Discrete and Continuous Dynamical Systems - Series B
Vol. 28, No. 1, January 2023, pp. 765–807
doi:10.3934/dcdsb.2022098

STOCHASTIC DIFFERENTIAL EQUATION WITH PIECEWISE CONTINUOUS ARGUMENTS: MARKOV PROPERTY, INVARIANT MEASURE AND NUMERICAL APPROXIMATION

Chuchu Chen, Jialin Hong and Yulan Lu*

LSEC, ICMSEC, Academy of Mathematics and Systems Science
Chinese Academy of Sciences, Beijing 100049, China
School of Mathematical Sciences
University of Chinese Academy of Sciences, Beijing 100049, China

(Communicated by Arnulf Jentzen)

Abstract. For the stochastic differential equation with piecewise continuous arguments, multiplicative noises and dissipative drift coefficients, we show that the solution at integer time is a Markov chain and admits a unique invariant measure. In order to numerically preserve the invariant measure, we apply the backward Euler method to the equation, and prove that the numerical solution at integer time is also a Markov chain and possesses a unique numerical invariant measure. By establishing several a priori estimations, we present the time-independent weak error analysis for the method via Malliavin calculus, which implies that the numerical invariant measure converges to the original one with weak order 1. Numerical experiments verify the theoretical analysis.

1. **Introduction.** In this paper, we consider the following stochastic differential equation (SDE) with piecewise continuous arguments (PCAs)

$$\begin{cases} dX(t) = f(X(t), X([t]))dt + g(X(t), X([t]))dB(t), & t > 0, \\ X(0) = x \in \mathbb{R}^d, \end{cases} \tag{1}$$

where $[\cdot]$ denotes the greatest-integer function, $f : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, $g : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^{d\times r}$ and $B(t)$ is an r-dimensional Brownian motion defined on a filtered complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$. Many phenomena in physics, biology, engineering and other fields can be modeled by differential equations with PCAs, such as elastic systems impelled by a Geneva wheel, population model with PCAs, machinery driven by servo units (see e.g., [7, 20, 23]). In practical circumstances, stochasticity is common, and thus SDEs with PCAs arise and attract lots of attention; see [12] for the application in neural networks, [15] for the application in control theory, etc. In fact, SDEs with PCAs present a hybrid of continuous and discrete dynamical

2020 *Mathematics Subject Classification.* Primary: 60H35, 37M25; Secondary: 65C30.
Key words and phrases. Invariant measure, Markov chain, weak convergence, backward Euler method, stochastic differential equations with piecewise continuous arguments.
This work is funded by National Natural Science Foundation of China (No. 11971470, No. 11871068, No. 12031020, No. 12022118, No. 12026428), by Youth Innovation Promotion Association CAS, and by National Postdoctoral Program for Innovative Talents (No. BX20180347).
* Corresponding author: Yulan Lu.

systems. Therefore, this type of equations possesses properties of both stochastic differential and difference equations, and exhibits complex and extraordinary dynamical behavior.

Invariant measure is a key characteristic in describing the long-time dynamical behavior. To our best knowledge, there is no result on the invariant measures of both Eq. (1) and numerical approximations. The aim of this paper is to investigate the invariant measures of a Markov chain based on an exact discrete-time sampling of Eq. (1) and backward Euler (BE) approximations. We concern with the following questions.

(I) Does an exact discrete-time sampling of Eq. (1) admit an invariant measure? If so, is it unique?

(II) If the previous Markov chain admits a unique invariant measure, does the BE method reproduce a unique numerical invariant measure?

(III) Does the numerical invariant measure, if it exists, converge to the original one?

For the stochastic functional differential equation (SFDE) with either continuous or discrete delay arguments, it is well known that the solution is non-Markovian because of the dependence on the history. However, the segment process of the SFDE with continuous arguments is proved to be Markovian. And the invariant measure of the SFDE with continuous arguments has been studied extensively (see, e.g., [2, 24] and references therein). Different from the equation with continuous arguments, for Eq. (1) whose arguments are discrete, we prove that the restriction of its solution at integer time, namely $\{X(k)\}_{k\in\mathbb{N}}$, is a time-homogeneous Markov chain. This reveals the influence of the discrete arguments and reflects the characteristic of difference dynamics of Eq. (1). By proving the exponential convergence of $\{X(k)\}_{k\in\mathbb{N}}$ in mean-square sense and the continuous dependence on the initial value of $\{X(k)\}_{k\in\mathbb{N}}$, we then obtain that the Markov chain $\{X(k)\}_{k\in\mathbb{N}}$ is exponentially ergodic with a unique invariant measure π.

Taking the divergence of explicit Euler method without the linear growth condition on drift coefficients into consideration, we apply the BE method to discretize Eq. (1). Denoting Y_k the BE approximation of $X(k)$, we show that $\{Y_k\}_{k\in\mathbb{N}}$ is also a time-homogenous Markov chain. The uniform boundedness of $\{Y_k\}_{k\in\mathbb{N}}$ in mean-square sense and the continuous dependence on the initial value guarantee the existence and uniqueness of the numerical invariant measure π^δ, δ is the step-size of the BE method. Moreover, the distribution of Y_k converges exponentially to π^δ as k tends to infinity, which means that the BE approximation $\{Y_k\}_{k\in\mathbb{N}}$ preserves the exponential ergodicity of the induced Markov chain $\{X(k)\}_{k\in\mathbb{N}}$ of Eq. (1). The error between π and π^δ is estimated via deducing the weak error between $X(k)$ and Y_k, which is required not only to be independent of k but also to decay exponentially.

By deriving several uniform a priori estimations and presenting the weak error analysis via Malliavin calculus, we show that π^δ converges to π with order 1 which coincides with the weak convergence order of the BE method.

This paper is organized as follows. In Section 2, some notations are introduced and the solution of Eq. (1) at integer time is proved to be a time-homogeneous Markov chain as well as exponentially ergodic with a unique invariant measure. In Section 3, we apply the BE method to Eq. (1) and prove that the BE approximation at integer time preserves the exponential ergodicity with a unique numerical invariant measure. The time-independent weak error of the solutions together with the error between invariant measures are given in Section 4. In Section 5, numerical experiments are presented to verify the theoretical results.

2. **Notation and invariant measure of the solution.** To begin with, we introduce some notations. Let $(\mathbb{R}^d, \langle\cdot,\cdot\rangle, \|\cdot\|)$ be the d-dimensional real Euclidean space. Given a matrix $A \in \mathbb{R}^{d\times r}$, its trace norm is defined as $\|A\| := \sqrt{\text{trace}(A^TA)}$. Denote by $C_b(\mathbb{R}^d)$ (resp. $B_b(\mathbb{R}^d)$) the Banach space of all uniformly continuous and bounded mappings (resp. Borel bounded mappings) $\varphi : \mathbb{R}^d \to \mathbb{R}$ endowed with the norm $\|\varphi\|_0 = \sup_{x\in\mathbb{R}^d} |\varphi(x)|$. For any $k \in \mathbb{N}$, $C_b^k(\mathbb{R}^d)$ is the subspace of $C_b(\mathbb{R}^d)$ consisting of all functions with bounded partial derivatives $D_x^i\varphi(x)$ for $1 \le i \le k$ and with the norm $\|\varphi\|_k = \|\varphi\|_0 + \sum_{i=1}^k \sup_{x\in\mathbb{R}^d} \|D_x^i\varphi(x)\|$. The notation $\mathscr{P}(\mathbb{R}^d)$ denotes the family of all probability measures on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$. For $a, b \in \mathbb{R}$, we denote $\max(a, b)$ and $\min(a, b)$ by $a \vee b$ and $a \wedge b$, respectively. We denote by 1_D the indicative function of a set D.

Now, we make the following assumptions on the drift and diffusion coefficients.

Assumption 2.1. *There exist $\lambda_1, \lambda_2, \lambda_3 > 0$ such that for any $x_1, y_1, x_2, y_2 \in \mathbb{R}^d$,*

$$\langle x_1 - x_2, f(x_1, y) - f(x_2, y)\rangle \le -\lambda_1\|x_1 - x_2\|^2, \tag{2}$$

$$\|f(x, y_1) - f(x, y_2)\|^2 \le \lambda_2 \|y_1 - y_2\|^2 \tag{3}$$

and

$$\|g(x_1, y_1) - g(x_2, y_2)\|^2 \le \lambda_3(\|x_1 - x_2\|^2 + \|y_1 - y_2\|^2). \tag{4}$$

From Assumption 2.1, for any $x, y \in \mathbb{R}^d$, we have

$$\begin{aligned} 2\langle x, f(x, y)\rangle =& 2\langle x - 0, f(x, y) - f(0, y)\rangle + 2\langle x, f(0, y)\rangle \\ \le& -(2\lambda_1 - 1)\|x\|^2 + 2\lambda_2\|y\|^2 + 2\|f(0, 0)\|^2 \end{aligned} \tag{5}$$

and

$$\begin{aligned} \|g(x, y)\|^2 \le& 2\|g(x, y) - g(0, 0)\|^2 + 2\|g(0, 0)\|^2 \\ \le& 2\lambda_3\|x\|^2 + 2\lambda_3\|y\|^2 + 2\|g(0, 0)\|^2. \end{aligned} \tag{6}$$

Under Assumption 2.1, Eq. (1) admits a unique global solution $\{X(t)\}_{t\ge0}$ and the solution is mean-square stable and almost sure stable (see, e.g., [13, 15, 12]). To demonstrate the dynamics of $\{X(k)\}_{k\in\mathbb{N}}$, for any $x \in \mathbb{R}^d$ and $B \in \mathscr{B}(\mathbb{R}^d)$, we define

$$P(x, B) = \mathbb{P}\{X(1) \in B|X(0) = x\} \quad \text{and} \quad P_k(x, B) = \mathbb{P}\{X(k) \in B|X(0) = x\}.$$

Unless otherwise specified, we write $X^{k,x}(t)$ in lieu of $X(t)$ to highlight the initial value $X(k) = x$. Let us first verify that $\{X(k)\}_{k\in\mathbb{N}}$ is indeed a Markov chain.

Theorem 2.2. *Suppose that Assumption 2.1 hold. Then $\{X(k)\}_{k\in\mathbb{N}}$ is a time-homogeneous Markov chain with the transition probability kernel $P(x, B)$.*

Proof. We divide this proof into two parts.

(i) Time-homogeneity. For $k, k' \in \mathbb{N}$, if $X(k') = x$, then

$$\begin{aligned} &X^{k',x}(k + k') \\ =& x + \int_{k'}^{k+k'} f(X^{k',x}(s), X^{k',x}([s]))ds + \int_{k'}^{k+k'} g(X^{k',x}(s), X^{k',x}([s]))dB(s) \\ =& x + \int_0^k f(X^{k',x}(u + k'), X^{k',x}([u] + k'))du \\ &+ \int_0^k g(X^{k',x}(u + k'), X^{k',x}([u] + k'))d\widetilde{B}(u), \end{aligned}$$

where $\widetilde{B}(u) = B(u+k') - B(k')$, $u \geq 0$. In addition, if $X(0) = x$, then

$$X^{0,x}(k) = x + \int_0^k f(X^{0,x}(u), X^{0,x}([u]))du + \int_0^k g(X^{0,x}(u), X^{0,x}([u]))dB(u).$$

Since $\widetilde{B}(u)$ and $B(u)$ have the same distribution, by the weak uniqueness of the solution for Eq. (1), we obtain that $X^{k',x}(k+k')$ and $X^{0,x}(k)$ are identical in distribution. Hence

$$\mathbb{P}\{X(k+k') \in B | X(k') = x\} = \mathbb{P}\{X(k) \in B | X(0) = x\}$$

for any $B \in \mathscr{B}(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$, which means that $\{X^{0,x}(k)\}_{k\in\mathbb{N}}$ is time-homogeneous.

(ii) Markov property. Define $\mathscr{G}_{t,s} = \sigma\{B(u) - B(s),\ s \leq u \leq t\} \cup \mathscr{N}$, where $s, t > 0$ and $\mathscr{N}$ denotes the collection of all $\mathbb{P}$-null sets in $\mathscr{F}$. The property of Brownian motion yields that $\mathscr{F}_s$ is independent of $\mathscr{G}_{t,s}$. For $k \in \mathbb{N}$, let $B^k(t) := B(t) - B(k)$, $t \geq k$. Then $B^k(t)$ is $\mathscr{F}_t \cap \mathscr{G}_{t,k}$-measurable. Fix $y \in \mathbb{R}^d$. Replacing $B(t)$ by $B^k(t)$ in Eq. (1), we get the unique solution $\{X^{k,y}(t)\}_{t\geq k}$, which is adapted to $\{\mathscr{F}_t \cap \mathscr{G}_{t,k}\}_{t\geq k}$. Thus, $X^{k,y}(k+k')$ is independent of $\mathscr{F}_k$ for any $k, k' \in \mathbb{N}$.

For any fixed $k, k' \in \mathbb{N}$, define $\Psi : \mathbb{R}^d \times \Omega \to \mathbb{R}^d$, $(y, \omega) \mapsto X^{k,y}(k+k', \omega)$. We claim that Ψ is $\mathscr{B}(\mathbb{R}^d) \otimes \mathscr{G}_{k+k',k}$-measurable. By Itô's formula, Assumption 2.1 yields

$$\begin{aligned}
&\mathbb{E}\left\|X^{k,z}(t) - X^{k,y}(t)\right\|^2 \\
&= \|z-y\|^2 + \mathbb{E}\int_k^t \left\|g(X^{k,z}(s), X^{k,z}([s])) - g(X^{k,y}(s), X^{k,y}([s]))\right\|^2 ds \\
&\quad + 2\mathbb{E}\int_k^t \left\langle X^{k,z}(s) - X^{k,y}(s), f(X^{k,z}(s), X^{k,z}([s])) - f(X^{k,y}(s), X^{k,y}([s]))\right\rangle ds \\
&\leq \|z-y\|^2 - (2\lambda_1 - \lambda_3 - 1)\mathbb{E}\int_k^t \left\|X^{k,z}(s) - X^{k,y}(s)\right\|^2 ds \\
&\quad + (\lambda_2 + \lambda_3)\mathbb{E}\int_k^t \left\|X^{k,z}([s]) - X^{k,y}([s])\right\|^2 ds \\
&\leq \|z-y\|^2 + \lambda \int_k^t \sup_{k\leq u\leq s} \mathbb{E}\left\|X^{k,z}(u) - X^{k,y}(u)\right\|^2 ds,
\end{aligned}$$

where $\lambda := |2\lambda_1 - \lambda_3 - 1| + \lambda_2 + \lambda_3$. By Gronwall's inequality, we have

$$\mathbb{E}\left\|X^{k,z}(t) - X^{k,y}(t)\right\|^2 \leq e^{\lambda(t-k)} \|z-y\|^2, \tag{7}$$

which implies that Ψ is continuous in probability with respect to y, i.e., for any $\varepsilon > 0$

$$\mathbb{P}\left\{\omega \in \Omega : \left\|X^{k,z}(k+k', \omega) - X^{k,y}(k+k', \omega)\right\| > \varepsilon\right\} \to 0, \qquad \text{as } z \to y.$$

Theorem 3.1 in [8] implies that there is a modification $\widetilde{\Psi}$ of Ψ that is $\mathscr{B}(\mathbb{R}^d) \otimes \mathscr{G}_{k+k',k}$-measurable. Therefore, $\varphi(X^{0,x}(k+k'))$ is $\mathscr{B}(\mathbb{R}^d) \otimes \mathscr{G}_{k+k',k}$-measurable for any $\varphi \in B_b(\mathbb{R}^d)$, where the uniqueness of the solution to Eq. (1) is used, i.e.,

$$X^{0,x}(k+k') = X^{k,X^{0,x}(k)}(k+k') = \widetilde{\Psi}(X^{0,x}(k)), \qquad a.s.$$

Combining the fact that $X^{0,x}(k)$ is $\mathscr{F}_k$-measurable, we have

$$\mathbb{E}\left[\varphi(X^{k,X^{0,x}(k)}(k+k')) | \mathscr{F}_k\right] = \mathbb{E}\left[\varphi(X^{k,y}(k+k')) | \mathscr{F}_k\right]\Big|_{y=X^{0,x}(k)}$$

and

$$\mathbb{E}\left[\varphi(X(k+k')) | \mathscr{F}_k\right] = \mathbb{E}\left[\varphi(X^{k,y}(k+k'))\right]\Big|_{y=X(k)} = \mathbb{E}\left[\varphi(X(k+k')) | X(k)\right].$$

The proof is completed. □

Theorem 2.3. *Under Assumption 2.1, if* $\lambda_1 - \lambda_2 - 2\lambda_3 - 1 > 0$*, then the Markov chain* $\{X(k)\}_{k\in\mathbb{N}}$ *admits a unique invariant measure* π *and there exist two positive constants* $C_1 := C_1(\|\varphi\|_1, \lambda_1, \lambda_2, \lambda_3)$ *and* $\nu := \nu(\lambda_1, \lambda_2, \lambda_3)$ *independent of* k *and* x *such that*

$$\left|\mathbb{E}\varphi(X^{0,x}(k)) - \int_{\mathbb{R}^d} \varphi(x)\pi(dx)\right| \le C_1 e^{-\nu k}(1 + \|x\|), \quad \forall\ \varphi \in C_b^1(\mathbb{R}^d). \tag{8}$$

Proof. **(i) Existence of invariant measures.** Let $\alpha := 2\lambda_1 - 2\lambda_3 - 1$, $\beta := 2(\lambda_2 + \lambda_3)$ and $\gamma := 2(\|f(0,0)\|^2 + \|g(0,0)\|^2)$, then $\alpha > 0$ and $\frac{\beta}{\alpha} < 1$ since $\lambda_1 - \lambda_2 - 2\lambda_3 - 1 = \frac{1}{2}(\alpha - \beta - 1) > 0$. Let $\{\tilde{B}(t)\}_{t\ge0}$ be another Brownian motion defined on $(\Omega, \mathcal{F}, \mathbb{P})$, independent of $\{B(t)\}_{t\ge0}$, and define

$$\bar{B}(t) = \begin{cases} B(t), & t \ge 0, \\ \tilde{B}(-t), & t < 0 \end{cases}$$

with the filtration $\bar{\mathcal{F}}_t := \sigma\{\bar{B}(s), s \le t\}$, $t \in \mathbb{R}$. For any $k \in \mathbb{N}$ and $x \in \mathbb{R}^d$, we consider the following equation

$$\begin{cases} dX(t) = f(X(t), X([t]))dt + g(X(t), X([t]))d\bar{B}(t), & t \ge -k, \\ X(-k) = x. \end{cases} \tag{9}$$

It can be verified that (9) admits a unique solution under Assumption 2.1. In what follows, we show the existence of invariant measure through three steps.

Step 1. A priori estimate

For any $k \in \mathbb{N}$ and $t > -k$, applying Itô's formula to $e^{\alpha t}\|X^{-k,x}(t)\|^2$, (5)-(6) lead to

$$\begin{aligned}
& e^{\alpha t}\mathbb{E}\|X^{-k,x}(t)\|^2 \\
=& e^{\alpha[t]}\mathbb{E}\|X^{-k,x}([t])\|^2 + \alpha\mathbb{E}\int_{[t]}^t e^{\alpha s}\|X^{-k,x}(s)\|^2 ds \\
+& \mathbb{E}\int_{[t]}^t e^{\alpha s}\left(2\left\langle X^{-k,x}(s), f(X^{-k,x}(s), X^{-k,x}([s]))\right\rangle + \|g(X^{-k,x}(s), X^{-k,x}([s]))\|^2\right) ds \\
\le& \left(e^{\alpha[t]} + \frac{\beta}{\alpha}\left(e^{\alpha t} - e^{\alpha[t]}\right)\right)\mathbb{E}\|X^{-k,x}([t])\|^2 + \frac{\gamma}{\alpha}\left(e^{\alpha t} - e^{\alpha[t]}\right).
\end{aligned}$$

Hence

$$\mathbb{E}\|X^{-k,x}(t)\|^2 \le \left(\frac{\beta}{\alpha} + \left(1 - \frac{\beta}{\alpha}\right)e^{-\alpha\{t\}}\right)\mathbb{E}\|X^{-k,x}([t])\|^2 + \frac{\gamma}{\alpha}\left(1 - e^{-\alpha\{t\}}\right), \tag{10}$$

where $\{t\} = t - [t] \in [0,1)$. Define $r : [0,1) \to (0,1]$ by $r(\{t\}) = \frac{\beta}{\alpha} + \left(1 - \frac{\beta}{\alpha}\right)e^{-\alpha\{t\}}$. Since $\lim_{t\to\bar{k}^-}\{t\} = 1$ for any $\bar{k} \in \mathbb{N}$, we extend the domain of r to $[0,1]$ and define $r(1) := \lim_{t\to\bar{k}^-} r(\{t\}) = \frac{\beta}{\alpha} + \left(1 - \frac{\beta}{\alpha}\right)e^{-\alpha}$. Let $F = \frac{\gamma}{\alpha}$. Then Fatou's lemma leads to

$$\begin{aligned}
\mathbb{E}\|X^{-k,x}(\bar{k})\|^2 =& \mathbb{E}\lim_{t\to\bar{k}^-}\|X^{-k,x}(t)\|^2 \\
&\le \lim_{t\to\bar{k}^-}\mathbb{E}\|X^{-k,x}(t)\|^2 \le r(1)\mathbb{E}\|X^{-k,x}(\bar{k} - 1)\|^2 + F,
\end{aligned}$$

which implies

$$\begin{aligned}\mathbb{E}\|X^{-k,x}(t)\|^2 \leq & r(\{t\})r(1)\mathbb{E}\|X^{-k,x}([t]-1)\|^2 + r(\{t\})F + F \\ \leq & \cdots \\ \leq & r(\{t\})e^{([t]+k)\log r(1)}\|x\|^2 + \frac{1-r(1)^{[t]+k}}{1-r(1)}r(\{t\})F + F \\ \leq & \frac{1}{r(1)}e^{(t+k)\log r(1)}\|x\|^2 + \frac{1}{1-r(1)}F + F.\end{aligned}$$

Since $\log r(1) < 0$, there exists a positive constant C independent of k and t such that

$$\sup_{k\in\mathbb{N}} \mathbb{E}\|X^{-k,x}(t)\|^2 \leq \frac{1}{r(1)}\|x\|^2 + \frac{1}{1-r(1)}F + F \leq C(1+\|x\|^2). \tag{11}$$

Step 2. For any $k_1, k_2 \in \mathbb{N}$, $-k_1 < -k_2 \leq t < \infty$, let $Z(t) = X^{-k_1,x}(t) - X^{-k_2,x}(t)$, then

$$\begin{aligned}Z(t) =& Z(-k_2) + \int_{-k_2}^{t} \Big(f(X^{-k_1,x}(s), X^{-k_1,x}([s])) - f(X^{-k_2,x}(s), X^{-k_2,x}([s]))\Big)\, ds \\ &+ \int_{-k_2}^{t} \Big(g(X^{-k_1,x}(s), X^{-k_1,x}([s])) - g(X^{-k_2,x}(s), X^{-k_2,x}([s]))\Big)\, dB(s).\end{aligned}$$

Similar to Step 1, applying Itô's formula to $e^{\alpha t}\mathbb{E}\|Z(t)\|^2$, Assumption 2.1 leads to

$$\mathbb{E}\|Z(t)\|^2 \leq r_1(\{t\})\mathbb{E}\,\|Z([t])\|^2,$$

where the function r_1 is defined similarly to r in step 1 with $r_1(\{t\}) = \frac{\lambda_2+\lambda_3}{2\lambda_1-\lambda_3-1} + \left(1 - \frac{\lambda_2+\lambda_3}{2\lambda_1-\lambda_3-1}\right) e^{-(2\lambda_1-\lambda_3-1)\{t\}}$ for $\{t\} \in [0,1)$ and $r_1(1) := \lim_{t\to k^-} r_1(\{t\})$, $k \in \mathbb{N}$. Then
we derive

$$\begin{aligned}\mathbb{E}\|Z(t)\|^2 \leq & \frac{1}{r_1(1)}e^{(t+k_2)\log r_1(1)}\mathbb{E}\left\|X^{-k_1,x}(-k_2) - x\right\|^2 \\ \leq & Ce^{(t+k_2)\log r_1(1)}\left(1+\|x\|^2\right).\end{aligned}$$

In particular,

$$\mathbb{E}\left\|X^{-k_1,x}(0) - X^{-k_2,x}(0)\right\|^2 \leq Ce^{k_2\log r_1(1)}\left(1+\|x\|^2\right),$$

which implies that $\{X^{-k,x}(0)\}_{k\in\mathbb{N}}$ is a Cauchy sequence in $L^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^d)$. Therefore, there exists $\eta^x \in L^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^d)$ such that

$$\lim_{k\to+\infty} \mathbb{E}\left\|X^{-k,x}(0) - \eta^x\right\|^2 = 0. \tag{12}$$

Moreover, following the similar procedure, we obtain

$$\mathbb{E}\left\|X^{-k,x}(0) - X^{-k,y}(0)\right\|^2 \leq \frac{1}{r_1(1)}e^{k\log r_1(1)}\|x-y\|^2. \tag{13}$$

Combining (12) and (13), we get

$$\begin{aligned}&\mathbb{E}\,\|\eta^x - \eta^y\|^2 \\ =&\mathbb{E}\left\|\eta^x - X^{-k,x}(0) + X^{-k,x}(0) - X^{-k,y}(0) + X^{-k,y}(0) - \eta^y\right\|^2 \\ \leq& 3\lim_{k\to\infty}\mathbb{E}\left(\left\|\eta^x - X^{-k,x}(0)\right\|^2 + \left\|X^{-k,x}(0) - X^{-k,y}(0)\right\|^2 + \left\|X^{-k,y}(0) - \eta^y\right\|^2\right) \\ =&0.\end{aligned}$$

This means that η^x is independent of the initial value x, which is thus denoted by η. Furthermore,

$$\mathbb{E}\left\|X^{-k_2,x}(0)-\eta\right\|^2=\lim_{k_1\to+\infty}\mathbb{E}\left\|X^{-k_2,x}(0)-X^{-k_1,x}(0)\right\|^2\le Ce^{k_2\log r_1(1)}\left(1+\|x\|^2\right),$$

which indicates that $X^{-k,x}(0)$ converges to η in distribution as $k\to\infty$. Since $X^{-k,x}(0)$ and $X^{0,x}(k)$ possess the same distribution, by the definition of convergence in distribution, the transition probabilities $P_k(x,\cdot)=\mathbb{P}\{X(k)\in\cdot|X(0)=x\}$ weakly converges to $\mathbb{P}\circ\eta^{-1}(\cdot)$ as $k\to\infty$.

Step 3. Denoting by $\pi:=\mathbb{P}\circ\eta^{-1}$ the probability measure induced by η, we claim that π is an invariant measure. In fact, for any $B\in\mathscr{B}(\mathbb{R}^d)$, the Chapman-Kolmogorov equation leads to

$$\begin{aligned}\pi(B)&=\int_{\mathbb{R}^d}1_B(y)\pi(dy)=\lim_{k\to+\infty}\int_{\mathbb{R}^d}1_B(y)P_{k+1}(x,dy)\\&=\lim_{k\to+\infty}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}1_B(y)P(z,dy)P_k(x,dz)\\&=\lim_{k\to+\infty}\int_{\mathbb{R}^d}P(z,B)P_k(x,dz)=\int_{\mathbb{R}^d}P(z,B)\pi(dz).\end{aligned}$$

(ii) Uniqueness of the invariant measure. Since $P_k(x,\cdot)$ weakly converges to π as $k\to\infty$, for any $B\in\mathscr{B}(\mathbb{R}^d)$, $x\in\mathbb{R}^d$ and $k\in\mathbb{N}$, we get

$$\pi(B)=\int_{\mathbb{R}^d}1_B(y)\pi(dy)=\lim_{k\to+\infty}\int_{\mathbb{R}^d}1_B(y)P_k(x,dy)=\lim_{k\to+\infty}P_k(x,B).$$

Assume that $\tilde{\pi}\in\mathscr{P}(\mathbb{R}^d)$ is another invariant measure of $\{X(k)\}_{k\in\mathbb{N}}$, then for any $B\in\mathscr{B}(\mathbb{R}^d)$ and $k\in\mathbb{N}$,

$$\tilde{\pi}(B)=\int_{\mathbb{R}^d}P_k(x,B)\tilde{\pi}(dx).$$

Letting $k\to\infty$, we obtain

$$\tilde{\pi}(B)=\lim_{k\to\infty}\int_{\mathbb{R}^d}P_k(x,B)\tilde{\pi}(dx)=\pi(B),$$

which implies that π is the unique invariant measure of $\{X(k)\}_{k\in\mathbb{N}}$.

(iii) For any $\varphi\in C_b^1(\mathbb{R}^d)$, (2) and $\pi=\mathbb{P}\circ\eta^{-1}$ lead to

$$\begin{aligned}\left|\mathbb{E}\varphi(X^{0,x}(k))-\int_{\mathbb{R}^d}\varphi(x)\pi(dx)\right|&\le\mathbb{E}\left|\varphi(X^{0,x}(k))-\varphi(\eta)\right|\le\|\varphi\|_1\cdot\mathbb{E}\left\|X^{0,x}(k)-\eta\right\|\\&\le C_1e^{-\nu k}(1+\|x\|),\end{aligned}$$

where $C_1=\frac{\|\varphi\|_1}{\sqrt{r_1(1)}}$ and $\nu=-\frac{1}{2}\log r_1(1)$. The proof is completed. □

Remark 1. Condition (2) implies the contractive property of f, however, it can be weaken in certain sense. If we impose a mild assumption, i.e., there exist $\lambda_1>\lambda_2>0$ such that

$$2\langle x,f(x,y)\rangle+\|g(x,y)\|^2\le-\lambda_1\|x\|^2+\lambda_2\|y\|^2,\quad\forall\ x,y\in\mathbb{R}^d,$$

then $\{X(k)\}_{k\in\mathbb{N}}$ still admits an invariant measure. To ensure the uniqueness of the invariant measure, we need some additional restriction on diffusion coefficients, such as ellipticity, i.e.,

$$g(x,y)g(x,y)^\top>0,\quad x,y\in\mathbb{R}^d.$$

And

$$\langle x_1 - x_2, f(x_1, y_1) - f(x_2, y_2)\rangle \leq M(\|x_1 - x_2\|^2 + \|y_1 - y_2\|^2), \quad \forall\, x_1, y_1, x_2, y_2 \in \mathbb{R}^d,$$

$$\|g(x, y)\|^2 \leq M(1 + \|x\|^2 + \|y\|^2), \quad \forall\, x, y \in \mathbb{R}^d,$$

for some $M > 0$. By the asymptotic coupling method and the Girsanov transformation used in [21], the Markov chain $\{X(k)\}_{k\in\mathbb{N}}$ can be proved to be strong Feller and irreducible. Thus $\{X(k)\}_{k\in\mathbb{N}}$ admits a unique invariant measure.

Besides a priori estimate in Theorem 2.3, we also present the uniform boundedness of $X(t)$ in $2p$-th ($p \geq 1$) moment, which is crucial to estimate the time-independent weak error of the numerical method. Throughout this paper, we use C to denote a generic constant, which may be different from line to line.

Lemma 2.4. *Let Assumption 2.1 hold and $p \geq 1$. If $\lambda_1 - \lambda_2 - 2\lambda_3 - 1 > 4\lambda_3(p-1)$, then there exists a positive constant $C := C(\lambda_1, \lambda_2, \lambda_3, p)$ independent of t such that*

$$\mathbb{E}\|X(t)\|^{2p} \leq C(1 + \|x\|^{2p}). \tag{14}$$

Proof. Theorem 2.3 (i) implies that the assertion (14) holds for the case $p = 1$. Thus, it suffices to consider the case $p > 1$, which is proved by the induction.

We assume that there exists $C > 0$ independent of t such that (14) holds for all $p' = 1, 2, \cdots, p-1$, then we show $\mathbb{E}\|X(t)\|^{2p} \leq C(1 + \|x\|^{2p})$. Applying Itô's formula to $\mathbb{E}\|X(t)\|^{2p}$, (5) and (6) lead to

$$\begin{aligned}
\mathbb{E}\|X(t)\|^{2p} \leq& \|x\|^{2p} + 2p\mathbb{E}\int_0^t \|X(s)\|^{2p-2}\langle X(s), f(X(s), X([s]))\rangle ds \\
&+ p(2p-1)\mathbb{E}\int_0^t \|X(s)\|^{2p-2}\|g(X(s), X([s]))\|^2 ds \\
\leq& \|x\|^{2p} - p(2\lambda_1 - 2\lambda_3 - 1 - 4\lambda_3(p-1))\mathbb{E}\int_0^t \|X(s)\|^{2p} ds \qquad (15)\\
&+ 2p\left(\lambda_2 + \lambda_3(2p-1)\right)\mathbb{E}\int_0^t \|X(s)\|^{2p-2}\|X([s])\|^2 ds \\
&+ 2p\left(\|f(0,0)\|^2 + (2p-1)\|g(0,0)\|^2\right)\mathbb{E}\int_0^t \|X(s)\|^{2p-2} ds.
\end{aligned}$$

Using Young's inequality and the assumption $\mathbb{E}\|X(t)\|^{2(p-1)} \leq C(1 + \|x\|^{2(p-1)})$, we obtain

$$\begin{aligned}
\mathbb{E}\|X(t)\|^{2p} \leq& \|x\|^{2p} - \alpha_1(p)\int_0^t \mathbb{E}\|X(s)\|^{2p} ds \\
&+ \int_0^t \left(\gamma_1(p)(1 + \|x\|^{2(p-1)}) + \beta_1(p)\mathbb{E}\|X([s])\|^{2p}\right) ds,
\end{aligned}$$

where $\alpha_1(p) = 2\lambda_1 p - 2\lambda_2 p + 2\lambda_2 - p - 2\lambda_3(2p-1)^2$, $\beta_1(p) = 2\lambda_2 + 2\lambda_3(2p-1)$ and $\gamma_1(p) = 2p\left(\|f(0,0)\|^2 + (2p-1)\|g(0,0)\|^2\right)C$. In addition,

$$\begin{aligned}
\mathbb{E}\|X(t)\|^{2p} \leq& \|x\|^{2p} - \alpha_1(p)\int_0^t \mathbb{E}\|X(s)\|^{2p} ds \\
&+ \int_0^t \left(\gamma_1(p)(1 + \|x\|^{2(p-1)}) + \beta_1(p)\sup_{0\leq r\leq s}\mathbb{E}\|X(r)\|^{2p}\right) ds.
\end{aligned}$$

According to [11, Lemma 8.1], we have

$$\begin{aligned}&\sup_{0\le s\le t}\mathbb{E}\|X(s)\|^{2p}\\ \le&\|x\|^{2p}+\int_0^t e^{-\alpha_1(p)(t-s)}\left(\gamma_1(p)(1+\|x\|^{2(p-1)})+\beta_1(p)\sup_{0\le r\le s}\mathbb{E}\|X(r)\|^{2p}\right)ds.\end{aligned}$$

Due to $2\lambda_1-2\lambda_2-4\lambda_3-1>8\lambda_3(p-1)$, it can be verified that $\alpha_1(p)>\beta_1(p)>0$. Thus, [11, Lemma 8.2] leads to

$$\mathbb{E}\|X(t)\|^{2p}\le\frac{\gamma_1(p)(1+\|x\|^{2(p-1)})+\alpha_1(p)\|x\|^{2p}}{\alpha_1(p)-\beta_1(p)}\le C(1+\|x\|^{2p}),$$

where $C:=\frac{2\gamma_1(p)+\alpha_1(p)}{\alpha_1(p)-\beta_1(p)}$. The proof is completed. □

Remark 2. The inequality $\lambda_1-\lambda_2-2\lambda_3>4p(\lambda_3-1)$ is a sufficient condition which ensures the uniform boundedness of the $2p$th moments of the exact solution. For some p that violates this inequality, the uniform boundedness of the $2p$th moments of the exact solution may also hold.

Remark 3. If the equation is driven by additive noise, i.e. $g(x,y)=G$ which is a constant matrix, the condition under which the $2p$-th moment estimate (14) in Lemma 2.4 holds is changed to be $\lambda_1-\lambda_2-2\lambda_3-1>0$.

3. **Invariant measure of backward Euler method.** Numerical methods for the SDE with PCAs are widely studied. For example, [17] studies the convergence and stability of the Euler–Maruyama method for the equation with globally Lipschitz and linearly growing coefficients. For the non-globally Lipschitz case, the convergence and stability of some implicit numerical methods such as split-step θ method, one-leg θ method are studied in [13, 14, 25] and references therein. Here, we apply the BE method to numerically solve Eq. (1), and investigate whether the BE approximations at integer time reproduce a unique numerical invariant measure. For numerical approximations of invariant measures, there are fruitful results on that for the SDE and the stochastic partial differential equation, such as [4, 5, 9, 1, 10].

Let $\delta=\frac{1}{m}$ be the given step-size with integer $m\ge 1$. Grid points t_n are defined as $t_n=n\delta,\ n=0,1,\cdots$. The backward Euler (BE) method for (1) is given by

$$X_{n+1}=X_n+\delta f(X_{n+1},X_{[n\delta]m})+g(X_n,X_{[n\delta]m})\Delta B_n,$$

where $X_0=x$, $\Delta B_n=B(t_{n+1})-B(t_n)$, X_n (resp. $X_{[n\delta]m}$) is the approximation to $X(t_n)$ (resp. $X([t_n])$). Since, for arbitrary $n=0,1,2,\cdots$, there exist $k\in\mathbb{N}$ and $l=0,1,2,\cdots,m-1$ such that $n=km+l$, the BE method can be written as

$$X_{km+l+1}=X_{km+l}+\delta f(X_{km+l+1},X_{km})+g(X_{km+l},X_{km})\Delta B_{km+l}. \tag{16}$$

Under the condition (2), the BE method admits a unique solution $\{X_{km+l}:l=0,1,\cdots,m-1,k\in\mathbb{N}\}$ for all step-sizes. Rewrite (16) as

$$X_{km+l+1}-\delta f(X_{km+l+1},X_{km})=X_{km+l}+g(X_{km+l},X_{km})\Delta B_{km+l}. \tag{17}$$

For any $a\in\mathbb{R}^d$ and $\delta\in(0,1)$, define the mapping $G:\mathbb{R}^d\to\mathbb{R}^d$, $x\mapsto x-\delta f(x,a)$. Then G admits its inverse function $G^{-1}:\mathbb{R}^d\to\mathbb{R}^d$ based on [19, Theorem 6.4.4]. Moreover, the numerical solution X_{km+l+1} satisfies

$$X_{km+l+1}=G^{-1}(X_{km+l}+g(X_{km+l},X_{km})\Delta B_{km+l}) \tag{18}$$

for all $k\in\mathbb{N}$ and $l=0,1,2,\cdots,m-1$.

In order to investigate whether the BE method inherits the Markov property and admits a unique numerical invariant measure, we denote by $Y_k := X_{km}$ the solution of BE method at $t = k$, $k \in \mathbb{N}$ and define

$$P^\delta(x, B) = \mathbb{P}\{Y(1) \in B|Y(0) = x\} \quad \text{and} \quad P_k^\delta(x, B) = \mathbb{P}\{Y(k) \in B|Y(0) = x\},$$

where $x \in \mathbb{R}^d$ and $B \in \mathscr{B}(\mathbb{R}^d)$. Similar to $X^{0,x}(k)$, we write $Y_k^{0,x}$ in lieu of Y_k to highlight the initial value $Y_0 = x$. Now, let us proceed to show the Markov property of $\{Y_k\}_{k\in\mathbb{N}}$.

Theorem 3.1. *Assume that Assumption 2.1 hold. Then $\{Y_k\}_{k\in\mathbb{N}}$ is a time-homogeneous Markov chain with the transition probability $P^\delta(x, B)$.*

Proof. **(i) Time-homogeneity.** For $k \in \mathbb{N}$, if $Y_k = x$, i.e., $X_{km} = x$, then from (18), it follows

$$X_{km+1}^{km,x} = G^{-1}(x + g(x, x)\Delta B_{km}).$$

Define the mapping $G_1 : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d$, $(y, z) \mapsto G^{-1}(y + g(y, x)z)$, then $X_{km+1}^{km,x} = G_1(x, \Delta B_{km})$. In addition, if $Y_0 = X_0 = x$, then

$$X_1^{0,x} = G^{-1}(x + g(x, x)\Delta B_0) = G_1(x, \Delta B_0).$$

Since ΔB_{km} and ΔB_0 are identical in distribution, $X_{km+1}^{km,x}$ and $X_1^{0,x}$ possess the same distribution.

Again, by (18), we have

$$\begin{aligned} X_{km+2}^{km,x} =& G^{-1}(X_{km+1}^{km,x} + g(X_{km+1}^{km,x}, x)\Delta B_{km+1}) \\ =& G^{-1}(G_1(x, \Delta B_{km})) + g(G_1(x, \Delta B_{km}), x)\Delta B_{km+1}) \end{aligned}$$

and

$$X_2^{0,x} = G^{-1}(X_1^{0,x} + g(X_1^{0,x}, x)\Delta B_1) = G^{-1}(G_1(x, \Delta B_0)) + g(G_1(x, \Delta B_0), x)\Delta B_1).$$

Therefore, there exists a function $G_2 : \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^d$ such that

$$X_{km+2}^{km,x} = G_2(x, \Delta B_{km}, \Delta B_{km+1})$$

and

$$X_2^{0,x} = G_2(x, \Delta B_0, \Delta B_1).$$

Since $(\Delta B_{km}, \Delta B_{km+1})$ and $(\Delta B_0, \Delta B_1)$ have the same distribution, $X_{km+2}^{km,x}$ and $X_2^{0,x}$ are identical in distribution.

By the same procedure as above, there exists a function G_m such that

$$Y_{k+1}^{k,x} = X_{(k+1)m}^{km,x} = G_m(x, \Delta B_{km}, \Delta B_{km+1}, \cdots, \Delta B_{km+m-1}) \tag{19}$$

and

$$Y_1^{0,x} = X_m^{0,x} = G_m(x, \Delta B_0, \Delta B_1, \cdots, \Delta B_{m-1}).$$

Since $(\Delta B_{km}, \Delta B_{km+1}, \cdots, \Delta B_{km+m-1})$ and $(\Delta B_0, \Delta B_1, \cdots, \Delta B_{m-1})$ have the same distribution, $Y_{k+1}^{k,x}$ and $Y_1^{0,x}$ are also identical in distribution. Hence

$$\mathbb{P}\{Y_{k+1} \in B|Y_k = x\} = \mathbb{P}\{Y_1 \in B|Y_0 = x\}$$

for any $B \in \mathscr{B}(\mathbb{R}^d)$. Further, for any $k, k' \in \mathbb{N}$, we have

$$\mathbb{P}\{Y_{k+k'} \in B|Y_{k'} = x\} = \mathbb{P}\{Y_k \in B|Y_0 = x\},$$

which implies the time-homogeneous property.

(ii) Markov property. By the uniqueness of the numerical solution of (16), we have

$$Y_{k+1}^{0,x} = X_{(k+1)m}^{0,x} = X_{(k+1)m}^{km,X_{km}^{0,x}} = Y_{k+1}^{k,Y_k^{0,x}}, \qquad a.s.$$

For $k \in \mathbb{N}$, define $\bar{\mathcal{G}}_{k+1,k} := \sigma\{\Delta B_{km+l}, l = 0, 1, 2, \cdots, m-1\}$. Then $\bar{\mathcal{G}}_{k+1,k}$ is independent of $\mathcal{F}_k$. From (19), for fixed $y \in \mathbb{R}^d$, we know that $Y_{k+1}^{k,y}$ is $\bar{\mathcal{G}}_{k+1,k}$-measurable, and thus is independent of $\mathcal{F}_k$. Using similar techniques as Step 2 in Theorem 2.2, we obtain that $Y_{k+1}^{k,\cdot}$ is $\mathcal{B}(\mathbb{R}^d) \otimes \bar{\mathcal{G}}_{k+1,k}$-measurable. Since $Y_k^{0,x}$ is $\mathcal{F}_k$-measurable,

$$\mathbb{E}\left[\varphi(Y_{k+1})|\mathcal{F}_k\right] = \mathbb{E}\left[\varphi(Y_{k+1}^{k,Y_k})|\mathcal{F}_k\right] = \mathbb{E}\left[\varphi(Y_{k+1}^{k,y})\right]\Big|_{y=Y_k} = \mathbb{E}\left[\varphi(Y_{k+1})|Y_k\right],$$

which is the required Markov property. Further, the same procedure yields

$$\mathbb{E}\left[\varphi(Y_{k+k'})|\mathcal{F}_k\right] = \mathbb{E}\left[\varphi(Y_{k+k'})|Y_k\right].$$

The proof is completed. □

Remark 4. Though SDEs with PCAs are a special class of equations with delay, the hybrid characteristic of discrete and continuous dynamic systems shows some difference in the analysis of the properties of exact solutions and numerical methods with the stochastic delay differential equation.

Taking the Markov property for instance, we directly analyze this property of the numerical solution $\{Y_k\}_{k\in\mathbb{N}}$ for the SDE with PCAs, while it is considered for the corresponding segment process of the continuous formulation of the numerical solution for the stochastic delay differential equation. Moreover, compared with the stochastic delay differential equation, numerical methods are generally required some additional algebraic stability when solving the SDE with PCAs.

Before we present the existence and uniqueness of the invariant measure of $\{Y_k\}_{k\in\mathbb{N}}$, we prepare several lemmas, including the mean-square boundedness and the dependence on initial data of the numerical solution.

Lemma 3.2. *For a nonnegative sequence Z_{km+l+1}, if there exist $\alpha > \beta > 0$, $\gamma > 0$ such that $1 - \alpha\delta > 0$ and*

$$Z_{km+l+1} \le (1 - \alpha\delta)Z_{km+l} + \beta\delta Z_{km} + \gamma\delta \tag{20}$$

for $k \in \mathbb{N}$, $l = 0, 1, \cdots, m-1$, then

$$Z_{km+l+1} \le \left(\frac{\beta}{\alpha} + \left(1 - \frac{\beta}{\alpha}\right)e^{-\alpha(l+1)\delta}\right)Z_{km} + \frac{\gamma}{\alpha}. \tag{21}$$

Proof. From $1 - \alpha\delta > 0$ and (20), it follows

$$\begin{aligned} Z_{km+l+1} \le& (1 - \alpha\delta)^{l+1}Z_{km} + \frac{\beta}{\alpha}\left(1 - (1 - \alpha\delta)^{l+1}\right)Z_{km} + \frac{\gamma}{\alpha}\left(1 - (1 - \alpha\delta)^{l+1}\right) \\ =& \left(\frac{\beta}{\alpha} + \left(1 - \frac{\beta}{\alpha}\right)(1 - \alpha\delta)^{l+1}\right)Z_{km} + \frac{\gamma}{\alpha}\left(1 - (1 - \alpha\delta)^{l+1}\right) \\ \le& \left(\frac{\beta}{\alpha} + \left(1 - \frac{\beta}{\alpha}\right)e^{-\alpha(l+1)\delta}\right)Z_{km} + \frac{\gamma}{\alpha}, \end{aligned} \tag{22}$$

where in the last step we use $1 - \frac{\beta}{\alpha} > 0$. This completes the proof. □

Lemma 3.3. *Let conditions in Lemma 2.4 hold. Then there exists a positive constant $C := C(\lambda_1, \lambda_2, \lambda_3, p)$ independent of δ, k and l such that*

$$\mathbb{E}\|X_{km+l+1}\|^{2p} \le C(1 + \|x\|^{2p}) \tag{23}$$

for $k \in \mathbb{N}$, $l = 0, 1, \cdots, m-1$ *and* $\delta \in (0, \delta_0)$ *with* δ_0 *being sufficiently small.*

Proof. **Case 1.** If $p = 1$, then taking the inner product of (16) with X_{km+l+1}, we get

$$\begin{aligned} &\|X_{km+l+1}\|^2 - \|X_{km+l}\|^2 + \|X_{km+l+1} - X_{km+l}\|^2 \\ =&2\delta \langle X_{km+l+1}, f(X_{km+l+1}, X_{km})\rangle + 2\langle X_{km+l+1}, g(X_{km+l}, X_{km})\Delta B_{km+l}\rangle . \end{aligned} \tag{24}$$

From (5) and (6), it follows

$$\begin{aligned} &\mathbb{E}\|X_{km+l+1}\|^2 \\ \le&\mathbb{E}\|X_{km+l}\|^2 + \delta\mathbb{E}\|g(X_{km+l}, X_{km})\|^2 + 2\delta\mathbb{E}\langle X_{km+l+1}, f(X_{km+l+1}, X_{km})\rangle \\ \le&(1 + 2\lambda_3\delta)\mathbb{E}\|X_{km+l}\|^2 - \delta(2\lambda_1 - 1)\mathbb{E}\|X_{km+l+1}\|^2 \\ &+ 2\delta(\lambda_2 + \lambda_3)\mathbb{E}\|X_{km}\|^2 + 2\delta(\|f(0,0)\|^2 + \|g(0,0)\|^2). \end{aligned} \tag{25}$$

Let $\alpha_2 := \frac{2\lambda_1 - 2\lambda_3 - 1}{1+(2\lambda_1-1)\delta}$, $\beta_2 := \frac{2(\lambda_2+\lambda_3)}{1+(2\lambda_1-1)\delta}$ and $\gamma_2 := \frac{2(\|f(0,0)\|^2+\|g(0,0)\|^2)}{1+(2\lambda_1-1)\delta}$. Then

$$\mathbb{E}\|X_{km+l+1}\|^2 \le (1 - \alpha_2\delta)\mathbb{E}\|X_{km+l}\|^2 + \beta_2\delta\mathbb{E}\|X_{km}\|^2 + \gamma_2\delta. \tag{26}$$

Since $\lambda_1 - \lambda_2 - 2\lambda_3 - 1 > 0$, we have $0 < \alpha_2\delta < 1$ and $\frac{\beta_2}{\alpha_2} < 1$ for any $\delta \in (0,1)$. By Lemma 3.2, (26) yields

$$\begin{aligned} \mathbb{E}\|X_{km+l+1}\|^2 \le& \left(\frac{\beta_2}{\alpha_2} + \left(1 - \frac{\beta_2}{\alpha_2}\right)e^{-\alpha_2\delta(l+1)}\right)\mathbb{E}\|X_{km}\|^2 + \frac{\gamma_2}{\alpha_2} \\ =:& \, r_2(l)\mathbb{E}\|X_{km}\|^2 + \frac{\gamma_2}{\alpha_2}. \end{aligned} \tag{27}$$

Here $0 < r_2(l) < 1$ for $l = 0, 1, 2, \cdots, m-1$. If $l = m-1$, then

$$\mathbb{E}\left\|X_{(k+1)m}\right\|^2 \le r_2(m-1)\mathbb{E}\|X_{km}\|^2 + \frac{\gamma_2}{\alpha_2}.$$

Let $F_1 := \left(\frac{\beta_2}{\alpha_2} + \left(1 - \frac{\beta_2}{\alpha_2}\right)e^{-(2\lambda_1 - 2\lambda_3 - 1)}\right)^{-1}$, $F_2 := \left(1 - \frac{\beta_2}{\alpha_2} - \left(1 - \frac{\beta_2}{\alpha_2}\right)e^{-\frac{2\lambda_1 - 2\lambda_3 - 1}{2\lambda_1}}\right)^{-1}$. Then

$$\begin{aligned} \mathbb{E}\|X_{km+l+1}\|^2 \le& r_2(l)(r_2(m-1))^k\|x\|^2 + \frac{\gamma_2}{\alpha_2} \\ &+ \frac{\gamma_2}{\alpha_2}r_2(l)\left(1 + r_2(m-1) + \cdots + (r_2(m-1))^k\right) \\ \le& \frac{1}{r_2(m-1)}e^{(km+l+1)\delta\log r_2(m-1)}\|x\|^2 + \frac{\gamma_2}{\alpha_2}\cdot\frac{2 - r_2(m-1)}{1 - r_2(m-1)} \\ \le& F_1\|x\|^2 + \frac{\gamma_2}{\alpha_2}(1 + F_2) \\ \le& \left(F_1 + \frac{\gamma_1}{\alpha_2}(1 + F_2)\right)(1 + \|x\|^2), \end{aligned} \tag{28}$$

where we use $\delta \in (0,1)$ and $0 < r_2(l) < 1$ for $l = 0, 1, 2, \cdots, m-1$.

Case 2. For $p > 1$, we show the assertion (23) by induction. Since $\lambda_1 - \lambda_2 - 2\lambda_3 - 1 > 0$, Case 1 implies that $\mathbb{E}\|X_{km+l+1}\|^{2p'} \le C(1 + \|x\|^2)$ holds with $p' = 1$. Multiplying (24) by $\|X_{km+l+1}\|^2$ and taking expectation yield $LHS = RHS$, where

$$\begin{aligned} &LHS \\ =&\mathbb{E}\|X_{km+l+1}\|^4 - \mathbb{E}\|X_{km+l+1}\|^2\|X_{km+l}\|^2 + \mathbb{E}\|X_{km+l+1}\|^2\|X_{km+l+1} - X_{km+l}\|^2 \\ =&\frac{1}{2}\mathbb{E}\left(\|X_{km+l+1}\|^4 - \|X_{km+l}\|^4 + \left(\|X_{km+l+1}\|^2 - \|X_{km+l}\|^2\right)^2\right) \\ &+ \mathbb{E}\|X_{km+l+1}\|^2\|X_{km+l+1} - X_{km+l}\|^2 \end{aligned}$$

and

$$\begin{aligned}
RHS =&2\delta\mathbb{E}\left\|X_{km+l+1}\right\|^2\left\langle X_{km+l+1}, f(X_{km+l+1}, X_{km})\right\rangle \\
&+2\mathbb{E}\left\|X_{km+l+1}\right\|^2\left\langle X_{km+l+1}, g(X_{km+l}, X_{km})\Delta B_{km+l}\right\rangle \\
=&2\delta\mathbb{E}\left\|X_{km+l+1}\right\|^2\left\langle X_{km+l+1}, f(X_{km+l+1}, X_{km})\right\rangle \\
&+2\mathbb{E}\left\|X_{km+l+1}\right\|^2\left\langle X_{km+l+1}-X_{km+l}, g(X_{km+l}, X_{km})\Delta B_{km+l}\right\rangle \\
&+2\mathbb{E}\left\|X_{km+l+1}\right\|^2\left\langle X_{km+l}, g(X_{km+l}, X_{km})\Delta B_{km+l}\right\rangle \\
=:&R_1+R_2+R_3.
\end{aligned}$$

For term R_1, by (5) and $2ab \le a^2+b^2$, we have

$$\begin{aligned}
R_1 \le& -(2\lambda_1-1)\,\delta\mathbb{E}\left\|X_{km+l+1}\right\|^4+2\lambda_2\delta\mathbb{E}\left\|X_{km+l+1}\right\|^2\left\|X_{km+l}\right\|^2 \\
&+\|f(0,0)\|^2\delta\mathbb{E}\left\|X_{km+l+1}\right\|^2 \\
\le& -(2\lambda_1-1-\lambda_2)\,\delta\mathbb{E}\left\|X_{km+l+1}\right\|^4+\lambda_2\delta\mathbb{E}\left\|X_{km+l}\right\|^4 \\
&+2\|f(0,0)\|^2\delta\mathbb{E}\left\|X_{km+l+1}\right\|^2.
\end{aligned}$$

For terms R_2 and R_3, according to (6) and $2ab \le \epsilon a^2+\frac{1}{\epsilon}b^2$, $\epsilon>0$, we obtain

$$\begin{aligned}
R_2 \le&\mathbb{E}\|X_{km+l+1}\|^2\|X_{km+l+1}-X_{km+l}\|^2+\mathbb{E}\|X_{km+l}\|^2\|g(X_{km+l},X_{km})\Delta B_{km+l}\|^2 \\
&+\mathbb{E}\left(\|X_{km+l+1}\|^2-\|X_{km+l}\|^2\right)\|g(X_{km+l},X_{km})\Delta B_{km+l}\|^2 \\
\le&\mathbb{E}\|X_{km+l+1}\|^2\|X_{km+l+1}-X_{km+l}\|^2+\frac{\epsilon_1}{4}\mathbb{E}\left(\|X_{km+l+1}\|^2-\|X_{km+l}\|^2\right)^2 \\
&+\frac{1}{\epsilon_1}\mathbb{E}\|g(X_{km+l},X_{km})\Delta B_{km+l}\|^4+\mathbb{E}\|X_{km+l}\|^2\|g(X_{km+l},X_{km})\Delta B_{km+l}\|^2 \\
\le&\mathbb{E}\|X_{km+l+1}\|^2\|X_{km+l+1}-X_{km+l}\|^2+\frac{\epsilon_1}{4}\mathbb{E}\left(\|X_{km+l+1}\|^2-\|X_{km+l}\|^2\right)^2 \\
&+\frac{3}{\epsilon_1}\delta^2\mathbb{E}(2\lambda_3\|X_{km+l}\|^2+2\lambda_3\|X_{km}\|^2+2\|g(0,0)\|^2)^2 \\
&+2\lambda_3\delta\mathbb{E}\|X_{km+l}\|^4+2\lambda_3\delta\mathbb{E}\|X_{km+l}\|^2\|X_{km}\|^2+2\|g(0,0)\|^2\delta\mathbb{E}\|X_{km+l}\|^2 \\
=&\mathbb{E}\|X_{km+l+1}\|^2\|X_{km+l+1}-X_{km+l}\|^2+\frac{\epsilon_1}{4}\mathbb{E}\left(\|X_{km+l+1}\|^2-\|X_{km+l}\|^2\right)^2 \\
&+\left(\frac{24\lambda_3^2\delta^2}{\epsilon_1}+3\lambda_3\delta\right)\mathbb{E}\|X_{km+l}\|^4+\left(\frac{24\lambda_3^2\delta^2}{\epsilon_1}+\lambda_3\delta\right)\mathbb{E}\|X_{km}\|^4+\frac{12\delta^2}{\epsilon_1}\|g(0,0)\|^4 \\
&+\frac{24\lambda_3\delta^2}{\epsilon_1}\|g(0,0)\|^2\mathbb{E}(\|X_{km+l}\|^2+\|X_{km}\|^2)+2\|g(0,0)\|^2\delta\mathbb{E}\|X_{km+l}\|^2
\end{aligned}$$

and

$$\begin{aligned}
R_3 =&2\mathbb{E}(\|X_{km+l+1}\|^2-\|X_{km+l}\|^2)\langle X_{km+l}, g(X_{km+l},X_{km})\rangle\Delta B_{km+l} \\
\le&\epsilon_2\mathbb{E}(\|X_{km+l+1}\|^2-\|X_{km+l}\|^2)^2+\frac{1}{\epsilon_2}\mathbb{E}(\langle X_{km+l}, g(X_{km+l},X_{km})\rangle\Delta B_{km+l})^2 \\
\le&\epsilon_2\mathbb{E}(\|X_{km+l+1}\|^2-\|X_{km+l}\|^2)^2+\frac{3\lambda_3\delta}{\epsilon_2}\mathbb{E}\|X_{km+l}\|^4 \\
&+\frac{\lambda_3\delta}{\epsilon_2}\mathbb{E}\|X_{km}\|^4+\frac{2\|g(0,0)\|^2\delta}{\epsilon_2}\mathbb{E}\|X_{km+l}\|^2.
\end{aligned}$$

Substituting R_1, R_2 and R_3 into RHS, we get

$$RHS \le -(2\lambda_1-1-\lambda_2)\delta\mathbb{E}\|X_{km+l+1}\|^4+\mathbb{E}\|X_{km+l+1}\|^2\|X_{km+l+1}-X_{km+l}\|^2$$

$$+\left(\frac{\epsilon_1}{4}+\epsilon_2\right)\mathbb{E}\left(\|X_{km+l+1}\|^2-\|X_{km+l}\|^2\right)^2 \qquad (29)$$
$$+\left(\frac{24\lambda_3^2\delta^2}{\epsilon_1}+3\lambda_3\delta+\lambda_2\delta+\frac{3\lambda_3\delta}{\epsilon_2}\right)\mathbb{E}\|X_{km+l}\|^4$$
$$+\left(\frac{24\lambda_3^2\delta^2}{\epsilon_1}+\lambda_3\delta+\frac{\lambda_3\delta}{\epsilon_2}\right)\mathbb{E}\|X_{km}\|^4+C\delta(1+\|x\|^2),$$

where $\mathbb{E}\|X_{km+l+1}\|^2 \le C(1+\|x\|^2)$, $k\in\mathbb{N}$, $l=0,1,\cdots,m-1$ are used and C is independent of δ, k and l. Let $\epsilon_2+\frac{\epsilon_1}{4}=\frac{1}{2}$, $\alpha_3=\frac{2(2\lambda_1-1-\lambda_2)-\frac{48\lambda_3^2\delta}{\epsilon_1}-2\lambda_2-6\lambda_3-\frac{6\lambda_3}{\epsilon_2}}{1+2(2\lambda_1-1-\lambda_2)\delta}$ and $\beta_3=\frac{\frac{48\lambda_3^2\delta}{\epsilon_1}+2\lambda_3+\frac{2\lambda_3}{\epsilon_2}}{1+2(2\lambda_1-1-\lambda_2)\delta}$. Recall that $LHS=RHS$,

$$\begin{aligned}&(1+2(2\lambda_1-1-\lambda_2)\delta)\mathbb{E}\|X_{km+l+1}\|^4\\ \le&\left(1+\frac{48\lambda_3^2\delta^2}{\epsilon_1}+6\lambda_3\delta+\frac{6\lambda_3\delta}{\epsilon_2}+2\lambda_2\delta\right)\mathbb{E}\|X_{km+l}\|^4\\ &+\left(\frac{48\lambda_3^2\delta^2}{\epsilon_1}+2\lambda_3\delta+\frac{2\lambda_3\delta}{\epsilon_2}\right)\mathbb{E}\|X_{km}\|^4+C\delta(1+\|x\|^2),\end{aligned}$$

which implies

$$\mathbb{E}\|X_{km+l+1}\|^4 \le (1-\alpha_3\delta)\mathbb{E}\|X_{km+l}\|^4+\beta_3\delta\mathbb{E}\|X_{km}\|^4+C\delta(1+\|x\|^2).$$

Up to now, it suffices to show that

$$\alpha_3-\beta_3=\frac{2\left(2\lambda_1-1-2\lambda_2-4\lambda_3\right)-\frac{96\lambda_3^2\delta}{\epsilon_1}-\frac{8\lambda_3}{\epsilon_2}}{1+2(2\lambda_1-1-\lambda_2)\delta}>0.$$

Since $1+2(2\lambda_1-1-\lambda_2)\delta>0$, we just need

$$2\lambda_1-1-2\lambda_2-4\lambda_3>\frac{48\lambda_3^2\delta}{\epsilon_1}+\frac{4\lambda_3}{\epsilon_2}.$$

Note that the condition $\lambda_1-1-\lambda_2-2\lambda_3>4(p-1)\lambda_3$ equals to $2\lambda_1-1-2\lambda_2-4\lambda_3>8(p-1)\lambda_3+1$, and $8(p-1)\lambda_3+1>8\lambda_3+1$. We choose $\epsilon_2=\frac{8\lambda_3}{16\lambda_3+1}$, then

$$\frac{4\lambda_3}{\epsilon_2}=8\lambda_3+\frac{1}{2},$$

and there exists $\delta'>0$ such that

$$\frac{48\lambda_3^2\delta}{\epsilon_1}\le\frac{1}{2},\qquad \forall\ \delta\in(0,\delta'),$$

which leads to

$$\frac{48\lambda_3^2\delta}{\epsilon_1}+\frac{4\lambda_3}{\epsilon_2}\le 8\lambda_3+1<2\lambda_1-1-2\lambda_2-4\lambda_3,$$

i.e., $\alpha_3-\beta_3>0$. Therefore, there exists $C>0$ independent of δ, k and l such that

$$\mathbb{E}\|X_{km+l+1}\|^4\le C(1+\|x\|^4)$$

for all $k\in\mathbb{N}$ and $l=0,1,\cdots,m-1$. This implies that $\mathbb{E}\|X_{km+l+1}\|^{2p'}\le C(1+\|x\|^{2p'})$ holds with $p'=2$.

By repeating the same procedure as the case $p'=2$, there exist $\delta_0>0$ and $C>0$ independent of δ, k and l such that $\mathbb{E}\|X_{km+l+1}\|^{2p'}\le C(1+\|x\|^{2p'})$ for $p'=3,4,\cdots,p$, which completes the proof. □

Corollary 1. *Under Assumption 2.1, if* $\lambda_1-\lambda_2-2\lambda_3-1>0$*, then, for any* $\delta\in(0,1)$*, there exists a constant* $C_2>0$ *independent of* δ *and* k *such that* $\{Y_k^{0,x}\}_{k\in\mathbb{N}}$ *satisfies*

$$\sup_{k\in\mathbb{N}}\mathbb{E}\left\|Y_k^{0,x}\right\|^2\le C_2(1+\|x\|^2). \tag{30}$$

Denote $\alpha_4:=\frac{2\lambda_1-\lambda_3-1}{1+(2\lambda_1-1)\delta}$, $\beta_4:=\frac{\lambda_2+\lambda_3}{1+(2\lambda_1-1)\delta}$ and $r_3(l):=\frac{\beta_4}{\alpha_4}+\left(1-\frac{\beta_4}{\alpha_4}\right)e^{-\alpha_4(l+1)\delta}$, $l=0,1,2,\cdots,m-1$. By $\lambda_1-\lambda_2-2\lambda_3-1>0$, we get $0<\alpha_4\delta<1$, $\frac{\beta_4}{\alpha_4}<1$ and thus $0<r_3(l)<1$ for $l=0,1,2,\cdots,m-1$ and $\delta\in(0,1)$. In what follows, we show the continuous dependence on initial data of $\{Y_k\}_{k\in\mathbb{N}}$.

Lemma 3.4. *Let Assumption 2.1 hold. If* $\lambda_1-\lambda_2-2\lambda_3-1>0$*, then, for any* $\delta\in(0,1)$ *and any two initial values* $x,y\in\mathbb{R}^d$ *with* $x\neq y$*, the solutions generated by the BE method satisfy*

$$\mathbb{E}\left\|Y_{k+1}^{0,x}-Y_{k+1}^{0,y}\right\|^2\le\frac{1}{r_3(m-1)}e^{(k+1)\log r_3(m-1)}\|x-y\|^2. \tag{31}$$

Proof. For any initial data $x,y\in\mathbb{R}^d$, Assumption 2.1 leads to

$$\begin{aligned}
&\mathbb{E}\left\|X_{km+l+1}^{0,x}-X_{km+l+1}^{0,y}\right\|^2\\
=&\mathbb{E}\left\|X_{km+l}^{0,x}-X_{km+l}^{0,y}\right\|^2-\delta^2\mathbb{E}\left\|f(X_{km+l+1}^{0,x},X_{km}^{0,x})-f(X_{km+l+1}^{0,y},X_{km}^{0,y})\right\|^2\\
&+2\delta\mathbb{E}\left\langle X_{km+l+1}^{0,x}-X_{km+l+1}^{0,y},f(X_{km+l+1}^{0,x},X_{km}^{0,x})-f(X_{km+l+1}^{0,y},X_{km}^{0,y})\right\rangle\\
&+\mathbb{E}\left\|\left(g(X_{km+l}^{0,x},X_{km}^{0,x})-g(X_{km+l}^{0,y},X_{km}^{0,y})\right)\Delta B_{km+l}\right\|^2\\
\le&(1+\lambda_3\delta)\mathbb{E}\left\|X_{km+l}^{0,x}-X_{km+l}^{0,y}\right\|^2+(\lambda_2+\lambda_3)\delta\mathbb{E}\left\|X_{km}^{0,x}-X_{km}^{0,y}\right\|^2\\
&-(2\lambda_1-1)\delta\mathbb{E}\left\|X_{km+l+1}^{0,x}-X_{km+l+1}^{0,y}\right\|^2.
\end{aligned}$$

Therefore

$$\begin{aligned}
&\mathbb{E}\left\|X_{km+l+1}^{0,x}-X_{km+l+1}^{0,y}\right\|^2\\
\le&(1-\alpha_4\delta)\mathbb{E}\left\|X_{km+l}^{0,x}-X_{km+l}^{0,y}\right\|^2+\beta_4\delta\mathbb{E}\left\|X_{km}^{0,x}-X_{km}^{0,y}\right\|^2\\
\le&\left(\frac{\beta_4}{\alpha_4}+\left(1-\frac{\beta_4}{\alpha_4}\right)(1-\alpha_4\delta)^{l+1}\right)\mathbb{E}\left\|X_{km}^{0,x}-X_{km}^{0,y}\right\|^2\\
\le&r_3(l)\mathbb{E}\left\|X_{km}^{0,x}-X_{km}^{0,y}\right\|^2.
\end{aligned} \tag{32}$$

If $l=m-1$, then

$$\mathbb{E}\left\|X_{(k+1)m}^{0,x}-X_{(k+1)m}^{0,y}\right\|^2\le r_3(m-1)\mathbb{E}\left\|X_{km}^{0,x}-X_{km}^{0,y}\right\|^2.$$

Therefore, (32) yields

$$\begin{aligned}
\mathbb{E}\left\|X_{km+l+1}^{0,x}-X_{km+l+1}^{0,y}\right\|^2\le&r_3(l)\mathbb{E}\left\|X_{km}^{0,x}-X_{km}^{0,y}\right\|^2\\
\le&\frac{1}{r_3(m-1)}e^{(km+l+1)\delta\log r_3(m-1)}\|x-y\|^2
\end{aligned} \tag{33}$$

and

$$\begin{aligned}\mathbb{E}\left\|Y_{k+1}^{0,x}-Y_{k+1}^{0,y}\right\|^2 =&\mathbb{E}\left\|X_{(k+1)m}^{0,x}-X_{(k+1)m}^{0,y}\right\|^2\\ \leq&\ \frac{1}{r_3(m-1)}e^{(k+1)\log r_3(m-1)}\|x-y\|^2.\end{aligned}\tag{34}$$

The proof is completed. ◻

Now, we are in a position to show the existence and uniqueness of the BE method's invariant measure.

Theorem 3.5. *Under the assumption of Theorem 2.3, for any $\delta\in(0,1)$, the Markov chain $\{Y_k^{0,x}\}_{k\in\mathbb{N}}$ admits a unique invariant measure π^δ and there exist two positive constants $C_3:=C_3(\|\varphi\|_1,\lambda_1,\lambda_2,\lambda_3)$ and $\bar{\nu}:=\bar{\nu}(\lambda_1,\lambda_2,\lambda_3)$ independent of x, δ and k such that*

$$\left|\mathbb{E}\varphi(Y_k^{0,x})-\int_{\mathbb{R}^d}\varphi(x)\pi^\delta(dx)\right|\leq C_3(1+\|x\|)e^{-\bar{\nu}k},\quad \forall\ \varphi\in C_b^1(\mathbb{R}^d).\tag{35}$$

Proof. According to the Chebyshev's inequality and (30), it follows that the transition probability measure $\{P_k^\delta(x,\cdot)\}_{k\in\mathbb{N}}$ is tight. Define

$$\mu_K^\delta(B)=\frac{1}{K}\sum_{k=0}^{K-1}P_k^\delta(x,B),\quad \forall\ B\in\mathscr{B}(\mathbb{R}^d),\ K\in\mathbb{N},$$

then $\mu_K^\delta(\cdot)\in\mathscr{P}(\mathbb{R}^d)$ for any $K\in\mathbb{N}$ and $\{\mu_K^\delta(\cdot)\}_{K\geq 0}$ is tight. By the Prokhorov theorem [6], $\{\mu_K^\delta(\cdot)\}_{K\in\mathbb{N}}$ is weakly relatively compact, i.e., there exists $\pi^\delta\in\mathscr{P}(\mathbb{R}^d)$ such that the subsequence $\{\mu_{K_i}^\delta\}_{K_i\in\mathbb{N}}$ is weakly convergent to π^δ as $K_i\to\infty$. Now, we prove that π^δ is an invariant measure. Indeed, for any $\varphi\in B_b(\mathbb{R}^d)$, $k\in\mathbb{N}$, we have

$$\begin{aligned}&\int_{\mathbb{R}^d}P_{k'}^\delta\varphi(y)\pi^\delta(dy)\\ =&\lim_{K_i\to\infty}\int_{\mathbb{R}^d}P_{k'}^\delta\varphi(y)\mu_{K_i}^\delta(dy)=\lim_{K_i\to\infty}\frac{1}{K_i}\sum_{k=0}^{K_i-1}\int_{\mathbb{R}^d}P_{k'}^\delta\varphi(y)P_k^\delta(x,dy)\\ =&\lim_{K_i\to\infty}\frac{1}{K_i}\sum_{k=0}^{K_i-1}\mathbb{E}\left[\varphi(Y_{k+k'}^{0,x})\right]=\lim_{K_i\to\infty}\frac{1}{K_i}\sum_{k=k'}^{K_i+k'-1}\mathbb{E}\left[\varphi(Y_k^{0,x})\right]\\ =&\lim_{K_i\to\infty}\frac{1}{K_i}\left(\sum_{k=0}^{K_i-1}\mathbb{E}\left[\varphi(Y_k^{0,x})\right]+\sum_{k=K_i}^{K_i+k'-1}\mathbb{E}\left[\varphi(Y_k^{0,x})\right]-\sum_{k=0}^{k'-1}\mathbb{E}\left[\varphi(Y_k^{0,x})\right]\right)\\ =&\lim_{K_i\to\infty}\frac{1}{K_i}\sum_{k=0}^{K_i-1}\mathbb{E}\left[\varphi(Y_k^{0,x})\right]=\int_{\mathbb{R}^d}\varphi(y)\pi^\delta(dy),\end{aligned}$$

where $P_{k'}^\delta\varphi(y)=\mathbb{E}\left[\varphi(Y_{k'}^{0,y})\right]$. Below, we show the uniqueness of the invariant measure. For any $x,\ y\in\mathbb{R}^d$ and $\varphi\in C_b^1$, (34) leads to

$$\begin{aligned}&\lim_{K\to\infty}\left|\frac{1}{K}\sum_{k=0}^{K-1}\int_{\mathbb{R}^d}\varphi(z)P_k^\delta(x,dz)-\frac{1}{K}\sum_{k=0}^{K-1}\int_{\mathbb{R}^d}\varphi(z)P_k^\delta(y,dz)\right|\\ =&\lim_{K\to\infty}\left|\frac{1}{K}\sum_{k=0}^{K-1}\mathbb{E}\left[\varphi(Y_k^{0,x})\right]-\frac{1}{K}\sum_{k=0}^{K-1}\mathbb{E}\left[\varphi(Y_k^{0,y})\right]\right|\end{aligned}$$

$$
\begin{aligned}
&\le \lim_{K\to\infty} \frac{\|\varphi\|_1}{K} \sum_{k=0}^{K-1} \mathbb{E}\left|Y_k^{0,x} - Y_k^{0,y}\right| \\
&\le \frac{\|x-y\|}{\sqrt{r_3(m-1)}} \cdot \lim_{K\to\infty} \frac{\|\varphi\|_1}{K} \sum_{k=0}^{K-1} e^{\frac{k}{2}\log r_3(m-1)} \\
&= \frac{\|x-y\|}{\sqrt{r_3(m-1)}} \cdot \lim_{K\to\infty} \frac{\|\varphi\|_1}{K} \cdot \frac{1-\exp(\frac{K}{2}\log r_3(m-1))}{1-r_3(m-1)} \\
&= 0,
\end{aligned}
$$

which means that π^δ is independent of the initial value x. In addition, if $\tilde{\pi}^\delta \in \mathscr{P}(\mathbb{R}^d)$ is another invariant measure of $\{Y_k^{0,x}\}_{k\in\mathbb{N}}$, then

$$
\begin{aligned}
\int_{\mathbb{R}^d} \varphi(z)\tilde{\pi}^\delta(dz) &= \int_{\mathbb{R}^d} P_k^\delta \varphi(z)\tilde{\pi}^\delta(dz) = \frac{1}{K_i}\sum_{k=0}^{K_i-1} \int_{\mathbb{R}^d} P_k^\delta \varphi(z)\tilde{\pi}^\delta(dz) \\
&= \int_{\mathbb{R}^d} \left(\lim_{K_i\to\infty} \frac{1}{K_i} \sum_{k=0}^{K_i-1} \int_{\mathbb{R}^d} \varphi(y) P_k^\delta(z,dy) \right) \tilde{\pi}^\delta(dz) \\
&= \int_{\mathbb{R}^d} \varphi(y)\pi^\delta(dy).
\end{aligned}
$$

Therefore, $\pi^\delta = \tilde{\pi}^\delta$, which means that π^δ is the unique invariant measure for $\{Y_k^{0,x}\}_{k\in\mathbb{N}}$.

Moreover, for any $x \in \mathbb{R}^d$ and $\varphi \in C_b^1(\mathbb{R}^d)$, Lemma 3.4 yields

$$
\begin{aligned}
&\left|\mathbb{E}\varphi(Y_k^{0,x}) - \int_{\mathbb{R}^d} \varphi(y)\pi^\delta(dy)\right| = \left|\mathbb{E}\varphi(Y_k^{0,x}) - \int_{\mathbb{R}^d} P_k^\delta \varphi(y)\pi^\delta(dy)\right| \\
&= \left|\int_{\mathbb{R}^d} \mathbb{E}\varphi(Y_k^{0,x})\pi^\delta(dy) - \int_{\mathbb{R}^d} \mathbb{E}\varphi(Y_k^{0,y})\pi^\delta(dy)\right| \le \|\varphi\|_1 \int_{\mathbb{R}^d} \mathbb{E}\left\|Y_k^{0,x} - Y_k^{0,y}\right\| \pi^\delta(dy) \\
&\le \frac{1}{\sqrt{r_3(m-1)}} e^{k/2\log r_3(m-1)} \int_{\mathbb{R}^d} \|x-y\|\,\pi^\delta(dy) \le C_3(1+\|x\|)e^{-\bar{\nu}k},
\end{aligned}
$$

where $\bar{\nu} = -\frac{1}{2}\log r_3(m-1)$. We complete the proof. □

4. **Approximation of invariant measure.** In this section, we aim to estimate the error between invariant measures π and π^δ, i.e.,

$$
\left|\int_{\mathbb{R}^d} \phi(x)\pi(dx) - \int_{\mathbb{R}^d} \phi(x)\pi^\delta(dx)\right|.
$$

Unless otherwise specified, we assume that $B(t)$ is a 1-dimensional Brownian motion for simplicity throughout this section. According to the uniqueness of the invariant measures π and π^δ, we know that both the Markov chains $\{X(k)\}_{k\in\mathbb{N}}$ and $\{Y_k\}_{k\in\mathbb{N}}$ are ergodic, that is

$$
\lim_{K\to\infty} \frac{1}{K}\sum_{k=0}^{K-1} \mathbb{E}\phi(X(k)) = \int_{\mathbb{R}^d} \phi(x)\pi(dx) \text{ and } \lim_{K\to\infty} \frac{1}{K}\sum_{k=0}^{K-1} \mathbb{E}\phi(Y_k) = \int_{\mathbb{R}^d} \phi(x)\pi^\delta(dx).
$$

Therefore,

$$\left|\int_{\mathbb{R}^d}\phi(x)\pi(dx)-\int_{\mathbb{R}^d}\phi(x)\pi^{\delta}(dx)\right|=\left|\lim_{K\to\infty}\frac{1}{K}\sum_{k=0}^{K-1}\left[\mathbb{E}\phi(X(k))-\sum_{k=0}^{K-1}\mathbb{E}\phi(Y_k)\right]\right|$$
$$\leq\lim_{K\to\infty}\frac{1}{K}\sum_{k=0}^{K-1}\left|\mathbb{E}\phi(X(k))-\mathbb{E}\phi(Y_k)\right|.$$

From the inequality above, it is observed that the error between π and π^{δ} can be estimated by the weak error of the BE method. So, we contribute on the time-independent weak convergence analysis of the BE method and then give the approximation between π and π^{δ}.

4.1. **A priori estimates.** Suppose that the Fréchet partial derivatives of f and g exist, then the definition of Fréchet derivatives and Assumption 2.1 yield

$$\xi^T\frac{\partial f}{\partial x}(x,y)\xi\leq-\lambda_1\|\xi\|^2,\quad\left\|\frac{\partial f}{\partial y}(x,y)\xi\right\|^2\leq\lambda_2\|\xi\|^2,\quad\forall\ \xi\in\mathbb{R}^d\tag{36}$$

and

$$\left\|\frac{\partial g}{\partial x}(x,y)\xi\right\|^2\vee\left\|\frac{\partial g}{\partial y}(x,y)\xi\right\|^2\leq\lambda_3\|\xi\|^2,\quad\forall\ \xi\in\mathbb{R}^d.\tag{37}$$

Denote $\mathcal{D}$ and D by the Malliavin differentiation operator and the Fréchet differentiation operator, respectively. We get the uniform estimation of the Fréchet derivative of $X^{i,\eta}(t)$ as follows.

Lemma 4.1. *Assume that the Fréchet derivatives of f and g exist and the conditions in Lemma 2.4 are satisfied. Then, for fixed $i\in\mathbb{N}$, there exist two positive constants $C:=C(\lambda_1,\lambda_2,\lambda_3,p)$ and $\nu_1:=\nu_1(\lambda_1,\lambda_2,\lambda_3,p)$ independent of t such that*

$$\mathbb{E}\left\|DX^{i,\eta}(t)\xi\right\|^{2p}\leq Ce^{-\nu_1(t-i)}\|\xi\|^{2p},\quad t\geq i,\tag{38}$$

where $\xi\in\mathbb{R}^d$ and $\eta\in L^{2p}(\Omega,\mathbb{R}^d;\mathcal{F}_i)$.

Proof. For any $\xi\in\mathbb{R}^d$, we have

$$\begin{aligned}&DX^{i,\eta}(t)\xi=DX^{i,\eta}([t])\xi\\&+\int_{[t]}^t\left(\frac{\partial f}{\partial x}(X^{i,\eta}(s),X^{i,\eta}([t]))DX^{i,\eta}(s)\xi+\frac{\partial f}{\partial y}(X^{i,\eta}(s),X^{i,\eta}([t]))DX^{i,\eta}([t])\xi\right)ds\\&+\int_{[t]}^t\left(\frac{\partial g}{\partial x}(X^{i,\eta}(s),X^{i,\eta}([t]))DX^{i,\eta}(s)\xi+\frac{\partial g}{\partial y}(X^{i,\eta}(s),X^{i,\eta}([t]))DX^{i,\eta}([t])\xi\right)dB(s).\end{aligned}$$

Denote $H(t):=DX^{i,\eta}(t)\xi$. For any $\alpha>0$, applying Itô's formula to $e^{2\alpha pt}\|H(t)\|^{2p}$, we obtain

$$\begin{aligned}&e^{2\alpha pt}\mathbb{E}\|H(t)\|^{2p}\\\leq&e^{2\alpha p[t]}\mathbb{E}\|H([t])\|^{2p}+2\alpha p\mathbb{E}\int_{[t]}^te^{2\alpha ps}\|H(s)\|^{2p}ds\\&+2p\mathbb{E}\int_{[t]}^te^{2\alpha ps}\|H(s)\|^{2(p-1)}\left\langle H(s),\frac{\partial f}{\partial x}(X^{i,\eta}(s),X^{i,\eta}([t]))H(s)\right\rangle ds\\&+2p\mathbb{E}\int_{[t]}^te^{2\alpha ps}\|H(s)\|^{2(p-1)}\left\langle H(s),\frac{\partial f}{\partial y}(X^{i,\eta}(s),X^{i,\eta}([t]))H([t])\right\rangle ds\end{aligned}$$

$$+ 2p(2p-1)\mathbb{E}\int_{[t]}^{t} e^{2\alpha ps}\|H(s)\|^{2(p-1)}\left\|\frac{\partial g}{\partial x}(X^{i,\eta}(s), X^{i,\eta}([t]))H(s)\right\|^2 ds$$
$$+ 2p(2p-1)\mathbb{E}\int_{[t]}^{t} e^{2\alpha ps}\|H(s)\|^{2(p-1)}\left\|\frac{\partial g}{\partial y}(X^{i,\eta}(s), X^{i,\eta}([t]))H([t])\right\|^2 ds.$$

Taking $2\alpha_5 p = 2\lambda_1 p - 2\lambda_3(2p-1)^2 - p - \lambda_2(p-1)$ and $\beta_5 = \lambda_2 + 2\lambda_3(2p-1)$, then $2\alpha_5 p > \beta_5 p$ since $\lambda_1 - \lambda_2 - 2\lambda_3 - 1 > 4\lambda_3(p-1)$. From (36) and Young's inequality, it follows that

$$\begin{aligned}
&e^{2\alpha_5 pt}\mathbb{E}\|H(t)\|^{2p}\\
\le& e^{2\alpha_5 p[t]}\mathbb{E}\|H([t])\|^{2p} + (2\alpha_5 - 2\lambda_1 + 2\lambda_3(2p-1)+1)\,p\mathbb{E}\int_{[t]}^{t} e^{2\alpha_5 ps}\|H(s)\|^{2p} ds\\
&+ (\lambda_2 p + 2\lambda_3 p(2p-1))\,\mathbb{E}\int_{[t]}^{t} e^{2\alpha_5 ps}\|H(s)\|^{2(p-1)}\|H([t])\|^2 ds\\
\le& e^{2\alpha_5 p[t]}\mathbb{E}\|H([t])\|^{2p} + \beta_5 p\mathbb{E}\int_{[t]}^{t} e^{2\alpha_5 ps}\|H([t])\|^{2p} ds\\
=&\left(e^{2\alpha_5 p[t]} + \frac{\beta_5}{2\alpha_5}\left(e^{2\alpha_5 pt} - e^{2\alpha_5 p[t]}\right)\right)\mathbb{E}\|H([t])\|^{2p}.
\end{aligned}$$

Hence

$$\mathbb{E}\left\|DX^{i,\eta}(t)\xi\right\|^{2p} \le r_4(\{t\})\mathbb{E}\left\|DX^{i,\eta}([t])\xi\right\|^{2p},$$

where r_4 is defined similarly to r in the proof of Theorem 2.3 with $r_4(\{t\}) = \frac{\beta_5}{2\alpha_5} + \left(1 - \frac{\beta_5}{2\alpha_5}\right)e^{-2\alpha_5 p\{t\}}$ for $\{t\} \in [0,1)$ and $r_4(1) := \lim_{t\to k^-}$, $k \in \mathbb{N}$. Then $0 < r_4 \le 1$ and

$$\mathbb{E}\left\|DX^{i,\eta}(k)\xi\right\|^{2p} \le \lim_{t\to k^-}\mathbb{E}\left\|DX^{i,\eta}(t)\xi\right\|^{2p} \le r_4(1)\mathbb{E}\left\|DX^{i,\eta}(k-1)\xi\right\|^{2p}.$$

Therefore

$$\mathbb{E}\left\|DX^{i,\eta}(t)\xi\right\|^{2p} \le r_4(\{t\})r_4(1)^{[t]-i}\mathbb{E}\|\xi\|^{2p} \le Ce^{-\nu_1(t-i)}\mathbb{E}\|\xi\|^{2p},$$

where $C = \frac{1}{r_4(1)}$ and $\nu_1 = -\log r_4(1)$. The proof is completed. □

Next, we show the uniform estimate of the Malliavin derivative of X_{km+l+1}, $k \in \mathbb{N}$, $l = 0, 1, \cdots, m-1$.

Lemma 4.2. *Let the conditions in Lemma 2.4 hold. Then there exists $C > 0$ independent of δ, k and l such that*

$$\mathbb{E}\|\mathcal{D}_u X_{km+l+1}\|^{2p} \le C \tag{39}$$

for all $k \in \mathbb{N}$, $l = 0, 1, \cdots, m-1$ and $\delta \in (0, \tilde{\delta}_0)$ with $\tilde{\delta}_0$ being sufficiently small.

Proof. From (16), the Malliavin derivative of X_{km+l+1} is

$$\begin{aligned}
&\mathcal{D}_u X_{km+l+1}\\
=&\mathcal{D}_u X_{km+l}1_{\{u<t_{km+l}\}} + g(X_{km+l}, X_{km})1_{\{t_{km+l}\le u<t_{km+l+1}\}}\\
&+ \delta\frac{\partial f}{\partial x}(X_{km+l+1}, X_{km})\mathcal{D}_u X_{km+l+1} + \delta\frac{\partial f}{\partial y}(X_{km+l+1}, X_{km})\mathcal{D}_u X_{km}\\
&+ \frac{\partial g}{\partial x}(X_{km+l}, X_{km})\mathcal{D}_u X_{km+l}\Delta B_{km+1} + \frac{\partial g}{\partial y}(X_{km+l}, X_{km})\mathcal{D}_u X_{km}\Delta B_{km+1}
\end{aligned} \tag{40}$$

for $0 \le u < t_{km+l+1}$. If $u \ge t_{km+l+1}$, then $\mathcal{D}_u X_{km+l+1} = 0$. Thus we only consider the case of $0 \le u < t_{km+l+1}$ in the following.

Case 1. If $t_{km+l} \le u < t_{km+l+1}$, then (40) becomes

$$\mathcal{D}_u X_{km+l+1} = g(X_{km+l}, X_{km}) + \delta \frac{\partial f}{\partial x}(X_{km+l+1}, X_{km}) \mathcal{D}_u X_{km+l+1}. \tag{41}$$

Multiplying (41) by $\mathcal{D}_u X_{km+l+1}$, (36) leads to

$$\begin{aligned} \|\mathcal{D}_u X_{km+l+1}\|^2 =& \langle \mathcal{D}_u X_{km+l+1}, g(X_{km+l}, X_{km}) \rangle \\ &+ \delta \left\langle \mathcal{D}_u X_{km+l+1}, \frac{\partial f}{\partial x}(X_{km+l+1}, X_{km}) \mathcal{D}_u X_{km+l+1} \right\rangle \\ \le& \left(\frac{1}{2} - \lambda_1 \delta \right) \|\mathcal{D}_u X_{km+l+1}\|^2 + \frac{1}{2} \mathbb{E} \|g(X_{km+l}, X_{km})\|^2 . \end{aligned} \tag{42}$$

Then multiplying (42) by $\|\mathcal{D}_u X_{km+l+1}\|^{2(p-1)}$ and taking expectation, Young's inequality leads to

$$\begin{aligned} &\mathbb{E} \|\mathcal{D}_u X_{km+l+1}\|^{2p} \\ \le& \left(\frac{1}{2} - \lambda_1 \delta \right) \mathbb{E} \|\mathcal{D}_u X_{km+l+1}\|^{2p} + \frac{1}{2} \mathbb{E} \|\mathcal{D}_u X_{km+l+1}\|^{2(p-1)} \|g(X_{km+l}, X_{km})\|^2 \\ \le& \left(\frac{3}{4} - \lambda_1 \delta \right) \mathbb{E} \|\mathcal{D}_u X_{km+l+1}\|^{2p} + \frac{1}{2p} \left(\frac{2(p-1)}{p} \right)^{p-1} \mathbb{E} \|g(X_{km+l}, X_{km})\|^{2p} , \end{aligned}$$

which implies

$$\mathbb{E} \|\mathcal{D}_u X_{km+l+1}\|^{2p} \le \frac{2}{(1 + 4\lambda_1 \delta)\, p} \left(\frac{2(p-1)}{p} \right)^{p-1} \mathbb{E} \|g(X_{km+l}, X_{km})\|^{2p} .$$

Since $\mathbb{E}\|X_{km+l}\|^{2p} \le C$, and g satisfies the global Lipschitz condition, we get (39) when $u \in [t_{km+l}, t_{km+l+1})$.

Case 2. If $t_{km} \le u < t_{km+l}$, then (40) becomes

$$\begin{aligned} \mathcal{D}_u X_{km+l+1} - \mathcal{D}_u X_{km+l} =& \delta \frac{\partial f}{\partial x}(X_{km+l+1}, X_{km}) \mathcal{D}_u X_{km+l+1} \\ &+ \frac{\partial g}{\partial x}(X_{km+l}, X_{km}) \mathcal{D}_u X_{km+l} \Delta B_{km+1}. \end{aligned} \tag{43}$$

We prove the assertion (39) by induction. Let us first show that $\mathbb{E} \|\mathcal{D}_u X_{km+l+1}\|^{2p'} \le C$ with $p' = 1$. Multiplying (43) by $\mathcal{D}_u X_{km+l+1}$ leads to

$$\begin{aligned} &\frac{1}{2} \left(\mathbb{E} \|\mathcal{D}_u X_{km+l+1}\|^2 - \mathbb{E} \|\mathcal{D}_u X_{km+l}\|^2 + \mathbb{E} \|\mathcal{D}_u X_{km+l+1} - \mathcal{D}_u X_{km+l}\|^2 \right) \\ =& \delta \mathbb{E} \left\langle \mathcal{D}_u X_{km+l+1}, \frac{\partial f}{\partial x}(X_{km+l+1}, X_{km}) \mathcal{D}_u X_{km+l+1} \right\rangle \\ &+ \mathbb{E} \left\langle \mathcal{D}_u X_{km+l+1} - \mathcal{D}_u X_{km+l}, \frac{\partial g}{\partial x}(X_{km+l}, X_{km}) \mathcal{D}_u X_{km+l} \Delta B_{km+1} \right\rangle \\ \le& -\lambda_1 \delta \mathbb{E} \|\mathcal{D}_u X_{km+l+1}\|^2 + \frac{1}{2} \mathbb{E} \|\mathcal{D}_u X_{km+l+1} - \mathcal{D}_u X_{km+l}\|^2 + \frac{1}{2} \lambda_3 \delta \mathbb{E} \|\mathcal{D}_u X_{km+l}\|^2 , \end{aligned} \tag{44}$$

which implies

$$(1 + 2\lambda_1 \delta) \mathbb{E} \|\mathcal{D}_u X_{km+l+1}\|^2 \le (1 + \lambda_3 \delta) \mathbb{E} \|\mathcal{D}_u X_{km+l}\|^2 .$$

Since $\lambda_1 - 1 - \lambda_2 - 2\lambda_3 > 0$, and for any u, there exist $n \in \mathbb{N}$ and $w \in \{0, 1, \cdots, m-1\}$ such that $u \in [t_{nm+w}, t_{nm+w+1})$. Combining Case 1, we obtain

$$\mathbb{E}\left\|\mathcal{D}_u X_{km+l+1}\right\|^2 \le \mathbb{E}\left\|\mathcal{D}_u X_{km+l}\right\|^2 \le \cdots \le \mathbb{E}\left\|\mathcal{D}_u X_{nm+w+1}\right\|^2 \le C.$$

If $p' = 2$, without taking expectation in (44), multiplying (44) by $\left\|\mathcal{D}_u X_{km+l+1}\right\|^2$ and then taking expectation, Young's inequality leads to

$$\begin{aligned}
&\frac{1}{2}\left(\mathbb{E}\left\|\mathcal{D}_u X_{km+l+1}\right\|^4 - \mathbb{E}\left\|\mathcal{D}_u X_{km+l}\right\|^4 + \mathbb{E}\left(\left\|\mathcal{D}_u X_{km+l+1}\right\|^2 - \left\|\mathcal{D}_u X_{km+l}\right\|^2\right)\right)\\
&\quad + \mathbb{E}\left\|\mathcal{D}_u X_{km+l+1}\right\|^2 \left\|\mathcal{D}_u X_{km+l+1} - \mathcal{D}_u X_{km+l}\right\|^2\\
=&2\delta\mathbb{E}\left\|\mathcal{D}_u X_{km+l+1}\right\|^2 \left\langle \mathcal{D}_u X_{km+l+1}, \frac{\partial f}{\partial x}(X_{km+l+1}, X_{km})\mathcal{D}_u X_{km+l+1}\right\rangle\\
&\quad + 2\mathbb{E}\left\|\mathcal{D}_u X_{km+l+1}\right\|^2 \times\\
&\qquad\qquad \left\langle \mathcal{D}_u X_{km+l+1} - \mathcal{D}_u X_{km+l}, \frac{\partial g}{\partial x}(X_{km+l}, X_{km})\mathcal{D}_u X_{km+l}\Delta B_{km+1}\right\rangle\\
&\quad + 2\mathbb{E}\left(\left\|\mathcal{D}_u X_{km+l+1}\right\|^2 - \left\|\mathcal{D}_u X_{km+l}\right\|^2\right) \times\\
&\qquad\qquad \left\langle \mathcal{D}_u X_{km+l}, \frac{\partial g}{\partial x}(X_{km+l}, X_{km})\mathcal{D}_u X_{km+l}\Delta B_{km+1}\right\rangle\\
=:&\tilde{R}_1 + \tilde{R}_2 + \tilde{R}_3.
\end{aligned}$$

From (36), $\tilde{R}_1$ satisfies

$$\tilde{R}_1 \le -2\lambda_1\delta\mathbb{E}\left\|\mathcal{D}_u X_{km+l+1}\right\|^4.$$

For $\tilde{R}_2$ and $\tilde{R}_3$, Young's inequality and (36) lead to

$$\begin{aligned}
\tilde{R}_2 \le&\mathbb{E}\left\|\mathcal{D}_u X_{km+l+1}\right\|^2 \left\|\mathcal{D}_u X_{km+l+1} - \mathcal{D}_u X_{km+l}\right\|^2\\
&+ \mathbb{E}\left\|\mathcal{D}_u X_{km+l+1}\right\|^2 \left\|\frac{\partial g}{\partial x}(X_{km+l}, X_{km})\mathcal{D}_u X_{km+l}\Delta B_{km+1}\right\|^2\\
\le&\mathbb{E}\left\|\mathcal{D}_u X_{km+l+1}\right\|^2 \left\|\mathcal{D}_u X_{km+l+1} - \mathcal{D}_u X_{km+l}\right\|^2 + \lambda_3\delta\mathbb{E}\left\|\mathcal{D}_u X_{km+l}\right\|^4\\
&+ \frac{1}{2}\epsilon_1\mathbb{E}\left(\left\|\mathcal{D}_u X_{km+l+1}\right\|^2 - \left\|\mathcal{D}_u X_{km+l}\right\|^2\right)^2\\
&+ \frac{1}{2\epsilon_1}\mathbb{E}\left\|\frac{\partial g}{\partial x}(X_{km+l}, X_{km})\mathcal{D}_u X_{km+l}\Delta B_{km+1}\right\|^4\\
\le&\mathbb{E}\left\|\mathcal{D}_u X_{km+l+1}\right\|^2 \left\|\mathcal{D}_u X_{km+l+1} - \mathcal{D}_u X_{km+l}\right\|^2\\
&+ \left(\lambda_3\delta + \frac{3\lambda_3^3\delta^2}{2\epsilon_1}\right)\mathbb{E}\left\|\mathcal{D}_u X_{km+l}\right\|^4 + \frac{1}{2}\epsilon_1\mathbb{E}\left(\left\|\mathcal{D}_u X_{km+l+1}\right\|^2 - \left\|\mathcal{D}_u X_{km+l}\right\|^2\right)^2
\end{aligned}$$

and

$$\begin{aligned}
\tilde{R}_3 \le&\epsilon_2\mathbb{E}\left(\left\|\mathcal{D}_u X_{km+l+1}\right\|^2 - \left\|\mathcal{D}_u X_{km+l}\right\|^2\right)^2\\
&+ \frac{1}{\epsilon_2}\mathbb{E}\left\langle \mathcal{D}_u X_{km+l}, \frac{\partial g}{\partial x}(X_{km+l}, X_{km})\mathcal{D}_u X_{km+l}\Delta B_{km+1}\right\rangle^2\\
\le&\epsilon_2\mathbb{E}\left(\left\|\mathcal{D}_u X_{km+l+1}\right\|^2 - \left\|\mathcal{D}_u X_{km+l}\right\|^2\right)^2 + \frac{\lambda_3\delta}{\epsilon_2}\mathbb{E}\left\|\mathcal{D}_u X_{km+l}\right\|^4.
\end{aligned}$$

Let $\epsilon_2 + \frac{1}{2}\epsilon_1 = \frac{1}{2}$, then

$$(1+4\lambda_1\delta)\,\mathbb{E}\,\|\mathcal{D}_u X_{km+l+1}\|^4 \le \left(1+2\left(1+\frac{1}{\epsilon_2}\right)\lambda_3\delta + \frac{3\lambda_3^3\delta^2}{\epsilon_1}\right)\mathbb{E}\,\|\mathcal{D}_u X_{km+l}\|^4 .$$

Since $\lambda_1 - \lambda_2 - 1 - 2\lambda_3 > 0$, there exists $\delta' > 0$ such that $\mathbb{E}\,\|\mathcal{D}_u X_{km+l+1}\|^4 \le \mathbb{E}\,\|\mathcal{D}_u X_{km+l}\|^4$ for all $\delta \in (0,\delta')$. Following the same procedure as the case $p'=1$, there exists $C>0$ independent of δ, k and l such that

$$\mathbb{E}\,\|\mathcal{D}_u X_{km+l+1}\|^4 \le C.$$

The inequality above shows $\mathbb{E}\|\mathcal{D}_u X_{km+l+1}\|^{2p'} \le C$ with $p'=2$. Repeating the procedure as above, we prove (39) when $u \in [t_{km}, t_{km+l})$.

Case 3. If $u < t_{km}$, then

$$\begin{aligned}
&\mathcal{D}_u X_{km+l+1} - \mathcal{D}_u X_{km+l}\\
&= \delta\frac{\partial f}{\partial x}(X_{km+l+1},X_{km})\mathcal{D}_u X_{km+l+1} + \delta\frac{\partial f}{\partial y}(X_{km+l+1},X_{km})\mathcal{D}_u X_{km}\\
&+ \frac{\partial g}{\partial x}(X_{km+l},X_{km})\mathcal{D}_u X_{km+l}\Delta B_{km+1} + \frac{\partial g}{\partial y}(X_{km+l},X_{km})\mathcal{D}_u X_{km}\Delta B_{km+1}.
\end{aligned} \tag{45}$$

Multiplying (45) by $\mathcal{D}_u X_{km+l+1}$ and taking expectation, we have $LHS = RHS$, where

$$LHS = \frac{1}{2}\left(\mathbb{E}\|\mathcal{D}_u X_{km+l+1}\|^2 - \mathbb{E}\,\|\mathcal{D}_u X_{km+l}\|^2 + \mathbb{E}\|\mathcal{D}_u X_{km+l+1} - \mathcal{D}_u X_{km+l}\|^2\right) \tag{46}$$

and

$$\begin{aligned}
&RHS\\
=&\delta\mathbb{E}\left\langle \mathcal{D}_u X_{km+l+1}, \frac{\partial f}{\partial x}(X_{km+l+1},X_{km})\mathcal{D}_u X_{km+l+1} + \frac{\partial f}{\partial y}(X_{km+l+1},X_{km})\mathcal{D}_u X_{km}\right\rangle\\
+&\mathbb{E}\left\langle \mathcal{D}_u X_{km+l+1}, \frac{\partial g}{\partial x}(X_{km+l},X_{km})\mathcal{D}_u X_{km+l} + \frac{\partial g}{\partial y}(X_{km+l},X_{km})\mathcal{D}_u X_{km}\right\rangle \Delta B_{km+1}.
\end{aligned} \tag{47}$$

By (36) and Young's inequality, RHS yields

$$\begin{aligned}
RHS \le& \left(-\lambda_1 + \frac{1}{2}\right)\delta\mathbb{E}\,\|\mathcal{D}_u X_{km+l+1}\|^2 + \frac{1}{2}\delta\mathbb{E}\left\|\frac{\partial f}{\partial y}(X_{km+l+1},X_{km})\mathcal{D}_u X_{km}\right\|^2\\
&+ \mathbb{E}\left\langle \mathcal{D}_u X_{km+l+1} - \mathcal{D}_u X_{km+l}, \frac{\partial g}{\partial x}(X_{km+l},X_{km})\mathcal{D}_u X_{km+l}\Delta B_{km+1}\right\rangle\\
&+ \mathbb{E}\left\langle \mathcal{D}_u X_{km+l+1} - \mathcal{D}_u X_{km+l}, \frac{\partial g}{\partial y}(X_{km+l},X_{km})\mathcal{D}_u X_{km}\Delta B_{km+1}\right\rangle\\
\le& \left(-\lambda_1 + \frac{1}{2}\right)\delta\mathbb{E}\,\|\mathcal{D}_u X_{km+l+1}\|^2 + \frac{1}{2}\mathbb{E}\,\|\mathcal{D}_u X_{km+l+1} - \mathcal{D}_u X_{km+l}\|^2\\
&+ \left(\frac{1}{2}\lambda_2 + \lambda_3\right)\delta\mathbb{E}\,\|\mathcal{D}_u X_{km}\|^2 + \lambda_3\delta\mathbb{E}\,\|\mathcal{D}_u X_{km+l}\|^2,
\end{aligned}$$

which implies

$$\mathbb{E}\|\mathcal{D}_u X_{km+l+1}\|^2 \le (1-\alpha_6\delta)\,\mathbb{E}\,\|\mathcal{D}_u X_{km+l}\|^2 + \beta_6\delta\mathbb{E}\,\|\mathcal{D}_u X_{km}\|^2,$$

where $\alpha_6 = \frac{2\lambda_1 - 1 - 2\lambda_3}{1+2\lambda_1\delta-\delta}$ and $\beta_6 = \frac{\lambda_2 + 2\lambda_3}{1+2\lambda_1\delta-\delta}$. Since $\alpha_6 > \beta_6 > 0$ and $0 < \frac{\beta_6}{\alpha_6} + \left(1-\frac{\beta_6}{\alpha_6}\right)e^{-\alpha_6(l+1)\delta} < 1$, $l = 0,1,\cdots,m-1$, under the condition $\lambda_1 - 1 - \lambda_2 - 2\lambda_3 > 0$, we get

$$\mathbb{E}\|\mathscr{D}_u X_{km+l+1}\|^2 \le \left(\frac{\beta_6}{\alpha_6} + \left(1 - \frac{\beta_6}{\alpha_6}\right) e^{-\alpha_6 (l+1)\delta}\right) \mathbb{E}\,\|\mathscr{D}_u X_{km}\|^2 .$$

If $l = m-1$, then $\mathbb{E}\|\mathscr{D}_u X_{(k+1)m}\|^2 \le \left(\frac{\beta_6}{\alpha_6} + \left(1 - \frac{\beta_6}{\alpha_6}\right) e^{-\alpha_6}\right) \mathbb{E}\,\|\mathscr{D}_u X_{km}\|^2$. Since for any $u \ge 0$, there exists $n \in \mathbb{N}$ such that $u \in [n-1, n)$. Combining the result in Case 2, we obtain

$$\begin{aligned}
&\mathbb{E}\|\mathscr{D}_u X_{km+l+1}\|^2 \\
\le &\left(\frac{\beta_6}{\alpha_6} + \left(1 - \frac{\beta_6}{\alpha_6}\right) e^{-\alpha_6 (l+1)\delta}\right) \left(\frac{\beta_6}{\alpha_6} + \left(1 - \frac{\beta_6}{\alpha_6}\right) e^{-\alpha}\right)^{k-n} \mathbb{E}\,\|\mathscr{D}_u X_{nm}\|^2 \\
\le &\mathbb{E}\,\|\mathscr{D}_u X_{nm}\|^2 \le C.
\end{aligned}$$

The inequality above shows $\mathbb{E}\|\mathscr{D}_u X_{km+l+1}\|^{2p'} \le C$ with $p' = 1$. Following the same procedure as in Case 2, there exists $C > 0$ independent of δ, k and l such that $\mathbb{E}\|\mathscr{D}_u X_{km+l+1}\|^{2p'} \le C$ for all $\delta \in (0, \delta_0'')$ with $\delta_0'' > 0$ sufficiently small and $p' = 2, 3, \cdots, p$.

Choosing $\tilde{\delta_0} = \min\{\delta_0', \delta_0''\}$, then (39) holds for all $\delta \in (0, \tilde{\delta_0})$. The proof is completed. ◻

For a general function $h : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$, $(x, y) \mapsto h(x, y)$, we use $D_1^{(n)} h(x, y)$ and $D_2^{(n)} h(x, y)$ to denote the partial derivatives with order n of h with respect to the vectors x and y, respectively. We require that the derivatives satisfy some polynomial growth conditions.

Assumption 4.3. *For any x, x', $y \in \mathbb{R}^d$, there exist two positive constants K and q such that*

$$\|f(x, y) - f(x', y)\|^2 \le K\,(1 + \|x\|^q + \|x'\|^q)\,\|x - x'\|^2.$$

Assumption 4.4. *Assume that f and g have all continuous partial derivatives up to order 2. For any x, x', y, y', ξ and $\eta \in \mathbb{R}^d$, there exist two positive constants K and q such that, for $i = 1, 2$,*

$$\begin{aligned}
\left\|D_i^{(1)} f(x, y)\xi - D_i^{(1)} f(x', y)\xi\right\|^2 &\le K\,(1 + \|x\|^q + \|x'\|^q)\,\|x - x'\|^2 \|\xi\|^2, \\
\left\|D_i^{(1)} f(x, y)\xi - D_i^{(1)} f(x, y')\xi\right\|^2 &\le K\|y - y'\|^2 \|\xi\|^2, \\
\left\|D_i^{(2)} f(x, y)(\xi, \eta) - D_i^{(2)} f(x', y)(\xi, \eta)\right\|^2 &\le K\,(1 + \|x\|^q + \|x'\|^q)\,\|x - x'\|^2 \|\xi\|^2 \|\eta\|^2, \\
\left\|D_i^{(2)} f(x, y)(\xi, \eta) - D_i^{(2)} f(x, y')(\xi, \eta)\right\|^2 &\le K\|y - y'\|^2 \|\xi\|^2 \|\eta\|^2, \\
\left\|D_i^{(1)} g(x, y)\xi - D_i^{(1)} g(x', y')\xi\right\|^2 &\le K\left(\|x - x'\|^2 + \|y - y'\|^2\right) \|\xi\|^2, \\
\left\|D_i^{(2)} g(x, y)(\xi, \eta) - D_i^{(2)} g(x', y')(\xi, \eta)\right\|^2 &\le K\left(\|x - x'\|^2 + \|y - y'\|^2\right) \|\xi\|^2 \|\eta\|^2.
\end{aligned}$$

Assumption 4.4 implies

$$\begin{aligned}
\left\|\frac{\partial^2 f}{\partial x^2}(x, y)(\xi, \eta)\right\|^2 \vee \left\|\frac{\partial^2 f}{\partial y \partial x}(x, y)(\xi, \eta)\right\|^2 &\le K\,(1 + 2\|x\|^q)\,\|\xi\|^2 \|\eta\|^2, \\
\left\|\frac{\partial^2 f}{\partial x \partial y}(x, y)(\xi, \eta)\right\|^2 \vee \left\|\frac{\partial^2 f}{\partial y^2}(x, y)(\xi, \eta)\right\|^2 &\le K\|\xi\|^2 \|\eta\|^2
\end{aligned}$$

and

$$\left\|\frac{\partial^2 g}{\partial x^2}(x,y)(\xi,\eta)\right\|^2 \vee \left\|\frac{\partial^2 g}{\partial x\partial y}(x,y)(\xi,\eta)\right\|^2 \vee \left\|\frac{\partial^2 g}{\partial y^2}(x,y)(\xi,\eta)\right\|^2 \vee \left\|\frac{\partial^2 g}{\partial y\partial x}(x,y)(\xi,\eta)\right\|^2 \le K\|\xi\|^2\|\eta\|^2.$$

Similarly, if f and g have all continuous partial derivatives up to order 3, then

$$\left\|\frac{\partial^3 f}{\partial x^3}(x,y)(\xi,\eta,\gamma)\right\|^2 \le K\left(1+2\|x\|^q\right)\|\xi\|^2\|\eta\|^2\|\gamma\|^2$$

and

$$\left\|\frac{\partial^3 f}{\partial x^2\partial y}(x,y)(\xi,\eta,\gamma)\right\|^2 \vee \left\|\frac{\partial^3 f}{\partial x\partial y^2}(x,y)(\xi,\eta,\gamma)\right\|^2 \le K\|\xi\|^2\|\eta\|^2\|\gamma\|^2.$$

Lemma 4.5. *Let the conditions in Lemma 2.4 hold. Assume that the Fréchet derivatives of f and g exist and $\mathbb{E}\|\mathscr{D}_u\eta\|^{2p} < \infty$, then there exist two positive constants C and ν_2 independent of t such that*

$$\mathbb{E}\left\|\mathscr{D}_u X^{i,\eta}(t)\right\|^{2p} \le Ce^{-\nu_2(t-u\vee i)}\left(1+\mathbb{E}\|\mathscr{D}_u\eta\|^{2p}\right), \quad t\ge i. \tag{48}$$

Proof. Since

$$X^{i,\eta}(t)=\eta+\int_i^t f(X^{i,\eta}(s),X^{i,\eta}([s]))ds+\int_i^t g(X^{i,\eta}(s),X^{i,\eta}([s]))dB(s),$$

we have $\mathscr{D}_u X^{i,\eta}(t)=0$ for $u\ge t$, and for $u<t$,

$$\begin{aligned}\mathscr{D}_u X^{i,\eta}(t) =&\mathscr{D}_u\eta 1_{\{u<i\}}+g(X^{i,\eta}(u),X^{i,\eta}([u]))1_{\{i\le u<t\}}\\&+\int_u^t \frac{\partial f}{\partial x}(X^{i,\eta}(s),X^{i,\eta}([s]))\mathscr{D}_u X^{i,\eta}(s)1_{[i,t]}(s)ds\\&+\int_u^t \frac{\partial f}{\partial y}(X^{i,\eta}(s),X^{i,\eta}([s]))\mathscr{D}_u X^{i,\eta}([s])1_{[i,t]}(s)ds\\&+\int_u^t \frac{\partial g}{\partial x}(X^{i,\eta}(s),X^{i,\eta}([s]))\mathscr{D}_u X^{i,\eta}(s)1_{[i,t]}(s)dB(s)\\&+\int_u^t \frac{\partial g}{\partial y}(X^{i,\eta}(s),X^{i,\eta}([s]))\mathscr{D}_u X^{i,\eta}([s])1_{[i,t]}(s)dB(s).\end{aligned} \tag{49}$$

Case 1. If $i\le u<t$, by denoting $I(t):=\mathscr{D}_u X^{i,\eta}(t)$, then

$$\begin{aligned}I(t) =&I(u)+\int_u^t \frac{\partial f}{\partial x}(X^{i,\eta}(s),X^{i,\eta}([s]))I(s)ds+\int_u^t \frac{\partial f}{\partial y}(X^{i,\eta}(s),X^{i,\eta}([s]))I([s])ds\\&+\int_u^t \frac{\partial g}{\partial x}(X^{i,\eta}(s),X^{i,\eta}([s]))I(s)dB(s)+\int_u^t \frac{\partial g}{\partial y}(X^{i,\eta}(s),X^{i,\eta}([s]))I([s])dB(s),\end{aligned}$$

where $I(u)=g(X^{i,\eta}(u),X^{i,\eta}([u]))$. If $[t]\le u<t$, then $\mathscr{D}_u X^{i,\eta}([t])=0$. And Itô's formula leads to

$$\begin{aligned}\mathbb{E}\|I(t)\|^{2p} \le&\mathbb{E}\|I(u)\|^{2p}+2p\mathbb{E}\int_u^t \|I(s)\|^{2(p-1)}\left\langle I(s),\frac{\partial f}{\partial x}(X^{i,\eta}(s),X^{i,\eta}([s]))I(s)\right\rangle ds\\&+p(2p-1)\mathbb{E}\int_u^t \|I(s)\|^{2(p-1)}\left\|\frac{\partial g}{\partial x}(X^{i,\eta}(s),X^{i,\eta}([s]))I(s)\right\|^2 ds\\\le&\mathbb{E}\|I(u)\|^{2p}-\left(2\lambda_1-\lambda_3(2p-1)\right)p\int_u^t \mathbb{E}\|I(s)\|^{2p}ds.\end{aligned}$$

From [11, Lemma 8.2], it follows

$$\mathbb{E}\,\|I(t)\|^{2p} \le e^{-(2\lambda_1-\lambda_3(2p-1))p(t-u)}\mathbb{E}\,\|I(u)\|^{2p}.$$

And if $i \le u < [t]$, for any $\alpha > 0$, applying Itô's formula to $e^{2\alpha pt}\|I(t)\|^{2p}$, we obtain

$$\begin{aligned}
&e^{2\alpha pt}\mathbb{E}\,\|I(t)\|^{2p}\\
\le& e^{2\alpha p[t]}\mathbb{E}\,\|I([t])\|^{2p} + 2\alpha p\mathbb{E}\int_{[t]}^{t} e^{2\alpha ps}\mathbb{E}\,\|I(s)\|^{2p}\,ds\\
&+ 2p\mathbb{E}\int_{[t]}^{t} e^{2\alpha ps}\|I(s)\|^{2(p-1)}\left\langle I(s), \frac{\partial f}{\partial x}(X^{i,\eta}(s), X^{i,\eta}([s]))I(s)\right\rangle ds\\
&+ 2p\mathbb{E}\int_{[t]}^{t} e^{2\alpha ps}\|I(s)\|^{2(p-1)}\left\langle I(s), \frac{\partial f}{\partial y}(X^{i,\eta}(s), X^{i,\eta}([s]))I([s])\right\rangle ds\\
&+ 2p(2p-1)\mathbb{E}\int_{[t]}^{t} e^{2\alpha ps}\|I(s)\|^{2(p-1)}\left\|\frac{\partial g}{\partial x}(X^{i,\eta}(s), X^{i,\eta}([s]))I(s)\right\|^2 ds\\
&+ 2p(2p-1)\mathbb{E}\int_{[t]}^{t} e^{2\alpha ps}\|I(s)\|^{2(p-1)}\left\|\frac{\partial g}{\partial y}(X^{i,\eta}(s), X^{i,\eta}([s]))I([s])\right\|^2 ds.
\end{aligned}$$

Following the same procedure as Lemma 4.1, we get

$$\begin{aligned}
\mathbb{E}\,\|I(t)\|^{2p} \le& Ce^{-\nu_1(t-[u]-1)}\mathbb{E}\,\|I([u]+1)\|^{2p}\\
\le& Ce^{-\nu_1(t-[u]-1)}\cdot e^{-(2\lambda_1-\lambda_3(2p-1))p([u]+1-u)}\mathbb{E}\,\|I(u)\|^{2p}\\
=& Ce^{-\nu_2(t-u)}\mathbb{E}\,\|I(u)\|^{2p},
\end{aligned}$$

where $\nu_2 = \min\{\nu_1, (2\lambda_1 - \lambda_3(2p-1))p\}$.

Case 2. If $u < i$, then $I(u) = \mathcal{D}_u\eta$ and

$$\mathbb{E}\,\|I(t)\|^{2p} \le Ce^{-\nu_2(t-i)}\mathbb{E}\,\|I(u)\|^{2p}.$$

Since $\mathbb{E}\|X(t)\|^{2p} \le C$ for all $t > 0$ and C is independent of t, we obtain

$$\mathbb{E}\,\|I(u)\|^{2p} \le \mathbb{E}\,\|\mathcal{D}_u\eta\|^{2p} + \mathbb{E}\left\|g(X^{i,\eta}(u), X^{i,\eta}([u]))\right\|^{2p} \le C\left(1 + \mathbb{E}\,\|\mathcal{D}_u\eta\|^{2p}\right).$$

Hence

$$\mathbb{E}\left\|\mathcal{D}_uX^{i,\eta}(t)\right\|^{2p} \le Ce^{-\nu_2(t-u\vee i)}\left(1 + \mathbb{E}\,\|\mathcal{D}_u\eta\|^{2p}\right),$$

where C is independent of t. We complete the proof □

Lemma 4.6. *Let conditions in Lemma 2.4 with $p \ge 4$, and Assumptions 4.3- 4.4 hold. Then there exist two positive constants C and ν_3 independent of t such that, for fixed $i \in \mathbb{N}$,*

$$\begin{aligned}
&\mathbb{E}\left\|\mathcal{D}_uDX^{i,\eta}(t)\xi\right\|^{2p'}\\
&\le Ce^{-\nu_3(t-u\vee i)}\|\xi\|^{2p'} + Ce^{-\nu_1([t]-i-1)p'/p}\Big(1 + \mathbb{E}\,\|\mathcal{D}_u\eta\|^{2p'}\Big)\|\xi\|^{2p'},
\end{aligned}$$

where $t \ge i$, $1 \le p' \le \min\{\frac{p}{4}, \frac{p}{q}\}$, $\xi \in \mathbb{R}^d$ and $\mathbb{E}\,\|\mathcal{D}_u\eta\|^{2p} < \infty$.

Proof. For any $\xi \in \mathbb{R}^d$, denote $J(t) := \mathcal{D}_uDX^{i,\eta}(t)\xi$, then

$$\begin{aligned}
&J(t)\\
=&\frac{\partial g}{\partial x}(X^{i,\eta}(u), X^{i,\eta}([u]))H(u)1_{\{i\le u<t\}} + \frac{\partial g}{\partial y}(X^{i,\eta}(u), X^{i,\eta}([u]))H([u])1_{\{i\le u<t\}}
\end{aligned}$$

$$
\begin{aligned}
&+\int_u^t \frac{\partial^2 f}{\partial x^2}(X^{i,\eta}(s),X^{i,\eta}([s]))(H(s),I(s))1_{[i,t]}(s)ds\\
&+\int_u^t \frac{\partial^2 f}{\partial x\partial y}(X^{i,\eta}(s),X^{i,\eta}([s]))(H(s),I([s]))1_{[i,t]}(s)ds\\
&+\int_u^t \frac{\partial f}{\partial x}(X^{i,\eta}(s),X^{i,\eta}([s]))J(s)1_{[i,t]}(s)ds\\
&+\int_u^t \frac{\partial^2 f}{\partial y\partial x}(X^{i,\eta}(s),X^{i,\eta}([s]))(H([s]),I(s))1_{[i,t]}(s)ds\\
&+\int_u^t \frac{\partial f}{\partial y}(X^{i,\eta}(s),X^{i,\eta}([s]))J([s])1_{[i,t]}(s)ds\\
&+\int_u^t \frac{\partial^2 f}{\partial y^2}(X^{i,\eta}(s),X^{i,\eta}([s]))(H([s]),I([s]))1_{[i,t]}(s)ds\\
&+\int_u^t \frac{\partial^2 g}{\partial x^2}(X^{i,\eta}(s),X^{i,\eta}([s]))(H(s),I(s))1_{[i,t]}(s)dB(s)\\
&+\int_u^t \frac{\partial^2 g}{\partial x\partial y}(X^{i,\eta}(s),X^{i,\eta}([s]))(H(s),I([s]))1_{[i,t]}(s)dB(s)\\
&+\int_u^t \frac{\partial g}{\partial x}(X^{i,\eta}(s),X^{i,\eta}([s]))J(s)1_{[i,t]}(s)dB(s)\\
&+\int_u^t \frac{\partial^2 g}{\partial y\partial x}(X^{i,\eta}(s),X^{i,\eta}([s]))(H([s]),I(s))1_{[i,t]}(s)dB(s)\\
&+\int_u^t \frac{\partial g}{\partial y}(X^{i,\eta}(s),X^{i,\eta}([s]))J([s])1_{[i,t]}(s)dB(s)\\
&+\int_u^t \frac{\partial^2 g}{\partial y^2}(X^{i,\eta}(s),X^{i,\eta}([s]))(H([s]),I([s]))1_{[i,t]}(s)dB(s)
\end{aligned}
$$

for $u < t$. Since $J(t) = 0$ for $u \geq t$, we only consider the case $u < t$.

Case 1. If $u \geq [t]$, then $[s] = [t]$, $I([t]) = 0$, $J(u) = \frac{\partial g}{\partial x}(X^{i,\eta}(u), X^{i,\eta}([u]))H(u) + \frac{\partial g}{\partial y}(X^{i,\eta}(u), X^{i,\eta}([u]))H([u])$ and $J([t]) = 0$. From Itô's formula, it follows

$$
\begin{aligned}
&\mathbb{E}\,\|J(t)\|^{2p'}\\
\leq&\mathbb{E}\,\|J(u)\|^{2p'} + 2p'\mathbb{E}\int_u^t \|J(s)\|^{2(p'-1)}\left\langle J(s), \frac{\partial^2 f}{\partial x^2}(X^{i,\eta}(s),X^{i,\eta}([t]))(H(s),I(s))\right\rangle ds\\
&+2p'\mathbb{E}\int_u^t \|J(s)\|^{2(p'-1)}\left\langle J(s), \frac{\partial f}{\partial x}(X^{i,\eta}(s),X^{i,\eta}([t]))J(s)\right\rangle ds\\
&+2p'\mathbb{E}\int_u^t \|J(s)\|^{2(p'-1)}\left\langle J(s), \frac{\partial^2 f}{\partial y\partial x}(X^{i,\eta}(s),X^{i,\eta}([t]))(H([t]),I(s))\right\rangle ds\\
&+2p'(2p'-1)\mathbb{E}\int_u^t \|J(s)\|^{2(p'-1)}\left\|\frac{\partial g}{\partial x}(X^{i,\eta}(s),X^{i,\eta}([t]))J(s)\right\|^2 ds\\
&+4p'(2p'-1)\mathbb{E}\int_u^t \|J(s)\|^{2(p'-1)}\left\|\frac{\partial^2 g}{\partial x^2}(X^{i,\eta}(s),X^{i,\eta}([t]))(H(s),I(s))\right\|^2 ds\\
&+4p'(2p'-1)\mathbb{E}\int_u^t \|J(s)\|^{2(p'-1)}\left\|\frac{\partial^2 g}{\partial y\partial x}(X^{i,\eta}(s),X^{i,\eta}([t]))(H([t]),I(s))\right\|^2 ds.
\end{aligned}
$$

And Young's inequality yields

$$\mathbb{E}\,\|J(t)\|^{2p'}$$

$$
\begin{aligned}
&\leq \mathbb{E}\,\|J(u)\|^{2p'} - (2\lambda_1 - 2\lambda_3(2p'-1) - 4\epsilon_1 - 8(2p'-1)\epsilon_2)\,p'\mathbb{E}\int_u^t \|J(s)\|^{2p'}\,ds \\
&\quad + \left(\frac{2p'-1}{2p'\epsilon_1}\right)^{2p'-1} \mathbb{E}\int_u^t \left\|\frac{\partial^2 f}{\partial x^2}(X^{i,\eta}(s), X^{i,\eta}([t]))(H(s), I(s))\right\|^{2p'} ds \\
&\quad + \left(\frac{2p'-1}{2p'\epsilon_1}\right)^{2p'-1} \mathbb{E}\int_u^t \left\|\frac{\partial^2 f}{\partial y \partial x}(X^{i,\eta}(s), X^{i,\eta}([t]))(H([t]), I(s))\right\|^{2p'} ds \\
&\quad + 2(2p'-1)\left(\frac{p'-1}{p'\epsilon_2}\right)^{p'-1} \mathbb{E}\int_u^t \left\|\frac{\partial^2 g}{\partial x^2}(X^{i,\eta}(s), X^{i,\eta}([t]))(H(s), I(s))\right\|^{2p'} ds \\
&\quad + 2(2p'-1)\left(\frac{p'-1}{p'\epsilon_2}\right)^{p'-1} \mathbb{E}\int_u^t \left\|\frac{\partial^2 g}{\partial y \partial x}(X^{i,\eta}(s), X^{i,\eta}([t]))(H([t]), I(s))\right\|^{2p'} ds.
\end{aligned}
$$

Taking $\epsilon_1 = \frac{1}{2}\lambda_3$ and $\epsilon_2 = \frac{2p'}{4(2p'-1)}\lambda_3$ and using the estimates of the partial derivatives of f and g with order 1 and 2, we obtain

$$
\begin{aligned}
&\mathbb{E}\,\|J(t)\|^{2p'} \\
&\leq \mathbb{E}\,\|J(u)\|^{2p'} - 2\left(\lambda_1 - 2\lambda_3 - 4\lambda_3(p'-1)\right)p'\mathbb{E}\int_u^t \|J(s)\|^{2p'}\,ds \\
&\quad + \left(\frac{2p'-1}{\lambda_3 p'}\right)^{2p'-1} K^{p'}\int_u^t \mathbb{E}\left(\left(1 + 2\left\|X^{i,\eta}(s)\right\|^q\right)^{p'} \|H(s)\|^{2p'}\|I(s)\|^{2p'}\right) ds \\
&\quad + \left(\frac{2p'-1}{\lambda_3 p'}\right)^{2p'-1} K^{p'}\int_u^t \mathbb{E}\left(\left(1 + 2\left\|X^{i,\eta}(s)\right\|^q\right)^{p'} \|H([t])\|^{2p'}\|I(s)\|^{2p'}\right) ds \\
&\quad + 2(2p'-1)\left(\frac{2(p'-1)(2p-1)}{p'^2\lambda_3}\right)^{p'-1} K^{p'}\int_u^t \mathbb{E}\left(\|H(s)\|^{2p'}\,\|I(s))\|^{2p'}\right) ds \\
&\quad + 2(2p'-1)\left(\frac{2(p'-1)(2p-1)}{p'^2\lambda_3}\right)^{p'-1} K^{p'}\int_u^t \mathbb{E}\left(\|H([t])\|^{2p'}\,\|I(s))\|^{2p'}\right) ds.
\end{aligned}
$$

Using Hölder inequality, $qp' \leq p$ and $4p' \leq p$, Lemmas 4.1 and 4.5 lead to

$$
\begin{aligned}
&\mathbb{E}\left(\left(1 + 2\left\|X^{i,\eta}(s)\right\|^q\right)^{p'} \|H(s)\|^{2p'}\|I(s)\|^{2p'}\right) \\
&\leq \left(\mathbb{E}\left(1 + 2\left\|X^{i,\eta}(s)\right\|^q\right)^{2p'}\right)^{\frac{1}{2}} \left(\mathbb{E}\|H(s)\|^{8p'}\right)^{\frac{1}{4}} \left(\mathbb{E}\|I(s)\|^{8p'}\right)^{\frac{1}{4}} \\
&\leq \left(\mathbb{E}\left(1 + 2\left\|X^{i,\eta}(s)\right\|^q\right)^{2p'}\right)^{\frac{1}{2}} \left(\mathbb{E}\|H(s)\|^{2p}\right)^{\frac{p'}{p}} \left(\mathbb{E}\|I(s)\|^{2p}\right)^{\frac{p'}{p}} \\
&\leq C\left(1 + \mathbb{E}\,\|\mathcal{D}_u\eta\|^{2p'}\right)\|\xi\|^{2p'} e^{-\nu_1(s-i)p'/p} e^{-\nu_2(s-u)p'/p} \\
&\leq C\left(1 + \mathbb{E}\,\|\mathcal{D}_u\eta\|^{2p'}\right)\|\xi\|^{2p'} e^{-\nu_1([t]-i)p'/p}.
\end{aligned}
$$

Similarly,

$$
\mathbb{E}\left(\|H(s)\|^{2p'}\|I(s)\|^{2p'}\right) \leq Ce^{-\nu_1([t]-i)p'/p}\left(1 + \mathbb{E}\,\|\mathcal{D}_u\eta\|^{2p'}\right)\|\xi\|^{2p'},
$$

$$
\begin{aligned}
&\mathbb{E}\left(\left(1 + 2\left\|X^{i,\eta}(s)\right\|^q\right)^{p'}\|H([t])\|^{2p'}\|I(s)\|^{2p'}\right) \\
&\leq Ce^{-\nu_1([t]-i)p'/p}\left(1 + \mathbb{E}\,\|\mathcal{D}_u\eta\|^{2p'}\right)\|\xi\|^{2p'}
\end{aligned}
$$

and

$$\mathbb{E}\left(\|H([t])\|^{2p'}\|I(s)\|^{2p'}\right) \le Ce^{-\nu_1([t]-i)p'/p}\left(1+\mathbb{E}\|\mathscr{D}_u\eta\|^{2p'}\right)\|\xi\|^{2p'}.$$

Let $\alpha_7 = \lambda_1 - 2\lambda_3 - 4\lambda_3(p'-1)$, then $2\alpha_7 p' > 0$ by the condition $\lambda_1 - \lambda_2 - 1 - 2\lambda_3 \ge 4\lambda_3(p'-1)$. Therefore, [11, Lemma 8.2] leads to

$$\mathbb{E}\|J(t)\|^{2p'} \le e^{-2\alpha_7 p'(t-u)}\mathbb{E}\|J(u)\|^{2p'} + Ce^{-\nu_1([t]-i)p'/p}\left(1+\mathbb{E}\|\mathscr{D}_u\eta\|^{2p'}\right)\|\xi\|^{2p'}.$$

Case 2. If $i \le u < [t]$, applying Itô's formula to $e^{2\alpha p'}\|J(t)\|^{2p'}$, $\alpha > 0$, we have

$$\begin{aligned}
&e^{2\alpha p' t}\mathbb{E}\|J(t)\|^{2p'}\\
\le& e^{2\alpha p'[t]}\mathbb{E}\|J([t])\|^{2p'} + 2\alpha p'\mathbb{E}\int_{[t]}^t e^{2\alpha p' s}\|J(s)\|^{2p'}\,ds\\
&+ 2p'\mathbb{E}\int_{[t]}^t e^{2\alpha p' s}\|J(s)\|^{2(p'-1)}\left\langle J(s), \frac{\partial f}{\partial x}(X^{i,\eta}(s), X^{i,\eta}([t]))J(s)\right\rangle ds\\
&+ 2p'\mathbb{E}\int_{[t]}^t e^{2\alpha p' s}\|J(s)\|^{2(p'-1)}\left\langle J(s), \frac{\partial f}{\partial y}(X^{i,\eta}(s), X^{i,\eta}([t]))J([t])\right\rangle ds\\
&+ 2p'\mathbb{E}\int_{[t]}^t e^{2\alpha p' s}\|J(s)\|^{2(p'-1)}\langle J(s), \mathscr{A}(s)\rangle\,ds\\
&+ p'(2p'-1)\mathbb{E}\int_{[t]}^t e^{2\alpha p' s}\|J(s)\|^{2(p'-1)}\Bigg\|\frac{\partial g}{\partial x}(X^{i,\eta}(s), X^{i,\eta}([t]))J(s)\\
&\qquad + \frac{\partial g}{\partial y}(X^{i,\eta}(s), X^{i,\eta}([t]))J([t]) + \mathscr{B}(s)\Bigg\|^2 ds,
\end{aligned}$$

where

$$\begin{aligned}
\mathscr{A}(s) =& \frac{\partial^2 f}{\partial x^2}(X^{i,\eta}(s), X^{i,\eta}([t]))(H(s), I(s)) + \frac{\partial^2 f}{\partial x\partial y}(X^{i,\eta}(s), X^{i,\eta}([t]))(H(s), I([t]))\\
&+ \frac{\partial^2 f}{\partial y\partial x}(X^{i,\eta}(s), X^{i,\eta}([t]))(H([t]), I(s)) + \frac{\partial^2 f}{\partial y^2}(X^{i,\eta}(s), X^{i,\eta}([t]))(H([t]), I([t]))
\end{aligned}$$

and

$$\begin{aligned}
\mathscr{B}(s) =& \frac{\partial^2 g}{\partial x^2}(X^{i,\eta}(s), X^{i,\eta}([t]))(H(s), I(s)) + \frac{\partial^2 g}{\partial x\partial y}(X^{i,\eta}(s), X^{i,\eta}([t]))(H(s), I([t]))\\
&+ \frac{\partial^2 g}{\partial y\partial x}(X^{i,\eta}(s), X^{i,\eta}([t]))(H([t]), I(s)) + \frac{\partial^2 g}{\partial y^2}(X^{i,\eta}(s), X^{i,\eta}([t]))(H([t]), I([t])).
\end{aligned}$$

By the estimates of all the partial derivatives of f and g up to order 2 and Young's inequality, we get

$$\begin{aligned}
&e^{2\alpha p' t}\mathbb{E}\|J(t)\|^{2p'}\\
\le& e^{2\alpha p'[t]}\mathbb{E}\|J([t])\|^{2p'}\\
&+ (2\alpha - 2\lambda_1 + 1 + 2\epsilon_1 + 2\lambda_3(2p'-1))\,p'\mathbb{E}\int_{[t]}^t e^{2\alpha p' s}\|J(s)\|^{2p'}\,ds\\
&+ p'\mathbb{E}\int_{[t]}^t e^{2\alpha p' s}\|J(s)\|^{2(p'-1)}\left\|\frac{\partial f}{\partial y}(X^{i,\eta}(s), X^{i,\eta}([t]))J([t])\right\|^2 ds\\
&+ \left(\frac{2p'-1}{2p'\epsilon_1}\right)^{2p'-1}\mathbb{E}\int_{[t]}^t e^{2\alpha p' s}\|\mathscr{A}(s)\|^{2p'}\,ds
\end{aligned}$$

$$
\begin{aligned}
&+ 4p'(2p'-1)\mathbb{E}\int_{[t]}^{t} e^{2\alpha p's}\,\|J(s)\|^{2(p'-1)}\,\|\mathscr{B}(s)\|^2\,ds\\
&+ 4p'(2p'-1)\mathbb{E}\int_{[t]}^{t} e^{2\alpha p's}\,\|J(s)\|^{2(p'-1)}\left\|\frac{\partial g}{\partial y}(X^{i,\eta}(s),X^{i,\eta}([t]))J([t])\right\|^2 ds\\
\le& e^{2\alpha p'[t]}\mathbb{E}\,\|J([t])\|^{2p'} + (\lambda_2 + 4\lambda_3(2p'-1))\,\mathbb{E}\int_{[t]}^{t} e^{2\alpha p's}\,\|J([t])\|^{2p'}\,ds\\
&+ (2\alpha p' - \tilde{\lambda})\mathbb{E}\int_{u}^{t} e^{2\alpha p's}\,\|J(s)\|^{2p'}\,ds + \left(\frac{2p'-1}{2p'\epsilon_1}\right)^{2p'-1}\mathbb{E}\int_{[t]}^{t} e^{2\alpha p's}\,\|\mathscr{A}(s)\|^{2p'}\,ds\\
&+ 2(2p'-1)\left(\frac{p'-1}{p'\epsilon_2}\right)^{p'-1}\mathbb{E}\int_{[t]}^{t} e^{2\alpha p's}\,\|\mathscr{B}(s)\|^{2p'}\,ds,
\end{aligned}
$$

where $\tilde{\lambda} = (2\lambda_1 - 1 - 2\epsilon_1 - 6\lambda_3(2p'-1) - 4(2p'-1)\epsilon_2 - \lambda_2)\,p' + \lambda_2 + 4\lambda_3(2p'-1)$. Taking $2\alpha_8 p' = \tilde{\lambda}$, $\beta_7 = \lambda_2 + 4\lambda_3(2p'-1)$, $\epsilon_1 = \frac{1}{4}$ and $\epsilon_2 = \frac{1}{8(2p'-1)}$, then $2\alpha_8 p' = 2\lambda_1 p' - 2p' - \lambda_2 p' - 6\lambda_3(2p'-1)p' + \lambda_2 + 4\lambda_3(2p'-1)$, and

$$
\begin{aligned}
e^{2\alpha_8 p't}\mathbb{E}\,\|J(t)\|^{2p'} \le& e^{2\alpha_8 p'[t]}\mathbb{E}\,\|J([t])\|^{2p'} + \beta_7\mathbb{E}\int_{[t]}^{t} e^{2\alpha_8 p's}\,\|J([t])\|^{2p'}\,ds\\
&+ C\int_{[t]}^{t} e^{2\alpha_8 p's}\mathbb{E}\left((1+2\|X^{i,\eta}(s)\|^{q})^{p'}\,\|H(s)\|^{2p'}\,\|I(s)\|^{2p'}\right)ds\\
&+ C\int_{[t]}^{t} e^{2\alpha_8 p's}\mathbb{E}\left(\|H(s)\|^{2p'}\,\|I([t])\|^{2p'}\right)ds\\
&+ C\int_{[t]}^{t} e^{2\alpha_8 p's}\mathbb{E}\left((1+2\|X^{i,\eta}(s)\|^{q})^{p'}\,\|H([t])\|^{2p'}\,\|I(s)\|^{2p'}\right)ds\\
&+ C\int_{[t]}^{t} e^{2\alpha_8 p's}\mathbb{E}\left(\|H([t])\|^{2p'}\,\|I([t])\|^{2p'}\right)ds,
\end{aligned}
$$

where C is independent of t. Similarly to the case $u \ge [t]$, we get

$$
\begin{aligned}
e^{2\alpha_8 p't}\mathbb{E}\,\|J(t)\|^{2p'} \le& e^{2\alpha_8 p'[t]}\mathbb{E}\,\|J([t])\|^{2p'} + \beta_7\mathbb{E}\int_{[t]}^{t} e^{2\alpha_8 p's}\,\|J([t])\|^{2p'}\,ds\\
&+ Ce^{-\nu_1([t]-i)p'/p}\left(1+\mathbb{E}\,\|\mathscr{D}_u\eta\|^{2p'}\right)\|\xi\|^{2p'}\int_{[t]}^{t} e^{2\alpha_8 p's}ds\\
\le& \left(\frac{\beta_7}{2\alpha_8 p'} + \left(1-\frac{\beta_7}{2\alpha_8 p'}\right)e^{-2\alpha_8 p'\{t\}}\right)e^{2\alpha_8 p't}\mathbb{E}\,\|J([t])\|^{2p'}\\
&+ Ce^{-\nu_1([t]-i)p'/p}\left(1+\mathbb{E}\,\|\mathscr{D}_u\eta\|^{2p'}\right)\|\xi\|^{2p'}\left(e^{2\alpha_8 p't} - e^{2\alpha_8 p'[t]}\right),
\end{aligned}
$$

which implies

$$
\mathbb{E}\,\|J(t)\|^{'} \le r_5(\{t\})\mathbb{E}\,\|J([t])\|^{2p'} + Ce^{-\nu_1([t]-i)p'/p}\left(1+\mathbb{E}\,\|\mathscr{D}_u\eta\|^{2p'}\right)\|\xi\|^{2p'},
$$

where the function r_5 is defined similarly to r in the proof of Theorem 2.3 with $r_5(\{t\}) = \frac{\beta_7}{2\alpha_8 p'} + \left(1-\frac{\beta_7}{2\alpha_8 p'}\right)e^{-2\alpha_8 p'\{t\}}$ for $\{t\} \in [0,1)$ and $r_5(1) := \lim_{t\to k^-} r_5(\{t\})$, $k \in \mathbb{N}$. Since $\lambda_1 - \lambda_2 - 1 - 2\lambda_3 > 4\lambda_3(p-1)$, we have $2\alpha_8 p' > \beta_7$ and $0 < r_5 \le 1$. Similarly to the proof of Lemma 4.1, we obtain

$$
\mathbb{E}\,\|J(t)\|^{2p'} \le r_5(\{t\})r_5(1)^{[t]-[u]-1}\mathbb{E}\,\|J([u]+1)\|^{2p'}
$$

$$+C\left(1+\mathbb{E}\,\|\mathscr{D}_u\eta\|^{2p'}\right)\|\xi\|^{2p'}e^{-\nu_1([t]-i)p'/p}\times\left(\sum_{j=1}^{[t]-[u]-1}r_5(1)^{j-1}e^{\nu_1 jp'/p}\right)$$
$$+C\left(1+\mathbb{E}\,\|\mathscr{D}_u\eta\|^{2p'}\right)\|\xi\|^{2p'}e^{-\nu_1([t]-i)p'/p}.$$

Without loss of generality, we assume $r_5(1)e^{\nu_1 p'/p}<1$, then

$$\begin{aligned}\mathbb{E}\,\|J(t)\|^{2p'}\leq& r_5(\{t\})r_5(1)^{[t]-[u]-1}\mathbb{E}\,\|J([u]+1)\|^{2p'}\\&+C\left(1+\mathbb{E}\,\|\mathscr{D}_u\eta\|^{2p'}\right)\|\xi\|^{2p'}e^{-\nu_1([t]-i-1)p'/p}\\\leq&\frac{1}{r_5(1)}e^{(t-[u]-1)\log r_5(1)}e^{-2\alpha_7 p'([u]+1-u)}\mathbb{E}\,\|J(u)\|^{2p'}\\&+Ce^{-\nu_1([t]-i-1)p'/p}\left(1+\mathbb{E}\,\|\mathscr{D}_u\eta\|^{2p'}\right)\|\xi\|^{2p'}\\\leq&Ce^{-\nu_3(t-u)}\mathbb{E}\,\|J(u)\|^{2p'}+Ce^{-\nu_1([t]-i)p'/p}\left(1+\mathbb{E}\,\|\mathscr{D}_u\eta\|^{2p'}\right)\|\xi\|^{2p'},\end{aligned}$$

where $\nu_3=\min\{-\log r_5(1),2\alpha_7 p'\}$. By the estimates and the uniform boundedness of $X(s)$, we have

$$\mathbb{E}\|J(u)\|^{2p'}\leq C\|\xi\|^{2p'}.$$

Therefore

$$\mathbb{E}\,\|J(t)\|^{2p'}\leq Ce^{-\nu_3(t-u)}\|\xi\|^{2p'}+Ce^{-\nu_1([t]-i)p'/p}\left(1+\mathbb{E}\,\|\mathscr{D}_u\eta\|^{2p'}\right)\|\xi\|^{2p'}.$$

Case 3. If $u<i$, then similar to Case 2, we have

$$\mathbb{E}\,\|J(t)\|^{2p'}\leq Ce^{-\nu_3(t-i)}\|\xi\|^{2p'}+Ce^{-\nu_1([t]-i)p'/p}\left(1+\mathbb{E}\,\|\mathscr{D}_u\eta\|^{2p'}\right)\|\xi\|^{2p'}.$$

The proof is completed. □

Lemma 4.7. *Let f and g have continuous partial derivatives up to order 3. Suppose that conditions in Lemma 2.4 with $p\geq 4$, Assumptions 4.3-4.4, and $\mathbb{E}\,\|\mathscr{D}_u\eta\|^{2p}<\infty$ hold. Then*

$$\begin{aligned}\mathbb{E}\left\|\mathscr{D}_w\mathscr{D}_u X^{i,\eta}(t)\xi\right\|^{2p'}\leq& Ce^{-\nu_3(t-w\vee u\vee i)}\mathbb{E}\|\mathscr{D}_w\mathscr{D}_u\eta\|^{2p'}\\&+Ce^{-\nu_2([t]-w\vee u\vee i)p'/p}\left(1+\mathbb{E}\,\|\mathscr{D}_u\eta\|^{2p'}+\mathbb{E}\,\|\mathscr{D}_w\eta\|^{2p'}\right)\end{aligned}\tag{50}$$

and

$$\begin{aligned}\mathbb{E}\left\|\mathscr{D}_w\mathscr{D}_u DX^{i,\eta}(t)\xi\right\|^{2p'}\leq& Ce^{-\nu_3(t-w\vee u\vee i)}\mathbb{E}\|\mathscr{D}_w\mathscr{D}_u\eta\|^{2p'}\\&+Ce^{-\nu_2([t]-w\vee u\vee i)p'/p}\left(1+\mathbb{E}\,\|\mathscr{D}_u\eta\|^{2p'}+\mathbb{E}\,\|\mathscr{D}_w\eta\|^{2p'}\right)\end{aligned}\tag{51}$$

for any $1\leq p'\leq\min\{\frac{p}{4},\frac{p}{q}\}$, $\xi\in\mathbb{R}^d$.

4.2. **Errors of π and π^δ.** Let us first derive the weak error of $X(k)$ and Y_k.

Theorem 4.8. *Let conditions in Lemma 2.4 with $p\geq 4$, and Assumptions 4.3-4.4 hold. Then there exists $C:=C(\|\phi\|_3,\lambda_1,\lambda_2,\lambda_3,p,q,K)>0$ independent of k,δ such that*

$$\left|\mathbb{E}\phi(X^{0,x}(k))-\mathbb{E}\phi(Y_k^{0,x})\right|\leq C\delta,\quad\forall\ \phi\in C_b^3$$

for any $\delta\in(0,\delta_1)$ with $\delta_1>0$ sufficiently small.

Proof. For any $\phi \in C_b^3$,

$$
\begin{aligned}
&\left|\mathbb{E}\phi(X^{0,x}(k)) - \mathbb{E}\phi(Y_k^{0,x})\right| \\
=&\left|\sum_{i=0}^{k-1}\left(\mathbb{E}\phi(X^{i+1,X^{0,x}_{(i+1)m}}(k)) - \mathbb{E}\phi(X^{i,X^{0,x}_{im}}(k))\right)\right| \\
=&\left|\sum_{i=0}^{k-1}\left(\mathbb{E}\phi(X^{i+1,X^{0,x}_{(i+1)m}}(k)) - \mathbb{E}\phi(X^{i+1,X^{i,X^{0,x}_{im}}(i+1)}(k))\right)\right| \\
\le&\sum_{i=0}^{k-1}\left|\mathbb{E}\phi(X^{i+1,X^{0,x}_{(i+1)m}}(k)) - \mathbb{E}\phi(X^{i+1,X^{i,X^{0,x}_{im}}(i+1)}(k))\right| \\
=&\sum_{i=0}^{k-1}\left|\mathbb{E}\int_0^1 D(\phi\circ X)_{i+1}(k,\theta)\left(X^{0,x}_{(i+1)m} - X^{i,X^{0,x}_{im}}(i+1)\right)d\theta\right|,
\end{aligned}
$$

where $D(\phi\circ X)_{i+1}(k,\theta) = D(\phi\circ X)(k; i+1, \theta X^{0,x}_{(i+1)m} + (1-\theta)X^{i,X^{0,x}_{im}}(i+1))$. From

$$
\begin{aligned}
&X^{0,x}_{(i+1)m} - X^{i,X^{0,x}_{im}}(i+1) = X^{i,X^{0,x}_{im}}_{(i+1)m} - X^{i,X^{0,x}_{im}}(i+1) \\
=&\sum_{l=0}^{m-1}\int_{t_{im+l}}^{t_{im+l+1}}\left(f(X_{im+l+1}, X^{0,x}_{im}) - f(X^{i,X^{0,x}_{im}}(s), X^{0,x}_{im})\right)ds \\
&+\sum_{l=0}^{m-1}\int_{t_{im+l}}^{t_{im+l+1}}\left(g(X_{im+l}, X^{0,x}_{im}) - g(X^{i,X^{0,x}_{im}}(s), X^{0,x}_{im})\right)dB(s),
\end{aligned}
$$

it follows that

$$
\begin{aligned}
&\left|\mathbb{E}\phi(X^{0,x}(k)) - \mathbb{E}\phi(Y_k^{0,x})\right| \\
\le&\sum_{i=0}^{k-1}\sum_{l=0}^{m-1}\left|\mathbb{E}\int_0^1 D(\phi\circ X)_{i+1}(k,\theta)\cdot\int_{t_{im+l}}^{t_{im+l+1}}\left(f(X_{im+l+1}, X^{0,x}_{im}) - f(X_{im+l}, X^{0,x}_{im})\right)ds d\theta\right| \\
&+\sum_{i=0}^{k-1}\sum_{l=0}^{m-1}\left|\mathbb{E}\int_0^1 D(\phi\circ X)_{i+1}(k,\theta)\cdot\int_{t_{im+l}}^{t_{im+l+1}}\left(f(X^{i,X^{0,x}_{im}}(s), X^{0,x}_{im}) - f(X_{im+l}, X^{0,x}_{im})\right)ds d\theta\right| \\
&+\sum_{i=0}^{k-1}\sum_{l=0}^{m-1}\left|\mathbb{E}\int_0^1 D(\phi\circ X)_{i+1}(k,\theta)\cdot\int_{t_{im+l}}^{t_{im+l+1}}\left(g(X^{i,X^{0,x}_{im}}(s), X^{0,x}_{im}) - g(X_{im+l}, X^{0,x}_{im})\right)dB(s)d\theta\right| \\
=:&\sum_{i=0}^{k-1}\sum_{l=0}^{m-1}(I_1 + I_2 + I_3).
\end{aligned}
$$

Denote $X_{im+l+\tau} := \tau X_{im+l+1} + (1-\tau)X_{im+l}$, $\tau \in [0,1]$. Then the estimate of I_1 is

$$
\begin{aligned}
I_1 =&\delta\left|\mathbb{E}\int_0^1\int_0^1 D(\phi\circ X)_{i+1}(k,\theta)\cdot\frac{\partial f}{\partial x}(X_{im+l+\tau}, X^{0,x}_{im})\,(X_{im+l+1} - X_{im+l})\,d\tau d\theta\right| \\
\le&\delta^2\int_0^1\int_0^1\left|\mathbb{E}\left(D(\phi\circ X)_{i+1}(k,\theta)\cdot\frac{\partial f}{\partial x}(X_{im+l+\tau}, X^{0,x}_{im})f(X_{im+l+1}, X^{0,x}_{im})\right)\right|d\tau d\theta \\
&+\delta\int_0^1\int_0^1\left|\mathbb{E}\left(D(\phi\circ X)_{i+1}(k,\theta)\cdot\frac{\partial f}{\partial x}(X_{im+l+\tau}, X^{0,x}_{im})g(X_{im+l}, X^{0,x}_{im})\Delta B_{im+l}\right)\right|d\tau d\theta \\
=:&I_{11} + I_{12}.
\end{aligned}
$$

Denote $X_{i+1}(k,\theta) = X^{i+1,\theta X^{0,x}_{(i+1)m}+(1-\theta)X^{i,X^{0,x}_{im}}(i+1)}(k)$. The chain rule of the Fréchet derivative leads to

$$D(\phi\circ X)_{i+1}(k,\theta) = D\phi(X_{i+1}(k,\theta))\cdot DX_{i+1}(k,\theta).$$

From $\phi\in C_b^1$, Hölder inequality and Lemma 4.1, it follows

$$\begin{aligned}
I_{11} \le& C\delta^2\int_0^1\int_0^1\left\|\mathbb{E}\left(DX_{i+1}(k,\theta)\cdot\frac{\partial f}{\partial x}(X_{im+l+\tau},X^{0,x}_{im})f(X_{im+l+1},X^{0,x}_{im})\right)\right\|d\tau d\theta\\
\le& C\delta^2\int_0^1\int_0^1\left(\mathbb{E}\left\|DX_{i+1}(k,\theta)\cdot\frac{\partial f}{\partial x}(X_{im+l+\tau},X^{0,x}_{im})f(X_{im+l+1},X^{0,x}_{im})\right\|^2\right)^{\frac12}d\tau d\theta\\
\le& Ce^{-\frac12\nu_1(k-i-1)}\delta^2\int_0^1\left(\mathbb{E}\left\|\frac{\partial f}{\partial x}(X_{im+l+\tau},X^{0,x}_{im})f(X_{im+l+1},X^{0,x}_{im})\right\|^2\right)^{\frac12}d\tau\\
\le& Ce^{-\frac12\nu_1(k-i-1)}\delta^2\int_0^1\Bigg(\mathbb{E}\bigg(2K(1+\|X_{im+l+\tau}\|^q)\|X_{im+l+\tau}\|^2\left\|f(X_{im+l+1},X^{0,x}_{im})\right\|^2\\
&+4K\|X^{0,x}_{im}\|^2\left\|f(X_{im+l+1},X^{0,x}_{im})\right\|^2+4\left\|\frac{\partial f}{\partial x}(0,0)f(X_{im+l+1},X^{0,x}_{im})\right\|^2\bigg)\Bigg)^{\frac12}d\tau.
\end{aligned}$$

By the L^{2p} $(p\ge 4)$ uniform boundedness of the numerical solution and Assumptions 4.3-4.4,

$$I_{11}\le Ce^{-\frac12\nu_1(k-i-1)}\delta^2.$$

For I_{12}, the duality formula of Malliavin derivative [18, P. 43] leads to

$$\begin{aligned}
I_{12} =& \delta\int_0^1\int_0^1\left\|\mathbb{E}\int_{t_{im+l}}^{t_{im+l+1}}\mathscr{D}_uD(\phi\circ X)_{i+1}(k,\theta)\cdot\frac{\partial f}{\partial x}(X_{im+l+\tau},X^{0,x}_{im})g(X_{im+l},X^{0,x}_{im})du\right\|d\tau d\theta\\
\le& \delta\int_0^1\int_0^1\int_{t_{im+l}}^{t_{im+l+1}}\left\|\mathbb{E}\left(\mathscr{D}_uD(\phi\circ X)_{i+1}(k,\theta)\cdot\frac{\partial f}{\partial x}(X_{im+l+\tau},X^{0,x}_{im})g(X_{im+l},X^{0,x}_{im})\right)\right\|dud\tau d\theta.
\end{aligned}$$

Then by the chain rule for Fréchet derivatives, the chain rule and the product rule for Malliavin derivatives [18, P. 37], we obtain

$$\begin{aligned}
\mathscr{D}_uD(\phi\circ X)_{i+1}(k,\theta)\xi =& \mathscr{D}_u\left(D\phi(X_{i+1}(k,\theta))\cdot DX_{i+1}(k,\theta)\xi\right)\\
=& (\mathscr{D}_uX_{i+1}(k,\theta))^\top\cdot D^2\phi(X_{i+1}(k,\theta))DX_{i+1}(k,\theta)\xi\\
&+D\phi(X_{i+1}(k,\theta))\cdot\mathscr{D}_uDX_{i+1}(k,\theta)\xi.
\end{aligned}$$

Thus, Lemmas 4.1-4.6, Assumptions 2.1, 4.3 and $\phi\in C_b^2$ lead to

$$\begin{aligned}
&I_{12}\\
\le& \delta\int_0^1\int_0^1\int_{t_{im+l}}^{t_{im+l+1}}\bigg\|\mathbb{E}\bigg((\mathscr{D}_uX_{i+1}(k,\theta))^\top\cdot D^2\phi(X_{i+1}(k,\theta))\cdot DX_{i+1}(k,\theta)\times\\
&\qquad\frac{\partial f}{\partial x}(X_{im+l+\tau},X^{0,x}_{im})g(X_{im+l},X^{0,x}_{im})\bigg)\bigg\|dud\tau d\theta\\
&+\delta\int_0^1\int_0^1\int_{t_{im+l}}^{t_{im+l+1}}\bigg\|\mathbb{E}\bigg(D\phi(X_{i+1}(k,\theta))\cdot\mathscr{D}_uDX_{i+1}(k,\theta)\times\\
&\qquad\frac{\partial f}{\partial x}(X_{im+l+\tau},X^{0,x}_{im})g(X_{im+l},X^{0,x}_{im})\bigg)\bigg\|dud\tau d\theta\\
\le& C\delta\int_0^1\int_0^1\int_{t_{im+l}}^{t_{im+l+1}}\left\|\mathbb{E}\left(\mathscr{D}_uX_{i+1}(k,\theta)^\top\cdot DX_{i+1}(k,\theta)\cdot\frac{\partial f}{\partial x}(X_{im+l+\tau},X^{0,x}_{im})g(X_{im+l},X^{0,x}_{im})\right)\right\|dud\tau d\theta\\
&+C\delta\int_0^1\int_0^1\int_{t_{im+l}}^{t_{im+l+1}}\left\|\mathbb{E}\left(\mathscr{D}_uDX_{i+1}(k,\theta)\cdot\frac{\partial f}{\partial x}(X_{im+l+\tau},X^{0,x}_{im})g(X_{im+l},X^{0,x}_{im})\right)\right\|dud\tau d\theta\\
\le& C\delta\int_0^1\int_0^1\int_{t_{im+l}}^{t_{im+l+1}}\left(\mathbb{E}\|\mathscr{D}_uX_{i+1}(k,\theta)\|^2\right)^{\frac12}\times
\end{aligned}$$

$$\left(\mathbb{E}\left\|DX_{i+1}(k,\theta)\cdot\frac{\partial f}{\partial x}(X_{im+l+\tau},X_{im}^{0,x})g(X_{im+l},X_{im}^{0,x})\right\|^2\right)^{\frac{1}{2}}dud\tau d\theta$$

$$\begin{aligned}&+C\delta\int_0^1\int_0^1\int_{t_{im+l}}^{t_{im+l+1}}\left(\mathbb{E}\left\|\mathcal{D}_uDX_{i+1}(k,\theta)\cdot\frac{\partial f}{\partial x}(X_{im+l+\tau},X_{im}^{0,x})g(X_{im+l},X_{im}^{0,x})\right\|^2\right)^{\frac{1}{2}}dud\tau d\theta\\ &\le Ce^{-\frac{1}{2}(\nu_1+\nu_2)(k-i-1)}\delta^2+\left(Ce^{-\frac{1}{2}\nu_3(k-i-1)}+Ce^{-\frac{1}{2p}\nu_1(k-i-1)}\right)\delta^2\\ &\le Ce^{-\frac{1}{2}\nu_4(k-i-1)}\delta^2,\end{aligned}$$

where $\nu_4=\min\{\nu_3,\frac{1}{p}\nu_1\}$ and $u\le i+1$ are used.

Next, we estimate I_2. Itô's formula implies

$$\begin{aligned}I_2\le&\left|\mathbb{E}\int_0^1\int_{t_{im+l}}^{t_{im+l+1}}D(\phi\circ X)_{i+1}(k,\theta)\cdot\left(f(X^{i,X_{im}^{0,x}}(s),X_{im}^{0,x})-f(X^{i,X_{im}^{0,x}}(t_{im+l}),X_{im}^{0,x})\right)dsd\theta\right|\\ &+\delta\left|\mathbb{E}\int_0^1D(\phi\circ X)_{i+1}(k,\theta)\cdot\left(f(X^{i,X_{im}^{0,x}}(t_{im+l}),X_{im}^{0,x})-f(X_{im+l},X_{im}^{0,x})\right)d\theta\right|\\ \le&\left|\mathbb{E}\int_0^1\int_{t_{im+l}}^{t_{im+l+1}}D(\phi\circ X)_{i+1}(k,\theta)\cdot\int_{t_{im+l}}^s\frac{\partial f}{\partial x}(X^{i,X_{im}^{0,x}}(u),X_{im}^{0,x})f(X^{i,X_{im}^{0,x}}(u),X_{im}^{0,x})dudsd\theta\right|\\ &+\frac{1}{2}\left|\mathbb{E}\int_0^1\int_{t_{im+l}}^{t_{im+l+1}}D(\phi\circ X)_{i+1}(k,\theta)\cdot\int_{t_{im+l}}^s\frac{\partial^2 f}{\partial x^2}(X^{i,X_{im}^{0,x}}(u),X_{im}^{0,x})\right.\\ &\qquad\left.(g(X^{i,X_{im}^{0,x}}(u),X_{im}^{0,x}),g(X^{i,X_{im}^{0,x}}(u),X_{im}^{0,x}))dudsd\theta\right|\\ &+\left|\mathbb{E}\int_0^1\int_{t_{im+l}}^{t_{im+l+1}}D(\phi\circ X)_{i+1}(k,\theta)\cdot\int_{t_{im+l}}^s\frac{\partial f}{\partial x}(X^{i,X_{im}^{0,x}}(u),X_{im}^{0,x})g(X^{i,X_{im}^{0,x}}(u),X_{im}^{0,x})dB(u)dsd\theta\right|\\ &+\delta\left|\mathbb{E}\int_0^1D(\phi\circ X)_{i+1}(k,\theta)\cdot\left(f(X^{i,X_{im}^{0,x}}(t_{im+l}),X_{im}^{0,x})-f(X_{im+l},X_{im}^{0,x})\right)d\theta\right|\\ =:&\,I_{21}+I_{22}+I_{23}+I_{24}.\end{aligned}$$

By Assumptions 2.1 and 4.3, the L^{2p} $(p\ge4)$ uniform boundedness of $X^{i,\eta}(t)$ and the chain rule for Fréchet derivatives, Hölder's inequality and Lemma 4.1 yield

$$\begin{aligned}I_{21}\le&\int_0^1\int_{t_{im+l}}^{t_{im+l+1}}\int_{t_{im+l}}^s\left(\mathbb{E}\left|D(\phi\circ X)_{i+1}(k,\theta)\cdot\frac{\partial f}{\partial x}(X^{i,X_{im}^{0,x}}(u),X_{im}^{0,x})f(X^{i,X_{im}^{0,x}}(u),X_{im}^{0,x})\right|^2\right)^{\frac{1}{2}}dudsd\theta\\ \le&\,Ce^{-\frac{1}{2}\nu_1(k-i-1)}\delta^2\end{aligned}$$

and

$$\begin{aligned}I_{22}\le&\frac{1}{2}\int_0^1\int_{t_{im+l}}^{t_{im+l+1}}\int_{t_{im+l}}^s\left(\mathbb{E}\left\|D(\phi\circ X)_{i+1}(k,\theta)\cdot\frac{\partial^2 f}{\partial x^2}(X^{i,X_{im}^{0,x}}(u),X_{im}^{0,x})\right.\right.\\ &\qquad\left.\left.(g(X^{i,X_{im}^{0,x}}(u),X_{im}^{0,x}),g(X^{i,X_{im}^{0,x}}(u),X_{im}^{0,x}))\right\|^2\right)^{\frac{1}{2}}dudsd\theta\\ \le&\,Ce^{-\frac{1}{2}\nu_1(k-i-1)}\delta^2.\end{aligned}$$

The estimates of I_{23} and I_{24} are similar to that of I_{12},

$$I_{23}\le Ce^{-\frac{1}{2}\nu_4(k-i-1)}\delta^2\quad\text{and}\quad I_{24}\le Ce^{-\nu_5(k-i-1)}\delta^2,$$

where $\nu_5>0$ is a constant. The estimate of I_3 is as follows. Itô's formula leads to

$$\begin{aligned}I_3\le&\left|\mathbb{E}\int_0^1D(\phi\circ X)_{i+1}(k,\theta)\cdot\int_{t_{im+l}}^{t_{im+l+1}}\int_{t_{im+l}}^s\frac{\partial g}{\partial x}(X^{i,X_{im}^{0,x}}(u),X_{im}^{0,x})f(X^{i,X_{im}^{0,x}}(u),X_{im}^{0,x})dudB(s)d\theta\right|\\ &+\frac{1}{2}\left|\mathbb{E}\int_0^1D(\phi\circ X)_{i+1}(k,\theta)\cdot\int_{t_{im+l}}^{t_{im+l+1}}\int_{t_{im+l}}^s\frac{\partial^2 g}{\partial x^2}(X^{i,X_{im}^{0,x}}(u),X_{im}^{0,x})\right.\\ &\qquad\left.(g(X^{i,X_{im}^{0,x}}(u),X_{im}^{0,x}),g(X^{i,X_{im}^{0,x}}(u),X_{im}^{0,x}))dudB(s)d\theta\right|\\ &+\left|\mathbb{E}\int_0^1D(\phi\circ X)_{i+1}(k,\theta)\cdot\int_{t_{im+l}}^{t_{im+l+1}}\int_{t_{im+l}}^s\frac{\partial g}{\partial x}(X^{i,X_{im}^{0,x}}(u),X_{im}^{0,x})g(X^{i,X_{im}^{0,x}}(u),X_{im}^{0,x})dB(u)dB(s)d\theta\right|\\ =:&\,I_{31}+I_{32}+I_{33}.\end{aligned}$$

The estimates of I_{31} and I_{32} are similar to that of I_{12},

$$\begin{aligned}I_{31} =&\left|\int_0^1\int_{t_{im+l}}^{t_{im+l+1}}\mathbb{E}\left(\mathcal{D}_sD(\phi\circ X)_{i+1}(k,\theta)\cdot\int_{t_{im+l}}^{s}\frac{\partial g}{\partial x}(X^{i,X^{0,x}_{im}}(u),X^{0,x}_{im})f(X^{i,X^{0,x}_{im}}(v),X^{0,x}_{im})du\right)dsd\theta\right|\\ \le&\int_0^1\int_{t_{im+l}}^{t_{im+l+1}}\int_{t_{im+l}}^{s}\mathbb{E}\left|\mathcal{D}_sD(\phi\circ X)_{i+1}(k,\theta)\cdot\frac{\partial g}{\partial x}(X^{i,X^{0,x}_{im}}(u),X^{0,x}_{im})f(X^{i,X^{0,x}_{im}}(u),X^{0,x}_{im})\right|dudsd\theta\\ \le& Ce^{-\frac{1}{2}\nu_4(k-i-1)}\delta^2\end{aligned}$$

and

$$\begin{aligned}I_{32} =&\frac{1}{2}\left|\int_0^1\int_{t_{im+l}}^{t_{im+l+1}}\mathbb{E}\left(\mathcal{D}_sD(\phi\circ X)_{i+1}(k,\theta)\cdot\int_{t_{im+l}}^{s}\frac{\partial^2 g}{\partial x^2}(X^{i,X^{0,x}_{im}}(u),X^{0,x}_{im})\right.\right.\\ &\left.\left.(g(X^{i,X^{0,x}_{im}}(u),X^{0,x}_{im}),g(X^{i,X^{0,x}_{im}}(u),X^{0,x}_{im}))\right)dudsd\theta\right|\\ \le&\frac{1}{2}\int_0^1\int_{t_{im+l}}^{t_{im+l+1}}\int_{t_{im+l}}^{s}\mathbb{E}\left|\mathcal{D}_sD(\phi\circ X)_{i+1}(k,\theta)\cdot\frac{\partial^2 g}{\partial x^2}(X^{i,X^{0,x}_{im}}(u),X^{0,x}_{im})\right.\\ &\left.(g(X^{i,X^{0,x}_{im}}(u),X^{0,x}_{im}),g(X^{i,X^{0,x}_{im}}(u),X^{0,x}_{im}))\right|dudsd\theta\\ \le& Ce^{-\frac{1}{2}\nu_4(k-i-1)}\delta^2.\end{aligned}$$

For I_{33}, we obtain

$$\begin{aligned}&I_{33}\\ =&\left|\int_0^1\int_{t_{im+l}}^{t_{im+l+1}}\mathbb{E}\left(\mathcal{D}_sD(\phi\circ X)_{i+1}(k,\theta)\cdot\int_{t_{im+l}}^{s}\frac{\partial g}{\partial x}(X^{i,X^{0,x}_{im}}(u),X^{0,x}_{im})g(X^{i,X^{0,x}_{im}}(u),X^{0,x}_{im})dB(u)\right)dsd\theta\right|\\ =&\left|\int_0^1\int_{t_{im+l}}^{t_{im+l+1}}\int_{t_{im+l}}^{s}\mathbb{E}\left(\mathcal{D}_u\mathcal{D}_sD(\phi\circ X)_{i+1}(k,\theta)\cdot\frac{\partial g}{\partial x}(X^{i,X^{0,x}_{im}}(u),X^{0,x}_{im})g(X^{i,X^{0,x}_{im}}(u),X^{0,x}_{im})\right)dudsd\theta\right|\\ \le&\int_0^1\int_{t_{im+l}}^{t_{im+l+1}}\int_{t_{im+l}}^{s}\mathbb{E}\left|\mathcal{D}_u\mathcal{D}_sD(\phi\circ X)_{i+1}(k,\theta)\cdot\frac{\partial g}{\partial x}(X^{i,X^{0,x}_{im}}(u),X^{0,x}_{im})g(X^{i,X^{0,x}_{im}}(u),X^{0,x}_{im})\right|dudsd\theta.\end{aligned}$$

Taking Malliavin derivative $\mathcal{D}_u$ on $\mathcal{D}_sD(\phi\circ X)_{i+1}(k,\theta)\xi$ yields

$$\begin{aligned}&\mathcal{D}_u\mathcal{D}_sD(\phi\circ X)_{i+1}(k,\theta)\xi\\ =&\mathcal{D}_u\Big((\mathcal{D}_sX_{i+1}(k,\theta))^\top\cdot D^2\phi(X_{i+1}(k,\theta))\cdot DX_{i+1}(k,\theta)\xi\\ &\quad+D\phi(X_{i+1}(k,\theta))\cdot\mathcal{D}_sDX_{i+1}(k,\theta)\xi\Big)\\ =&(\mathcal{D}_u\mathcal{D}_sX_{i+1}(k,\theta))^\top\cdot D^2\phi(X_{i+1}(k,\theta))\cdot DX_{i+1}(k,\theta)\xi\\ &+D^3\phi(X_{i+1}(k,\theta))(DX_{i+1}(k,\theta)\xi,\mathcal{D}_sX_{i+1}(k,\theta),\mathcal{D}_uX_{i+1}(k,\theta))\\ &+(\mathcal{D}_sX_{i+1}(k,\theta))^\top\cdot D^2\phi(X_{i+1}(k,\theta)\cdot\mathcal{D}_uDX_{i+1}(k,\theta)\xi\\ &+(\mathcal{D}_uX_{i+1}(k,\theta))^\top\cdot D^2\phi(X_{i+1}(k,\theta))\cdot\mathcal{D}_sDX_{i+1}(k,\theta)\xi\\ &+D\phi(X_{i+1}(k,\theta))\cdot\mathcal{D}_u\mathcal{D}_sDX_{i+1}(k,\theta)\xi.\end{aligned}$$

By the estimates of Lemmas 4.1-4.7, there exists $\nu_6>0$ such that

$$I_{33}\le Ce^{-\nu_6(k-i-1)}\delta^2.$$

Combining the estimates of I_1, I_2 and I_3, we conclude that there exists $\nu>0$ such that

$$I_1+I_2+I_3\le Ce^{-\nu(k-i-1)}\delta^2,$$

which implies that

$$\sum_{i=0}^{k-1}\sum_{l=0}^{m-1}(I_1+I_2+I_3)\le C\delta\sum_{i=0}^{k-1}e^{-\nu(k-i-1)}=C\delta\frac{e^\nu-e^{-\nu(k-1)}}{e^\nu-1}\le C\delta.$$

Here C is independent of k and $m\delta = 1$ is used. The proof is completed. □

Now, the uniform weak convergence of the BE method implies the weak error between invariant measures.

Theorem 4.9. *Let conditions in Theorem 4.8 hold. Then there exists a positive constant* $C := C(\|\phi\|_3, \lambda_1, \lambda_2, \lambda_3, p, q, K)$ *independent of* δ *such that*

$$\left| \int_{\mathbb{R}^d} \phi(x)\pi(dx) - \int_{\mathbb{R}^d} \phi(x)\pi^{\delta}(dx) \right| \leq C\delta, \quad \forall\ \phi \in C_b^3$$

for any $\delta \in (0, \delta_1)$ *with* $\delta_1 > 0$ *sufficiently small.*

5. Numerical simulations.

In this section, we present three examples to verify the theoretical results.

Example 1. Consider the following 1-dimensional equation with additive noise

$$\begin{cases} dX(t) = (-\theta_1 X(t) + \theta_2 X([t]))dt + dB(t) \\ X(0) = x, \end{cases} \tag{52}$$

where $\theta_1 > 0$ and $\theta_2 \in \mathbb{R}$. If $t \in [k, k+1)$, $k \in \mathbb{N}$, then the solution of (52) is

$$X(t) = X(k)\left(e^{-\theta_1(t-k)} + \frac{\theta_2}{\theta_1}\left(1 - e^{-\theta_1(t-k)}\right)\right) + \int_k^t e^{-\theta_1(t-s)} dB(s). \tag{53}$$

It can be seen that the solution obeys Gaussian distribution. And the expectation of the solution is

$$\mathbb{E}X(t) = x\left(\frac{\theta_2}{\theta_1} + \left(1 - \frac{\theta_2}{\theta_1}\right)e^{-\theta_1}\right)^k \left(\frac{\theta_2}{\theta_1} + \left(1 - \frac{\theta_2}{\theta_1}\right)e^{-\theta_1(t-k)}\right). \tag{54}$$

Denote $\mu : [0, 1) \to (-\infty, \infty)$ and $\sigma : [0, 1) \to (0, 1]$ by $\mu(\{t\}) = \frac{\theta_2}{\theta_1} + \left(1 - \frac{\theta_2}{\theta_1}\right)e^{-\theta_1\{t\}}$ and $\sigma(\{t\}) = \frac{1}{2\theta_1}\left(1 - e^{-2\theta_1\{t\}}\right)$. Defining $\mu(1) := \lim_{t\to k^-} \mu(\{t\})$ and $\sigma(1) := \lim_{t\to k^-} \sigma(\{t\})$, we have $\mu(1) = \frac{\theta_2}{\theta_1} + \left(1 - \frac{\theta_2}{\theta_1}\right)e^{-\theta_1}$ and $\sigma(1) = \frac{1}{2\theta_1}\left(1 - e^{-2\theta_1}\right)$. Therefore the variance of the solution is

$$\mathrm{Var}(X(t)) = \left(\frac{1 - \mu(1)^{2k}}{1 - \mu(1)^2}\sigma(1)\right)\mu(\{t\})^2 + \sigma(\{t\}). \tag{55}$$

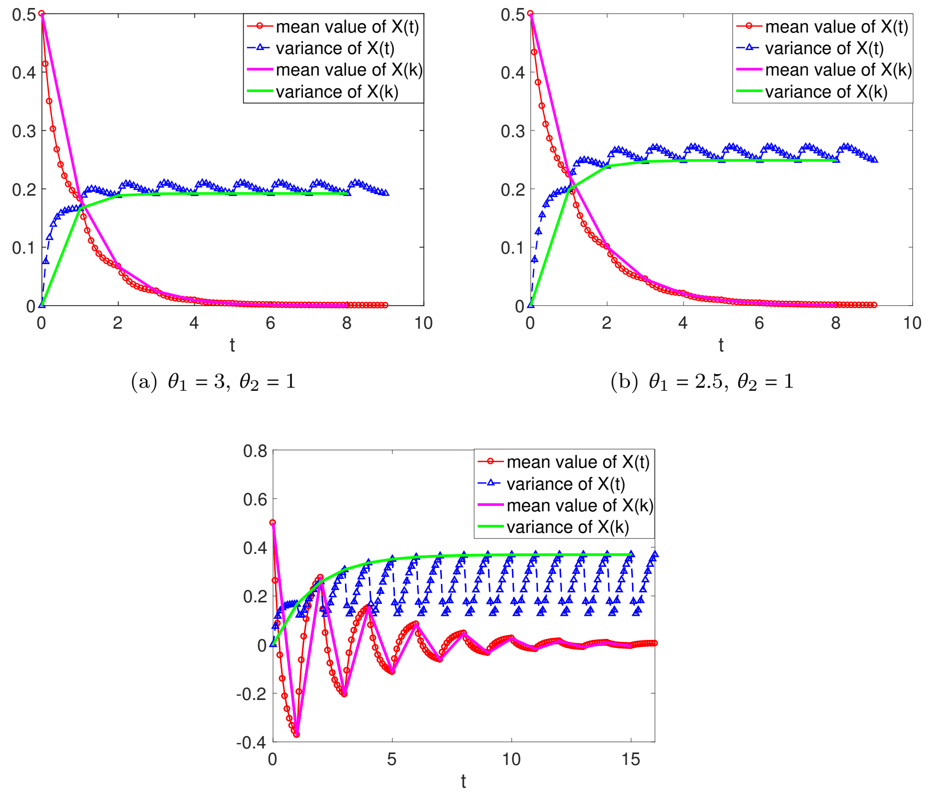
Especially, the expectation and variance of $X(t)$ at the integral time $t = k$ are, respectively,

$$\mathbb{E}X(k) = x\mu(1)^k \text{ and } \mathrm{Var}(X(k)) = \frac{1 - \mu(1)^{2k}}{1 - \mu(1)^2}\sigma(1). \tag{56}$$

The sufficient and necessary condition under which $X(k)$ may admit a stationary distribution is

$$|\mu(1)| < 1 \Leftrightarrow -\frac{1 + e^{-\theta_1}}{1 - e^{-\theta_1}}\theta_1 < \theta_2 < \theta_1.$$

Firstly, we verify that the solution $\{X(t)\}_{t\geq 0}$ does not admit a stationary distribution while the chain $\{X(k)\}_{k\in\mathbb{N}}$ does. Let the initial value $x = 0.5$. Fig. 1 shows the expectations and variances of both $X(t)$ and $X(k)$ with three different parameters which satisfy $|\mu(1)| < 1$. It can be seen that the variances of the solution $\{X(t)\}_{t\geq 0}$ are not convergent as t tends to infinity, though the expectations of $\{X(t)\}_{t\geq 0}$ converge to zero. However, both the expectations and variances of the chain $\{X(k)\}_{k\in\mathbb{N}}$ converge as k goes to infinity. This means that the chain $\{X(k)\}_{k\in\mathbb{N}}$ admits a stationary Gaussian distribution. Comparing Fig. 1 (a) with (b), we



(a) $\theta_1 = 3,\ \theta_2 = 1$

(b) $\theta_1 = 2.5,\ \theta_2 = 1$

(c) $\theta_1 = 3,\ \theta_2 = -2.5$

FIGURE 1. The expectations and variances of $X(t)$ and $X(k)$

observe that the distribution of $\{X(k)\}_{k\in\mathbb{N}}$ converges more rapidly for larger θ_1, which implies that the convergence rate increases as the dissipativity increases.

Fig. 2 shows the evolution of the expectations and variances of $X(t)$ for several different θ_2. We take the parameter $\theta_1 = 3$. The pink lines are the expectation and the variance of $X(t)$ with $\theta_2 = 0$ (i.e., the expectation and the variance of the corresponding OU process) in Fig. 2(a)(c) and Fig. 2(b)(d), respectively. We can observe that the expectation and the variance of $X(t)$ with sufficiently small θ_2 are close to that of the corresponding OU process.

Next the weak convergence order of the BE method is tested. In fact, the solution of (52) can be expressed as

$$X(t) = x\mu(1)^k\mu(\{t\}) + \sum_{i=1}^{k}\mu(\{t\})\mu(1)^i\int_{i-1}^{i}e^{-\theta_1(i-s)}dB(s) + \int_k^t e^{-\theta_1(t-s)}dB(s).$$

Let $T = 5$. We create 1000 discretized Brownian paths over $[0,T]$ with a small step-size $\bar{\delta} = 2^{-11}$ and approximate the stochastic integral in the exact solution above using the Euler method with this small step-size. We also compute the numerical solutions of the BE method using 4 different step-sizes $\delta = 2^{-6}, 2^{-7}, 2^{-8}, 2^{-9}$ on the same Brownian path at $T = 5$. Moreover, we choose 4 different functions

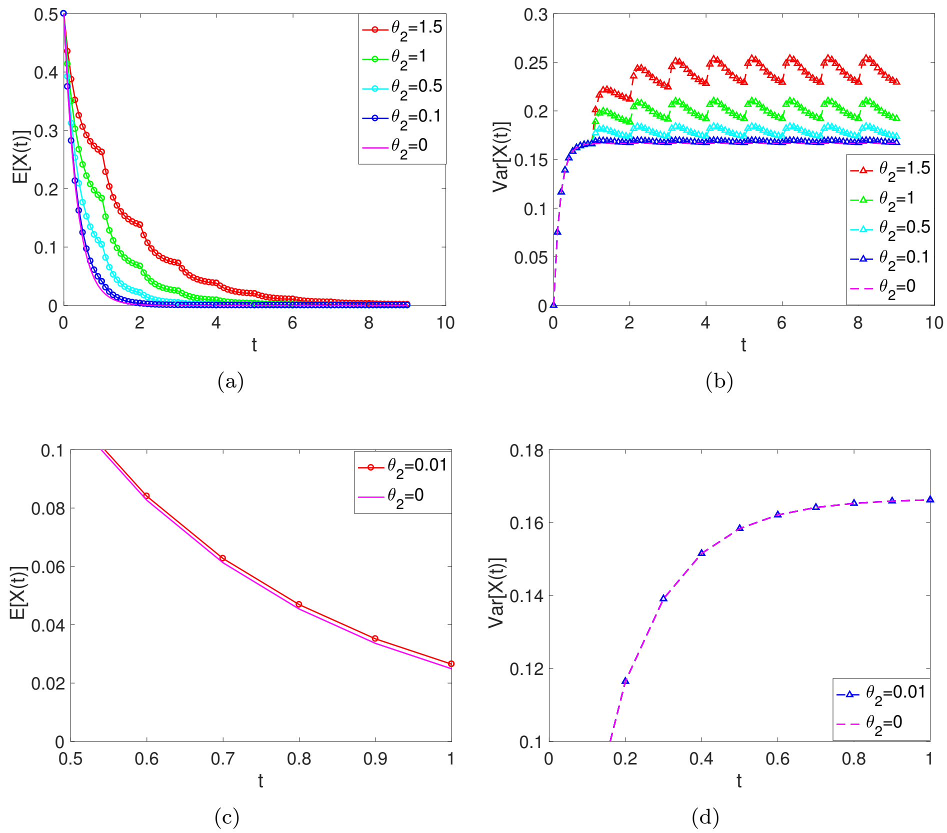
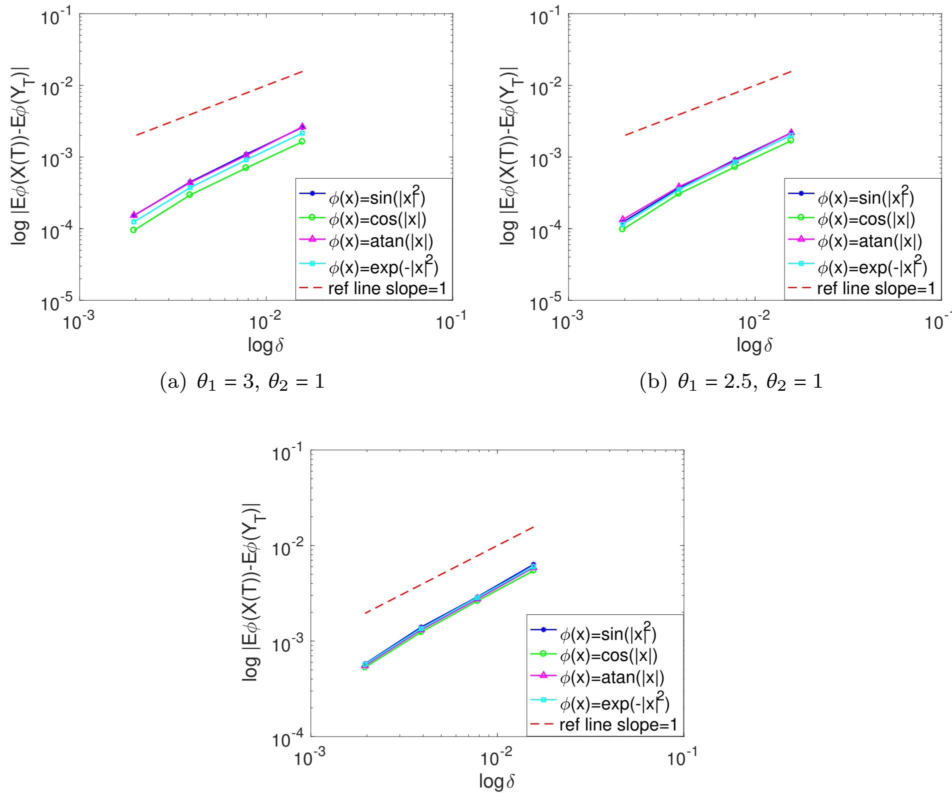


FIGURE 2. The expectations (left column) and variances (right column) of $X(t)$ with different θ_2

$\phi(x) = \sin(|x|^2)$, $\phi(x) = \cos(|x|)$, $\phi(x) = \arctan(|x|)$ and $\phi(x) = e^{-|x|^2}$ as the test functions for weak convergence. Fig. 3 plots the weak errors $|\mathbb{E}\phi(X(T)) - \mathbb{E}\phi(Y_T)|$ against δ on a log-log scale, where $X(T)$ and Y_T denote the exact and numerical solutions at the endpoint T, respectively. The red dashed line represents a reference line with slope 1. From Fig. 3, it is observed that the BE method is convergent with weak order 1.

Then we consider the longtime behavior of the Markov chain $\{Y_k\}_{k\in\mathbb{N}}$. Theorem 3.15 shows that $\mathbb{E}\phi(Y_k^{0,x})$ converges exponentially to the "spatial" average of ϕ with different initial data, i.e., Y_k is strongly mixing, and this implies the ergodicity of Y_k. In this test, we let $\theta_1 = 3$ and $\theta_2 = 1$ and choose three test functions (a) $\phi(x) = \arctan(|x|)$, (b) $\phi(x) = \cos(|x|)$ and (c) $\phi(x) = \sin(|x|^2)$ to compute $\mathbb{E}\phi(Y_k^{0,x})$. Fig. 4 shows the mean value of $\phi(Y_k^{0,x})$ with 5 different initial data. As can be seen from the figure, for each ϕ, $\mathbb{E}\phi(Y_k^{0,x})$ converges exponentially to the spatial average of ϕ with respect to the invariant measure.



(a) $\theta_1 = 3,\ \theta_2 = 1$

(b) $\theta_1 = 2.5,\ \theta_2 = 1$

(c) $\theta_1 = 3,\ \theta_2 = -2.5$

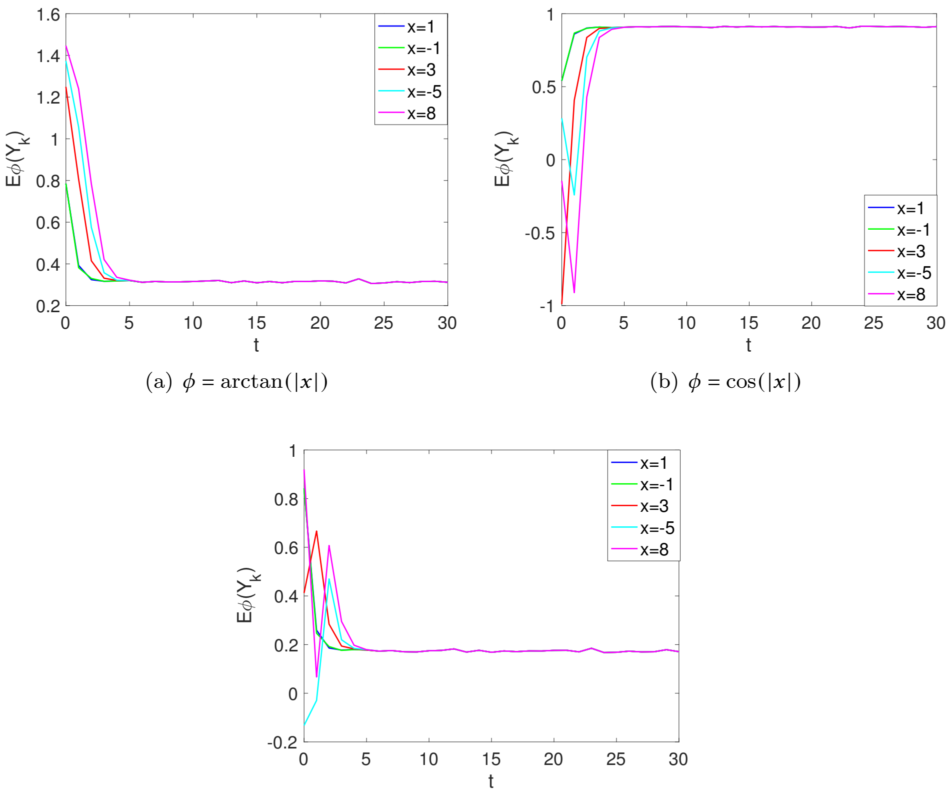
Figure 3. Order of weak convergence of BE method

Example 2. Consider the following 1-dimensional nonlinear SDE with PCAs driven by multiplicative noise

$$\begin{cases} dX(t) = (-X(t)^3 - 10X(t) + 2X([t]) + 1)dt + (aX(t) + bX([t]))dB(t) \\ X_0 = x, \end{cases} \tag{57}$$

where $x = 2$ and a, b are two parameters. Firstly, we verify the weak convergence of the BE method on a finite time interval $[0, T]$. Let $T = 6$ and we create 2000 discretized Brownian paths over $[0, T]$ with a small step-size $\bar{\delta} = 2^{-11}$. Since the exact solution can not be obtained, we use the numerical solution of the split-step backward Euler method with $\bar{\delta} = 2^{-11}$ as the "exact solution". We also compute the numerical solutions of the BE method using 4 different step-sizes $\delta = 2^{-6}, 2^{-7}, 2^{-8}, 2^{-9}$ on the same Brownian path. Let $X(T)$ and Y_T denote the exact and numerical solutions at the endpoint T, respectively. And three sets of a, b are tested. Fig. 5 plots the weak errors $|\mathbb{E}\phi(X(T)) - \mathbb{E}\phi(Y_T)|$ against δ on a log-log scale with 4 different kinds of test functions $\phi(x) = \sin(|x|^2 + \pi/2)$, $\phi(x) = \cos(|x|)$, $\phi(x) = \arctan(|x|^2)$ and $\phi(x) = e^{-|x|^2}$. The red dashed line represents a reference line with slope 1. As can be observed from Fig. 5, the BE method converges in the weak sense with order 1.

Finally the longtime behavior of the Markov chain $\{Y_k\}_{k\in\mathbb{N}}$ is considered. In this simulation, we take $a = 1$ and $b = 1$ for example and choose three test functions (a)



(a) $\phi = \arctan(|x|)$ (b) $\phi = \cos(|x|)$

(c) $\phi = \sin(|x|^2)$

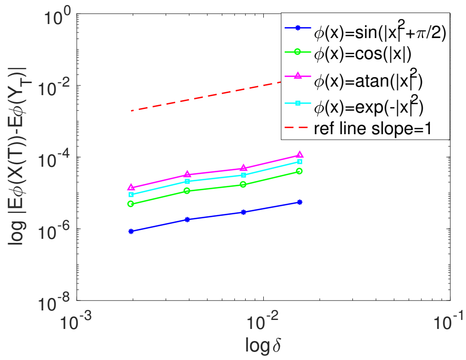
FIGURE 4. The evolution of $\mathbb{E}\phi(Y_k)$ started from different initial data

$\phi(x) = \arctan(|x|)$, (b) $\phi(x) = \sin(|x|^2)$ and (c) $\phi(x) = e^{-|x|^2}$. Fig. 8 plots the mean value of $\phi(Y_k^{0,x})$ with 5 different initial data. It is observed that, for each ϕ, $\mathbb{E}\phi(Y_k^{0,x})$ is exponentially convergent as k tends to infinity, which verifies theoretical results.

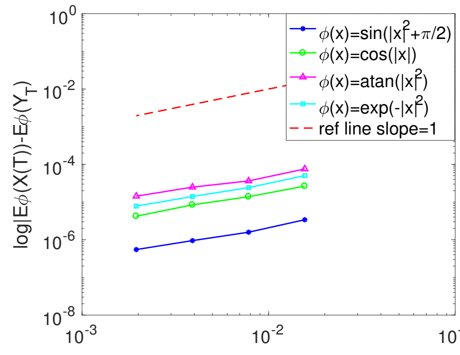
Example 3. We are also interested in the following SDE with PCAs whose coefficients do not satisfy the assumptions of our main theorems

$$\begin{cases} dX(t) = (-X(t)^3 - 8X(t) + aX([t])^2)dt + b\big|X([t])\big|^{1.1} dB(t) \\ X_0 = x. \end{cases} \tag{58}$$

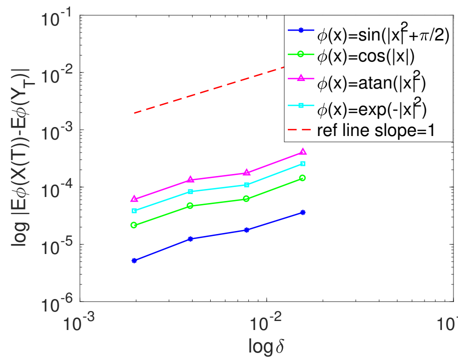
Let $x = 5$, $T = 5$ and the discretized Brownian paths be 1000. Since the exact solution can not be obtained, we use the numerical solution of the split-step backward Euler method with $\bar{\delta} = 2^{-11}$ as the "exact solution". We also compute the numerical solutions of the BE method using 4 different step-sizes $\delta = 2^{-6}, 2^{-7}, 2^{-8}, 2^{-9}$ on the same Brownian path. Choosing the test functions $\phi(x) = \sin(|x|)$, $\phi(x) = \cos(|x|)$, $\phi(x) = \arctan(|x|^2)$ and $\phi(x) = e^{-|x|^2}$, we plot the weak errors $|\mathbb{E}\phi(X(T)) - \mathbb{E}\phi(Y_T)|$ against δ on a log-log scale in Fig. 7. It seems that the weak convergence order of the BE method is slightly smaller than 1. Taking $a = -1$, $b = -1$ and $T = 20$ and choosing respectively three test functions (a) $\phi(x) = \arctan(|x|)$, (b) $\phi(x) = \sin(|x|)$ and (c) $\phi(x) = e^{-|x|^2}$, we plot the mean value of $\phi(Y_k^{0,x})$ with 5 different initial data



(a) $a = 1,\ b = 0$



(b) $a = 0,\ b = 1$



(c) $a = 1,\ b = 1$

FIGURE 5. Order of weak convergence of the BE method

in Fig. 8. We observe that $\mathbb{E}\phi(Y_k^{0,x})$ converges exponentially as k tends to infinity. These numerical simulations show the coincidence with our theoretical results. We may say that our theory suits for a broader class of SDEs with PCAs than we claimed, and the study for the optimal assumptions is left for the future work.

Acknowledgments. The authors thank the referees and the editor for their valuable comments and helpful suggestions.

REFERENCES

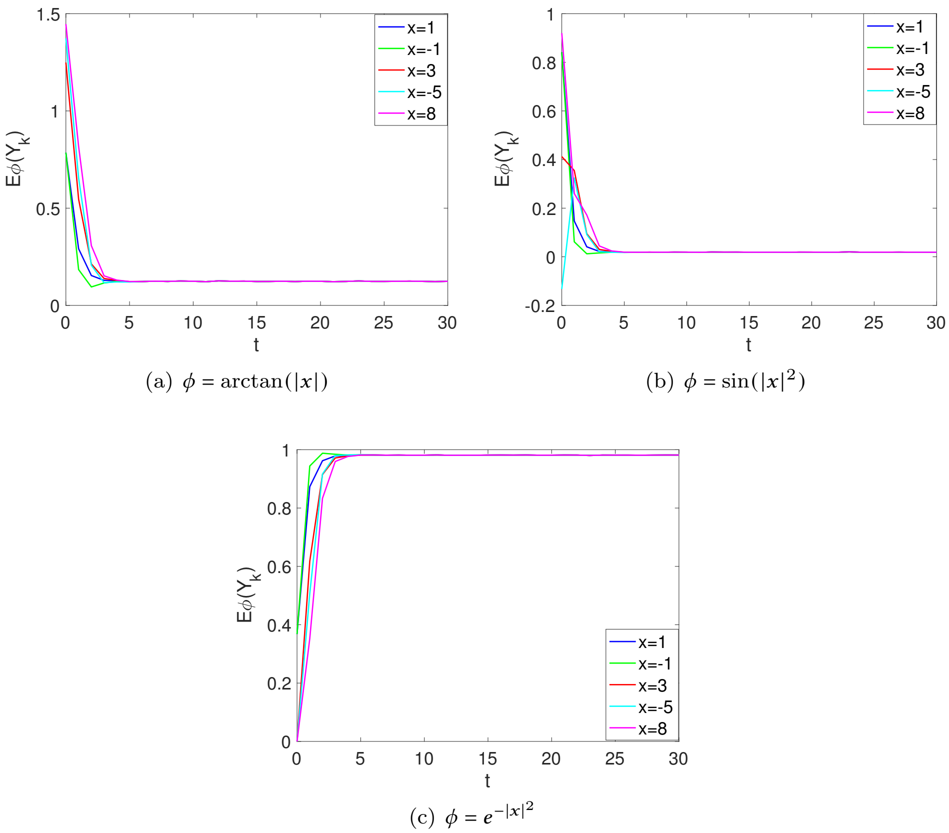
[1] A. Abdulle, G. Vilmart and K. C. Zygalakis, High order numerical approximation of the invariant measure of ergodic SDEs, *SIAM J. Numer. Anal.*, **52** (2014), 1600–1622.

[2] J. Bao, G. Yin and C. Yuan, Ergodicity for functional stochastic differential equations and applications, *Nonlinear Anal.*, **98** (2014), 66–82.

[3] C.-E. Bréhier, Approximation of the invariant measure with an Euler scheme for stochastic PDEs driven by space-time white noise, *Potential Anal.*, **40** (2014), 1–40.

[4] C. Chen, J. Hong and X. Wang, Approximation of invariant measure for damped stochastic nonlinear Schrödinger equation via an ergodic numerical scheme, *Potential Anal.*, **46** (2017), 323–367.

[5] J. Cui, J. Hong and L. Sun, Weak convergence and invariant measure of a full discretization for non-globally Lipschitz parabolic SPDE, *Stochastic Process. Appl.*, **134** (2021), 55–93.

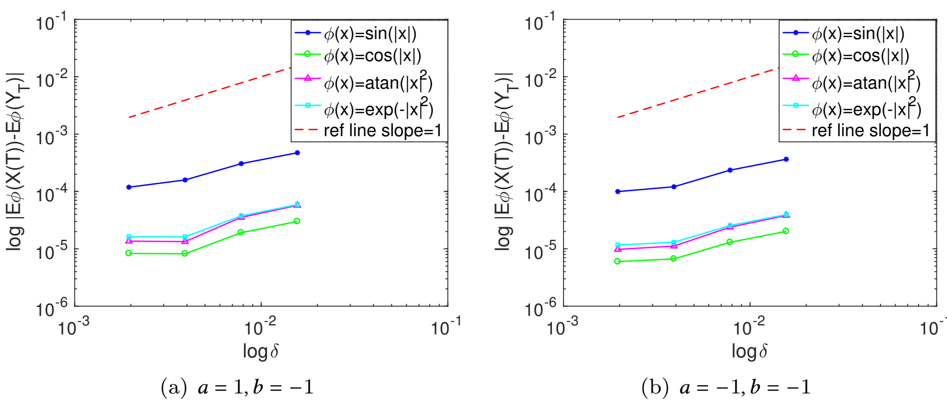


(a) $\phi = \arctan(|x|)$

(b) $\phi = \sin(|x|^2)$

(c) $\phi = e^{-|x|^2}$

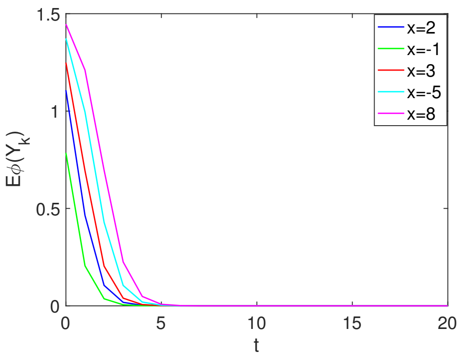
FIGURE 6. The evolution of $\mathbb{E}\phi(Y_k)$ started from different initial data



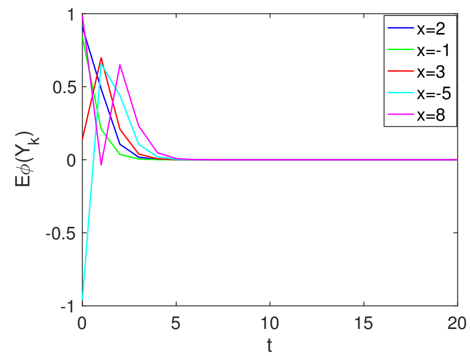
(a) $a = 1, b = -1$

(b) $a = -1, b = -1$

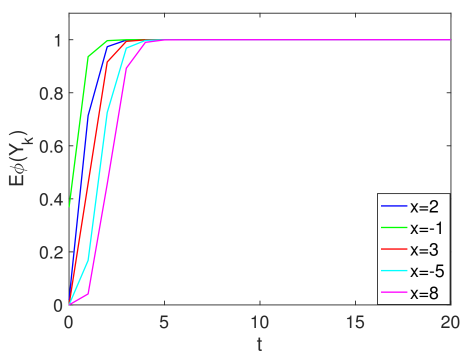
FIGURE 7. Order of weak convergence of the BE method



(a) $\phi = \arctan(|x|)$



(b) $\phi = \sin(|x|)$



(c) $\phi = \exp(-|x|^2)$

FIGURE 8. The evolution of $\mathbb{E}\phi(Y_k)$ started from different initial data

[6] G. Da Prato, *An Introduction to Infinite-Dimensional Analysis*, Universitext, Springer-Verlag, Berlin, 2006.

[7] L. Dai, *Nonlinear Dynamics of Piecewise Constant Systems and Implementation of Piecewise Constant Arguments*, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2008.

[8] D. Gusak, A. Kukush, A. Kulik, Y. Mishura and A. Pilipenko, *Theory of Stochastic Processes: With Applications to Financial Mathematics and Risk Theory*, Springer, New York, 2010.

[9] J. Hong, L. Sun and X. Wang, High order conformal symplectic and ergodic schemes for the stochastic Langevin equation via generating functions, *SIAM J. Numer. Anal.*, **55** (2017), 3006–3029.

[10] J. Hong and X. Wang, *Invariant Measures for Stochastic Nonlinear Schrödinger Equations: Numerical Approximations and Symplectic Structures*, Lecture Notes in Mathematics, 2251, Springer, Singapore, 2019.

[11] K. Itô and M. Nisio, On stationary solutions of a stochastic differential equation, *J. Math. Kyoto Univ.*, **4** (1964), 1–75.

[12] L. Liu, A. Wu, Z. Zeng and T. Huang, Global mean square exponential stability of stochastic neural networks with retarded and advanced argument, *Neurocomputing*, **247** (2017), 156–164.

[13] Y. L. Lu, M. H. Song and M. Z. Liu, Convergence and stability of the split-step theta method for stochastic differential equations with piecewise continuous arguments, *J. Comput. Appl. Math.*, **317** (2017), 55–71.

[14] Y. Lu, M. Song and M. Liu, Convergence rate and stability of the split-step theta method for stochastic differential equations with piecewise continuous arguments, *Discrete Contin. Dyn. Syst. Ser. B*, **24** (2019), 695–717.

[15] X. Mao, Almost sure exponential stabilization by discrete-time stochastic feedback control, *IEEE Trans. Automat. Control*, **61** (2016), 1619–1624.

[16] J. C. Mattingly, A. M. Stuart and M. V. Tretyakov, Convergence of numerical time-averaging and stationary measures via Poisson equations, *SIAM J. Numer. Anal.*, **48** (2010), 552–577.

[17] M. Milošević, The Euler–Maruyama approximation of solutions to stochastic differential equations with piecewise constant arguments, *J. Comput. Appl. Math.*, **298** (2016), 1–12.

[18] D. Nualart, *The Malliavin Calculus and Related Topics*, Springer-Verlag, Berlin, 2006.

[19] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York-London 1970.

[20] I. Ozturk and F. Bozkurt, Stability analysis of a population model with piecewise constant arguments, *Nonlinear Anal. Real World Appl.*, **12** (2011), 1532–1545.

[21] J. Ren, J. Wu and X. Zhang, Exponential ergodicity of non-Lipschitz multivalued stochastic differential equations, *Bull. Sci. Math.*, **134** (2010), 391–404.

[22] D. Talay, Second-order discretization schemes of stochastic differential systems for the computation of the invariant law, *Rapports de Recherche, Institut National de Recherche en Informatique et en Automatique*, (1987), 13–26.

[23] J. Wiener, *Generalized Solutions of Functional Differential Equations*, World Scientific Publishing Co., Inc., River Edge, NJ, 1993.

[24] F. Wu, G. Yin and H. Mei, Stochastic functional differential equations with infinite delay: Existence and uniqueness of solutions, solution maps, Markov properties, and ergodicity, *J. Differential Equations*, **262** (2017), 1226–1252.

[25] Y. Xie and C. Zhang, A class of stochastic one-parameter methods for nonlinear SFDEs with piecewise continuous arguments, *Appl. Numer. Math.*, **135** (2019), 1–14.

Received February 2020; 1st revision July 2021; 2nd revision November 2021; early access June 2022.

E-mail address: chenchuchu@lsec.cc.ac.cn
E-mail address: hjl@lsec.cc.ac.cn
E-mail address: yulanlu@lsec.cc.ac.cn