

$$
e^{\delta T}\mathbb{E}\left|X_T\right|^p \leq \frac{2C_{p,\sigma}}{p\delta\varepsilon^{\frac{p}{2}}}e^{\delta T} + e^{K_1 T}\left(|x|^p - \frac{2C_{p,\sigma}}{p\delta\varepsilon^{\frac{p}{2}}}\right) + \frac{2K_1C_{p,\sigma}}{p(\delta - K_1)\delta\varepsilon^{\frac{p}{2}}}(e^{\delta T} - e^{K_1 T}). \tag{2.11}
$$

Then (2.9) follows easily. □

Proposition 2.5. *Assume Conditions (1) and (3). Then for any* $p \in \mathbb{N}$, $\frac{1}{\tau}\int_0^\tau\int_{\mathbb{R}^d}|x|^p q(s,x)dxds \leq C_p$, *where* C_p *is determined from Proposition 2.4.*

Proof. For the density function of transition kernel $P(t+s, s, x, \cdot)$, there exists a constant C such that $|p(t+s,s,x,y)| \leq C$, for any $t \geq 1$. Then by dominated convergent theorem and Theorem 3.7 in [13], for any compact set $K \subset \mathbb{R}^d$, we have

$$
\begin{aligned}
&\frac{1}{\tau}\int_0^\tau \lim_{n\to\infty}\mathbb{E}\left(\left|X^{s,x}_{n\tau+s}\right|^p 1_K(X^{s,x}_{n\tau+s})\right) ds = \frac{1}{\tau}\int_0^\tau \lim_{n\to\infty}\int_K |y|^p p(n\tau+s,s,x,y)dyds \\
=&\frac{1}{\tau}\int_0^\tau\int_K |y|^p \lim_{n\to\infty} p(n\tau+s,s,x,y)dyds = \frac{1}{\tau}\int_0^\tau\int_K |y|^p q(s,y)dyds.
\end{aligned}
$$

Thus the average of periodic measure possesses finite moments of any order on any compact set K from the estimates in Proposition 2.4. Note the bound can be independent of K. The result follows from taking limit $K \uparrow \mathbb{R}^d$ and Fatou's Lemma. □

Consider the sequence $\{X_{t_n}\}_{n\in\mathbb{N}}$ with $t_n = n\tau$. We prove

Proposition 2.6. *Assume Conditions (1) and (3), then there exists a constant $r > 1$ and a ball $B(0,R)$, such that,*

$$\sup_{x\in B(0,R)^c} \mathbb{E}\left[r\left|X_{t_{n+1}}\right|^2 - \left|X_{t_n}\right|^2 \middle| X_{t_n} = x \right] < 0.$$

Proof. Apply the result (2.11) in Proposition 2.4 with $p = 2$, we have

$$\mathbb{E}\left[r\left|X_{t_{n+1}}\right|^2 - \left|X_{t_n}\right|^2 \middle| X_{t_n} = x \right] \leq \left(|x|^2 e^{(K_1-\delta)\tau} + C_{\beta,\sigma}(1 - e^{(K_1-\delta)\tau}) \right) r - |x|^2 .$$

In order to make the right hand side of the above negative, we need $(1 - re^{(K_1-\delta)\tau})\,|x|^2 > rC_{\beta,\sigma}(1-e^{(K_1-\delta)\tau})$. As $K_1-\delta < 0$, there always exists a constant r to ensure $1-re^{(K_1-\delta)\tau} > 0$ for the given period τ. Then the ball $B(0,R)$ is determined by taking $R > \sqrt{\frac{rC_{\beta,\sigma}(1-e^{(K_1-\delta)\tau})}{1-re^{(K_1-\delta)\tau}}}$. □

Let the function $\phi \in C_p^\infty$ and $\mathcal{U}(t+s,s,x) = \mathbb{E}\phi(X_{t+s}^{s,x})$. Then $\mathcal{U}$ satisfies the PDE (2.4). Considering the spatial differentiation of the solution with respect to x, Kunita showed in [22] that the function $\mathcal{U}(t+s,s,x)$ satisfies that for any order $n \in \mathbb{N}$, there exists an integer $r_n \in \mathbb{N}$ such that for any $T > 0$, $\exists C_n(t) > 0$,

$$|D^n\mathcal{U}(t+s,s,x)| \leq C_n(T)(1+|x|^{r_n}),\ \forall t < T. \tag{2.12}$$

From Proposition 2.5, the average of periodic measure possesses finite moments of any order. Together with (2.12), we have that the initial condition ϕ and $D^n\mathcal{U}(t+s,s,x)$ belong to $L^2(\mathbb{R}^{d+1},\bar{\tilde{\rho}})$.

Note that the function $\frac{1}{\tau}\int_0^\tau \mathcal{U}(t+s,s,x)ds$ has the same spatial derivatives as $\frac{1}{\tau}\int_0^\tau \mathcal{U}(t+s,s,x)ds - \frac{1}{\tau}\int_0^\tau\int_{\mathbb{R}^d}\phi(x)q(s,x)dxds$. Without loss of generality, in the following sections, we assume that

$$\int_0^\tau \int_{\mathbb{R}^d} \phi(x)q(s,x)dxds = 0. \tag{2.13}$$

Note when $\tilde{\phi}(\tilde{x}) = \phi(x)$, we have $\int_{\mathbb{S}} \tilde{\phi}(\tilde{x})d\bar{\tilde{\rho}}(\tilde{x}) = \frac{1}{\tau}\int_0^\tau\int_{\mathbb{R}^d}\phi(x)q(s,x)dxds$, where $\bar{\tilde{\rho}}$ is the average of lifted periodic measure, which is the invariant measure of the lifted Markov semigroup. It is easy to know that $\bar{\tilde{\rho}}(d\tilde{x}) = q(s,x)dxds$

For simplicity, in the following sections, we may often write $\tilde{\mathcal{U}}(t)$ or $\mathcal{U}(t+s,s)$ to represent the function $\mathcal{U}(t+s,s,x)$. We also often write b to represent $b(s,x)$ as we have the uniform conditions for the function and any order of its derivatives in Condition (1). The operators ∂_i, ∂_{ij}, ∇ and D^k on function $\mathcal{U}(t+s,s,x)$ always refer to derivatives with respect to spatial coordinates. The derivatives with respect to initial time will stay as $\frac{\partial}{\partial s}$.

3 Exponential decay of initial distribution and spatial derivatives

3.1 Estimates on the average of $\mathcal{U}(t+s,s)$ on a ball

We always assume (2.13) in this section unless otherwise stated.

Lemma 3.1. *Assume Conditions (1), (2) and (3). Then for any ball B, there exist strictly positive constants C and λ such that for any $t > 0$ and any $x \in B$, $\mathcal{U}$ defined with ϕ satisfying (2.13) has the following estimate:* $\frac{1}{\tau}\int_0^\tau |\mathcal{U}(t+s,s,x)|\,ds \le C\exp(-\lambda t)$.

Proof. First we apply mathematical induction to obtain that for any $p \in \mathbb{N}^+$, there exist constants $C_p > 0$, $\gamma_p > 0$ such that

$$\frac{1}{\tau}\int_0^\tau \int_{\mathbb{R}^d} |D^p\mathcal{U}(t+s,s,x)|^2\, q(s,x)dxds \le C_p\exp(-\gamma_p t). \tag{3.1}$$

We start to prove the basis step, when $p = 1$. Consider the Markov chain $\{X_{t_n}\}_{n\in\mathbb{N}}$ with $t_n = n\tau$. In [13], it was proved that the transition kernel $P(s, s+k\tau, x, \Gamma)$ is irreducible. With the result of Proposition 2.6, one can find some compact set K and a constant $\beta > 0$ such that for any $x \in K^c$, we have

$$\mathbb{E}\left|X_t^{0,x}\right|^2 - \frac{1}{r}\mathbb{E}x^2 < 0,$$

where $1/r < 1$. Now we take $V(x) = x^2$ as the norm-like function and from Proposition 2.4, we obtain that the norm-like function $V(x) = x^2$ is finite on the compact set K. Combining the above results, we have that

$$(P(t_1,0)V)(x) = \mathbb{E}X_{t_1}^2 \le \frac{1}{r}x^2 + \beta = \frac{1}{r}V(x) + \beta,$$

where β is a positive number. Thus the condition of Theorem 3.7 in [13] is satisfied. So the Markov chain $\{X_{t_n}\}_{n\in\mathbb{N}}$ is geometrically ergodic. That is for those $\phi \in C_p^\infty$ with the assumption (2.13), there exist strictly positive constants C and λ such that for any n,

$$\frac{1}{\tau}\int_0^\tau \int_{\mathbb{R}^d} \left|\mathbb{E}\phi(X_{t_n+s}^{s,x})\right| q(s,x)dxds \le Ce^{-\lambda t_n}. \tag{3.2}$$

As function ϕ has at most polynomial growth at infinity, we have $|\phi(x)| \le C|x|^N$ for some integer $N \in \mathbb{N}$. By Proposition 2.4, there exist $C_0 > 0$, $\gamma > 0$ such that

$$|\mathcal{U}(t+s,s,x)| \le C_0(1+|x|^N\exp(-\gamma t)). \tag{3.3}$$

Applying (3.3) and (3.2), together with Proposition 2.5, we have that for any n,

$$\begin{aligned}
&\frac{1}{\tau}\int_0^\tau \int_{\mathbb{R}^d} |\mathcal{U}(t_n+s,s,x)|^2 q(s,x)dxds \\
\le&\frac{C_0}{\tau}\int_0^\tau \int_{\mathbb{R}^d} |\mathcal{U}(t_n+s,s,x)|\,(1+|x|^N\exp(-\gamma t_n))q(s,x)dxds \\
\le&C_0C\exp(-\lambda t_n) + \frac{C_0^2\exp(-\gamma t_n)}{\tau}\int_0^\tau \int_{\mathbb{R}^d} |x|^N q(s,x)dxds \\
\le&C_1\exp(-\lambda_1 t_n).
\end{aligned} \tag{3.4}$$

In the following, we prove that the function $t \to \frac{1}{\tau}\int_0^\tau\int_{\mathbb{R}^d}|\mathcal{U}(t+s,s,x)|^2 q(s,x)dxds$ is monotonic. For this, note

$$\frac{d}{dt}|\mathcal{U}(t+s,s,x)|^2 = \tilde{\mathcal{L}}\,|\mathcal{U}(t+s,s,x)|^2 - a_{ij}\partial_i\mathcal{U}(t+s,s,x)\partial_j\mathcal{U}(t+s,s,x),$$

and

$$\frac{1}{\tau}\int_0^\tau \int_{\mathbb{R}^d} \tilde{\mathcal{L}}\,|\mathcal{U}(t+s,s,x)|^2 q(s,x)dxds = \frac{1}{\tau}\int_0^\tau \int_{\mathbb{R}^d} |\mathcal{U}(t+s,s,x)|^2\,\tilde{\mathcal{L}}^* q(s,x)dxds = 0.$$

It turns out from the elliptic condition (2) that

$$\begin{aligned}
&\frac{d}{dt}\left(\frac{1}{\tau}\int_0^\tau\int_{\mathbb{R}^d}|\mathcal{U}(t+s,s,x)|^2 q(s,x)dxds\right)\\
=&\frac{1}{\tau}\int_0^\tau\int_{\mathbb{R}^d}\tilde{\mathcal{L}}\,|\mathcal{U}(t+s,s,x)|^2 q(s,x)dxds\\
&-\frac{1}{\tau}\int_0^\tau\int_{\mathbb{R}^d}a_{ij}(x)\partial_i\mathcal{U}(t+s,s,x)\partial_j\mathcal{U}(t+s,s,x)q(s,x)dxds\\
\leq&-\frac{\alpha}{\tau}\int_0^\tau\int_{\mathbb{R}^d}|\nabla\mathcal{U}(t+s,s,x)|^2 q(s,x)dxds\leq 0.
\end{aligned}$$

This implies that $\frac{1}{\tau}\int_0^\tau\int_{\mathbb{R}^d}|\mathcal{U}(t+s,s,x)|^2 q(s,x)dxds$ is decreasing in t. Thus by (3.4), we have that for any t,

$$\frac{1}{\tau}\int_0^\tau\int_{\mathbb{R}^d}|\mathcal{U}(t+s,s,x)|^2 q(s,x)dxds\leq C_2\exp(-\lambda_1 t). \tag{3.5}$$

The above shows that the exponential contraction of $\mathcal{U}(t+s,s,x)$ under the average of periodic measure holds for any t.

On the other hand, by Condition (2) we have

$$\frac{d}{dt}|\mathcal{U}(t+s,s,x)|^2-\tilde{\mathcal{L}}\,|\mathcal{U}(t+s,s,x)|^2\leq-\alpha\,|\nabla\mathcal{U}(t+s,s,x)|^2. \tag{3.6}$$

Multiplying the above inequality with $e^{\delta t}$, and integrating both sides with respect to the average periodic measure $\bar{\tilde{\rho}}$ and time t, together with Corollary 2.3, we obtain for arbitrary $T>0$,

$$\begin{aligned}
&\int_0^T e^{\delta t}\int_{\mathbb{S}}\frac{d}{dt}|\mathcal{U}(t+s,s,x)|^2 d\bar{\tilde{\rho}}dt+\int_0^T\alpha e^{\delta t}\int_{\mathbb{S}}|\nabla\mathcal{U}(t+s,s,x)|^2 d\bar{\tilde{\rho}}dt\\
\leq&\int_0^T e^{\delta t}\int_{\mathbb{S}}\tilde{\mathcal{L}}\,|\mathcal{U}(t+s,s,x)|^2 d\bar{\tilde{\rho}}dt=0.
\end{aligned}\tag{3.7}$$

Here $\mathbb{S}=[0,\tau)\times\mathbb{R}^d$. Integration by parts on the first term of (3.7) gives us

$$\begin{aligned}
&\int_0^T e^{\delta t}\int_{\mathbb{S}}\frac{d}{dt}|\mathcal{U}(t+s,s,x)|^2 d\bar{\tilde{\rho}}dt\\
=&e^{\delta T}\int_{\mathbb{S}}|\mathcal{U}(T+s,s,x)|^2 d\bar{\tilde{\rho}}-\int_{\mathbb{S}}|\mathcal{U}(s,s,x)|^2 d\bar{\tilde{\rho}}-\delta\int_0^T e^{\delta t}\int_{\mathbb{S}}|\mathcal{U}(t+s,s,x)|^2 d\bar{\tilde{\rho}}dt,
\end{aligned}$$

where we have the initial condition that $\mathcal{U}(s,s,x)=\tilde{\phi}(\tilde{x})$. By Proposition 2.4 and $\phi\in C_p^\infty$ of the function ϕ, we have a constant $C_3>0$ such that $\int_{\mathbb{S}}\left|\tilde{\phi}(\tilde{x})\right|^2 d\bar{\tilde{\rho}}<C_3$. Consider (3.5) and take $\delta<\lambda_1$,

$$\delta\int_0^T e^{\delta t}\int_{\mathbb{S}}|\mathcal{U}(t+s,s,x)|^2 d\bar{\tilde{\rho}}dt\leq\delta\int_0^T e^{\delta t}C_2 e^{(-\lambda_1 t)}dt=\frac{C_2\delta}{\lambda_1-\delta}(1-e^{(\delta-\lambda_1)T})\leq C_4$$

for a constant $C_4>0$. Applying these results to (3.7), we obtain that for any T and any $s\in[0,\tau)$,

$$\begin{aligned}
&\int_0^T e^{\delta t}\int_{\mathbb{S}}|\nabla\mathcal{U}(t+s,s,x)|^2 d\bar{\tilde{\rho}}dt\\
\leq&\frac{1}{\alpha}\left(\int_{\mathbb{S}}|\mathcal{U}(s,s,x)|^2 d\bar{\tilde{\rho}}+\delta\int_0^T e^{\delta t}\int_{\mathbb{S}}|\mathcal{U}(t+s,s,x)|^2 d\bar{\tilde{\rho}}dt\right)\leq C_5,
\end{aligned}\tag{3.8}$$

where $C_5 = C_3 + C_4$. Now let's consider $|\nabla\mathcal{U}(t+s,s,x)|^2$ and note that

$$
\begin{aligned}
&\frac{d}{dt}|\nabla\mathcal{U}(t+s,s,x)|^2 - \tilde{\mathcal{L}}|\nabla\mathcal{U}(t+s,s,x)|^2 \\
=&-a_{ij}(\partial_{ik}\mathcal{U}(t+s,s,x))(\partial_{jk}\mathcal{U}(t+s,s,x)) + 2(\partial_k b_i)(\partial_k\mathcal{U}(t+s,s,x))(\partial_i\mathcal{U}(t+s,s,x)) \\
&+(\partial_k a_{ij})(\partial_k\mathcal{U}(t+s,s,x))(\partial_{ij}\mathcal{U}(t+s,s,x)).
\end{aligned}
$$

Applying Young's inequality with ε, we have

$$
\begin{aligned}
&(\partial_k a_{ij})(\partial_k\mathcal{U}(t+s,s,x))(\partial_{ij}\mathcal{U}(t+s,s,x)) \\
\le&\frac{\varepsilon}{2}(\partial_{ij}\mathcal{U}(t+s,s,x)))^2 + \frac{(\partial_k a_{ij}\partial_k\mathcal{U}(t+s,s,x)))^2}{2\varepsilon}.
\end{aligned}
$$

From Conditions (1) and (2), we can choose ε small enough such that $-\alpha + \frac{\varepsilon}{2} < 0$. It turns out that there exist strictly positive constants C_6 and C_7 such that

$$
\begin{aligned}
&\frac{d}{dt}|\nabla\mathcal{U}(t+s,s,x)|^2 - \tilde{\mathcal{L}}|\nabla\mathcal{U}(t+s,s,x)|^2 \\
\le&-C_6\left|D^2\mathcal{U}(t+s,s,x)\right|^2 + C_7|\nabla\mathcal{U}(t+s,s,x)|^2.
\end{aligned}
\tag{3.9}
$$

We choose $\gamma < \delta$ and multiply $e^{\gamma t}$ on both sides of the above inequality. It follows that

$$
\begin{aligned}
&\int_0^T e^{\gamma t}\int_{\mathbb{S}}\frac{d}{dt}|\nabla\mathcal{U}(t+s,s,x)|^2 d\bar{\bar{\rho}}dt - \int_0^T e^{\gamma t}\int_{\mathbb{S}}\tilde{\mathcal{L}}|\nabla\mathcal{U}(t+s,s,x)|^2 d\bar{\bar{\rho}}dt \\
\le&-C_6\int_0^T e^{\gamma t}\int_{\mathbb{S}}\left|D^2\mathcal{U}(t+s,s,x)\right|^2 d\bar{\bar{\rho}}dt + C_7\int_0^T e^{\gamma t}\int_{\mathbb{S}}|\nabla\mathcal{U}(t+s,s,x)|^2 d\bar{\bar{\rho}}dt.
\end{aligned}
$$

Following Corollary 2.3, we see that $\int_0^T e^{\gamma t}\int_{\mathbb{S}}\tilde{\mathcal{L}}|\nabla\mathcal{U}(t+s,s,x)|^2 d\bar{\bar{\rho}}dt = 0$. Thus

$$
\int_0^T e^{\gamma t}\int_{\mathbb{S}}\frac{d}{dt}|\nabla\mathcal{U}(t+s,s,x)|^2 d\bar{\bar{\rho}}dt \le C_7\int_0^T e^{\gamma t}\int_{\mathbb{S}}|\nabla\mathcal{U}(t+s,s,x)|^2 d\bar{\bar{\rho}}dt.
$$

Now by integration by parts, we note that

$$
\begin{aligned}
&\int_0^T e^{\gamma t}\int_{\mathbb{S}}\frac{d}{dt}|\nabla\mathcal{U}(t+s,s,x)|^2 d\bar{\bar{\rho}}dt \\
=&e^{\gamma T}\int_{\mathbb{S}}|\nabla\mathcal{U}(T+s,s,x)|^2 d\bar{\bar{\rho}} - \int_{\mathbb{S}}|\nabla\mathcal{U}(s,s,x)|^2 d\bar{\bar{\rho}} - \gamma\int_0^T e^{\gamma t}\int_{\mathbb{S}}|\nabla\mathcal{U}(t+s,s,x)|^2 d\bar{\bar{\rho}}dt.
\end{aligned}
$$

Then apply (3.8) with γ small enough and the boundedness of $\int_{\mathbb{S}}|\nabla\tilde{\phi}|^2 d\bar{\bar{\rho}}$ to have

$$
\begin{aligned}
&e^{\gamma T}\int_{\mathbb{S}}|\nabla\mathcal{U}(T+s,s,x)|^2 d\bar{\bar{\rho}} \\
\le&\int_{\mathbb{S}}\left|\nabla\tilde{\phi}\right|^2 d\bar{\bar{\rho}} + (\gamma + C_7)\int_0^T e^{\gamma t}\int_{\mathbb{S}}|\nabla\mathcal{U}(t+s,s,x)|^2 d\bar{\bar{\rho}}dt \le C_8.
\end{aligned}
$$

Thus we obtained (3.1) for the case when $p = 1$. Now we continue to prove the induction step in the following content. Assume that for any $k \le m$, there exist strictly positive constants C_k and γ_k such that for any $t > 0$,

$$
\frac{1}{\tau}\int_0^\tau\int_{\mathbb{R}^d}\left|D^k\mathcal{U}(t+s,s,x)\right|^2 q(s,x)dxds \le C_k\exp(-\gamma_k t).
$$

Here we need to compare the expansion of the operators $\frac{d}{dt}$ and $\tilde{\mathcal{L}}$ in the following:

$$|D^m\mathcal{U}(t+s,s,x)|^2 = \sum_{|J|=m} (\partial_J \mathcal{U}(t+s,s,x))^2,$$

where J is the multi-index with length $|J| = m$. The multi-indices J_a and J_b are introduced for the following identity,

$$\begin{aligned}
&\frac{d}{dt}|D^m\mathcal{U}(t+s,s,x)|^2 - \tilde{\mathcal{L}}|D^m\mathcal{U}(t+s,s,x)|^2 \\
=& -a_{ij}\left(\partial_J\partial_i\mathcal{U}(t+s,s,x)\right)\left(\partial_J\partial_j\mathcal{U}(t+s,s,x)\right) \\
&+ \sum_{|J_a|+|J_b|\le 2m+1} \Phi^J_{J_a,J_b}\partial_{J_a}\mathcal{U}(t+s,s,x)\partial_{J_b}\mathcal{U}(t+s,s,x).
\end{aligned}$$

Here the notation $\Phi^J_{J_a,J_b}$ contains all the combinations of spatial derivatives on the functions a and b with respect multi-indices J_a and J_b under some specified J. It is obvious the length of J_a and J_b will not exceed $m+1$. The boundedness of each elements in $\Phi^J_{J_a,J_b}$ comes from Condition (1). Therefore we will always have the following result by Young's inequality,

$$\begin{aligned}
&\frac{d}{dt}|D^m\mathcal{U}(t+s,s,x)|^2 - \tilde{\mathcal{L}}|D^m\mathcal{U}(t+s,s,x)|^2 \\
\le& -C_1^m\left|D^{m+1}\mathcal{U}(t+s,s,x)\right|^2 + C_2^m\sum_{k\le m}\left|D^k\mathcal{U}(t+s,s,x)\right|^2.
\end{aligned}$$

Then we choose a strictly positive constant δ_{m+1} small enough to proceed as in (3.8). Multiplying $e^{\delta_{m+1}t}$ on both sides and integrating with respect to $\bar{\bar{\rho}}$, we will have

$$\int_0^\infty e^{\delta_{m+1}t}\left(\int_{\mathbb{R}^d}\left|D^{m+1}\mathcal{U}(t+s,s)\right|^2 d\bar{\bar{\rho}}\right)dt < \infty.$$

Consider a higher order

$$\begin{aligned}
&\frac{d}{dt}\left|D^{m+1}\mathcal{U}(t+s,s,x)\right|^2 - \tilde{\mathcal{L}}\left|D^{m+1}\mathcal{U}(t+s,s,x)\right|^2 \\
\le& -C_1^{m+1}\left|D^{m+2}\mathcal{U}(t+s,s,x)\right|^2 + C_2^{m+1}\sum_{k\le m+1}\left|D^k\mathcal{U}(t+s,s,x)\right|^2.
\end{aligned}$$

By choosing $\gamma_{m+1} < \delta_{m+1}$ and following the same procedure as above, we have (3.1) for the case when $p = m+1$. By induction principle, we proved the above result holds for any order of spatial derivatives of $\mathcal{U}(t+s,s,x)$.

By (2.7), we can conclude that $q(s,x) > 0$ as $p(s+\tau,s,y,x) > 0$ for any $s \in \mathbb{R}$ and $x, y \in \mathbb{R}^d$. We can also prove the continuity of $q(t,x)$ from the continuity of $p(t+\tau,t,y,x)$ in x. Thus the density function $q(s,x)$ is strictly positive continuous function on any ball $B = B(0,R)$. It turns out that there exists $C > 0$ such that

$$\frac{1}{\tau}\int_0^\tau \|\partial_J\mathcal{U}(t+s,s)\|^2_{L^2(B)}\,ds \le \frac{C}{\tau}\int_0^\tau\int_{\mathbb{R}^d}|\partial_J\mathcal{U}(t+s,s,x)|^2 q(s,x)dxds.$$

By the Sobolev embedding $W^{k,2}(B) \hookrightarrow C(B)$ for $k > \frac{d}{2}$ ([1]), we have that

$$\frac{1}{\tau}\int_0^\tau |\mathcal{U}(t+s,s,x)|\,ds \le \frac{1}{\tau}\int_0^\tau\int_B\left|D^k\mathcal{U}(t+s,s,x)\right|^2 dxds \le C_k\exp(-\lambda_k t),$$

for any $x \in B$, where $k > \frac{d}{2}$. The proof is completed. □

3.2 Estimates on the average of $\mathcal{U}(t+s,s)$ in $L^2(\pi_r)$

In Section 3.1, we obtained the exponential contraction of $\frac{1}{\tau}\int_0^\tau |\mathcal{U}(t+s,s,x)|\,ds$ in any ball B when we assumed $\int_0^\tau\int_{\mathbb{R}^d}\phi(x)q(s,x)dxds=0$. To consider the behaviour outside of the ball B, we need to introduce the weight $\pi_r(s,x)$ with some integer r determined later,

$$\pi_r(s,x)=1/(2+|x|^2+cos(\frac{2\pi s}{\tau}))^r.$$

We consider its gradient and partial derivatives with respect to time s,

$$\nabla\pi_r(s,x)=-\frac{2rx}{2+|x|^2+cos(\frac{2\pi s}{\tau})}\pi_r(s,x),\ \frac{\partial}{\partial s}\pi_r(s,x)=\frac{\frac{2\pi r}{\tau}sin(\frac{2\pi s}{\tau})}{2+|x|^2+cos(\frac{2\pi s}{\tau})}\pi_r(s,x).$$

In general, it is easy to see that for any multi-index J and any integer r, there exist smooth functions $\psi_{J,r}(s,x)$ and $\psi_{s,r}(s,x)$ such that,

$$\partial_J\pi_r(s,x)=\psi_{J,r}(s,x)\pi_r(s,x),\ \frac{\partial}{\partial s}\pi_r(s,x)=\psi_{s,r}(s,x)\pi_r(s,x),$$

where $\psi_{J,r}(s,x)\to 0$ and $\psi_{s,r}(s,x)\to 0$ when $|x|\to+\infty$.

Lemma 3.2. *Assume Conditions (1), (2) and (3), there exist strictly positive constants C and λ such that for any $t>0$, we have*

$$\frac{1}{\tau}\int_0^\tau\int_{\mathbb{R}^d}|\mathcal{U}(t+s,s,x)|^2\pi_r(s,x)dxds\le C\exp(-\lambda t).$$

Proof. Recall (2.12) to lead that for any integer $n\ge 0$, it is possible to choose an integer r_n such that, for any $0\le m\le n$, $t\ge 0$, we have $|D^m\mathcal{U}(t+s,s,x)|\,\pi_{r_n}(s,x)\in L^2(\mathbb{R}^d)$. We denote the multi-index I for the derivative ∂ with the length $|I|$. Consider the integer M_I defined by $|I|=[M_I-d/2]$, and the property of the weight π_r, we have that there exists an integer r_0 such that for any $t>0$, any $r\ge r_0$ and any $m\le M_I$, $D^m(\mathcal{U}(t+s,s,x)\pi_r(s,x))\in L^2(\mathbb{R}^d)$. It is easy to see the periodicity of the function $\mathcal{U}(t+s,s,x)\pi_r(s,x)$ with respect to the initial time s. Note any order of its spatial derivatives are also τ-periodic in s, so by integration by parts formula and periodicity,

$$\begin{aligned}
&\int_0^\tau\int_{\mathbb{R}^d}\frac{d}{dt}|\mathcal{U}(t+s,s)|^2\pi_r dxds\\
=&-\int_0^\tau\int_{\mathbb{R}^d}(\partial_i b_i)|\mathcal{U}(t+s,s)|^2\pi_r dxds-\int_0^\tau\int_{\mathbb{R}^d}b_i|\mathcal{U}(t+s,s)|^2(\partial_i\pi_r)dxds\\
&-\int_0^\tau\int_{\mathbb{R}^d}|\mathcal{U}(t+s,s)|^2\left(\frac{\partial}{\partial s}\pi_r\right)dxds\\
&-\int_0^\tau\int_{\mathbb{R}^d}(\partial_i a_{ij})\mathcal{U}(t+s,s)(\partial_j\mathcal{U}(t+s,s))\pi_r dxds\\
&-\int_0^\tau\int_{\mathbb{R}^d}a_{ij}(\partial_i\mathcal{U}(t+s,s))(\partial_j\mathcal{U}(t+s,s))\pi_r dxds\\
&-\int_0^\tau\int_{\mathbb{R}^d}a_{ij}\mathcal{U}(t+s,s)(\partial_j\mathcal{U}(t+s,s))(\partial_i\pi_r)dxds.
\end{aligned}$$

By Condition (2) and the property of the weight $\pi(s,x)$, we have that

$$\int_0^\tau\int_{\mathbb{R}^d}\frac{d}{dt}|\mathcal{U}(t+s,s)|^2\pi_r dxds$$

$$
\begin{aligned}
&\le -\int_0^\tau \int_{\mathbb{R}^d} (\partial_i b_i) |\mathcal{U}(t+s,s)|^2 \pi_r dxds + \int_0^\tau \int_{\mathbb{R}^d} \frac{2r \cdot x \cdot b(s,x)}{2+|x|^2+cos(\frac{2\pi s}{\tau})} |\mathcal{U}(t+s,s)|^2 \pi_r dxds \\
&\quad - \int_0^\tau \int_{\mathbb{R}^d} |\mathcal{U}(t+s,s)|^2 \psi_s \pi_r dxds - \alpha \int_0^\tau \int_{\mathbb{R}^d} |\nabla \mathcal{U}(t+s,s)|^2 \pi_r dxds \\
&\quad + \frac{1}{2}\int_0^\tau \int_{\mathbb{R}^d} (\partial_{ij} a_{ij}) |\mathcal{U}(t+s,s)|^2 \pi_r dxds + \frac{1}{2}\int_0^\tau \int_{\mathbb{R}^d} (\partial_i a_{ij}) |\mathcal{U}(t+s,s)|^2 \psi_{j,r}\pi_r dxds \\
&\quad + \frac{1}{2}\int_0^\tau \int_{\mathbb{R}^d} (\partial_j a_{ij}) |\mathcal{U}(t+s,s)|^2 \psi_{i,r}\pi_r dxds + \frac{1}{2}\int_0^\tau \int_{\mathbb{R}^d} a_{ij} |\mathcal{U}(t+s,s)|^2 \psi_{ij,r}\pi_r dxds \\
&= \int_0^\tau \int_{\mathbb{R}^d} \left(\Phi_{a,b}(s,x) + \Phi_\psi(s,x) + \frac{2r \cdot x \cdot b(s,x)}{2+|x|^2+cos(\frac{2\pi s}{\tau})} \right) |\mathcal{U}(t+s,s)|^2 \pi_r dxds \\
&\quad - \alpha \int_0^\tau \int_{\mathbb{R}^d} |\nabla \mathcal{U}(t+s,s)|^2 \pi_r dxds,
\end{aligned}
$$

where $\Phi_{a,b}$ is a bounded function depending on functions a, b and their derivatives, Φ_ψ is a function which could depend on functions $\psi_{I,r}$. It is easy to prove that $\Phi_{a,b}$ is independent of r. We also know that Φ_ψ tends to 0 when $|x|$ goes to ∞. Therefore, we choose $r \ge r_0$ large enough to obtain,

$$
\limsup_{|x|\to\infty} \left(\Phi_{a,b}(s,x) + \Phi_\psi(s,x) + \frac{2r \cdot x \cdot b(s,x)}{2+|x|^2+cos(\frac{2\pi s}{\tau})} \right) < 0. \tag{3.10}
$$

Now choosing the ball $B = B(0,R)$ with R being large enough, which depends on the integer r, we have the following result from (3.10),

$$
\begin{aligned}
&\int_0^\tau \int_{\mathbb{R}^d} \left(\Phi_{a,b}(s,x) + \Phi_\psi(s,x) + \frac{2r \cdot x \cdot b(s,x)}{2+|x|^2+cos(\frac{2\pi s}{\tau})} \right) |\mathcal{U}(t+s,s)|^2 \pi_r dxds \\
\le& C_1 \int_0^\tau \int_B |\mathcal{U}(t+s,s)|^2 \pi_r dxds - C_2 \int_0^\tau \int_{B^c} |\mathcal{U}(t+s,s)|^2 \pi_r dxds \\
\le& (C_1 + C_2) \int_0^\tau \int_B |\mathcal{U}(t+s,s)|^2 \pi_r dxds - C_2 \int_0^\tau \int_{\mathbb{R}^d} |\mathcal{U}(t+s,s)|^2 \pi_r dxds,
\end{aligned}
$$

where $C_1, C_2 > 0$ are constants. Therefore, by Lemma 3.1,

$$
\frac{d}{dt}\int_0^\tau \int_{\mathbb{R}^d} |\mathcal{U}(t+s,s)|^2 \pi_r dxds \le -C_2 \int_0^\tau \int_{\mathbb{R}^d} |\mathcal{U}(t+s,s)|^2 \pi_r dxds + C_3 \exp(-\lambda t).
$$

The result follows then from the Gronwall's inequality. □

3.3 Exponential decay of the spatial derivatives of the solution

Theorem 3.3. *Assume Conditions (1), (2) and (3), and $\phi \in C_p^\infty$. Then for any multi-index I, there exists an integer k_I, strictly positive constants Γ_I and γ_I such that $\frac{1}{\tau}\int_0^\tau |\partial_I \mathcal{U}(t+s,s,x)| ds \le \Gamma_I (1+|x|^{k_I}) \exp(-\gamma_I t)$.*

Proof. The process of the proof is similar to Lemma 3.1. We first apply induction method on each order of spatial derivatives of $\mathcal{U}(t+s,s,x)$. It guaranteed first the exponential contraction in any ball $B(0,R)$. Now we consider the behaviour outside of the ball to have,

$$
\begin{aligned}
&\int_{B^c} \tilde{\mathcal{L}} |\mathcal{U}(t+s,s)|^2 \pi_r d\tilde{x} \\
=& \int_{B^c} \left(\Phi_{a,b} + \Phi_\psi + \frac{2r \cdot x \cdot b}{2+|x|^2+cos(\frac{2\pi s}{\tau})} \right) |\mathcal{U}(t+s,s)|^2 \pi_r d\tilde{x} < 0,
\end{aligned}
$$

if we choose the ball large enough. Thus we have some positive C_0 and λ_0 such that

$$\int_{\mathbb{S}} \tilde{\mathcal{L}} \left|\mathcal{U}(t+s,s)\right|^2 \pi_r d\tilde{x} < \int_B \tilde{\mathcal{L}} \left|\mathcal{U}(t+s,s)\right|^2 \pi_r d\tilde{x} \leq C_0 \exp(-\lambda_0 t). \tag{3.11}$$

On the other hand, we integrate with respect to $d\tilde{x}$ with weight π_r, multiply $e^{\delta t}$ and integrate with respect to t from 0 to T to have

$$\begin{aligned} & e^{\delta T} \int_{\mathbb{S}} \left|\mathcal{U}(T+s,s)\right|^2 \pi_r d\tilde{x} + C \int_0^T e^{\delta t} \int_{\mathbb{S}} \left|\nabla \mathcal{U}(t+s,s)\right|^2 \pi_r d\tilde{x} dt \\ \leq & \int_{\mathbb{S}} \left|\tilde{\phi}(\tilde{x})\right|^2 \pi_r d\tilde{x} + \delta \int_0^T e^{\delta t} \int_{\mathbb{S}} \left|\mathcal{U}(t+s,s)\right|^2 \pi_r d\tilde{x} dt + \int_0^T e^{\delta t} \int_{\mathbb{S}} \tilde{\mathcal{L}} \left|\mathcal{U}(t+s,s)\right|^2 \pi_r d\tilde{x} dt. \end{aligned}$$

By the estimates (3.5) and (3.11), we can choose constant δ small enough to obtain

$$\int_0^T e^{\delta t} \int_{\mathbb{S}} \left|\nabla \mathcal{U}(t+s,s)\right|^2 \pi_r d\tilde{x} dt \leq C.$$

Similarly we consider (3.9) to have

$$\begin{aligned} & e^{\gamma T} \int_{\mathbb{S}} \left|\nabla \mathcal{U}(T+s,s)\right|^2 \pi_r d\tilde{x} + C \int_0^T e^{\gamma t} \int_{\mathbb{S}} \left|D^2 \mathcal{U}(t+s,s)\right|^2 \pi_r d\tilde{x} dt \\ \leq & \int_{\mathbb{S}} \left|\nabla \tilde{\phi}(\tilde{x})\right|^2 \pi_r d\tilde{x} + (\gamma + C_2) \int_0^T e^{\gamma t} \int_{\mathbb{S}} \left|\nabla \mathcal{U}(t+s,s)\right|^2 \pi_r d\tilde{x} dt \\ & + \int_0^T e^{\gamma t} \int_{\mathbb{S}} \tilde{\mathcal{L}} \left|\nabla \mathcal{U}(t+s,s)\right|^2 \pi_r d\tilde{x} dt, \end{aligned}$$

which gives us the conclusion that $\int_{\mathbb{S}} \left|\nabla \mathcal{U}(t+s,s)\right|^2 \pi_r d\tilde{x} \leq Ce^{-\gamma t}$. It is easy to repeat the process for any $m \in \mathbb{N}$ with positive constants C_m and γ_m to obtain $\int_{\mathbb{S}} \left|D^m \mathcal{U}(t+s,s)\right|^2 \pi_r d\tilde{x} \leq C_m e^{-\gamma_m t}$. Then we proved the conclusion of the theorem by the weighted Sobolev embedding Theorem with $\pi_r(s,x) d\tilde{x}$ instead of the the density function of average periodic measure $q(s,x) d\tilde{x}$. □

The following remark applies to general case without assumption (2.13).

Remark 3.4. *The proof of the previous theorem also gives us the result that there exist some integer $l \in \mathbb{N}$ and constants $\Gamma > 0$, $\gamma > 0$, such that for any t and x,*

$$\left| \frac{1}{\tau} \int_0^\tau \mathcal{U}(t+s,s,x) ds - \int_{\mathbb{S}} \tilde{\phi}(\tilde{x}) d\bar{\tilde{\rho}}(\tilde{x}) \right| \leq \Gamma(1 + |x|^l) \exp(-\gamma t). \tag{3.12}$$

4 Ergodicity for discretized semi-flows of Euler-Maruyama scheme

We consider Euler-Maruyama numerical scheme with step size $\Delta t = \frac{\tau}{N} > 0$ for SDE (1.1):

$$\widehat{X}^{-k\tau}_{-k\tau+(i+1)\Delta t} = \widehat{X}^{-k\tau}_{-k\tau+i\Delta t} + b(i\Delta t, \widehat{X}^{-k\tau}_{-k\tau+i\Delta t})\Delta t + \sigma(\widehat{X}^{-k\tau}_{-k\tau+i\Delta t})\Delta W_i, \tag{4.1}$$

with $\widehat{X}^{-k\tau}_{-k\tau} = x$, where $i = 0, 1, 2, \ldots$, $\Delta W_i = W_{-k\tau+(i+1)\Delta t} - W_{-k\tau+i\Delta t}$. There are several methods to generate the stochastic increment ΔW_i. But in order to obtain the ergodicity of numerical schemes, in this paper we apply Gaussian distribution in the approximation i.e. $\Delta W = \sqrt{\Delta t}\mathcal{N}(0,1)$. Denote transition probability

$$\hat{P}(-k\tau + i\Delta t, -k\tau, x, \Gamma) = \mathbb{P}\left(\widehat{X}^{-k\tau}_{-k\tau+i\Delta t} \in \Gamma\right).$$

It is easy to see that $\hat{P}(-k\tau + i\Delta t, -k\tau, x, \Gamma) = \hat{P}(i\Delta t, 0, x, \Gamma)$. One can easily extend the numerical scheme (4.1) to $\widehat{X}_{i\Delta t}^{j\Delta t}$, $i \geq j$, $j \in \mathbb{Z}$ with $\widehat{X}_{j\Delta t}^{j\Delta t} = x$ and its transition probability to $\hat{P}(i\Delta t, j\Delta t, x, \cdot)$, $i \geq j$, $j \in \mathbb{Z}$. The corresponding semigroup $\hat{\mathcal{T}}(i\Delta t, j\Delta t)$, $i \geq j$, $j \in \mathbb{Z}$ can be generated from the transition probability in a standard way. A measure-valued function $\hat{\rho} : \mathbb{Z} \to \mathbb{P}(\mathbb{R}^d)$ is called a periodic measure of the semigroup $\hat{\mathcal{T}}(i\Delta t, j\Delta t)$ if

$$\int_{\mathbb{R}^d} \hat{P}(i\Delta t, j\Delta t, x, \Gamma)\hat{\rho}_j(dx) = \hat{\rho}_i(\Gamma)$$

and

$$\hat{\rho}_{i+N} = \hat{\rho}_i$$

for all $i \in \mathbb{Z}$. Recall here $N = \tau/\Delta t$.

Remark 4.1. *By Condition (1), if the function $b(t, 0)$ is bounded for any $t > 0$, then b is of linear growth $|b(t, X_t)| \leq L|X_t| + C$, where $L, C > 0$.*

Remark 4.2. *Under Condition (3), the conclusion in the following proposition still holds for sufficient small step size $\Delta t < \Delta t_c$ with some $\Delta t_c > 0$ that may depend on the growth order of the test function. In order to obtain a uniform Δt_c, we consider the following slightly stronger condition. But in the case of one-dimension, Condition (3') is the same as Condition (3). This means that in the case of one-dimension and Condition (3), a uniform $\Delta t < \Delta t_c$ is obtained with respect to all polynomial growth test functions.*

Condition (3') *For all $i = 1, 2, \dots, d$, there exist constants $\beta_i > 0$ and $C_{\beta_i} > 0$ such that for any $t \in \mathbb{R}^+$ and any $x_i \in \mathbb{R}$, $x_i \cdot b_i(t, x) \leq -\beta_i |x_i|^2 + C_{\beta_i}$.*

Proposition 4.3. *Assume Conditions (1), (3') and the boundedness of $b(t, 0)$ for any $t > 0$, then for any integer p and any $\delta > 0$, there exist constants $C_p, \hat{C}_p, \gamma, \hat{\gamma} > 0$ such that for any $0 < \Delta t \leq \frac{2}{L+\delta}$, $x \in \mathbb{R}^d$ and $n \in \mathbb{N}$,*

$$\mathbb{E}\left|\widehat{X}_{-k\tau+n\Delta t}^{-k\tau}\right|^p \leq C_p\left(1 + |x|^p \exp(-\gamma p n\Delta t)\right),$$

and

$$\mathbb{E}\left[\left|\widehat{X}_{-k\tau+(i+1)\Delta t}^{-k\tau}\right|^p \middle| \hat{\mathcal{F}}_i\right] \leq (1 - \hat{\gamma}\Delta t)\left|\widehat{X}_{-k\tau+i\Delta t}^{-k\tau}\right|^p + \hat{C}_p, \tag{4.2}$$

where $\hat{\mathcal{F}}_i = \mathcal{F}_{-k\tau+i\Delta t}$.

Proof. We first consider the one-dimensional case. Condition (3'), which is the same as Condition (3) in this case, implies that for any $|x| > \sqrt{\frac{C_\beta}{\beta}}$, $x \cdot b(t, x) \leq -\beta|x|^2 + C_\beta < 0$. It then follows that when $x > \sqrt{\frac{C_\beta}{\beta}}$,

$$(1 - L\Delta t)x - C\Delta t \leq x + b(t, x)\Delta t \leq (1 - \beta\Delta t)x + \frac{C_\beta \Delta t}{x} \leq (1 - \beta\Delta t)x + \sqrt{\beta C_\beta}\Delta t.$$

Thus,

$$|x + b(t, x)\Delta t| \leq \max\{|1 - \beta\Delta t|, |1 - L\Delta t|\}|x| + C_1\Delta t, \tag{4.3}$$

where C_1 is independent of Δt. One can obtain the same result for $x < -\sqrt{\frac{C_\beta}{\beta}}$. It is not hard to verify that $L \geq \beta$. Then, for $\hat{\gamma}_1 = \min\{\beta, \delta\} > 0$, we can see that for $0 < \Delta t \leq \frac{2}{L+\delta}$,

$$\max\{|1 - \beta\Delta t|, |1 - L\Delta t|\} < 1 - \hat{\gamma}_1\Delta t. \tag{4.4}$$

It then follows from (4.3) and (4.4) that

$$|x+b(t,x)\Delta t| \leq |1-\hat{\gamma}_1\Delta t|\,|x| + C_1\Delta t. \tag{4.5}$$

Fix any sufficiently small $\hat{\varepsilon}>0$, choose ε_k such that $\frac{p-k}{p}\varepsilon_k^{\frac{p}{p-k}} = \hat{\varepsilon}^k$. Now for any given integer p, by Young's inequality with positive ε_k, which will be fixed later

$$|1-\hat{\gamma}_1\Delta t|^{p-k}\,|x|^{p-k}\,C_1^k \leq |1-\hat{\gamma}_1\Delta t|^p\,|x|^p\left(\frac{p-k}{p}\varepsilon_k^{\frac{p}{p-k}}\right) + (kC_1^p)/(p\varepsilon_k^{\frac{p}{k}}), \tag{4.6}$$

it then follows from (4.5) and (4.6) that

$$\begin{aligned}
&|x+b(t,x)\Delta t|^p \\
\leq &|1-\hat{\gamma}_1\Delta t|^p\,|x|^p + \sum_{k=1}^{p-1}\binom{p}{k}|1-\hat{\gamma}_1\Delta t|^{p-k}\,|x|^{p-k}\,C_1^k(\Delta t)^k + (C_1\Delta t)^p \\
\leq &\left(1+\sum_{k=1}^{p-1}\binom{p}{k}(\Delta t)^k\hat{\varepsilon}^k\right)|1-\hat{\gamma}_1\Delta t|^p\,|x|^p + \sum_{k=1}^{p-1}\binom{p}{k}(\Delta t)^k\frac{kC_1^p}{p\varepsilon_k^{\frac{p}{k}}} + (C_1\Delta t)^p.
\end{aligned}$$

Now we choose $\hat{\varepsilon}$ small enough to obtain

$$\left(1+\sum_{k=1}^{p-1}\binom{p}{k}(\Delta t)^k\hat{\varepsilon}^k\right)|1-\hat{\gamma}_1\Delta t|^p < (1+\hat{\varepsilon}\Delta t)^p\,|1-\hat{\gamma}_1\Delta t|^p \leq |1-\hat{\gamma}_2\Delta t|^p < 1,$$

with some constant $\hat{\gamma}_2>0$ being independent of Δt. Then for any fixed p,

$$|x+b(t,x)\Delta t|^p \leq |1-\hat{\gamma}_2\Delta t|^p\,|x|^p + C_2\Delta t,$$

with some constant C_2 independent of Δt. Denote by $A=\left\{x:|x|\leq\sqrt{\frac{C_\beta}{\beta}}\right\}$. If the conclusion of this proposition holds for any even p, one can obtain the result for odd p by

$$\begin{aligned}
\mathbb{E}\left|\widehat{X}_{-k\tau+n\Delta t}^{-k\tau}\right|^p \leq&\sqrt{\mathbb{E}\left|\widehat{X}_{-k\tau+n\Delta t}^{-k\tau}\right|^{2p}} \leq \sqrt{C_{2p}\left(1+|x|^{2p}\exp(-2\gamma pn\Delta t)\right)} \\
\leq&\sqrt{C_{2p}}\left(1+|x|^p\exp(-\gamma pn\Delta t)\right),
\end{aligned}$$

with coefficient $C_p=\sqrt{C_{2p}}$. Then we only consider the cases where p is even in the following. For this we apply the same argument using Young's inequality as above on (4.1) and conditional expectation to have

$$\begin{aligned}
&\mathbb{E}\left[\left|\widehat{X}_{-k\tau+(i+1)\Delta t}^{-k\tau}\right|^p I_{A^c}\left(\widehat{X}_{-k\tau+i\Delta t}^{-k\tau}\right)\middle|\mathcal{F}_i\right] \\
=&I_{A^c}\left(\widehat{X}_{-k\tau+i\Delta t}^{-k\tau}\right)\left[\left(\widehat{X}_{-k\tau+i\Delta t}^{-k\tau}+b(i\Delta t,\widehat{X}_{-k\tau+i\Delta t}^{-k\tau})\Delta t\right)^p\right. \\
&+\sum_{l=1}^{p-1}\binom{p}{l}\left(\widehat{X}_{-k\tau+i\Delta t}^{-k\tau}+b(i\Delta t,\widehat{X}_{-k\tau+i\Delta t}^{-k\tau})\Delta t\right)^{p-l}\left(\sigma(\widehat{X}_{-k\tau+i\Delta t}^{-k\tau})\right)^l\mathbb{E}\left[(\Delta W_i)^l\middle|\mathcal{F}_i\right] \\
&\left.+\left(\sigma(\widehat{X}_{-k\tau+i\Delta t}^{-k\tau})\right)^p\mathbb{E}\left[(\Delta W_i)^p\middle|\mathcal{F}_i\right]\right] \\
\leq&\left[(1+\hat{\varepsilon}\Delta t)^{\frac{p}{2}}\,|1-\hat{\gamma}_2\Delta t|^p\left|\widehat{X}_{-k\tau+i\Delta t}^{-k\tau}\right|^p + C_3\Delta t\right]I_{A^c}\left(\widehat{X}_{-k\tau+i\Delta t}^{-k\tau}\right)
\end{aligned}$$

$$\leq \left(|1-\hat{\gamma}_3\Delta t|^{\frac{p}{2}}\left|\widehat{X}^{-k\tau}_{-k\tau+i\Delta t}\right|^p + C_3\Delta t\right) I_{A^c}\left(\widehat{X}^{-k\tau}_{-k\tau+i\Delta t}\right),$$

with some constant C_3 independent of Δt, where $\hat{\varepsilon}$ is chosen small enough such that $(1+\hat{\varepsilon}\Delta t)^{\frac{p}{2}}|1-\hat{\gamma}_2\Delta t|^p \leq |1-\hat{\gamma}_3\Delta t|^{\frac{p}{2}} < 1$ for some $\hat{\gamma}_3>0$. Moreover by linear growth condition of b and $|1+L\Delta t| \leq 1+L\cdot\frac{2}{L}=3$, we can obtain

$$\mathbb{E}\left[\left|\widehat{X}^{-k\tau}_{-k\tau+(i+1)\Delta t}\right|^p I_A\left(\widehat{X}^{-k\tau}_{-k\tau+i\Delta t}\right)\middle|\hat{\mathcal{F}}_i\right] \leq \left(3^p\left(\frac{C_\beta}{\beta}\right)^{\frac{p}{2}}+C\right) I_A\left(\widehat{X}^{-k\tau}_{-k\tau+i\Delta t}\right),$$

where L is the bound of function b's first derivative (or coefficient of global Lipschitz). We combine the above two estimates to obtain

$$\mathbb{E}\left[\left|\widehat{X}^{-k\tau}_{-k\tau+(i+1)\Delta t}\right|^p\middle|\hat{\mathcal{F}}_i\right] \leq |1-\hat{\gamma}_3\Delta t|^{\frac{p}{2}}\left|\widehat{X}^{-k\tau}_{-k\tau+i\Delta t}\right|^p + \hat{C}_p,$$

with $\hat{C}_p = \max\left\{C_3\Delta t, 3^p\left(\frac{C_\beta}{\beta}\right)^{\frac{p}{2}}+C\right\}$ and $\hat{\gamma}=\hat{\gamma}_3$ Therefore,

$$\begin{aligned}
&|1-\hat{\gamma}_3\Delta t|^{-\frac{p}{2}n}\,\mathbb{E}\left[\left|\widehat{X}^{-k\tau}_{-k\tau+n\Delta t}\right|^p\right]\\
&=|x|^p+\sum_{i=1}^n\left\{|1-\hat{\gamma}_3\Delta t|^{-\frac{p}{2}i}\,\mathbb{E}\left|\widehat{X}^{-k\tau}_{-k\tau+i\Delta t}\right|^p - |1-\hat{\gamma}_3\Delta t|^{-\frac{p}{2}(i-1)}\,\mathbb{E}\left|\widehat{X}^{-k\tau}_{-k\tau+(i-1)\Delta t}\right|^p\right\}\\
&=|x|^p+\sum_{i=1}^n|1-\hat{\gamma}_3\Delta t|^{-\frac{p}{2}i}\,\mathbb{E}\left[\mathbb{E}\left(\left|\widehat{X}^{-k\tau}_{-k\tau+i\Delta t}\right|^p - |1-\hat{\gamma}_3\Delta t|^{\frac{p}{2}}\left|\widehat{X}^{-k\tau}_{-k\tau+(i-1)\Delta t}\right|^p\middle|\hat{\mathcal{F}}_{i-1}\right)\right]\\
&\leq |x|^p+\hat{C}_p\sum_{i=1}^n|1-\hat{\gamma}_3\Delta t|^{-\frac{p}{2}i} \leq |x|^p + C_p,
\end{aligned}$$

where $C_p = \frac{\hat{C}_p}{1-|1-\hat{\gamma}_3\Delta t|^{\frac{p}{2}}}$. Finally from $|1-\hat{\gamma}_3\Delta t| < \exp(-\hat{\gamma}_3\Delta t)$, we have that $\mathbb{E}\left[\left|\widehat{X}^{-k\tau}_{-k\tau+n\Delta t}\right|^p\right] \leq |x|^p\exp(-\frac{\hat{\gamma}_3}{2}pn\Delta t)+C_p$.

For the multi-dimensional case, we apply Condition (3') to have the estimations with coefficients $C_{i,p}$ and γ_i for each $i=1,2,\ldots,d$. Then the final conclusion follows. □

Proposition 4.4. *Assume the conditions in Proposition 4.3 and Condition (2), then the Euler-Maruyama scheme (4.1) is geometrically ergodic for all step-size* $0<\Delta t<\frac{2}{L}$, *i.e. there exists a periodic measure* $\hat{\rho}^{\Delta t}:\mathbb{Z}\to\mathcal{P}(\mathbb{R}^d)$ *such that*

$$\left\|\hat{P}(i\Delta t, -k\tau+i\Delta t, x,\cdot)-\hat{\rho}_i(\cdot)\right\|_{TV} \leq Ce^{-\delta k\tau}, \quad k\in\mathbb{N},$$

for some constants $C,\delta>0$.

Proof. To check the local Doeblin condition, we need to prove the transition kernel $\hat{P}(s+\tau,s,x,A) = \hat{P}(s+N\Delta t,s,x,A) = \hat{\mathbb{P}}\{\widehat{X}^s_{s+N\Delta t}\in A|\widehat{X}^s_s = x\}$ possesses a density function $\hat{p}(t,s,x,y)$ satisfying $\inf_{x,y\in K}\hat{p}(s+\tau,s,x,y)>0$ for some non-empty set $K\in\mathcal{B}$ with Lebesgue measure $\Lambda(K)>0$. To apply the result of Theorem 3.5 in [13] to prove the local Doeblin condition of the transition kernel, we only need to prove there is a non-empty compact set K, such that for any $i=0,\ldots,N-1$ and any non-empty open set Γ, we have

$$\inf_{x\in K}\hat{P}(s+(i+1)\Delta t, s+i\Delta t, x,\Gamma)>0. \tag{4.7}$$

Consider the numerical approximation in the time interval $[s+i\Delta t, s+(i+1)\Delta t]$. For simplicity, we denote by $s_i = s+i\Delta t$, $i=0,\dots,N-1$ and $\widehat{X}^i_t = \widehat{X}^{s,x}_{s_i+t}$. So

$$\widehat{X}^i_{\Delta t} - \widehat{X}^i_0 = b(s_i, \widehat{X}^i_0)\Delta t + \sigma(\widehat{X}^i_0)\Delta W_t, \tag{4.8}$$

where $\Delta W_t = W_{\Delta t} - W_0$. Let $\widehat{X}^{i,x}_{\Delta t}$ be defined by (4.8) conditioned on $\widehat{X}^i_0 = x$. Then the law of $\widehat{X}^{i,x}_{\Delta t}$ is $\hat{P}(s+(i+1)\Delta t, s+i\Delta t, x, \cdot) = \hat{P}(s_i+\Delta t, s_i, x, \cdot)$. Note $b(s_i, x)$ and $\sigma(x)$ are non-random and given, thus $\hat{P}(s_i+\Delta t, s_i, x, \cdot)$ is simply the Gaussian distribution with mean $x+b(s_i,x)\Delta t \in \mathbb{R}^d$ and covariance matrix $\sigma\sigma^T(x)\Delta t \in \mathbb{R}^{d\times d}$. The covariance matrix is uniformly non-degenerate, thus for any non-empty open set $\Gamma \in \mathbb{R}^d$, we have $\hat{P}(s_i+\Delta t, s_i, x, \Gamma) > 0$ and the function $x \mapsto \hat{P}(s_i+\Delta t, s_i, x, \Gamma)$ is continuous. Thus for any compact set $K \subset \mathbb{R}^d$, we have (4.7).

By Theorem 3.5 in [13], we obtain the local Doeblin condition of $\hat{P}(s+N\Delta t, s, x, \cdot)$. Note condition $0 < \Delta t < \frac{2}{L}$ implies there exists $\delta > 0$ such that $0 < \Delta t < \frac{2}{L+\delta}$. So Proposition 4.3 holds and estimate (4.2) implies Lyapunov condition with Lyapunov function $V(x) = x^2$. Then by Theorem 3.3 in [13], we deduce the ergodicity of the numerical scheme and the convergence to the periodic measure $\hat{\rho}^{\Delta t}$. □

Similar in the continuous time case, we can lift the discrete semi-flow and periodic measure to $\{0,1,2,\dots,N\} \times \mathbb{R}^d$ as follows

$$\widetilde{\widehat{X}}^{j\Delta t}_{i\Delta t} = (i, \widehat{X}^{j\Delta t}_{i\Delta t}), \quad i \geq j,$$

$$\tilde{\hat{P}}(i, (j,x), \{k\}\times\Gamma) = \delta_{(i+j \ mod \ N)}(k)\hat{P}((i+j)\Delta t, j\Delta t, x, \Gamma), \quad i \geq j,$$

$$\tilde{\hat{\rho}}^{\Delta t}_i(\{k\}\times\Gamma) = \delta_{(i \ mod \ N)}(k)\hat{\rho}^{\Delta t}_i(\Gamma), \quad i \in \mathbb{Z}.$$

Then $\widetilde{\widehat{X}}$ is a cocycle and $\tilde{\hat{P}}$ is the transition probability of $\widetilde{\widehat{X}}$. Moreover, $\tilde{\hat{\rho}}^{\Delta t}$ is the periodic measure of $\tilde{\hat{P}}$, i.e.

$$\sum_{l=0}^{N-1}\int_{\mathbb{R}^d} \tilde{\hat{P}}(i,(l,x),\{k\}\times\Gamma)\tilde{\hat{\rho}}^{\Delta t}_j(\{l\}\times dx) = \tilde{\hat{\rho}}^{\Delta t}_{i+j}(\{k\}\times\Gamma).$$

Define $\bar{\tilde{\hat{\rho}}}^{\Delta t} = \frac{1}{N}\sum_{j=0}^{N-1}\tilde{\hat{\rho}}^{\Delta t}_j$. Then it is easy to see that

$$\sum_{l=0}^{N-1}\int_{\mathbb{R}^d} \tilde{\hat{P}}(i,(l,x),\{k\}\times\Gamma)\bar{\tilde{\hat{\rho}}}^{\Delta t}(\{l\}\times dx) = \bar{\tilde{\hat{\rho}}}^{\Delta t}(\{k\}\times\Gamma),$$

i.e. $\bar{\tilde{\hat{\rho}}}$ is the invariant measure of $\tilde{\hat{P}}(i)$, $i \in \mathbb{N}$. Moreover, for any measurable function $\phi : \mathbb{R}^d \to \mathbb{R}$,

$$\begin{aligned}\sum_{k=0}^{N-1}\int_{\mathbb{R}^d}\phi(x)\bar{\tilde{\hat{\rho}}}^{\Delta t}(\{k\}\times dx) &= \sum_{k=0}^{N-1}\int_{\mathbb{R}^d}\phi(x)\frac{1}{N}\sum_{j=0}^{N-1}\delta_j(k)\hat{\rho}^{\Delta t}_j(dx) \\ &= \frac{1}{N}\sum_{k=0}^{N-1}\int_{\mathbb{R}^d}\phi(x)\hat{\rho}^{\Delta t}_k(dx) \\ &= \sum_{k=0}^{N-1}\int_{\mathbb{R}^d}\phi(x)\bar{\hat{\rho}}^{\Delta t}(dx),\end{aligned} \tag{4.9}$$

where $\bar{\hat{\rho}}^{\Delta t} = \frac{1}{N}\sum_{j=0}^{N-1}\hat{\rho}^{\Delta t}_j$.

5 Error estimate for the approximation to periodic measures

In autonomous systems, there are some established results in the ergodicity of numerical schemes (Mattingly, Stuart and Higham [23]; Grorud and Talay [16]; Talay [28], [29]). But in our non-autonomous model, due to the lacking of weakly mixing property, those approaches may not give immediately the error between $\int_{\mathbb{R}^d} \phi(x)\bar{\rho}(dx)$ and $\int_{\mathbb{R}^d} \phi(x)\bar{\hat{\rho}}^{\Delta t}(dx)$ over $[0,\tau]$. We develop the following approach using integration with respect to initial time s to obtain the error estimate of invariant measure. To approximate the average of periodic measure, we need to consider the long time behaviour of the SDE (1.1) by pullback of the initial time to $s-k\tau$. First we notice from the ergodic theory and ergodicity of $\{X_t^{s,x}\}_{t\geq s}$ and $\{\widehat{X}_{s+i\Delta t}^{s,x}\}_{i\geq 0}$, we have the following law of large numbers. Recall (2.5) and (4.9), so for ϕ with at most polynomial growth at infinity and as both of $X_t^{s,x}$ and $\widehat{X}_{s+i\Delta t}^{s,x}$ possess finite moments of any order, we have that for any $\varepsilon>0$, there exists a constant N' such that for all $n\geq N'$,

$$\left|\frac{1}{n}\sum_{k=1}^{n}\frac{1}{\tau}\int_0^\tau \mathbb{E}\phi(X_{s+k\tau}^{s,x})ds-\int_{\mathbb{R}^d}\phi(x)\bar{\rho}(dx)\right| \leq \varepsilon, \quad a.s., \tag{5.1}$$

$$\left|\frac{1}{n}\sum_{k=1}^{n}\frac{1}{\tau}\int_0^\tau \mathbb{E}\phi(\widehat{X}_{s+kN\Delta t}^{s,x})ds-\int_{\mathbb{R}^d}\phi(x)\bar{\hat{\rho}}^{\Delta t}(dx)\right| \leq \varepsilon, \quad a.s.. \tag{5.2}$$

Theorem 5.1. *Assume conditions in Proposition 4.4. Then for any step size $\Delta t=\tau/N$, $N\in\mathbb{N}$, satisfying $\Delta t<\frac{2}{L}$ and any function $\phi\in\mathcal{C}_p^\infty$, we have:*

$$\left|\int_{\mathbb{R}^d}\phi(x)\bar{\rho}(dx)-\int_{\mathbb{R}^d}\phi(x)\bar{\hat{\rho}}^{\Delta t}(dx)\right|=\mathcal{O}(\Delta t). \tag{5.3}$$

Proof. Define $\mathcal{U}(s+i\Delta t,s,x)=\mathbb{E}\phi(X_{s+i\Delta t}^{s,x})$. Then

$$\mathcal{U}(s,s-k\tau,x)=\mathbb{E}\phi(X_s^{s-k\tau,x}),\quad \mathcal{U}(s,s,\widehat{X}_s^{s-k\tau,x})=\phi(\widehat{X}_s^{s-k\tau,x}),\ a.s.. \tag{5.4}$$

By the periodicity of $\mathcal{U}(t,s,x)$ with respect to initial time s, i.e. $\mathcal{U}(t,s-k\tau,x)=\mathcal{U}(t+k\tau,s,x)$, it is always possible to move the initial time into $[0,\tau)$. Now we consider the following Itô-Taylor expansion:

$$\begin{aligned}\mathbb{E}\mathcal{U}(s+k\tau,s+(i+1)\Delta t,\widehat{X}_{(i+1)\Delta t}^{s,x}) =&\mathbb{E}\mathcal{U}(s+k\tau,s+(i+1)\Delta t,\widehat{X}_{i\Delta t}^{s,x})\\ &+\tilde{\mathcal{L}}(s+(i+1)\Delta t)\mathbb{E}\mathcal{U}(s+k\tau,s+(i+1)\Delta t,\widehat{X}_{i\Delta t}^{s,x})\Delta t+R_{1,i}^s(\Delta t)^2.\end{aligned} \tag{5.5}$$

Denote $s'=s+(i+1)\Delta t$ and $t'=k\tau-(i+1)\Delta t$. Then it is obvious that

$$\mathbb{E}\mathcal{U}(s+k\tau,s+i\Delta t,x)=\mathbb{E}\mathcal{U}((s'-\Delta t)+(t'+\Delta t),s'-\Delta t,x).$$

Therefore, we have the following Itô-Taylor expansion:

$$\begin{aligned}\mathbb{E}\mathcal{U}(s+k\tau,s+i\Delta t,\widehat{X}_{i\Delta t}^{s,x}) =&\mathbb{E}\mathcal{U}(s+k\tau,s+(i+1)\Delta t,\widehat{X}_{i\Delta t}^{s,x})\\ &+\left(\frac{\partial}{\partial t}-\frac{\partial}{\partial s}\right)\mathbb{E}\mathcal{U}(s+t,s,\widehat{X}_{i\Delta t}^{s,x})\Delta t\bigg|_{s=s',t=t'}+R_{2,i}^s(\Delta t)^2.\end{aligned} \tag{5.6}$$

The coefficients $R_{1,i}^s$ and $R_{2,i}^s$ have the following form:

$$\mathbb{E}\left[\psi(\widehat{X}_{i\Delta t}^{s,x})\cdot\partial_J\mathcal{U}\left(s+k\tau,s+i\Delta t,\widehat{X}_{i\Delta t}^{s,x}+\vartheta\left(\widehat{X}_{(i+1)\Delta t}^{s,x}-\widehat{X}_{i\Delta t}^{s,x}\right)\right)\right], \tag{5.7}$$

where $0 < \vartheta < 1$ and the function $\psi(x)$ is a product of functions b, σ and their derivatives. One can obtain the boundedness of $\psi(x)$ from Condition (1). Combining (5.5) and (5.6), we have

$$\begin{aligned}&\mathbb{E}\mathcal{U}(s+k\tau, s+(i+1)\Delta t, \widehat{X}^{s,x}_{(i+1)\Delta t}) - \mathbb{E}\mathcal{U}(s+k\tau, s+i\Delta t, \widehat{X}^{s,x}_{i\Delta t})\\ =&\left(\frac{\partial}{\partial s} - \frac{\partial}{\partial t} + \mathcal{L}(s)\right)\mathbb{E}\mathcal{U}(s+t,s,x)\Delta t\Big|_{s=s',t=t',x=\widehat{X}^{s,x}_{i+1}} + (R^s_{2,i} - R^s_{1,i})(\Delta t)^2. \end{aligned}\tag{5.8}$$

As $\frac{\partial}{\partial s} + \mathcal{L}(s) = \frac{\partial}{\partial t}$, we take summation on both sides of the above and from (5.4), periodicity of $\mathcal{U}$,

$$\begin{aligned}&\mathbb{E}\phi(\widehat{X}^{s-k\tau,x}_s) - \mathbb{E}\phi(X^{s-k\tau,x}_s)\\ =&\sum_{i=0}^{kN-1}\left(\mathbb{E}\mathcal{U}(s+k\tau, s+(i+1)\Delta t, \widehat{X}^{s,x}_{(i+1)\Delta t}) - \mathbb{E}\mathcal{U}(s+k\tau, s+i\Delta t, \widehat{X}^{s,x}_{i\Delta t})\right) = \sum_{i=0}^{kN-1} R^s_i(\Delta t)^2,\end{aligned}$$

where $R^s_i = R^s_{1,i} - R^s_{2,i}$. Combining this with Proposition 4.3 and Theorem 3.3, there exists a constant $\lambda > 0$ and an integer $l \in \mathbb{N}$, such that

$$\begin{aligned}&\left|\sum_{i=0}^{kN-1}\frac{1}{\tau}\int_0^\tau R^s_i ds\right|\\ \leq&\sum_{i=0}^{kN-1} C\mathbb{E}\left[\frac{1}{\tau}\int_0^\tau \left|\partial_J\mathcal{U}\left(s+k\tau, s+i\Delta t, \widehat{X}^{s,x}_{i\Delta t} + \vartheta\left(\widehat{X}^{s,x}_{(i+1)\Delta t} - \widehat{X}^{s,x}_{i\Delta t}\right)\right)\right| ds\right]\\ \leq&\frac{C}{\tau}\sup_{i\geq 0}\mathbb{E}\left(1 + \left|\widehat{X}^{-k\tau,x}_{-k\tau+i\Delta t}\right|^l + \left|\widehat{X}^{-k\tau,x}_{-k\tau+(i+1)\Delta t}\right|^l\right)\sum_{i=0}^{kN-1}\exp\left(-\lambda(kN-i)\Delta t\right)\\ \leq&\frac{1-e^{-\lambda kN\Delta t}}{1-e^{-\lambda\Delta t}}e^{-\lambda\Delta t}C_l\left(1+|x|^l\right).\end{aligned}$$

Let k go to infinity and Δt be small enough, we have

$$\sum_{i=0}^{+\infty}\frac{1}{\tau}\int_0^\tau |R^s_i|(\Delta t)^2 ds \leq \tilde{C}(1+|x|^l)(\Delta t).\tag{5.9}$$

This is then followed by

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^n\left|\frac{1}{\tau}\int_0^\tau \mathbb{E}\phi(\widehat{X}^{s-k\tau}_s(x))ds - \frac{1}{\tau}\int_0^\tau \mathcal{U}(s, s-k\tau, x)ds\right| \leq \tilde{C}(1+|x|^l)(\Delta t).$$

It then follows from (5.1) and a triangle inequality argument that

$$\left|\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^n\frac{1}{\tau}\int_0^\tau \phi\left(\widehat{X}^{s,x}_{s+kN\Delta t}\right)ds - \int_{\mathbb{S}}\tilde{\phi}(\tilde{x})d\bar{\tilde{\rho}}(\tilde{x})\right| = \mathcal{O}(\Delta t), \quad a.s.,\tag{5.10}$$

Recall (5.1) and (5.2). Then by a triangle inequality argument, we obtain

$$\left|\int_{\mathbb{R}^d}\phi(x)\bar{\rho}(dx) - \int_{\mathbb{R}^d}\phi(x)\bar{\tilde{\rho}}^{\Delta t}(dx)\right| \leq 2\varepsilon + \tilde{C}(1+|x|^l)\Delta t.\tag{5.11}$$

Thus (5.3) follows as the left hand side of (5.11) is independent x and ε. □

Remark 5.2. *Consider the modification of Benzi-Parisi-Sutera-Vulpiani's stochastic resonance model mentioned in the introduction, the coefficients of Linear growth is estimated as* $L \leq \frac{\partial b(t,x)}{\partial_x}\big|_{x=4} \leq 48$. *Hence the step size* $\Delta t < \frac{1}{24}$ *will satisfy our theorem.*

6 Numerical examples

In this section, we carry out some numerical experiments to support the theoretical results obtained in the last section. We give the error analysis for the numerical scheme of the average of periodic measures of two specific models arising in modelling daily temperature and climate dynamics respectively. For each example, we firstly generate discrete random periodic paths $\widehat{X}^{s,x}_{s+k\Delta t}$, $k=0,\dots,N-1$ and test the convergence with different initial values. The numerical error we calculate in this section is

$$\left|\frac{1}{N}\sum_{k=0}^{N-1}\phi\left(\widehat{X}^{s,x}_{s+k\Delta t}\right)-\int_{\mathbb{S}}\phi(x)d\bar{\rho}(x)\right|, \tag{6.1}$$

which consists of three parts of errors: those influenced by the finiteness of N, the discretization error of time integral on s in (5.10) and the error given in (5.10). Here $\phi\in C_p^\infty$. The main task is to estimate the error in (6.1) in terms of rate with respect to Δt. We choose large enough N to reduce its impact on the error. We also compare numerically the errors with different initial time s and find that the convergence of solutions to the random periodic paths in both models are very fast, as seen in Figure 2 and Figure 5, so the effect of the initial time and position to the overall error is negligible as we take the average over a large number of iterations.

To carry out numerical experiments, we use Python 3.8.6 on Linux Fedora 32 with 3.40 GHz Intel(R) Core(TM) i7-3770 CPU and RAM 32.00 GB. There are two cores having higher computing speed (2958.762 MHz and 2121.630 MHz) compared with others (1600 MHz). We would not feel surprise to notice some abnormal computing times.

Example 6.1. *To present the error of our approximation scheme, we study the following temperature model considered by F. Benth and J. Benth [3],*

$$dX_t=(a_0+a_1\cos(2\pi(t-a_2)/365)-\pi X_t)dt+\sigma dW_t,$$

with $a_0=6.4$, $a_1=10.4$, $a_2=-166$ and $\sigma=0.3$. From the discussion in [13], it is known that the periodic measure of this model exists and is a Gaussian distribution with mean

$$\frac{a_0}{\pi}+a_1\frac{k\sin(k(t-a_2))+\pi\cos(k(t-a_2))}{k^2+\pi^2},$$

and variance $\frac{\sigma^2}{2\pi}$, where $k=\frac{2\pi}{365}$. This is the case where the periodic measure is known explicitly. But a numerical experiment of calculating the numerical error is carried out here in order to verify the accuracy of our scheme. To simplify the calculation, we take the test function $\phi(x)=x^2$ as $\mathbb{E}\left[X^2\right]=\mathbb{E}\left[X\right]^2+Var[X]$. Under the average of periodic measure, one can derive the exact value

$$\int_{\mathbb{R}^d}x^2\bar{\rho}(dx)=\frac{a_0^2}{\pi^2}+\frac{a_1^2}{2(\pi^2+k^2)}+\frac{\sigma^2}{2\pi}.$$

On the other hand, conducting numerical approximation with Euler-Maruyama scheme, we obtain $\widehat{X}^{s,x}_{s+k\Delta t}$ for a range of different $\Delta t<2/\pi=0.63662$ varying from 0.2 to 0.004. We apply the Euler-Maruyama scheme with same length of time of 10000 periods for different step size Δt. The error is presented in Table 1 and as a log-log graph in Figure 1. Our numerical results show very good order 1 line fitting. Note the exact value $\int_{\mathbb{R}^d}x^2\bar{\rho}(dx)=10.266021$ (rounded off in 6 decimal places).

In Figure 2, we present two numerical approximations, $\widehat{X}^{0,-10}_s$ and $\widehat{X}^{0,10}_s$, to random periodic path with different initial values. These two trajectories are generated with the same realisation of noise. They merge before $t=2$ and show inconspicuous difference after that. We then

Step sizes	$\Delta t = 0.2$	$\Delta t = 0.1$	$\Delta t = 0.08$	$\Delta t = 0.05$	$\Delta t = 0.04$
Approximation	10.560044	10.383899	10.357855	10.318126	10.307453
Numerical error	0.294024	0.117879	0.091834	0.052105	0.041433
CPU(seconds)	175.41	371.39	465.71	756.28	951.37
Step sizes	$\Delta t = 0.02$	$\Delta t = 0.01$	$\Delta t = 0.008$	$\Delta t = 0.005$	$\Delta t = 0.004$
Approximation	10.283706	10.277765	10.275399	10.271213	10.270654
Numerical error	0.017686	0.011745	0.009379	0.005192	0.004634
CPU(seconds)	1600.47	3711.75	4407.75	6551.42	7600.27

Table 1: Numerical results of Example 6.1 where approximations and numerical errors are rounded off in 6 decimal places

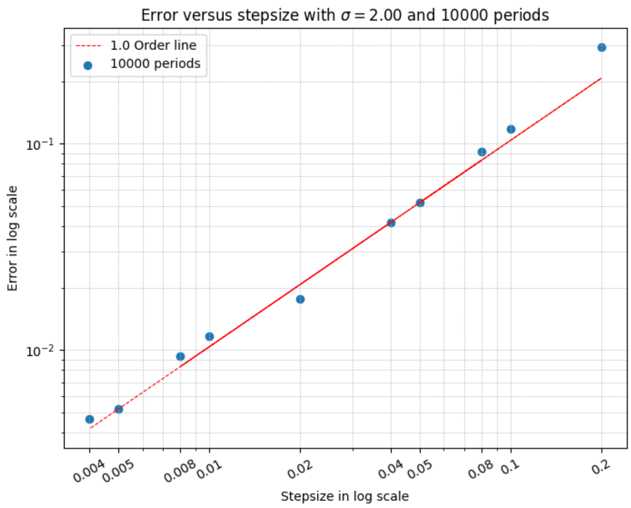


Figure 1: Error of approximation to the average of periodic measure versus step size in log-log graph (Example 6.1)

generate one numerical approximation to random periodic path with 10000 periods and step size $\Delta t = 0.01$. *We collect the points on time* $t = k\tau$ *to build the histogram in left hand side graph of Figure 3(a) compared with its theoretical result* ρ_0. *In Figure 3, we give 8 more comparisons between numerical approximations to the periodic measure and its theoretical results.*

In practice, we run the computation with 10 different step sizes simultaneously under "multiprocessing" package of Python with 7 cores of CPU. The above results of error analysis took 7600.286 seconds of computing time where the CPU time of each step size is given in Table 1. One can see the majority of time was consumed under the small step sizes such as $\Delta t = 0.004$ and $\Delta t = 0.005$.

If necessary, one can split the approximation of random periodic path with small step sizes into several jobs. This works well due to ergodicity and fast convergence to random periodic path under our scheme. We do not need it in Example 6.1 as the computing time is reasonably short. But the possibility to split the computation into several independent jobs plays a crucial role in the case when a model has a large period. For such a problem, we need to consider the long time behaviour of $N\tau$ where both N and τ are large. We will see that in the following example.

Example 6.2. *We consider Benzi-Parisi-Sutera-Vulpiani's climate dynamics model given by*

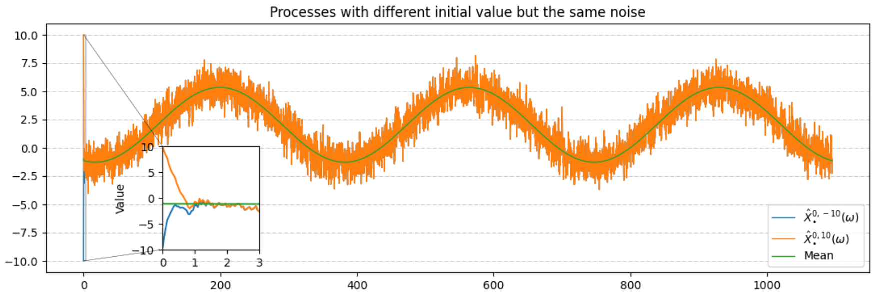
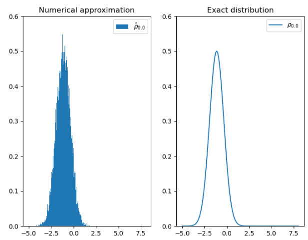
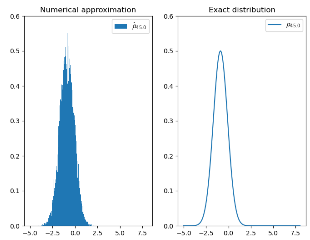


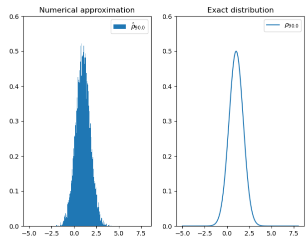
Figure 2: Paths of the temperature model (Example 6.1)



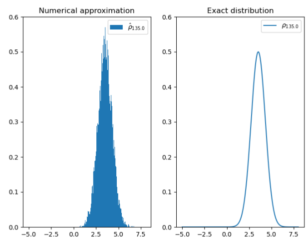
(a) $\hat{\rho}_0$ and ρ_0



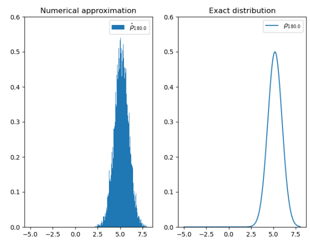
(b) $\hat{\rho}_{45}$ and ρ_{45}



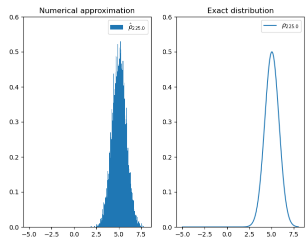
(c) $\hat{\rho}_{90}$ and ρ_{90}



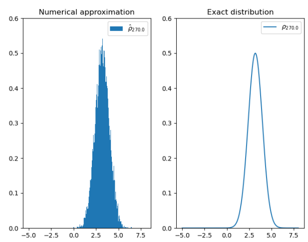
(d) $\hat{\rho}_{135}$ and ρ_{135}



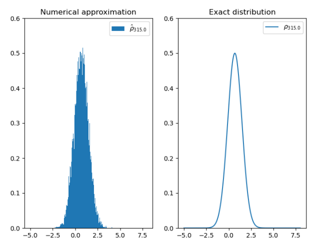
(e) $\hat{\rho}_{180}$ and ρ_{180}



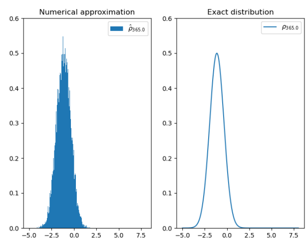
(f) $\hat{\rho}_{225}$ and ρ_{225}



(g) $\hat{\rho}_{270}$ and ρ_{270}



(h) $\hat{\rho}_{315}$ and ρ_{315}



(i) $\hat{\rho}_{365}$ and ρ_{365}

Figure 3: Comparisons between approximation of periodic measure with $\Delta t = 0.01$ and 10000 periods and its theoretical result (Example 6.1)

SDE (1.1) with $b(t,x) = x - x^3 + A\cos(Bt)$ *and* $\sigma(x) = \sigma$. *The coefficients are chosen as* $A = 0.12$, $B = 0.001$ *and* $\sigma = 0.285$ *as discussed in [6] and [14]. We make a time scaling by taking* $b(t,x) = 0.4\pi(x - x^3 + 0.12\cos(0.0004\pi t))$ *and* $\sigma(x) = 0.285 \times \sqrt{0.4\pi} \approx 0.3195$, *so the period* $\tau = 5000$ *in the system. Thus when we apply numerical approximation to the model, we can ensure our time step size dividing the period* $\tau = 5000$. *It is also mollified to satisfy the global Lipschitz assumption in this paper. For this, what we could do is to modify the function* $x - x^3$ *by a linear function* $128 - 47x$ *when* $x \geq 4$ *and* $-128 - 47x$ *when* $x \leq -4$ *and smooth this function by mollifier* $\eta_\varepsilon(x) = \eta(x)$, *where* $\eta(x) = \begin{cases} C\exp\left(\frac{1}{|x|^2-1}\right) & |x| \leq 1 \\ 0 & \text{otherwise} \end{cases}$ *and* C *is chosen such that* $\int \eta(x)dx = 1$. *But this adds a lot of computing time as integration of convolution is needed in every step of the computation.*

In our approximation, the drift term is $\hat{b}(t,x) = 0.4\pi((1-\exp(-\frac{50}{x^2}))(x-x^3)+0.12\cos(0.0004\pi t))$. *This function makes a very good approximation to function* $b(t,x)$ *when* $|x| \leq 4$ *though it is not the case globally. Note this function is Lipschitz and smooth. As we mentioned in the introduction, our modified model provides the same climate dynamics as the original one of Benzi-Parisi-Sutera-Vulpiani. Our numerical simulations presented in Figure 5 for the modified equation provide strong evidence that is the case as seen in the Figure 5 that the trajectory rarely goes outside of* $[-4,4]$. *In fact, during the approximation with 10000 periods, we did not find any point in the whole data set running out of this interval.*

Note this model does not have an explicit solution, so we cannot carry out the error analysis as we did in Example 6.1. To overcome this difficulty we use the approximation of solution with $\Delta t = 0.001$ *to replace our exact solution in the error analysis. To make the computation more efficient, we split the approximation involving* $\frac{5000\times 10000}{0.001} = 5 \times 10^{10}$ *iterations to 8 individual jobs of* 6.25×10^9 *iterations with independent Brownian motions. The results are shown in the left hand side table of Table 2. We then conduct the numerical experiment for step size varying from* $1/125$ *to* $1/50$. *The error is in the right hand side table of Table 2 and the log-log graph is presented in Figure 4. We carry out numerical simulation with 10000 periods for each step size and our total cost is 155547.50 seconds with 7 cores.*

We consider numerical simulation with $\Delta t = 0.01$ *to show the stochastic resonance phenomenon in Figure 5. The simulations start from different initial condition but the same realisation of noise in each sub-graph, where one can see the convergence to random periodic path is also very fast. Together with ergodicity, it provides the possibility of splitting 10000 periods into several independent approximations for the step size* $\Delta t = 0.001$. *Without the split, our total computing time would be about* 5×10^5 *seconds as 6 cores of CPU are idle for long time.*

We also generate the periodic measure approximations from two paths each with 10000 periods and the step size $\Delta t = 0.01$ *under two different realisations of noise. The distributions of periodic measure are presented in Figure 6. One can see the distributions produced by two different Brownian motions are very similar. There are some minor differences due to insufficient amount of data. If we utilise sufficiently large amount of computations, the differences will eventually disappear.*

We can apply time scaling on the model to rescale its period to a much smaller number. By doing this, the Lipschitz coefficient of the drift term will become very large. According to the upper bound of step size in Theorem 5.1, the total cost of approximation will remain the same as the step size has to be very small.

	Result	CPU(seconds)
1	1.0120872	71050.63
2	1.0121229	70247.79
3	1.0120366	70782.51
4	1.0121788	70421.53
5	1.0118910	70679.84
6	1.0117187	69755.99
7	1.0115235	67978.79
8	1.0118564	60069.49
Mean value of $\Delta t = 0.001$ is 1.0119269		

Step sizes	$\Delta t = 1/50$	$\Delta t = 1/64$	$\Delta t = 1/80$
Approximation	1.0104580	1.0106331	1.0109203
Numerical error	0.0014689	0.0012938	0.0010066
CPU(seconds)	28630.88	36177.61	45509.97
Step sizes	$\Delta t = 1/100$	$\Delta t = 1/125$	
Approximation	1.0111015	1.0112500	
Numerical error	0.0008254	0.0006769	
CPU(seconds)	56691.58	49771.05	

Table 2: Numerical results of Example 6.2 where numerical results and errors are rounded off in 7 decimal places

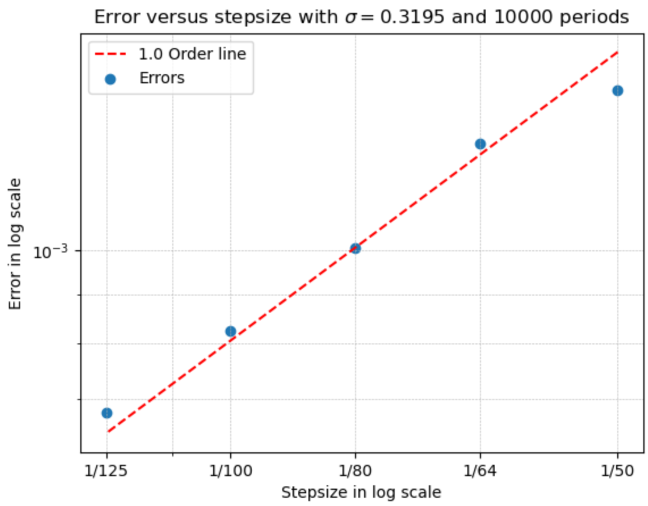


Figure 4: Error of approximation to average of periodic measure versus step size in log-log graph (Example 6.2)

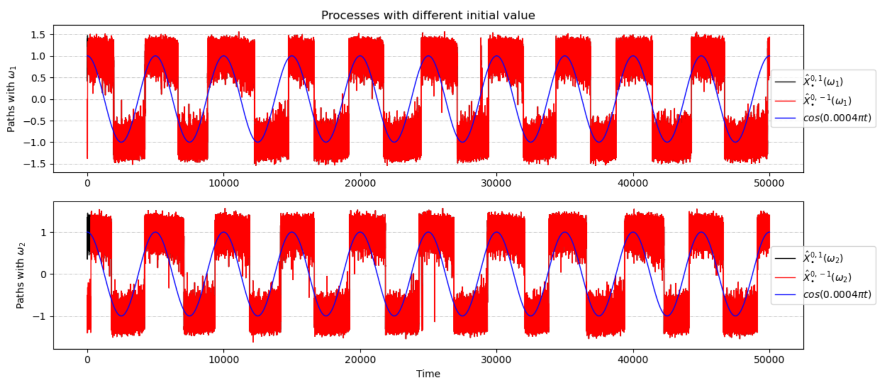
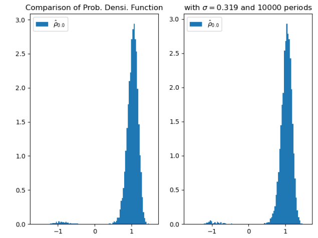
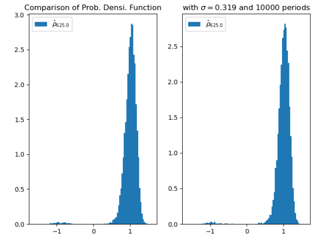


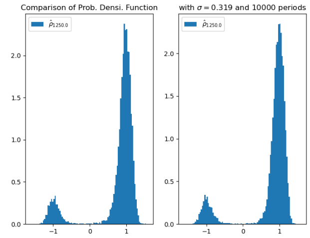
Figure 5: Paths of stochastic resonance model (Example 6.2)



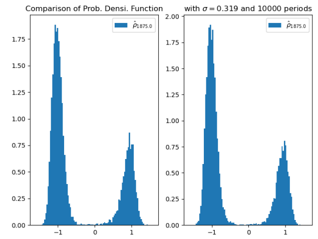
(a) Distribution of $\hat{\rho}_0$



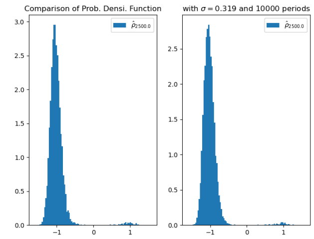
(b) Distribution of $\hat{\rho}_{625}$



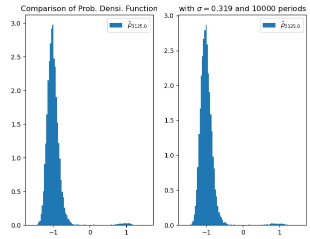
(c) Distribution of $\hat{\rho}_{1250}$



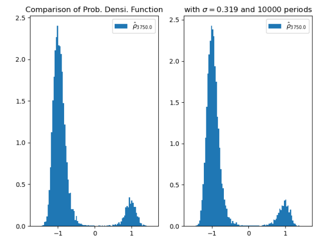
(d) Distribution of $\hat{\rho}_{1875}$



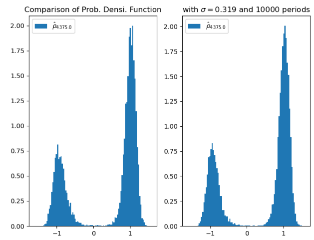
(e) Distribution of $\hat{\rho}_{2500}$



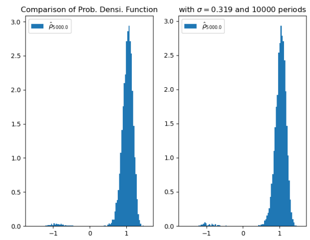
(f) Distribution of $\hat{\rho}_{3125}$



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



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



(i) Distribution of $\hat{\rho}_{5000}$

Figure 6: Approximations of periodic measure with $\Delta t = 0.01$ and 10000 periods (Example 6.2)

Acknowledgements

We would like to acknowledge the financial support of an EPSRC grant (ref. EP/S005293/2). We are very grateful to the referees for their constructive comments which led to significant improvements of this paper.

References

[1] R. A. Adams, J. J. F. Fournier, *Sobolev Spaces*, 2nd edn., Academic Press, China, 2009.

[2] P. W. Bates, K.N. Lu, B.X. Wang, Attractors of non-autonomous stochastic lattice systems in weighted spaces, *Phys. D*, 289 (2014), 32-50.

[3] F. E. Benth and J. Benth, The volatility of temperature and pricing of weather derivatives, *Quantitative Finance*, 7(5) (2007), 553-561.

[4] R. Benzi, G. Parisi, A. Sutera and A. Vulpiani, Stochastic resonance in climatic change, *Tellus*, 34 (1982), 10-16.

[5] M. D. Chekroun, E. Simonnet, M. Ghil, Stochastic climate dynamics: random attractors and time-dependent invariant measures, *Phys. D*, 240 (2011), 1685-1700.

[6] A. M. Cherubini, J. S. W. Lamb, M. Rasmussen, Y. Sato, A random dynamical systems perspective on stochastic resonance, *Nonlinearity*, 30 (2017), 2835-2853.

[7] M. Engel, C, Kuehn, A random dynamical systems perspective on isochronicity for stochastic oscillations, 2019, *arXiv:1911.08993*.

[8] C. Feng, Y. Liu, H. Zhao, Numerical approximation of random periodic solutions of stochastic differential equations, *Z. Angew. Math. Phys.*, 68 (2017) 119, pp1-32.

[9] C.R. Feng, B.Y. Qu and H.Z. Zhao, Random quasi-periodic paths and quasi-periodic measures of stochastic differential equations, *J. of Differential Equations*, Vol. 286 (2021), 119-163.

[10] C. Feng, Y. Wu, H. Zhao, Anticipating random periodic solutions–I. SDEs with multiplicative linear noise, *J. Funct. Anal.*, 271 (2016), 365-417.

[11] C. Feng, H. Zhao, Random periodic solutions of SPDEs via integral equations and Wiener-Sobolev compact embedding, *J. Funct. Anal.*, 262 (2012), 4377-4422.

[12] C. Feng, H. Zhao, Random periodic processes, periodic measures and ergodicity, *J. of Differential Equations*, 269 (2020), 7382-7428.

[13] C. Feng, H. Zhao, J. Zhong, Existence of geometric ergodic periodic measures of stochastic differential equations, Preprint, 2019, *arXiv:1904.08091.*

[14] C. Feng, H. Zhao, J. Zhong, Expected exit time for time-periodic stochastic differential equations and applications to stochastic resonance, *Physica D*, 417 (2021) 132815, pp. 1-18.

[15] C. Feng, H. Zhao, B. Zhou, Pathwise random periodic solutions of stochastic differential equations, *J. Differential Equations*, 251 (2011), 119-149.

[16] A. Grorud, D. Talay Approximation of Lyapunov exponents of nonlinear stochastic differential equations, *SIAM J. Appl. Math*, 56 (1996), 627-650.

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

[18] R. Höpfner, E. Löcherbach, M. Thieullen, Strongly degenerate time inhomogeneous SDEs: densities and support properties. Application to a Hodgkin-Huxley system with periodic input. *Bernoulli*, 23 (2017), 2587-2616.

[19] W. Huang, Z. Lian, K. Lu, Ergodic theory of random Anosov systems mixing on fibers, 2019, *arXiv:1612.08394.*

[20] A. Jentzen, P. Kloeden, Taylor expansions of solutions of stochastic partial differential equations with additive noise, *Ann. Probab.*, 38 (2010), 532-569.

[21] P. E. Kloeden, E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, New York, 1991.

[22] H. Kunita, Stochastic differential equations and stochastic flows of diffeomorphisms, *Ecole d'été de Probabilités de Saint-Flour XII-1982*, Lecture Notes in Mathematics, Springer, 1984.

[23] J. Mattingly, A.M. Stuart, D.J. Higham, Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise, *Stochastic Process. Appl.*, 101 (2002), 185-232.

[24] S. P. Meyn, R. L. Tweedie, *Markov Chains and Stochastic Stability*, Springer-Verlag, 1993.

[25] G. N. Milstein, *Numerical Integrations of Stochastic Differential Equations*, Kluwer, 1995.

[26] G. N. Milstein, M. V. Tretyakov, *Stochastic Numerics for Mathematical Physics*, Springer, 2004.

[27] E. Nummelin, *General irreducible Markov chains and non-negative operators*, Cambridge University Press, 1984.

[28] D. Talay, Second order discretization schemes of stochastic differential systems for the computation of the invariant law, *Stoch. Stoch. Rep.*, 29 (1990), 13-36.

[29] D. Talay, Approximation of upper Lyapunov exponents of bilinear stochastic differential systems, *SIAM J. Numer. Anal.*, 28 (1991), 1141-1164.

[30] D. Talay, L.Rubaro, Expansion of the global error for numerical schemes solving stochastic differential equations, *Stoch. Anal. Appl.*, 8 (1990), 483-509.

[31] A. Tocino, R. Ardanuy, Runge-Kutta methods for numerical solution of stochastic differential equations, *J. Comput. Appl. Math.*, 138 (2002), 219-241.

[32] R.L. Tweedie, Sufficient condition for ergodicity and recurrence of Markov chains on a general state space, *Stochastic Process. Appl.*, 3 (1975), 385-403.

[33] R.L. Tweedie, The existence of moments for stationary Markov chains, *J. Appl. Probab.*, 20 (1983), 191-196

[34] B.X. Wang, Existence, stability and bifurcation of random complete and periodic solutions of stochastic parabolic equations, *Nonlinear Anal.*, 103 (2014), 9-25.

[35] A. Yevik, H. Zhao, Numerical approximations to the stationary solutions of stochastic differential equations, *SIAM J. Numer. Anal.*, 49 (2011), 1397-1416.

[36] C. Yuan, X. Mao, Stability in distribution of numerical solutions for stochastic differential equations, *Stoch. Anal. Appl.*, 22 (2004), 1133-1150.

[37] H. Zhao, Z. Zheng, Random periodic solutions of random dynamical systems, *J. Differential Equations*, 246 (2009), 2020-2038.