Journal of Computational Mathematics
Vol.xx, No.x, 2024, 1–27.
http://www.global-sci.org/jcm
doi:10.4208/jcm.2311-m2023-0147

CENTRAL LIMIT THEOREM FOR TEMPORAL AVERAGE OF BACKWARD EULER-MARUYAMA METHOD*

Diancong Jin
School of Mathematics and Statistics,
Hubei Key Laboratory of Engineering Modeling and Scientific Computing,
Huazhong University of Science and Technology, Wuhan 430074, China
Email: jindc@hust.edu.cn

Abstract

This work focuses on the temporal average of the backward Euler-Maruyama (BEM) method, which is used to approximate the ergodic limit of stochastic ordinary differential equations (SODEs). We give the central limit theorem (CLT) of the temporal average of the BEM method, which characterizes its asymptotics in distribution. When the deviation order is smaller than the optimal strong order, we directly derive the CLT of the temporal average through that of original equations and the uniform strong order of the BEM method. For the case that the deviation order equals to the optimal strong order, the CLT is established via the Poisson equation associated with the generator of original equations. Numerical experiments are performed to illustrate the theoretical results. The main contribution of this work is to generalize the existing CLT of the temporal average of numerical methods to that for SODEs with super-linearly growing drift coefficients.

Mathematics subject classification: 60H35, 60F05, 60H10.
Key words: Central limit theorem, Temporal average, Ergodicity, Backward Euler-Maruyama method.

1. Introduction

Ergodic theory is a powerful tool to investigate the long-time dynamics and statistical properties of stochastic systems, which is widely applied in physics, biology and chemistry (see e.g. [8, 13, 30, 33]). A crucial problem in ergodic theory is to determine the ergodic measure and ergodic limit. Since explicit expressions of them are generally unavailable, one usually resorts to numerical methods to obtain their approximations. There have been lots of numerical methods which inherit the ergodicity or approximate the ergodic limit of original systems (see [1, 12, 14, 24, 27, 31] and references therein). In the aforementioned work, main efforts are made to analyze the error between the numerical invariant measure and the original one, and that between numerical temporal average and the ergodic limit.

Besides the convergence of the numerical temporal average in the moment sense, the asymptotics of its distribution is also an essential property. In recent several work, the central limit theorem (CLT) of the temporal average of some numerical methods is given, which characterizes the fluctuation of the numerical temporal average around ergodic limits of original systems in the sense of distribution. In [26], the CLT of the temporal average of the Euler-Maruyama (EM) method with decreasing step-size for ergodic stochastic ordinary differential equations (SODEs) is given. In addition, [23] proves the CLT and moderate deviation of the EM method

* Received July 19, 2023 / Revised version received September 15, 2023 / Accepted November 7, 2023 / Published online March 27, 2024 /

with a fixed step-size for SODEs. For a class of ergodic stochastic partial differential equations (SPDEs), [6] shows that the temporal average of a full discretization with fixed temporal and spatial step-sizes satisfies the CLT.

In the existing work, the CLT of numerical temporal average is established provided that coefficients of original equations are Lipschitz continuous. Compared with the Lipschitz case, stochastic systems with non-Lipschitz coefficients have more extensive applications in reality (see e.g. [3, 7, 9, 11] and references therein). For example, consider the overdamped Langevin equation

$$\mathrm{d}q(t) = -\nabla V\big(q(t)\big)\mathrm{d}t + \sqrt{2\beta^{-1}}\,\mathrm{d}W(t), \tag{1.1}$$

where $\{W(t), t \geq 0\}$ is a D-dimensional standard Brownian motion defined on a complete filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbf{P})$, and β^{-1} is the Boltzmann constant times the absolute temperature. The potential V is smooth and satisfies $\lim_{|q|\to+\infty} V(q) = +\infty$ (for example, one can take $D = 1$ and $V(q) = q^4/4 + q^2/2$, $q \in \mathbb{R}$). The Langevin equation describes the noise-induced transport in stochastic ratchets and dissipative particle dynamics. When the inertia of the particle is negligible compared with the damping force due to the friction, the trajectory of the Langevin equation is approximately described by (1.1) (see e.g. [17, 28, 29]). It is known that (1.1) admits a unique invariant measure (thus is ergodic) $\pi(\mathrm{d}q) = Z^{-1}e^{-\beta V(q)}\mathrm{d}q$ with $Z = \int_{\mathbb{R}} e^{-\beta V(q)}\mathrm{d}q$. Since the drift coefficient of (1.1) is non-Lipschitz, the existing results on the CLT of the numerical temporal average are not applicable to (1.1). In view of the above consideration, we are devoted to investigating the CLT of the numerical temporal average for general SODEs with non-Lipschitz coefficients.

In this work, we consider the following SODE:

$$\mathrm{d}X(t) = b\big(X(t)\big)\mathrm{d}t + \sigma\big(X(t)\big)\mathrm{d}W(t), \quad t > 0, \tag{1.2}$$

where W is the same one defined in (1.1). Here, $b : \mathbb{R}^d \to \mathbb{R}^d$ satisfies the strong dissipation condition and is allowed to grow super-linearly, and $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times D}$ is bounded and Lipschitz continuous (see Section 2 for the detailed assumptions on b and σ). Then, (1.2) admits a unique strong solution on $[0, +\infty)$ for any given deterministic initial value $X(0) \in \mathbb{R}^d$. It is shown in [20, Theorem 3.1] that (1.2) admits a unique invariant measure π and is thus ergodic, due to the strong dissipation condition on b. Our main purpose is to study the CLT of the temporal average of the backward Euler-Maruyama (BEM) method applied to (1.2). The reasons for the choice of the BEM method are as follows:

(1) The CLT of the numerical temporal average characterizes the long-time behavior of numerical solutions. Thus, one preference in the choice of the numerical method is that it should possess the long-time stability. The Euler–Maruyama (EM) method and the BEM method are used most frequently when simulating SODEs (see e.g. [10, p. 453]). Similar to the deterministic case, the BEM method shows a more excellent long-time stability than the EM method.

(2) When applied to SODEs with super-linearly growing coefficients, the EM method is known to diverge [15]. Other explicit numerical methods based on the Itô-Taylor expansion could suffer from the same fate, for SODEs with super-linearly growing coefficients [10]. Thus, as is pointed out by [21], for SODEs with super-linearly growing coefficients, one usually adopts implicit numerical methods or modified explicit numerical methods such as the adaptive time step size method, tamed method and the truncated method.

(3) In our arguments for the CLT, we need to ensure the p-th ($p>2$) moment boundedness in the infinite time horizon and exponential ergodicity. We will show that the BEM method fulfills the above requirements.

The authors in [20] discretize (1.2) by the BEM method (see (3.1)), and give the error between the numerical invariant measure π_τ and π with τ being the step-size. The above result together with the strong convergence of the BEM method in the infinite time horizon implies that the temporal average $\sum_{k=0}^{N-1} h(\bar{X}_k^x)/N$ converges to the ergodic limit

$$\pi(h) := \int_{\mathbb{R}^d} h(x)\pi(\mathrm{d}x), \quad h \in \mathbf{C}_b^1(\mathbb{R}^d)$$

in the sense of

$$\lim_{\tau\to 0}\lim_{N\to+\infty}\left|\frac{1}{N}\sum_{k=0}^{N-1}\mathbf{E}h\big(\bar{X}_k^x\big)-\pi(h)\right|=0, \tag{1.3}$$

where $\{\bar{X}_n^x\}_{n\geq 0}$ is the numerical solution generated by the BEM method with initial value $x\in\mathbb{R}^d$. The main result of this paper is the CLT for the following temporal average:

$$\Pi_{\tau,\alpha}(h)=\frac{1}{\tau^{-\alpha}}\sum_{k=0}^{\tau^{-\alpha}-1} h\big(\bar{X}_k^x\big), \quad \alpha\in(1,2], \quad h\in\mathbf{C}_b^4(\mathbb{R}^d),$$

where for convenience we always assume that $\tau^{-\alpha}$ is an integer in place of the step number N in (1.3). More precisely, we prove in Theorems 3.1 and 3.2 that the normalized temporal average $(\Pi_{\tau,\alpha}(h)-\pi(h))/\tau^{(\alpha-1)/2}$ converges to the normal distribution $\mathcal{N}(0,\pi(|\sigma^\top\nabla\varphi|^2))$ in distribution as $\tau\to 0$ for $\alpha\in(1,2)$ and $\alpha=2$. In fact, Theorem 3.1 indicates that the CLT holds for the temporal average of a class of numerical methods with global strong order $1/2$, for $\alpha\in(1,2)$. Here, φ is defined by (3.2) and solves the Poisson equation $\mathcal{L}\varphi=h-\pi(h)$ (see Lemma 4.2), with $\mathcal{L}$ being the generator of (1.2). We call the parameter $\tau^{(\alpha-1)/2}$ the deviation scale and $(\alpha-1)/2$ the deviation order (see Remark 3.1 for the reason of requiring $\alpha>1$).

The proof ideas of the CLT for $\Pi_{\tau,\alpha}(h)$ are different for $\alpha\in(1,2)$ and $\alpha=2$. In the case $\alpha\in(1,2)$, we directly derive the CLT for $\Pi_{\tau,\alpha}(h)$ in Theorem 3.1 by means of the CLT for (1.2) and the optimal strong order in the infinite time horizon of the BEM method, considering that the CLT for (1.2) is a classical result (see [2, Theorem 2.1]). The key of this proof lies in that the deviation order $(\alpha-1)/2$ is smaller than the optimal strong order $1/2$ for $\alpha\in(1,2)$, which does not apply to the case $\alpha=2$. In order to tackle the more subtle case $\alpha=2$, we follow the arguments in [23] and [6] to obtain the CLT for $\Pi_{\tau,2}(h)$. The main idea is to reformulate the normalized temporal average $(\Pi_{\tau,\alpha}(h)-\pi(h))/\tau^{(\alpha-1)/2}$ by means of the Poisson equation. This allows us to decompose $(\Pi_{\tau,\alpha}(h)-\pi(h))/\tau^{(\alpha-1)/(2)}$ into a martingale difference series sum converging to $\mathcal{N}(0,\pi(|\sigma^\top\nabla\varphi|^2))$ in distribution, and a negligible remainder converging to 0 in probability. In this proof, the p-th ($p>2$) moment boundedness of the BEM method in the infinite time horizon and the regularity of the solution to the Poisson equation are main difficulties and play important roles. On one hand, the p-th ($p>2$) moment boundedness of the BEM method in the infinite time horizon is required due to the super-linear growth of b, which is proved by the fine control for the iteration coefficient. On the other hand, the regularity of the solution to the Poisson equation is given by estimating the mean-square derivatives of the solution to (1.2) with respect to (w.r.t.) initial values. This argument is different from the case of additive noise in [6], where the exact solution is differentiable w.r.t. initial values almost surely.

To sum up, the contributions of this work are twofold. Firstly, we give the CLT for the temporal average of the BEM method, which generalizes the existing results to SODEs with super-linearly growing drift coefficients. Secondly, we prove the p-th ($p > 2$) moment boundedness of the BEM method in the infinite time horizon for the original equation. The rest of this paper is organized as follows. In Section 2, we give our assumptions and recall some basic properties of the exact solution. Section 3 presents our main results and proves the CLT for $\Pi_{\tau,\alpha}(h)$ with $\alpha \in (1, 2)$, and Section 4 gives the proof of the CLT for $\Pi_{\tau,2}(h)$. Some numerical tests are displayed to illustrate the theoretical results in Section 5. Finally, we give the conclusions and refer to future aspects in Section 6.

2. Preliminaries

In this section, we give our main assumptions on the coefficients of (1.2) and present some basic properties for (1.2). We begin with some notations. Denote by $|\cdot|$ the 2-norm of a vector or matrix, and by $\langle \cdot, \cdot \rangle$ the scalar product of two vectors. Let $d, m, k \in \mathbb{N}^+$ with $\mathbb{N}^+$ denoting the set of positive integers. For matrices $A, B \in \mathbb{R}^{d\times m}$, denote

$$\langle A, B\rangle_{HS} := \sum_{i=1}^{d}\sum_{j=1}^{m} A_{ij}B_{ij}, \quad \|A\|_{HS} := \sqrt{\langle A, A\rangle_{HS}}.$$

Let $\mathscr{B}(\mathbb{R}^d)$ stand for the set of all Borel sets of $\mathbb{R}^d$. Denote by $\mathcal{P}(\mathbb{R}^d)$ the space of all probability measures on $\mathbb{R}^d$. Denote

$$\mu(f) = \int_{\mathbb{R}^d} f(x)\mu(\mathrm{d}x)$$

for $\mu \in \mathcal{P}(\mathbb{R}^d)$ and μ-measurable function f. For convenience, we set $\mathcal{F}_t = \sigma(W(s), 0 \le s \le t)$ for $t \ge 0$. Moreover, $\xrightarrow{d}$ denotes the convergence in distribution of random variables and $\xrightarrow{w}$ denotes the weak convergence of probability measures in $\mathcal{P}(\mathbb{R}^d)$.

Denote by $\mathbf{C}(\mathbb{R}^d; \mathbb{R}^m)$ (respectively $\mathbf{C}^k(\mathbb{R}^d; \mathbb{R}^m)$) the space consisting of continuous (respectively k-th continuously differentiable) functions from $\mathbb{R}^d$ to $\mathbb{R}^m$. Let $\mathbf{C}_b^k(\mathbb{R}^d; \mathbb{R}^m)$ stand for the set of bounded and k-th continuously differentiable functions from $\mathbb{R}^d$ to $\mathbb{R}^m$ with bounded derivatives up to order k. Denote by $\mathbf{C}_b(\mathbb{R}^d; \mathbb{R}^m)$ the set of bounded and continuous functions from $\mathbb{R}^d$ to $\mathbb{R}^m$. When no confusion occurs, $\mathbf{C}(\mathbb{R}^d; \mathbb{R}^m)$ is simply written as $\mathbf{C}(\mathbb{R}^d)$, and $\mathbf{C}_b(\mathbb{R}^d; \mathbb{R}^m), \mathbf{C}^k(\mathbb{R}^d; \mathbb{R}^m)$ and $\mathbf{C}_b^k(\mathbb{R}^d; \mathbb{R}^m)$ are treated similarly. For $l \in \mathbb{N}^+$, denote by $Poly(l, \mathbb{R}^d)$ the set of functions growing polynomially with order l, i.e.

$$\begin{aligned} Poly(l, \mathbb{R}^d) := \big\{ g \in \mathbf{C}(\mathbb{R}^d; \mathbb{R}) : |g(x) - g(y)| \le K(g)(1 + |x|^{l-1} + |y|^{l-1})|x - y| \\ \text{for any } x, y \in \mathbb{R}^d \text{ and some } K(g) > 0 \big\}. \end{aligned}$$

For $f \in \mathbf{C}^k(\mathbb{R}^d; \mathbb{R})$, denote by $\nabla^k f(x)(\xi_1, \cdots, \xi_k)$ the k-th order Gâteaux derivative along the directions $\xi_1, \cdots, \xi_k \in \mathbb{R}^d$, i.e.

$$\nabla^k f(x)(\xi_1, \cdots, \xi_k) = \sum_{i_1,\cdots,i_k=1}^{d} \frac{\partial^k f(x)}{\partial x^{i_1} \cdots \partial x^{i_k}} \xi_1^{i_1} \cdots \xi_k^{i_k}.$$

For $f = (f_1, \cdots, f_m)^\top \in \mathbf{C}^k(\mathbb{R}^d; \mathbb{R}^m)$, denote

$$\nabla^k f(x)(\xi_1, \cdots, \xi_k) = \big(\nabla^k f_1(x)(\xi_1, \cdots, \xi_k), \cdots, \nabla^k f_m(x)(\xi_1, \cdots, \xi_k)\big)^\top.$$

The Gâteaux derivative for a matrix-valued function is defined as previously. For $f \in \mathbf{C}^k(\mathbb{R}^d;\mathbb{R})$, the notation $\nabla^k f(x)$ is viewed as a tensor, i.e. a multilinear form defined on $(\mathbb{R}^d)^{\otimes k}$. Denote by $\|\cdot\|_{\otimes}$ the norm of a tensor. Throughout this paper, let $K(a_1, a_2, \cdots, a_m)$ denote some generic constant dependent on the parameters $a_1, a_2, \cdots, a_m$ but independent of the step-size τ, which may vary for each appearance.

Let us first give the assumptions on b and σ.

Assumption 2.1. *There exist constants $L_1, L_2 \in (0, +\infty)$ such that*

$$\begin{aligned} &\|\sigma(u_1) - \sigma(u_2)\|_{HS} \le L_1 |u_1 - u_2|, \quad \forall\, u_1, u_2 \in \mathbb{R}^d, \\ &\|\sigma(u)\|_{HS} \le L_2, \qquad \forall\, u \in \mathbb{R}^d. \end{aligned}$$

Assumption 2.2. *There exist $c_1 > (15/2)L_1^2, L_3 > 0$ and $q \ge 1$ such that*

$$\begin{aligned} &\langle u_1 - u_2, b(u_1) - b(u_2)\rangle \le -c_1 |u_1 - u_2|^2, \qquad \forall\, u_1, u_2 \in \mathbb{R}^d, \\ &|b(u_1) - b(u_2)| \le L_3\big(1 + |u_1|^{q-1} + |u_2|^{q-1}\big)|u_1 - u_2|, \quad \forall\, u_1, u_2 \in \mathbb{R}^d. \end{aligned} \tag{2.1}$$

The above two assumptions ensure the well-posedness of (1.2) (see e.g. [20]). The generator of (1.2) is given by

$$\mathcal{L}f(x) = \langle \nabla f(x), b(x)\rangle + \frac{1}{2}\big\langle \nabla^2 f(x), \sigma(x)\sigma(x)^\top \big\rangle_{HS}, \quad f \in \mathbf{C}^2(\mathbb{R}^d;\mathbb{R}). \tag{2.2}$$

Notice that

$$\mathrm{trace}\big(\nabla^2 f(x)\sigma(x)\sigma(x)^\top\big) = \big\langle \nabla^2 f(x), \sigma(x)\sigma(x)^\top \big\rangle_{HS}.$$

As an immediate result of (2.1),

$$|b(u)| \le L_4(1 + |u|^q), \quad \forall\, u \in \mathbb{R}^d \tag{2.3}$$

for some $L_4 > 0$. In addition, it is straightforward to conclude from Assumptions 2.1-2.2 that for any $l_2 > 0$,

$$2\langle u_1 - u_2, b(u_1) - b(u_2)\rangle + 15\|\sigma(u_1) - \sigma(u_2)\|_{HS}^2 \le -L_5|u_1 - u_2|^2, \quad \forall\, u_1, u_2 \in \mathbb{R}^d, \tag{2.4}$$

$$2\langle u, b(u)\rangle + l_2\|\sigma(u)\|_{HS}^2 \le -c_1|u|^2 + \frac{1}{c_1}|b(0)|^2 + l_2 L_2^2, \qquad \forall\, u \in \mathbb{R}^d, \tag{2.5}$$

where $L_5 := 2c_1 - 15L_1^2$. Note that Assumptions 2.1-2.2 in this paper imply Assumptions 2.1-2.4 in [20], by taking $A = \varepsilon I_d, f(x) = b(x) + \varepsilon x$ and $g(x) = \sigma(x)$ in [20, Eq. (2)] with ε small enough. Thus, all conclusions in [20] apply to our case provided that Assumptions 2.1-2.2 hold.

In order to give the regularity of the solution to the Poisson equation, we need the following assumption.

Assumption 2.3. *Let $\sigma \in \mathbf{C}_b^4(\mathbb{R}^d)$ and $b \in \mathbf{C}^4(\mathbb{R}^d)$. In addition, there exist $q' \ge 1$ and $L_6 > 0$ such that for $i = 1, 2, 3, 4$,*

$$\|\nabla^i b(u)\|_{\otimes} \le L_6\big(1 + |u|^{q'}\big), \quad \forall\, u \in \mathbb{R}^d.$$

Remark 2.1. Under Assumptions 2.1-2.3, it holds that

$$2\langle v, \nabla b(u)v\rangle + 15\|\nabla\sigma(u)v\|_{HS}^2 \le -L_5|v|^2, \quad \forall\, u, v \in \mathbb{R}^d. \tag{2.6}$$

Let us give a short proof of (2.6). Under Assumptions 2.1-2.2, we can use (2.4) to get that for any $u, v \in \mathbb{R}^d$ and $t \in \mathbb{R}$,

$$2t\langle v, b(u+tv) - b(u)\rangle + 15\|\sigma(u+tv) - \sigma(u)\|_{HS}^2 \le -L_5 t^2 |v|^2.$$

Then the differentiability of b and σ in Assumption 2.3 allows us to apply the Taylor expansion to obtaining

$$b(u+tv) = b(u) + t\nabla b(u)v + \mathcal{O}(t^2), \quad \sigma(u+tv) = \sigma(u) + t\nabla\sigma(u)v + \mathcal{O}(t^2) \quad \text{as} \quad t \to 0.$$

In this way, it follows that

$$2t^2\langle v, \nabla b(u)v\rangle + 15t^2\|\nabla\sigma(u)v\|_{HS}^2 + \mathcal{O}(t^3) \le -L_5 t^2|v|^2,$$

which yields (2.6).

Next, we recall basic facts about the invariant measure and ergodicity. Denote by $X^{s,x}(t)$ the solution to (1.2) at time t, starting from $X(s) = x$. Especially, denote $X^x(t) := X^{0,x}(t)$. Let $\pi_t(x,\cdot)$ denote the transition probability of $\{X(t)\}_{t\ge 0}$, i.e. $\pi_t(x,A){=}\mathbf{P}(X^x(t){\in}A)$ for any $A{\in}\mathscr{B}(\mathbb{R}^d)$. For any $\phi{\in}\mathbf{B}_b(\mathbb{R}^d)$ and $t \ge 0$, define the operator $P_t : \mathbf{B}_b(\mathbb{R}^d) \to \mathbf{B}_b(\mathbb{R}^d)$ by

$$(P_t\phi)(x) := \mathbf{E}\phi\big(X^x(t)\big) = \int_{\mathbb{R}} \phi(y)\pi_t(x,\mathrm{d}y).$$

Then, $\{P_t\}_{t\ge 0}$ is a Markov semigroup on $\mathbf{B}_b(\mathbb{R}^d)$. Here, $\mathbf{B}_b(\mathbb{R}^d)$ is the space of all bounded and measurable functions on $\mathbb{R}^d$. A probability measure $\mu \in \mathcal{P}(\mathbb{R}^d)$ is called an invariant measure of $\{X(t)\}_{t\ge 0}$ or $\{P_t\}_{t\ge 0}$, if

$$\int_{\mathbb{R}^d} P_t\phi(x)\mu(\mathrm{d}x) = \int_{\mathbb{R}^d} \phi(x)\mu(\mathrm{d}x), \quad \forall\, \phi \in \mathbf{B}_b(\mathbb{R}^d), \quad t \ge 0. \tag{2.7}$$

Further, an invariant measure μ is called an ergodic measure of $\{X(t)\}_{t\ge 0}$ or $\{P_t\}_{t\ge 0}$, if for any $\phi \in \mathbf{L}^2(\mathbb{R}^d,\mu)$,

$$\lim_{T\to+\infty} \frac{1}{T}\int_0^T P_t\phi(x)\mathrm{d}t = \int_{\mathbb{R}^d} \phi(x)\mu(\mathrm{d}x) \quad \text{in } \mathbf{L}^2(\mathbb{R}^d,\mu),$$

where $\mathbf{L}^2(\mathbb{R}^d,\mu)$ is the space of all square integrable functions w.r.t. μ (see e.g. [13, Definition 1.6]). Especially, if μ is the unique invariant measure of $\{X(t)\}_{t\ge 0}$, then μ is also the ergodic measure. We refer readers to [13] for more details.

Proposition 2.1. *Let Assumptions* 2.1-2.2 *hold. Then the following hold:*

(1) *For any* $p \ge 1, \sup_{t\ge 0}\mathbf{E}|X^x(t)|^p \le K(p)(1+|x|^p)$.

(2) *For any* $t, s \ge 0, (\mathbf{E}|X^x(t) - X^x(s)|^2)^{1/2} \le K(1+|x|^q)|t-s|^{1/2}$.

(3) *For any* $t \ge 0, (\mathbf{E}|X^x(t) - X^y(t)|^2)^{1/2} \le |x-y|e^{-L_5 t/2}$.

The first and second conclusions come from [20, Propositions 3.1–3.2]. The third conclusion can be obtained by applying the Itô formula. In addition, [20, Theorem 3.1] gives the ergodicity for $\{X(t)\}_{t\ge 0}$.

Proposition 2.2. *Let Assumptions* 2.1-2.2 *hold. Then we have the following:*

(1) $\{X(t)\}_{t\geq 0}$ *admits a unique invariant measure* $\pi \in \mathcal{P}(\mathbb{R}^d)$.

(2) For any $p \geq 1$, $\pi(|\cdot|^p) < +\infty$.

(3) There is $\lambda_1 > 0$ *such that for any* $f \in Poly(l, \mathbb{R}^d)$, $l \geq 1$ *and* $t \geq 0$,

$$\big|\mathbf{E}f\big(X^x(t)\big) - \pi(f)\big| \leq K(f)(1+|x|^l)e^{-\lambda_1 t}. \tag{2.8}$$

Proof. It follows from [20, Theorem 3.1] that $\{X(t)\}_{t\geq 0}$ admits a unique invariant measure $\pi \in \mathcal{P}(\mathbb{R}^d)$, and $\pi_t(x,\cdot) \xrightarrow{w} \pi$, as $t \to +\infty$ for any $x \in \mathbb{R}^d$. Especially, $\pi_t(0,\cdot) \xrightarrow{w} \pi$, which implies that for any $M > 0$,

$$\begin{aligned}\int_{\mathbb{R}^d}(|x|^p \wedge M)\pi(\mathrm{d}x) &= \lim_{t\to+\infty}\int_{\mathbb{R}^d}(|x|^p \wedge M)\pi_t(0,\mathrm{d}x)\\ &\leq M \wedge \limsup_{t\to+\infty}\mathbf{E}|X^0(t)|^p \leq K,\end{aligned}$$

where we used $|\cdot|^p \wedge M \in \mathbf{C}_b(\mathbb{R}^d)$ and Proposition 2.1(1). Then the Fatou lemma gives

$$\pi(|\cdot|^p) = \int_{\mathbb{R}^d}|x|^p\,\pi(\mathrm{d}x) \leq \liminf_{M\to+\infty}\int_{\mathbb{R}^d}(|x|^p \wedge M)\pi(\mathrm{d}x) \leq K.$$

For any $M > 0$ and $f \in Poly(l, \mathbb{R}^d)$, it holds that $f \wedge M \in \mathbf{C}_b(\mathbb{R}^d)$. Accordingly, it follows from the definition of the invariant measure (see (2.7)) that

$$\pi(f \wedge M) = \int_{\mathbb{R}^d} P_t(f \wedge M)(y)\pi(\mathrm{d}y).$$

Thus, using Proposition 2.1(3), the Hölder inequality, the fact $|a \wedge b - a \wedge c| \leq |b - c|$ and the second conclusion, we conclude that for any $M > 0$,

$$\begin{aligned}&\big|\mathbf{E}\big(f\big(X^x(t)\big) \wedge M\big) - \pi(f \wedge M)\big|\\ =&\left|P_t(f \wedge M)(x) - \int_{\mathbb{R}^d} P_t(f \wedge M)(y)\pi(\mathrm{d}y)\right|\\ =&\left|\int_{\mathbb{R}^d}\big[P_t(f \wedge M)(x) - P_t(f \wedge M)(y)\big]\pi(\mathrm{d}y)\right|\\ \leq&\int_{\mathbb{R}^d}\big|\mathbf{E}\big(f\big(X^x(t)\big) \wedge M\big) - \mathbf{E}\big(f\big(X^y(t)\big) \wedge M\big)\big|\pi(\mathrm{d}y)\\ \leq&\int_{\mathbb{R}^d}\mathbf{E}\big|f\big(X^x(t)\big) - f\big(X^x(y)\big)\big|\pi(\mathrm{d}y)\\ \leq& K(f)\int_{\mathbb{R}^d}\Big(1 + \big(\mathbf{E}|X^x(t)|^{2l-2}\big)^{\frac12} + \big(\mathbf{E}|X^y(t)|^{2l-2}\big)^{\frac12}\Big)\big(\mathbf{E}|X^x(t) - X^y(t)|^2\big)^{\frac12}\pi(\mathrm{d}y)\\ \leq& K(f)e^{-\frac{L_5}{2}t}\int_{\mathbb{R}^d}(1 + |x|^{l-1} + |y|^{l-1})|x-y|\pi(\mathrm{d}y)\\ \leq& K(f)e^{-\frac{L_5}{2}t}(1 + |x|^l).\end{aligned}$$

The above formula and the monotone convergence theorem lead to (2.8), which completes the proof. □

3. Main Result

In this section, we give our main result, i.e. the CLT for the temporal average $\Pi_{\tau,\alpha}(h)$ of the BEM method used to approximate the ergodic limit $\pi(h)$. The BEM method has been widely applied to approximating SODEs or SPDEs with non-Lipschitz coefficients (see e.g. [4, 11, 22] and references therein). Let $\tau > 0$ be the temporal step-size. The BEM method for (1.2) reads

$$\bar{X}_{n+1} = \bar{X}_n + b(\bar{X}_{n+1})\tau + \sigma(\bar{X}_n)\Delta W_n, \quad n = 0, 1, 2, \ldots, \tag{3.1}$$

where $\Delta W_n := W(t_{n+1}) - W(t_n)$ with $t_n = n\tau$. We refer to [20, Lemma 2.1] for the unique solvability of the BEM method (3.1). In practice, we can use the fixed point algorithm to yield the numerical solution of (3.1), as is done in Section 5. More precisely, with $\bar{X}_n$ in hand, we calculate $\bar{X}_{n+1}$ by

$$\bar{X}_{n+1}^{(i)} = \bar{X}_n + b\big(\bar{X}_{n+1}^{(i-1)}\big)\tau + \sigma(\bar{X}_n)\Delta W_n, \quad i = 1, 2, \ldots$$

with $\bar{X}_{n+1}^{(0)} = \bar{X}_n$. The above iteration is terminated when $|\bar{X}_{n+1}^{(i)} - \bar{X}_{n+1}^{(i-1)}| < \varepsilon_0$ for some preset tolerance error ε_0, and $\bar{X}_{n+1}$ is endowed with the final iterative solution. We denote by $\bar{X}_n^{k,x}$ the solution to (3.1) at the n-th step provided $\bar{X}_k = x$. Especially, denote $\bar{X}_n^x := \bar{X}_n^{0,x}$, i.e. the solution to (3.1) with the initial value $x \in \mathbb{R}^d$.

The following results can be found in [20, Lemmas 4.1, 4.2, Theorem 4.2].

Proposition 3.1. *Let Assumptions* 2.1-2.2 *hold and* τ *sufficiently small. Then the following properties hold:*

(1) $\sup_{n\geq 0} \mathbf{E}|\bar{X}_n^x|^2 \leq K(1 + |x|^2)$.

(2) There is $\xi_1 > 0$ *such that for any* $n \geq 0$, $(\mathbf{E}|\bar{X}_n^x - \bar{X}_n^y|^2)^{1/2} \leq K|x - y|e^{-\xi_1 n\tau}$.

(3) $\sup_{n\geq 0} \mathbf{E}|X^x(t_n) - \bar{X}_n^x|^2 \leq K(x)\tau$.

Recall that the temporal average of the BEM method is

$$\Pi_{\tau,\alpha}(h) = \frac{1}{\tau^{-\alpha}} \sum_{k=0}^{\tau^{-\alpha}-1} h\big(\bar{X}_k^x\big), \quad \alpha \in (1, 2], \quad h \in \mathbf{C}_b^4(\mathbb{R}^d).$$

Define the function $\varphi : \mathbb{R}^d \to \mathbb{R}$ by

$$\varphi(x) = -\int_0^\infty \mathbf{E}\big(h\big(X^x(t)\big) - \pi(h)\big)\mathrm{d}t, \quad x \in \mathbb{R}^d, \tag{3.2}$$

which is indeed a solution to the Poisson equation $\mathcal{L}\varphi = h - \pi(h)$ due to Lemma 4.2. Then we have the following CLT for $\Pi_{\tau,\alpha}(h), \alpha \in (1, 2)$.

Theorem 3.1. *Let Assumptions* 2.1-2.3 *hold and* $h \in \mathbf{C}_b^4(\mathbb{R}^d)$.

(1) Let $\{Y_n\}_{n\geq 0}$ *be a numerical solution for* (1.2) *with global strong order* $1/2$ *in the infinite time horizon, i.e. there is* $K > 0$ *independent of* τ *such that*

$$\sup_{n\geq 0} \mathbf{E}|X(t_n) - Y_n|^2 \leq K\tau. \tag{3.3}$$

Then for any $\alpha \in (1, 2)$,

$$\frac{1}{\tau^{\frac{\alpha-1}{2}}}\left(\frac{1}{\tau^{-\alpha}} \sum_{k=0}^{\tau^{-\alpha}-1} h(Y_k) - \pi(h)\right) \xrightarrow{d} \mathcal{N}\left(0, \pi\big(|\sigma^\top \nabla\varphi|^2\big)\right) \quad \text{as} \quad \tau \to 0. \tag{3.4}$$

(2) For any $\alpha \in (1,2)$ *and* $x \in \mathbb{R}^d$,

$$\frac{1}{\tau^{\frac{\alpha-1}{2}}}\big(\Pi_{\tau,\alpha}(h)-\pi(h)\big) \xrightarrow{d} \mathcal{N}\big(0,\pi\big(|\sigma^\top\nabla\varphi|^2\big)\big) \quad \text{as} \quad \tau \to 0. \tag{3.5}$$

Proof. Let φ be that in (3.2). By Lemma 4.2, it holds that $\varphi \in \mathbf{C}^3(\mathbb{R}^d)$ and

$$\mathcal{L}\varphi = h - \pi(h).$$

It follows from [2, Theorem 2.1] that the CLT holds for (1.2), i.e.

$$\frac{1}{\sqrt{T}}\int_0^T \big(h\big(X(t)\big)-\pi(h)\big)\mathrm{d}t \xrightarrow{d} \mathcal{N}\big(0,-2\pi(\varphi\mathcal{L}\varphi)\big) \quad \text{as} \quad T \to \infty.$$

By (2.2) and a direct computation,

$$\varphi\mathcal{L}\varphi = \frac{1}{2}\mathcal{L}\varphi^2 - \frac{1}{2}|\sigma^\top\nabla\varphi|^2.$$

Since φ^2 belongs to the domain of $\mathcal{L}, \pi(\mathcal{L}\varphi^2)=0$ due to [2, Eq. (2.6)]. Combining the above relations, we have

$$\frac{1}{\sqrt{T}}\int_0^T \big(h\big(X(t)\big)-\pi(h)\big)\mathrm{d}t \xrightarrow{d} \mathcal{N}\big(0,\pi\big(|\sigma^\top\nabla\varphi|^2\big)\big) \quad \text{as} \quad T \to \infty. \tag{3.6}$$

Notice that

$$\begin{aligned}
&\frac{1}{\tau^{\frac{\alpha-1}{2}}}\left(\frac{1}{\tau^{-\alpha}}\sum_{k=0}^{\tau^{-\alpha}-1} h(Y_k)-\pi(h)\right)\\
&= \frac{1}{\tau^{\frac{\alpha-1}{2}}}\left(\frac{1}{\tau^{-\alpha}}\sum_{k=0}^{\tau^{-\alpha}-1} h(Y_k)-\tau^{\alpha-1}\int_0^{\tau^{1-\alpha}} h\big(X(t)\big)\mathrm{d}t\right)\\
&\quad+\tau^{\frac{\alpha-1}{2}}\int_0^{\tau^{1-\alpha}}\big(h\big(X(t)\big)-\pi(h)\big)\mathrm{d}t\\
&=: J_1(\tau)+J_2(\tau).
\end{aligned}$$

By (3.6) and $\alpha > 1$, $J_2(\tau) \xrightarrow{d} \mathcal{N}(0,\pi(|\sigma^\top\nabla\varphi|^2))$ as $\tau \to 0$. Denoting $N = \tau^{-\alpha}$, we use Proposition 2.1(2), (3.3) and $h \in \mathbf{C}_b^1(\mathbb{R}^d)$ to get

$$\begin{aligned}
\mathbf{E}|J_1(\tau)| &= \frac{1}{\tau^{\frac{\alpha-1}{2}}}\mathbf{E}\left|\frac{1}{N}\sum_{k=0}^{N-1}h(Y_k)-\frac{1}{N\tau}\sum_{k=0}^{N-1}\int_{k\tau}^{(k+1)\tau}h\big(X(t)\big)\mathrm{d}t\right|\\
&\le \frac{1}{\tau^{\frac{\alpha-1}{2}}}\frac{1}{N}\sum_{k=0}^{N-1}\mathbf{E}\big|h(Y_k)-h\big(X(t_k)\big)\big|\\
&\quad+\frac{1}{\tau^{\frac{\alpha-1}{2}}}\frac{1}{N\tau}\sum_{k=0}^{N-1}\int_{k\tau}^{(k+1)\tau}\mathbf{E}\big|h\big(X(t)\big)-h\big(X(t_k)\big)\big|\mathrm{d}t\\
&\le K(h)\frac{1}{\tau^{\frac{\alpha-1}{2}}}\sup_{k\ge 0}\big(\mathbf{E}|Y_k-X(t_k)|^2\big)^{\frac{1}{2}}\\
&\quad+K(h)\frac{1}{\tau^{\frac{\alpha-1}{2}}}\frac{1}{N\tau}\sum_{k=0}^{N-1}\int_{k\tau}^{(k+1)\tau}\big(\mathbf{E}|X(t)-X(t_k)|^2\big)^{\frac{1}{2}}\mathrm{d}t
\end{aligned}$$

$$\leq K(h)\frac{1}{\tau^{\frac{\alpha-1}{2}}}\tau^{\frac{1}{2}} = K(h)\tau^{\frac{2-\alpha}{2}}.$$

Thus, $\lim_{\tau\to 0}\mathbf{E}|J_1(\tau)| = 0$ due to $\alpha < 2$, which implies that $J_1(\tau)$ converges to 0 in probability. Thus, (3.4) follows by applying the Slutsky theorem (see e.g. [16, Theorem 18.8]).

Finally, (3.5) holds as a special case of (3.4) due to Proposition 3.1(3). Thus, the proof is complete. □

Remark 3.1. (1) It is observed that

$$\frac{1}{\tau^{\frac{\alpha-1}{2}}}\big(\Pi_{\tau,\alpha}(h) - \pi(h)\big) = \frac{1}{\tau^{\frac{1-\alpha}{2}}}\sum_{k=0}^{\tau^{-\alpha}-1}\big(h\big(\bar{X}_k^x\big) - \pi(h)\big)\tau,$$

which can be viewed as a numerical approximation of

$$\frac{1}{\sqrt{T}}\int_0^T \big(h\big(X^x(t)\big) - \pi(h)\big)\mathrm{d}t$$

with $T(\tau) = N\tau$ and $N = \tau^{-\alpha}$. Thus, $\alpha > 1$ is required such that $\lim_{\tau\to 0} T(\tau) = +\infty$, which coincides with the CLT for $\{X(t)\}_{t\geq 0}$.

(2) In fact, we give the CLT of the temporal average for a class of numerical methods satisfying (3.3) for $\alpha \in (1,2)$. For example, it follows from Theorem 3.1 and [18, Theorem 4.8] that the CLT holds for the truncated EM method, i.e. (3.4) holds with $\{Y_k\}$ being the numerical solution of the truncated EM method proposed in [18]. We also guess that there may be some non-ergodic numerical method whose temporal average satisfies the CLT in view of Theorem 3.1(1).

We close the section by presenting the CLT for $\Pi_{\tau,2}(h)$.

Theorem 3.2. *Let Assumptions* 2.1-2.3 *hold and* $h \in \mathbf{C}_b^4(\mathbb{R}^d)$*. Then for any* $x \in \mathbb{R}^d$*,*

$$\frac{1}{\sqrt{\tau}}\big(\Pi_{\tau,2}(h) - \pi(h)\big) \xrightarrow{d} \mathcal{N}\big(0, \pi\big(|\sigma^\top\nabla\varphi|^2\big)\big) \quad as \quad \tau \to 0.$$

As is pointed out in the introduction, the proof idea of Theorem 3.1 does not apply to the case $\alpha = 2$. Instead, we will use the Poisson equation $\mathcal{L}\varphi = h - \pi(h)$ to give a good decomposition of $\Pi_{\tau,2}(h)$, on basis of which the CLT of $\Pi_{\tau,2}(h)$ can be established. We postpone the proof of Theorem 3.2 to the next section.

4. Proof of Theorem 3.2

In this section, we will prove Theorem 3.2. The main strategy follows from [23] (see also [6]), whose main idea is to use the Poisson equation (4.25) to split $(\Pi_{\tau,2}(h)-\pi(h))/\sqrt{\tau}$ into a martingale difference series sum and a negligible remainder. Then, we will prove that the martingale difference series sum converges to a normal distribution in Lemma 4.3, and that the remainder converges to zero in probability in Lemma 4.4.

It should be mentioned that although the main idea is similar to that of [6] and [23], we still need to overcome the difficulties brought by the super-linearly growing drift coefficient and the multiplicative noise. The first difficulty is to give the p-th ($p > 2$) moment boundedness of the BEM method, which is overcome by a careful control of the iteration coefficient (see also the detailed statement below the proof of Theorem 4.1). The second difficulty lies in the regularity estimate of the Poisson equation. Note that [23, Lemma 3.1] is not applicable to our case due to the super-linearly growing coefficients. In addition, the argument of [6, Lemma 4.1] relies on the a.s. differentiability of the exact solution w.r.t. initial values, which applies to the case of additive noises. To give the regularity of the Poisson equation in the presence of multiplicative noises, we estimate the high-order moments of the mean-square derivatives of the exact solution w.r.t. initial values (see Lemma 4.1).

4.1. Auxiliary results

In this subsection, we give some auxiliary results for proving Theorem 3.2. We begin with the p-th ($p > 2$) moment boundedness of (3.1).

Theorem 4.1. *Suppose that Assumptions* 2.1-2.2 *hold. Then for any* $r \geq 1$ *and* $\tau \leq 1$,

$$\sup_{n\geq 0} \mathbf{E}\big|\bar{X}_n^x\big|^r \leq K(r)(1+|x|^r). \tag{4.1}$$

Proof. It is sufficient to show that for any positive integer p,

$$\sup_{n\geq 0} \mathbf{E}\big|\bar{X}_n^x\big|^{2p} \leq K(p)(1+|x|^{2p}), \tag{4.2}$$

in view of the Hölder inequality, which will be derived via mathematical induction.

By (3.1) and (2.5),

$$\begin{aligned}
&\big|\bar{X}_{n+1}^x\big|^2 - \big|\bar{X}_n^x\big|^2 + \big|\bar{X}_{n+1}^x - \bar{X}_n^x\big|^2 \\
&= 2\big\langle \bar{X}_{n+1}^x, \bar{X}_{n+1}^x - \bar{X}_n^x \big\rangle \\
&= 2\big\langle \bar{X}_{n+1}^x, b\big(\bar{X}_{n+1}^x\big)\big\rangle\tau + 2\big\langle \bar{X}_{n+1}^x - \bar{X}_n^x, \sigma\big(\bar{X}_n^x\big)\Delta W_n \big\rangle \\
&\quad + 2\big\langle \bar{X}_n^x, \sigma\big(\bar{X}_n^x\big)\Delta W_n\big\rangle \\
&\leq -c_1\tau\big|\bar{X}_{n+1}^x\big|^2 + K\tau + \big|\bar{X}_{n+1}^x - \bar{X}_n^x\big|^2 \\
&\quad + \big\|\sigma\big(\bar{X}_n^x\big)\big\|_{HS}^2 |\Delta W_n|^2 + 2\big\langle \bar{X}_n^x, \sigma\big(\bar{X}_n^x\big)\Delta W_n\big\rangle,
\end{aligned} \tag{4.3}$$

which together with the boundedness of σ yields

$$(1+c_1\tau)\big|\bar{X}_{n+1}^x\big|^2 \leq \big|\bar{X}_n^x\big|^2 + K\tau + L_2^2|\Delta W_n|^2 + 2\big\langle \bar{X}_n^x, \sigma\big(\bar{X}_n^x\big)\Delta W_n\big\rangle. \tag{4.4}$$

Noting that $\mathbf{E}\langle \bar{X}_n^x, \sigma(\bar{X}_n^x)\Delta W_n\rangle = 0$, we have

$$\mathbf{E}\big|\bar{X}_{n+1}^x\big|^2 \leq \frac{1}{1+c_1\tau}\mathbf{E}\big|\bar{X}_n^x\big|^2 + \frac{K\tau}{1+c_1\tau}. \tag{4.5}$$

By iteration, we arrive at

$$\mathbf{E}\big|\bar{X}_n^x\big|^2 \leq \frac{1}{(1+c_1\tau)^n}|x|^2 + K\tau\sum_{i=1}^{\infty}\frac{1}{(1+c_1\tau)^i} \leq |x|^2 + K.$$

Thus, (4.2) holds for $p = 1$. Now, we assume that

$$\sup_{n\geq 0} \mathbf{E}\big|\bar{X}_n^x\big|^{2(p-1)} \leq K(p)(1+|x|^{2(p-1)}), \quad p \geq 2. \tag{4.6}$$

It remains to prove

$$\sup_{n\geq 0}\mathbf{E}\big|\bar{X}_n^x\big|^{2p}\leq K(p)(1+|x|^{2p}).$$

In fact, using (4.4) and the inequality $(1+x)^\alpha\geq 1+\alpha x,\alpha\geq 1,x>-1$ leads to

$$(1+pc_1\tau)\big|\bar{X}_{n+1}^x\big|^{2p}\leq\Big(\big|\bar{X}_n^x\big|^2+2\big\langle\bar{X}_n^x,\sigma\big(\bar{X}_n^x\big)\Delta W_n\big\rangle+K\big(\tau+|\Delta W_n|^2\big)\Big)^p. \tag{4.7}$$

Notice that

$$\begin{aligned}&\Big(\big|\bar{X}_n^x\big|^2+2\big\langle\bar{X}_n^x,\sigma\big(\bar{X}_n^x\big)\Delta W_n\big\rangle+K\big(\tau+|\Delta W_n|^2\big)\Big)^p\\ =&\sum_{i_1=0}^{p}\sum_{i_2=0}^{p-i_1}C_p^{i_1}C_{p-i_1}^{i_2}2^{i_2}K^{p-(i_1+i_2)}\big|\bar{X}_n^x\big|^{2i_1}\big\langle\bar{X}_n^x,\sigma\big(\bar{X}_n^x\big)\Delta W_n\big\rangle^{i_2}\big(\tau+|\Delta W_n|^2\big)^{p-(i_1+i_2)}\\ =&\big|\bar{X}_n^x\big|^{2p}+\sum_{i_1=0}^{p-1}\sum_{i_2=0}^{p-i_1-1}C_p^{i_1}C_{p-i_1}^{i_2}2^{i_2}K^{p-(i_1+i_2)}S_{n,i_1,i_2}+\sum_{i=0}^{p-1}C_p^i2^{p-i}T_{n,i},\end{aligned}$$

where

$$\begin{aligned}S_{n,i_1,i_2}&:=\big|\bar{X}_n^x\big|^{2i_1}\big\langle\bar{X}_n^x,\sigma\big(\bar{X}_n^x\big)\Delta W_n\big\rangle^{i_2}\big(\tau+|\Delta W_n|^2\big)^{p-(i_1+i_2)}, && i_1\in[0,p-1],\quad i_2\in[0,p-i_1-1],\\ T_{n,i}&:=\big|\bar{X}_n^x\big|^{2i}\big\langle\bar{X}_n^x,\sigma\big(\bar{X}_n^x\big)\Delta W_n\big\rangle^{p-i}, && i\in[0,p-1].\end{aligned}$$

For any $i_1\in[0,p-1],i_2\in[0,p-i_1-1]$, it follows from the independence of ΔW_n and $\bar{X}_n^x$, the boundedness of σ, the Hölder inequality and (4.6) that for $\tau\leq 1$,

$$\begin{aligned}|\mathbf{E}S_{n,i_1,i_2}|&\leq K(p)\mathbf{E}\big|\bar{X}_n^x\big|^{2i_1+i_2}\mathbf{E}\Big[|\Delta W_n|^{i_2}\big(\tau+|\Delta W_n|^2\big)^{p-(i_1+i_2)}\Big]\\ &\leq K(p)\Big(\mathbf{E}\big|\bar{X}_n^x\big|^{2p-2}\Big)^{\frac{2i_1+i_2}{2p-2}}\tau\leq K(p)(1+|x|^{2p-2})\tau.\end{aligned}$$

Next we estimate $|\mathbf{E}T_{n,i}|$ for $i=0,\dots,p-1$. Notice that the property of conditional expectations (see e.g. [19, Lemma 2.6]) leads to

$$\begin{aligned}\mathbf{E}T_{n,p-1}&=\mathbf{E}\Big[\mathbf{E}\Big(\big|\bar{X}_n^x\big|^{2p-2}\big\langle\bar{X}_n^x,\sigma\big(\bar{X}_n^x\big)\Delta W_n\big\rangle\big|\,\mathcal{F}_{t_n}\Big)\Big]\\ &=\mathbf{E}\Big[\big(\mathbf{E}\left(|y|^{2p-2}\langle y,\sigma(y)\Delta W_n\rangle\right)\big)\big|_{y=\bar{X}_n^x}\Big]=0.\end{aligned}$$

For $i=0,\dots,p-2$, applying (4.6), the boundedness of σ and the Hölder inequality, we get

$$\begin{aligned}|\mathbf{E}T_{n,i}|&\leq K(p)\mathbf{E}\big|\bar{X}_n^x\big|^{p+i}\mathbf{E}|\Delta W_n|^{p-i}\\ &\leq K(p)\Big(\mathbf{E}\big|\bar{X}_n^x\big|^{2p-2}\Big)^{\frac{p+i}{2p-2}}\tau^{\frac{p-i}{2}}\leq K(p)(1+|x|^{2p-2})\tau.\end{aligned}$$

Combining the above formulas gives

$$\mathbf{E}\Big(\big|\bar{X}_n^x\big|^2+2\big\langle\bar{X}_n^x,\sigma\big(\bar{X}_n^x\big)\Delta W_n\big\rangle+K\big(\tau+|\Delta W_n|^2\big)\Big)^p\leq\mathbf{E}\big|\bar{X}_n^x\big|^{2p}+K(p)(1+|x|^{2p-2})\tau,$$

which along with (4.7) yields

$$\mathbf{E}\big|\bar{X}_{n+1}^x\big|^{2p}\leq\frac{1}{1+pc_1\tau}\mathbf{E}\big|\bar{X}_n^x\big|^{2p}+K(p)\frac{(1+|x|^{2p-2})\tau}{1+pc_1\tau}. \tag{4.8}$$

Then by iteration, we deduce

$$\mathbf{E}\big|\bar{X}_n^x\big|^{2p} \le \frac{1}{(1+pc_1\tau)^n}|x|^{2p} + K(p)(1+|x|^{2p-2})\tau\sum_{i=1}^{\infty}\frac{1}{(1+pc_1\tau)^i} \le K(p)(1+|x|^{2p}).$$

Thus, (4.2) holds by mathematical induction and the proof is complete. □

Notice that [21, Proposition 3.1] gives the p-th ($p>2$) moment boundedness of the BEM method in the finite time interval, where the conclusion relies on the application of the Gronwall inequality. The argument is not applicable to our case, because the Gronwall inequality fails to derive the moment boundedness in the infinite time horizon. Instead, we use the iteration argument and ensure the iteration coefficient strictly less than 1 in the proof of Theorem 4.1. In addition, we also note that [20] gives the second moment boundedness of the BEM method in the infinite time interval, i.e. Proposition 3.1(1). As is shown in (4.5), for the case $p=2$, when estimating $\mathbf{E}|\bar{X}_{n+1}^x|^2$, the effect of the Brownian increment ΔW_n is removed, while for the case $p>2$, we need to tackle the interaction between the Brownian increment and other terms. Thus, we must carefully tackle the right-hand side of (4.7).

Proposition 4.1. *Let Assumptions* 2.1-2.2 *hold and τ be sufficiently small. Then the BEM method* (3.1) *admits a unique invariant measure $\pi_\tau \in \mathcal{P}(\mathbb{R}^d)$. Moreover, for any $l\ge 1$, $f\in Poly(l,\mathbb{R}^d)$ and $n\ge 0$,*

$$\big|\mathbf{E}f\big(\bar{X}_n^x\big) - \pi_\tau(f)\big| \le K(f)(1+|x|^l)e^{-\xi_1 n\tau}, \quad x\in\mathbb{R}^d, \quad n\ge 0, \tag{4.9}$$

$$|\pi_\tau(f)-\pi(f)| \le K(f)\tau^{\frac{1}{2}}. \tag{4.10}$$

By means of Theorem 4.1, one can prove (4.9) similar to (2.8). Then (4.10) immediately follows as a result of (2.8), (4.9) and Proposition 3.1(3).

In order to prove the CLT for $\Pi_{\tau,2}(h)$, we need to give the regularity of φ. This can be done through a probabilistic approach by means of mean-square derivatives of $\{X^x(t)\}_{t\ge 0}$ w.r.t. the initial value x. For any $x, y_i\in\mathbb{R}^d, i=1,2,3,4$, denote by $\eta_{y_1}^x(t)$ the mean-square derivative of $X^x(t)$ along with the direction y_1, i.e.

$$\eta_{y_1}^x(t) = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\big(X^{x+\varepsilon y_1}(t) - X^x(t)\big)$$

in $\mathbf{L}^2(\Omega;\mathbb{R}^d)$. Further, denote

$$\eta_{y_1,y_2}^x(t) := \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\big(\eta_{y_1}^{x+\varepsilon y_2}(t) - \eta_{y_1}^x(t)\big)$$

in $\mathbf{L}^2(\Omega;\mathbb{R}^d)$, i.e. $\eta_{y_1,y_2}^x(t)$ is the second mean-square derivative of $X^x(t)$ along with the directions y_1 and y_2. $\eta_{y_1,y_2,y_3}^x(t)$ and $\eta_{y_1,y_2,y_3,y_4}^x(t)$ are defined similarly. We refer readers to [5, 32] for more details about the mean-square differentiability of SODEs w.r.t. initial values.

Lemma 4.1. *Suppose that Assumptions* 2.1-2.3 *hold. Then there exist C_1, $C_2>0$ and $\kappa_i>0$, $i=1,2,3$ such that for any $x, y_i\in\mathbb{R}^d$, $i=1,2,3,4$ and $t\ge 0$,*

$$\Big(\mathbf{E}\big|\eta_{y_1}^x(t)\big|^{16+\kappa_1}\Big)^{\frac{1}{16+\kappa_1}} \le C_1|y_1|e^{-C_2 t}, \tag{4.11}$$

$$\Big(\mathbf{E}\big|\eta_{y_1,y_2}^x(t)\big|^{8+\kappa_2}\Big)^{\frac{1}{8+\kappa_2}} \le C_1\big(1+|x|^{q'}\big)|y_1||y_2|e^{-C_2 t}, \tag{4.12}$$

$$\Big(\mathbf{E}\big|\eta_{y_1,y_2,y_3}^x(t)\big|^{4+\kappa_3}\Big)^{\frac{1}{4+\kappa_3}} \le C_1\big(1+|x|^{2q'}\big)|y_1||y_2||y_3|e^{-C_2 t}, \tag{4.13}$$

$$\Big(\mathbf{E}\big|\eta_{y_1,y_2,y_3,y_4}^x(t)\big|^{2}\Big)^{\frac{1}{2}} \le C_1\big(1+|x|^{3q'}\big)|y_1||y_2||y_3||y_4|e^{-C_2 t}. \tag{4.14}$$

Proof. Similarly to [5, Section 1.3.3], $\eta^x_{y_1}$ solves the following variational equation:

$$\begin{cases} \mathrm{d}\eta^x_{y_1}(t) = \nabla b\big(X^x(t)\big)\eta^x_{y_1}(t)\mathrm{d}t + \nabla\sigma\big(X^x(t)\big)\eta^x_{y_1}(t)\mathrm{d}W(t), \\ \eta^x_{y_1}(0) = y_1. \end{cases}$$

Notice that for any $p \geq 2$ and matrix A, it holds that $\nabla(|x|^p) = p|x|^{p-2}x$ and

$$\frac{1}{2}\mathrm{trace}\big(\nabla^2(|x|^p)AA^\top\big) \leq \frac{1}{2}p(p-1)|x|^{p-2}\|A\|^2_{HS}. \tag{4.15}$$

For any $\kappa \in (0,1)$ and $\lambda > 0$, by the Itô formula, (4.15), $\sigma \in \mathbf{C}^4_b(\mathbb{R}^d)$ and (2.6),

$$\begin{aligned} &\mathbf{E}\big(e^{\lambda t}\big|\eta^x_{y_1}(t)\big|^{16+\kappa}\big) \\ &\leq |y_1|^{16+\kappa} + \lambda\mathbf{E}\int_0^t e^{\lambda s}\big|\eta^x_{y_1}(s)\big|^{16+\kappa}\mathrm{d}s \\ &\quad + \frac{1}{2}(16+\kappa)\mathbf{E}\int_0^t e^{\lambda s}\big|\eta^x_{y_1}(s)\big|^{14+\kappa} \\ &\qquad\qquad \times\Big[2\big\langle\eta^x_{y_1}(s), \nabla b\big(X^x(s)\big)\eta^x_{y_1}(s)\big\rangle + (15+\kappa)\big\|\nabla\sigma\big(X^x(s)\big)\eta^x_{y_1}(s)\big\|^2_{HS}\Big]\mathrm{d}s \\ &\leq |y_1|^{16+\kappa} + \Big[\lambda + \Big(8+\frac{\kappa}{2}\Big)\big(-L_5 + \kappa L^2_\sigma\big)\Big]\,\mathbf{E}\int_0^t e^{\lambda s}\big|\eta^x_{y_1}(s)\big|^{16+\kappa}\mathrm{d}s, \end{aligned}$$

where $L_\sigma := \sup_{x\in\mathbb{R}^d}\|\nabla\sigma(x)\|_\otimes$. Letting $\kappa_1 < L_5/L^2_\sigma, \lambda_1$ small enough, we obtain

$$\mathbf{E}\big|\eta^x_{y_1}(t)\big|^{16+\kappa_1} \leq |y_1|^{16+\kappa_1}e^{-\lambda_1 t}, \quad \forall\, t \in [0,T],$$

which yields (4.11).

Secondly, similar to the argument for $\eta^x_{y_1}$, we have

$$\begin{cases} \mathrm{d}\eta^x_{y_1,y_2}(t) = \nabla b\big(X^x(t)\big)\eta^x_{y_1,y_2}(t)\mathrm{d}t + \nabla^2 b\big(X^x(t)\big)\big(\eta^x_{y_1}(t), \eta^x_{y_2}(t)\big)\mathrm{d}t \\ \qquad\qquad + \nabla\sigma\big(X^x(t)\big)\eta^x_{y_1,y_2}(t)\mathrm{d}W(t) + \nabla^2\sigma\big(X^x(t)\big)\big(\eta^x_{y_1}(t), \eta^x_{y_2}(t)\big)\mathrm{d}W(t), \\ \eta^x_{y_1,y_2}(0) = 0. \end{cases}$$

For any $\kappa, \lambda, \varepsilon_0 \in (0,1)$, again by the Itô formula, (4.15), $\sigma \in \mathbf{C}^4_b(\mathbb{R}^d)$ and the elementary inequality $(a+b)^2 \leq (1+\varepsilon_0)a^2 + (1+1/\varepsilon_0)b^2$ with $a, b \geq 0$, it holds that

$$\begin{aligned} &\mathbf{E}\big(e^{\lambda t}\big|\eta^x_{y_1,y_2}(t)\big|^{8+\kappa}\big) \\ &\leq \lambda\mathbf{E}\int_0^t e^{\lambda s}\big|\eta^x_{y_1,y_2}(s)\big|^{8+\kappa}\mathrm{d}s + (8+\kappa)\mathbf{E}\int_0^t e^{\lambda s}\big|\eta^x_{y_1,y_2}(s)\big|^{6+\kappa}\big\langle\eta^x_{y_1,y_2}(s), \nabla b\big(X^x(s)\big)\eta^x_{y_1,y_2}(s)\big\rangle\mathrm{d}s \\ &\quad + (8+\kappa)\mathbf{E}\int_0^t e^{\lambda s}\big|\eta^x_{y_1,y_2}(s)\big|^{6+\kappa}\big\langle\eta^x_{y_1,y_2}(s), \nabla^2 b\big(X^x(s)\big)\big(\eta^x_{y_1}(s), \eta^x_{y_2}(s)\big)\big\rangle\mathrm{d}s \\ &\quad + \frac{1}{2}(8+\kappa)(7+\kappa)\mathbf{E}\int_0^t e^{\lambda s}\big|\eta^x_{y_1,y_2}(s)\big|^{6+\kappa} \\ &\qquad\qquad \times\big\|\nabla\sigma\big(X^x(s)\big)\eta^x_{y_1,y_2}(s) + \nabla^2\sigma\big(X^x(s)\big)\big(\eta^x_{y_1}(s), \eta^x_{y_2}(s)\big)\big\|^2_{HS}\,\mathrm{d}s \\ &\leq \lambda\mathbf{E}\int_0^t e^{\lambda s}|\eta^x_{y_1,y_2}(s)|^{8+\kappa}\mathrm{d}s \\ &\quad + \frac{1}{2}(8+\kappa)\mathbf{E}\int_0^t e^{\lambda s}|\eta^x_{y_1,y_2}(s)|^{6+\kappa}\Big[2\big\langle\eta^x_{y_1,y_2}(s), \nabla b\big(X^x(s)\big)\eta^x_{y_1,y_2}(s)\big\rangle \end{aligned}$$

$$
\begin{aligned}
&\qquad\qquad + (7+\kappa)(1+\varepsilon_0)\big\|\nabla\sigma\big(X^x(s)\big)\eta^x_{y_1,y_2}(s)\big\|^2_{HS}\Big]\mathrm{d}s\\
&+ K(\kappa)\mathbf{E}\int_0^t e^{\lambda s}\big|\eta^x_{y_1,y_2}(s)\big|^{7+\kappa}\big\|\nabla^2 b\big(X^x(s)\big)\big\|_{\otimes}\big|\eta^x_{y_1}(s)\big|\,\big|\eta^x_{y_2}(s)\big|\mathrm{d}s\\
&+ K(\kappa,\varepsilon_0)\mathbf{E}\int_0^t e^{\lambda s}\big|\eta^x_{y_1,y_2}(s)\big|^{6+\kappa}\big|\eta^x_{y_1}(s)\big|^2\big|\eta^x_{y_2}(s)\big|^2\mathrm{d}s.
\end{aligned}
$$

Further, taking $\varepsilon_0 \ll 1$ and using (2.6), we get

$$
\begin{aligned}
&\mathbf{E}\big(e^{\lambda t}\big|\eta^x_{y_1,y_2}(t)\big|^{8+\kappa}\big)\\
&\le \Big(\lambda - \Big(4+\frac{\kappa}{2}\Big)L_5\Big)\,\mathbf{E}\int_0^t e^{\lambda s}\big|\eta^x_{y_1,y_2}(s)\big|^{8+\kappa}\mathrm{d}s\\
&\quad + K(\kappa)\mathbf{E}\int_0^t e^{\lambda s}\big|\eta^x_{y_1,y_2}(s)\big|^{7+\kappa}\big\|\nabla^2 b\big(X^x(s)\big)\big\|_{\otimes}\big|\eta^x_{y_1}(s)\big|\,\big|\eta^x_{y_2}(s)\big|\mathrm{d}s\\
&\quad + K(\kappa)\mathbf{E}\int_0^t e^{\lambda s}\big|\eta^x_{y_1,y_2}(s)\big|^{6+\kappa}\big|\eta^x_{y_1}(s)\big|^2\big|\eta^x_{y_2}(s)\big|^2\mathrm{d}s. \qquad (4.16)
\end{aligned}
$$

It follows from the Young inequality $ab \le \varepsilon a^p + K(\varepsilon)b^q$ with $a,b\ge 0$, $1/p+1/q=1$, $p,q>1$ and the Hölder inequality that for any ε, $\varepsilon'>0$,

$$
\begin{aligned}
&\mathbf{E}\left(\big|\eta^x_{y_1,y_2}(s)\big|^{7+\kappa}\big\|\nabla^2 b\big(X^x(s)\big)\big\|_{\otimes}\big|\eta^x_{y_1}(s)\big|\,\big|\eta^x_{y_2}(s)\big|\right)\\
&\le \varepsilon\mathbf{E}\big|\eta^x_{y_1,y_2}(s)\big|^{8+\kappa} + K(\varepsilon)\mathbf{E}\left[\left(\big\|\nabla^2 b\big(X^x(s)\big)\big\|_{\otimes}\big|\eta^x_{y_1}(s)\big|\,\big|\eta^x_{y_2}(s)\big|\right)^{8+\kappa}\right]\\
&\le \varepsilon\mathbf{E}\big|\eta^x_{y_1,y_2}(s)\big|^{8+\kappa} + K(\varepsilon)\left(\mathbf{E}\big|\eta^x_{y_1}(s)\big|^{(8+\kappa)(2+\varepsilon')}\right)^{\frac{1}{2+\varepsilon'}}\left(\mathbf{E}\big|\eta^x_{y_2}(s)\big|^{(8+\kappa)(2+\varepsilon')}\right)^{\frac{1}{2+\varepsilon'}}\\
&\quad\times\left(\mathbf{E}\big\|\nabla^2 b\big(X^x(s)\big)\big\|_{\otimes}^{(8+\kappa)(1+2/\varepsilon')}\right)^{\frac{\varepsilon'}{2+\varepsilon'}}.
\end{aligned}
$$

Taking sufficiently small κ and ε', from Assumption 2.3, Proposition 2.1(1) and (4.11) it follows that for any $\varepsilon>0$,

$$
\begin{aligned}
&\mathbf{E}\left(|\eta^x_{y_1,y_2}(s)|^{7+\kappa}\big\|\nabla^2 b(X^x(s))\big\|_{\otimes}\big|\eta^x_{y_1}(s)\big|\,\big|\eta^x_{y_2}(s)\big|\right)\\
&\le \varepsilon\mathbf{E}\big|\eta^x_{y_1,y_2}(s)\big|^{8+\kappa} + K(\varepsilon)\big(1+|x|^{(8+\kappa)q'}\big)|y_1|^{8+\kappa}|y_2|^{8+\kappa}e^{-Ks}. \qquad (4.17)
\end{aligned}
$$

Similarity, for any $\varepsilon>0$,

$$
\begin{aligned}
&\mathbf{E}\left(\big|\eta^x_{y_1,y_2}(s)\big|^{6+\kappa}\big|\eta^x_{y_1}(s)\big|^2\big|\eta^x_{y_2}(s)\big|^2\right)\\
&\le \varepsilon\mathbf{E}\big|\eta^x_{y_1,y_2}(s)\big|^{8+\kappa} + K(\varepsilon)\left(\mathbf{E}\big|\eta^x_{y_1}(s)\big|^{2(8+\kappa)}\right)^{\frac{1}{2}}\left(\mathbf{E}\big|\eta^x_{y_2}(s)\big|^{2(8+\kappa)}\right)^{\frac{1}{2}}\\
&\le \varepsilon\mathbf{E}\big|\eta^x_{y_1,y_2}(s)\big|^{8+\kappa} + K(\varepsilon)|y_1|^{8+\kappa}|y_2|^{8+\kappa}e^{-Ks}. \qquad (4.18)
\end{aligned}
$$

Plugging (4.17)–(4.18) into (4.16), and taking sufficiently small κ_2,λ_2 and ε, one has

$$
\mathbf{E}\left(e^{\lambda_2 t}\big|\eta^x_{y_1,y_2}(t)\big|^{8+\kappa_2}\right) \le -K\mathbf{E}\int_0^t e^{\lambda_2 s}\big|\eta^x_{y_1,y_2}(s)\big|^{8+\kappa_2}\mathrm{d}s + K\big(1+|x|^{(8+\kappa_2)q'}\big)|y_1|^{8+\kappa}|y_2|^{8+\kappa},
$$

which produces (4.12).

Further, $\eta^x_{y_1,y_2,y_3}$ solves the following SODE:

$$\begin{cases} \mathrm{d}\eta^x_{y_1,y_2,y_3}(t) = \nabla b\big(X^x(t)\big)\eta^x_{y_1,y_2,y_3}(t)\mathrm{d}t + \nabla^2 b\big(X^x(t)\big)\big(\eta^x_{y_1}(t),\eta^x_{y_2,y_3}(t)\big)\mathrm{d}t \\ \qquad + \nabla^2 b\big(X^x(t)\big)\big(\eta^x_{y_2}(t),\eta^x_{y_1,y_3}(t)\big)\mathrm{d}t + \nabla^2 b\big(X^x(t)\big)\big(\eta^x_{y_3}(t),\eta^x_{y_1,y_2}(t)\big)\mathrm{d}t \\ \qquad + \nabla^3 b\big(X^x(t)\big)\big(\eta^x_{y_1}(t),\eta^x_{y_2}(t),\eta^x_{y_3}(t)\big)\mathrm{d}t + \nabla\sigma\big(X^x(t)\big)\eta^x_{y_1,y_2,y_3}(t)\mathrm{d}W(t) \\ \qquad + \nabla^2\sigma\big(X^x(t)\big)\big(\eta^x_{y_1}(t),\eta^x_{y_2,y_3}(t)\big)\mathrm{d}W(t) + \nabla^2\sigma\big(X^x(t)\big)\big(\eta^x_{y_2}(t),\eta^x_{y_1,y_3}(t)\big)\mathrm{d}W(t) \\ \qquad + \nabla^2\sigma\big(X^x(t)\big)\big(\eta^x_{y_3}(t),\eta^x_{y_1,y_2}(t)\big)\mathrm{d}W(t) \\ \qquad + \nabla^3\sigma\big(X^x(t)\big)\big(\eta^x_{y_1}(t),\eta^x_{y_2}(t),\eta^x_{y_3}(t)\big)\mathrm{d}W(t), \\ \eta^x_{y_1,y_2,y_3}(0) = 0. \end{cases}$$

By the same argument for deriving (4.16), using Itô formula, (2.6) and $\sigma \in \mathbf{C}^4_b(\mathbb{R}^d)$, we have that for any $\kappa, \lambda \in (0,1)$,

$$\begin{aligned} &\mathbf{E}\left(e^{\lambda t}\big|\eta^x_{y_1,y_2,y_3}(t)\big|^{4+\kappa}\right) \\ &\leq \left(\lambda - \left(2+\frac{\kappa}{2}\right)L_5\right)\mathbf{E}\int_0^t e^{\lambda s}\big|\eta^x_{y_1,y_2,y_3}(s)\big|^{4+\kappa}\mathrm{d}s \\ &\quad + K(\kappa)\mathbf{E}\int_0^t e^{\lambda s}\big|\eta^x_{y_1,y_2,y_3}(s)\big|^{3+\kappa}\big\|\nabla^2 b(X^x(s))\big\|_{\otimes} \\ &\qquad\qquad \times\Big(\big|\eta^x_{y_1}(s)\big|\,\big|\eta^x_{y_2,y_3}(s)\big| + \big|\eta^x_{y_2}(s)\big|\,\big|\eta^x_{y_1,y_3}(s)\big| + \big|\eta^x_{y_3}(s)\big|\,\big|\eta^x_{y_1,y_2}(s)\big|\Big)\mathrm{d}s \\ &\quad + K(\kappa)\mathbf{E}\int_0^t e^{\lambda s}\big|\eta^x_{y_1,y_2,y_3}(s)\big|^{3+\kappa}\big\|\nabla^3 b(X^x(s))\big\|_{\otimes}\big|\eta^x_{y_1}(s)\big|\,\big|\eta^x_{y_2}(s)\big|\,\big|\eta^x_{y_3}(s)\big|\mathrm{d}s \\ &\quad + K(\kappa)\mathbf{E}\int_0^t e^{\lambda s}\big|\eta^x_{y_1,y_2,y_3}(s)\big|^{2+\kappa}\Big(\big|\eta^x_{y_1}(s)\big|^2\big|\eta^x_{y_2,y_3}(s)\big|^2 + \big|\eta^x_{y_2}(s)\big|^2\big|\eta^x_{y_1,y_3}(s)\big|^2 \\ &\qquad\qquad + \big|\eta^x_{y_3}(s)\big|^2\big|\eta^x_{y_1,y_3}(s)\big|^2 + \big|\eta^x_{y_1}(s)\big|^2\big|\eta^x_{y_2}(s)\big|^2\big|\eta^x_{y_3}(s)\big|^2\Big)\mathrm{d}s \\ &=: \left(\lambda - \left(2+\frac{\kappa}{2}\right)L_5\right)\mathbf{E}\int_0^t e^{\lambda s}\big|\eta^x_{y_1,y_2,y_3}(s)\big|^{4+\kappa}\mathrm{d}s + I_1(t) + I_2(t) + I_3(t). \end{aligned} \tag{4.19}$$

It follows from the Young inequality, Hölder inequality, (4.11)-(4.12), Assumption 2.3 and Proposition 2.1(1) that for sufficiently small $\kappa, \varepsilon, \varepsilon'$,

$$\begin{aligned} &\mathbf{E}\left(\big|\eta^x_{y_1,y_2,y_3}(s)\big|^{3+\kappa}\big\|\nabla^2 b\big(X^x(s)\big)\big\|_{\otimes}\big|\eta^x_{y_{\chi(1)}}(s)\big|\,\big|\eta^x_{y_{\chi(2)},y_{\chi(3)}}(s)\big|\right) \\ &\leq \varepsilon\mathbf{E}\big|\eta^x_{y_1,y_2,y_3}(s)\big|^{4+\kappa} + K(\varepsilon)\left(\mathbf{E}\big|\eta^x_{y_{\chi(1)}}(s)\big|^{(4+\kappa)(2+\varepsilon')}\right)^{\frac{1}{2+\varepsilon'}}\left(\mathbf{E}\big|\eta^x_{y_{\chi(2)},\chi(3)}(s)\big|^{(4+\kappa)(2+\varepsilon')}\right)^{\frac{1}{2+\varepsilon'}} \\ &\quad \times\left(\mathbf{E}\big\|\nabla^2 b\big(X^x(s)\big)\big\|_{\otimes}^{(4+\kappa)(1+\frac{2}{\varepsilon'})}\right)^{\frac{\varepsilon'}{2+\varepsilon'}} \\ &\leq \varepsilon\mathbf{E}\big|\eta^x_{y_1,y_2,y_3}(s)\big|^{4+\kappa} + K(\varepsilon)\big(1+|x|^{2q'(4+\kappa)}\big)(|y_1||y_2||y_3|)^{4+\kappa}e^{-Ks}, \end{aligned}$$

where $(\chi(1),\chi(2),\chi(3))$ is any permutation of $(1,2,3)$. Thus, for $\kappa,\lambda,\varepsilon \ll 1$,

$$I_1(t) \leq K(\kappa)\varepsilon\mathbf{E}\int_0^t e^{\lambda s}\big|\eta^x_{y_1,y_2,y_3}(s)\big|^{4+\kappa}\mathrm{d}s + K(\kappa,\varepsilon)\big(1+|x|^{2q'(4+\kappa)}\big)(|y_1||y_2||y_3|)^{4+\kappa}. \tag{4.20}$$

Similarly, it can be verified that for $\kappa,\lambda,\varepsilon \ll 1$,

$$I_2(t) \leq K\varepsilon\mathbf{E}\int_0^t e^{\lambda s}\big|\eta^x_{y_1,y_2,y_3}(s)\big|^{4+\kappa}\mathrm{d}s + K(\varepsilon)\big(1+|x|^{q'(4+\kappa)}\big)(|y_1||y_2||y_3|)^{4+\kappa}, \tag{4.21}$$

$$I_3(t) \leq K\varepsilon \mathbf{E}\int_0^t e^{\lambda s}\big|\eta^x_{y_1,y_2,y_3}(s)\big|^{4+\kappa}\mathrm{d}s + K(\varepsilon)\big(1+|x|^{q'(4+\kappa)}\big)(|y_1||y_2||y_3|)^{4+\kappa}. \tag{4.22}$$

Plugging (4.20)-(4.22) into (4.19) yields (4.13). Finally, by means of an analogous proof for (4.13), we obtain (4.14). Thus, the proof is finished. □

Lemma 4.2. *Let Assumptions* 2.1-2.3 *hold and* $h \in \mathbf{C}_b^4(\mathbb{R}^d)$. *Let* φ *be the function defined by* (3.2). *Then, for any* $x \in \mathbb{R}^d$,

$$|\varphi(x)| \leq K(1+|x|), \tag{4.23}$$

$$\|\nabla^i\varphi(x)\|_\otimes \leq K\big(1+|x|^{(i-1)q'}\big), \quad i = 1,2,3,4. \tag{4.24}$$

Moreover, φ *is a solution to the Poisson equation*

$$\mathcal{L}\varphi = h - \pi(h). \tag{4.25}$$

One can prove (4.23) by means of (2.8), and prove (4.24) based on Lemma 4.1. In addition, (4.25) follows by the Kolmogorov equation $\mathcal{L}u(t,x) = (\partial/\partial t)u(t,x)$ with $u(t,x) := \mathbf{E}h(X^x(t))$, $t \geq 0$. We refer to [6, Lemma 2.8] for a similar proof.

4.2. Detailed proof

In this part, we give the proof of Theorem 3.2. As is mentioned previously, we will split $(1/\sqrt{\tau})(\Pi_{\tau,2}(h)-\pi(h))$ into a martingale difference series sum and a negligible remainder, based on the Poisson equation (4.25).

Proof of Theorem 3.2. For the convenience of notations, we denote $m = \tau^{-2}$ with sufficiently small τ. By (4.25), we have

$$\begin{aligned}\frac{1}{\sqrt{\tau}}\big(\Pi_{\tau,2}(h)-\pi(h)\big) &= \tau^{-\frac{1}{2}}\frac{1}{m}\sum_{k=0}^{m-1}\big(h\big(\bar{X}^x_k\big)-\pi(h)\big) = \tau^{\frac{3}{2}}\sum_{k=0}^{m-1}\mathcal{L}\varphi\big(\bar{X}^x_k\big)\\ &= \tau^{\frac{1}{2}}\sum_{k=0}^{m-1}\Big(\mathcal{L}\varphi\big(\bar{X}^x_k\big)\tau - \big(\varphi\big(\bar{X}^x_{k+1}\big)-\varphi\big(\bar{X}^x_k\big)\big)\Big) + \tau^{\frac{1}{2}}\big(\varphi\big(\bar{X}^x_m\big)-\varphi(x)\big).\end{aligned}$$

Lemma 4.2 enables us to apply the Taylor expansion for φ:

$$\begin{aligned}\varphi\big(\bar{X}^x_{k+1}\big)-\varphi\big(\bar{X}^x_k\big) &= \big\langle\nabla\varphi\big(\bar{X}^x_k\big),\Delta\bar{X}^x_k\big\rangle + \frac{1}{2}\Big\langle\nabla^2\varphi\big(\bar{X}^x_k\big),\Delta\bar{X}^x_k\big(\Delta\bar{X}^x_k\big)^\top\Big\rangle_{HS}\\ &\quad+\frac{1}{2}\int_0^1(1-\theta)^2\nabla^3\varphi\big(\bar{X}^x_k+\theta\Delta\bar{X}^x_k\big)\big(\Delta\bar{X}^x_k,\Delta\bar{X}^x_k,\Delta\bar{X}^x_k\big)\mathrm{d}\theta,\end{aligned}$$

where

$$\Delta\bar{X}^x_k := b\big(\bar{X}^x_{k+1}\big)\tau + \sigma\big(\bar{X}^x_k\big)\Delta W_k, \quad k = 0,1,\ldots,m.$$

It follows from (2.2) and the above formulas that

$$\frac{1}{\sqrt{\tau}}\big(\Pi_{\tau,2}(h)-\pi(h)\big) = \mathcal{H}_\tau + \mathcal{R}_\tau,$$

where $\mathcal{H}_\tau$ and $\mathcal{R}_\tau$ are given by

$$\mathcal{H}_\tau := -\tau^{\frac{1}{2}}\sum_{k=0}^{m-1}\big\langle\nabla\varphi\big(\bar{X}^x_k\big),\sigma\big(\bar{X}^x_k\big)\Delta W_k\big\rangle, \quad \mathcal{R}_\tau = \sum_{i=1}^{6}R_{\tau,i} \tag{4.26}$$

with

$$R_{\tau,1} := \tau^{\frac{1}{2}}\big(\varphi\big(\bar{X}_m^x\big) - \varphi(x)\big), \tag{4.27}$$

$$R_{\tau,2} := -\tau^{\frac{3}{2}}\sum_{k=0}^{m-1}\big\langle \nabla\varphi\big(\bar{X}_k^x\big), b\big(\bar{X}_{k+1}^x\big) - b\big(\bar{X}_k^x\big)\big\rangle, \tag{4.28}$$

$$R_{\tau,3} := \frac{1}{2}\tau^{\frac{1}{2}}\sum_{k=0}^{m-1}\Big\langle \nabla^2\varphi\big(\bar{X}_k^x\big), \sigma\big(\bar{X}_k^x\big)\big(\tau I_D - \Delta W_k \Delta W_k^\top\big)\sigma\big(\bar{X}_k^x\big)^\top\Big\rangle_{HS}, \tag{4.29}$$

$$R_{\tau,4} := -\frac{1}{2}\tau^{\frac{5}{2}}\sum_{k=0}^{m-1}\Big\langle \nabla^2\varphi\big(\bar{X}_k^x\big), b\big(\bar{X}_{k+1}^x\big)b\big(\bar{X}_{k+1}^x\big)^\top\Big\rangle_{HS}, \tag{4.30}$$

$$R_{\tau,5} := -\tau^{\frac{3}{2}}\sum_{k=0}^{m-1}\Big\langle \nabla^2\varphi\big(\bar{X}_k^x\big), b\big(\bar{X}_{k+1}^x\big)\big(\sigma\big(\bar{X}_k^x\big)\Delta W_k\big)^\top\Big\rangle_{HS}, \tag{4.31}$$

$$R_{\tau,6} := -\frac{1}{2}\tau^{\frac{1}{2}}\sum_{k=0}^{m-1}\int_0^1 (1-\theta)^2\nabla^3\varphi\big(\bar{X}_k^x + \theta\Delta\bar{X}_k^x\big)\big(\Delta\bar{X}_k^x, \Delta\bar{X}_k^x, \Delta\bar{X}_k^x\big)\mathrm{d}\theta. \tag{4.32}$$

By Lemmas 4.3-4.4 below and the Slutsky theorem,

$$\frac{1}{\sqrt{\tau}}\big(\Pi_{\tau,2}(h) - \pi(h)\big) \quad \xrightarrow{d} \quad \mathcal{N}\big(0, \pi(|\sigma^\top\nabla\varphi|^2)\big) \quad \text{as} \quad \tau \ \to \ 0,$$

and the proof is complete. □

Lemma 4.3. *Suppose that Assumptions* 2.1-2.3 *hold. Then for any* $x \in \mathbb{R}^d$*,*

$$\mathcal{H}_\tau \quad \xrightarrow{d} \quad \mathcal{N}\Big(0, \pi\big(\big|\sigma^\top\nabla\varphi\big|^2\big)\Big) \quad \text{as} \quad \tau \ \to \ 0.$$

The proof of Lemma 4.3 can be derived based on the exponential ergodicity of the BEM method, i.e. (4.9), and [25, Theorem 2.3], which is similar to that of [23, Lemma 4.2] and thus is omitted.

Lemma 4.4. *Suppose that Assumptions* 2.1-2.3 *hold. Then for any* $x \in \mathbb{R}^d$, $\mathcal{R}_\tau \xrightarrow{\mathbf{P}} 0$ *as* $\tau \to 0$.

Proof. We will prove $\lim_{\tau\to 0}\mathbf{E}|\mathcal{R}_\tau| = 0$ to obtain the conclusion. Hereafter, we denote by $\mathbf{E}_i(\cdot)$ the conditional expectation $\mathbf{E}(\,\cdot\,|\,\mathcal{F}_{t_i})$, $i = 0, 1, 2, \ldots$.

Estimate of $\mathcal{R}_{\tau,1}$. By Theorem 4.1, (4.23) and (4.27),

$$\mathbf{E}|\mathcal{R}_{\tau,1}| \le K\tau^{\frac{1}{2}}\Big(1 + \sup_{n\ge 0}\mathbf{E}\big|\bar{X}_n^x\big|\Big) \le K(x)\tau^{\frac{1}{2}}.$$

Estimate of $\mathcal{R}_{\tau,2}$. By means of (2.3), Assumption 2.3, Theorem 4.1, (4.24) and the Hölder inequality, we have that for any $p \ge 1$, $i = 2, 3, 4$ and $j = 1, 2$,

$$\sup_{k\ge 0}\mathbf{E}\big|b\big(\bar{X}_k^x\big)\big|^p \le K\Big(1 + \sup_{k\ge 0}\mathbf{E}\big|\bar{X}_k^x\big|^{pq}\Big) \le K(1 + |x|^{pq}), \tag{4.33}$$

$$\sup_{k\ge 0}\mathbf{E}\big\|\nabla^j b\big(\bar{X}_k^x\big)\big\|_\otimes^p \le K\Big(1 + \sup_{k\ge 0}\mathbf{E}\big|\bar{X}_k^x\big|^{pq'}\Big) \le K(1 + |x|^{pq'}), \tag{4.34}$$

$$\sup_{k\ge 0}\mathbf{E}\big\|\nabla^i\varphi\big(\bar{X}_k^x\big)\big\|_\otimes^p \le K\Big(1 + \sup_{k\ge 0}\mathbf{E}\big|\bar{X}_k^x\big|^{(i-1)pq'}\Big) \le K(1 + |x|^{(i-1)pq'}). \tag{4.35}$$

Noting that

$$b\big(\bar X_{k+1}^x\big)-b\big(\bar X_k^x\big)=\nabla b\big(\bar X_k^x\big)\Delta\bar X_k^x+\int_0^1(1-\theta)\nabla^2 b\big(\bar X_k^x+\theta\Delta\bar X_k^x\big)\big(\Delta\bar X_k^x,\Delta\bar X_k^x\big)\mathrm{d}\theta,$$

one obtains from (4.28) that

$$\begin{aligned}\mathcal{R}_{\tau,2}=&-\tau^{\frac32}\sum_{k=0}^{m-1}\big\langle\nabla\varphi\big(\bar X_k^x\big),\nabla b\big(\bar X_k\big)\sigma\big(\bar X_k^x\big)\Delta W_k\big\rangle\\&-\tau^{\frac52}\sum_{k=0}^{m-1}\big\langle\nabla\varphi\big(\bar X_k^x\big),\nabla b\big(\bar X_k^x\big)b\big(\bar X_{k+1}^x\big)\big\rangle\\&-\tau^{\frac32}\sum_{k=0}^{m-1}\int_0^1(1-\theta)\big\langle\nabla\varphi(\bar X_k^x),\nabla^2 b\big(\bar X_k^x+\theta\Delta\bar X_k^x\big)\big(\Delta\bar X_k^x,\Delta\bar X_k^x\big)\big\rangle\mathrm{d}\theta\\=:&\,\mathcal{R}_{\tau,2}^1+\mathcal{R}_{\tau,2}^2+\mathcal{R}_{\tau,2}^3.\end{aligned}$$

By the property of conditional expectations, for $i<j$,

$$\begin{aligned}&\mathbf{E}\Big[\big\langle\nabla\varphi\big(\bar X_i^x\big),\nabla b\big(\bar X_i^x\big)\sigma\big(\bar X_i^x\big)\Delta W_i\big\rangle\big\langle\nabla\varphi\big(\bar X_j^x\big),\nabla b\big(\bar X_j^x\big)\sigma\big(\bar X_j^x\big)\Delta W_j\big\rangle\Big]\\=&\,\mathbf{E}\Big[\big\langle\nabla\varphi\big(\bar X_i^x\big),\nabla b\big(\bar X_i^x\big)\sigma\big(\bar X_i^x\big)\Delta W_i\big\rangle\big\langle\nabla\varphi\big(\bar X_j^x\big),\nabla b\big(\bar X_j^x\big)\sigma\big(\bar X_j^x\big)\mathbf{E}_j\big(\Delta W_j\big)\big\rangle\Big]=0.\end{aligned}$$

The above relation, combined with the boundedness of σ, (4.24) and (4.34), gives

$$\begin{aligned}\mathbf{E}\big|\mathcal{R}_{\tau,2}^1\big|^2&=\tau^3\sum_{k=0}^{m-1}\mathbf{E}\big\langle\nabla\varphi\big(\bar X_k^x\big),\nabla b\big(\bar X_k^x\big)\sigma\big(\bar X_k^x\big)\Delta W_k\big\rangle^2\\&\le K\tau^4\sum_{k=0}^{m-1}\mathbf{E}\big|\nabla b\big(\bar X_k^x\big)\big|^2\le K(x)\tau^2.\end{aligned}$$

Applying the Hölder inequality, (4.24) and (4.33)-(4.34), we have

$$\mathbf{E}\big|\mathcal{R}_{\tau,2}^2\big|\le K\tau^{\frac52}\sum_{k=0}^{m-1}\Big(\mathbf{E}\big|\nabla b\big(\bar X_k^x\big)\big|^2\Big)^{\frac12}\Big(\mathbf{E}\big|b\big(\bar X_{k+1}^x\big)\big|^2\Big)^{\frac12}\le K(x)\tau^{\frac12}.$$

Further, for any $p\ge1$ and $k\ge0$, it follows from the Minkowski inequality, (4.33) and the boundedness of σ that

$$\big(\mathbf{E}\big|\Delta\bar X_k^x\big|^p\big)^{\frac1p}\le\tau\left(\mathbf{E}\big|b\big(\bar X_{k+1}^x\big)\big|^p\right)^{\frac1p}+K\big(\mathbf{E}|\Delta W_k|^p\big)^{\frac1p}\le K(1+|x|^q)\tau^{\frac12}.\tag{4.36}$$

This together with the Hölder inequality, Assumption 2.3 and Theorem 4.1 yields

$$\mathbf{E}\big|\mathcal{R}_{\tau,2}^3\big|\le K\tau^{\frac32}\sum_{k=0}^{m-1}\Big(\mathbf{E}\big|\Delta\bar X_k^x\big|^4\Big)^{\frac12}\Big(1+\Big(\mathbf{E}\big|\Delta\bar X_k^x\big|^{2q'}\Big)^{\frac12}+\Big(\mathbf{E}\big|\bar X_k^x\big|^{2q'}\Big)^{\frac12}\Big)\le K(x)\tau^{\frac12}.$$

In this way, we get

$$\mathbf{E}|\mathcal{R}_{\tau,2}|\le\big(\mathbf{E}\big|\mathcal{R}_{\tau,2}^1\big|^2\big)^{\frac12}+\mathbf{E}\big|\mathcal{R}_{\tau,2}^2\big|+\mathbf{E}\big|\mathcal{R}_{\tau,2}^3\big|\le K(x)\tau^{\frac12}.$$

Estimate of $\mathcal{R}_{\tau,3}$. Notice that for $i<j$,

$$\begin{aligned}&\mathbf{E}\Big[\Big\langle\nabla^2\varphi\big(\bar X_i^x\big),\sigma\big(\bar X_i^x\big)\big(\tau I_D-\Delta W_i\Delta W_i^\top\big)\sigma\big(\bar X_i^x\big)^\top\Big\rangle_{HS}\\&\quad\times\Big\langle\nabla^2\varphi\big(\bar X_j^x\big),\sigma\big(\bar X_j^x\big)\big(\tau I_D-\Delta W_j\Delta W_j^\top\big)\sigma\big(\bar X_j^x\big)^\top\Big\rangle_{HS}\Big]\\&=\mathbf{E}\Big[\Big\langle\nabla^2\varphi\big(\bar X_i^x\big),\sigma\big(\bar X_i^x\big)\big(\tau I_D-\Delta W_i\Delta W_i^\top\big)\sigma\big(\bar X_i^x\big)^\top\Big\rangle_{HS}\\&\quad\times\Big\langle\nabla^2\varphi\big(\bar X_j^x\big),\sigma\big(\bar X_j^x\big)\mathbf{E}_j\big(\tau I_D-\Delta W_j\Delta W_j^\top\big)\sigma\big(\bar X_j^x\big)^\top\Big\rangle_{HS}\Big]=0.\end{aligned}\tag{4.37}$$

Combining (4.29), (4.37), the boundedness of σ and (4.35), we arrive at

$$\begin{aligned}\mathbf{E}|\mathcal{R}_{\tau,3}|^2&=\frac{\tau}{4}\sum_{k=0}^{m-1}\mathbf{E}\left\langle\nabla^2\varphi\big(\bar X_k^x\big),\sigma\big(\bar X_k^x\big)\left(\tau I_D-\Delta W_k\Delta W_k^\top\right)\sigma\big(\bar X_k^x\big)^\top\right\rangle_{HS}^2\\&\le K\tau\sum_{k=0}^{m-1}\mathbf{E}\left(\|\nabla^2\varphi\big(\bar X_k^x\big)\|_{HS}^2\big(\tau^2+|\Delta W_k|^4\big)\right)\\&\le K\tau\sum_{k=0}^{m-1}\left(\mathbf{E}\big\|\nabla^2\varphi\big(\bar X_k^x\big)\big\|_{HS}^4\right)^{\frac12}\left(\tau^2+\big(\mathbf{E}|\Delta W_k|^8\big)^{\frac12}\right)\le K(x)\tau.\end{aligned}$$

Estimate of $\mathcal{R}_{\tau,4}$. By (4.30), (4.33), (4.35) and the Hölder inequality,

$$\mathbf{E}|R_{\tau,4}|\le K\tau^{\frac52}\sum_{k=0}^{m-1}\left(\mathbf{E}\big|\nabla^2\varphi\big(\bar X_k^x\big)\big|^2\right)^{\frac12}\left(\mathbf{E}\big|b\big(\bar X_{k+1}^x\big)\big|^4\right)^{\frac12}\le K(x)\tau^{\frac52}m\le K(x)\tau.$$

Estimate of $\mathcal{R}_{\tau,5}$. We decompose $\mathcal{R}_{\tau,5}$ (see (4.31)) into $\mathcal{R}_{\tau,5}=\mathcal{R}_{\tau,5}^1+\mathcal{R}_{\tau,5}^2$ with

$$\begin{aligned}\mathcal{R}_{\tau,5}^1&:=-\tau^{\frac32}\sum_{k=0}^{m-1}\left\langle\nabla^2\varphi\big(\bar X_k^x\big),\big(b\big(\bar X_{k+1}^x\big)-b\big(\bar X_k^x\big)\big)\left(\sigma(\bar X_k^x)\Delta W_k\right)^\top\right\rangle_{HS},\\\mathcal{R}_{\tau,5}^2&:=-\tau^{\frac32}\sum_{k=0}^{m-1}\left\langle\nabla^2\varphi\big(\bar X_k^x\big),b\big(\bar X_k^x\big)\big(\sigma\big(\bar X_k^x\big)\Delta W_k\big)^\top\right\rangle_{HS}.\end{aligned}$$

By the Hölder inequality, (4.35), (2.1), Theorem 4.1 and (4.36),

$$\mathbf{E}\big|\mathcal{R}_{\tau,5}^1\big|\le K\tau^{\frac32}m\sup_{k\ge0}\left(\mathbf{E}\big|\nabla^2\varphi\big(\bar X_k^x\big)\big|^3\right)^{\frac13}\left(\mathbf{E}|\Delta W_k|^3\right)^{\frac13}\left(\mathbf{E}\big|b\big(\bar X_{k+1}^x\big)-b\big(\bar X_k^x\big)\big|^3\right)^{\frac13}\le K(x)\tau^{\frac12}.$$

Similar to (4.37), one has that for $i<j$,

$$\mathbf{E}\left[\left\langle\nabla^2\varphi\big(\bar X_i^x\big),b\big(\bar X_i^x\big)\big(\sigma\big(\bar X_i^x\big)\Delta W_i\big)^\top\right\rangle_{HS}\cdot\left\langle\nabla^2\varphi\big(\bar X_j^x\big),b\big(\bar X_j^x\big)\big(\sigma\big(\bar X_j^x\big)\Delta W_j\big)^\top\right\rangle_{HS}\right]=0.$$

The above formula, combined with (4.33), (4.35) and the Hölder inequality, yields

$$\mathbf{E}\big|\mathcal{R}_{\tau,5}^2\big|^2=\tau^3\sum_{k=0}^{m-1}\mathbf{E}\left\langle\nabla^2\varphi\big(\bar X_k^x\big),b\big(\bar X_k^x\big)\big(\sigma\big(\bar X_k^x\big)\Delta W_k\big)^\top\right\rangle_{HS}^2\le K(x)\tau^2.$$

Thus,

$$\mathbf{E}|\mathcal{R}_{\tau,5}|\le\mathbf{E}\big|\mathcal{R}_{\tau,5}^1\big|+\left(\mathbf{E}\big|\mathcal{R}_{\tau,5}^2\big|^2\right)^{\frac12}\le K(x)\tau^{\frac12}.$$

Estimate of $\mathcal{R}_{\tau,6}$. Plugging

$$\Delta\bar{X}_k^x = b\big(\bar{X}_{k+1}^x\big)\tau + \sigma\big(\bar{X}_k^x\big)\Delta W_k$$

into (4.32) gives

$$\mathcal{R}_{\tau,6} = \sum_{i=1}^{4} \mathcal{R}_{\tau,6}^{i}$$

with

$$\mathcal{R}_{\tau,6}^1 := -\frac{\tau^{\frac{7}{2}}}{2}\sum_{k=0}^{m-1}\int_0^1 (1-\theta)^2\nabla^3\varphi\big(\bar{X}_k^x+\theta\Delta\bar{X}_k^x\big)\Big(b\big(\bar{X}_{k+1}^x\big), b\big(\bar{X}_{k+1}^x\big), b\big(\bar{X}_{k+1}^x\big)\Big)\mathrm{d}\theta,$$

$$\mathcal{R}_{\tau,6}^2 := -\frac{3\tau^{\frac{5}{2}}}{2}\sum_{k=0}^{m-1}\int_0^1 (1-\theta)^2\nabla^3\varphi\big(\bar{X}_k^x+\theta\Delta\bar{X}_k^x\big)\Big(b\big(\bar{X}_{k+1}^x\big), b\big(\bar{X}_{k+1}^x\big), \sigma\big(\bar{X}_k^x\big)\Delta W_k\Big)\mathrm{d}\theta,$$

$$R_{\tau,6}^3 := -\frac{3\tau^{\frac{3}{2}}}{2}\sum_{k=0}^{m-1}\int_0^1 (1-\theta)^2\nabla^3\varphi\big(\bar{X}_k^x+\theta\Delta\bar{X}_k^x\big)\Big(b\big(\bar{X}_{k+1}^x\big), \sigma\big(\bar{X}_k^x\big)\Delta W_k, \sigma\big(\bar{X}_k^x\big)\Delta W_k\Big)\mathrm{d}\theta,$$

$$R_{\tau,6}^4 := -\frac{\tau^{\frac{1}{2}}}{2}\sum_{k=0}^{m-1}\int_0^1 (1-\theta)^2\nabla^3\varphi\big(\bar{X}_k^x+\theta\Delta\bar{X}_k^x\big)\Big(\sigma\big(\bar{X}_k^x\big)\Delta W_k, \sigma\big(\bar{X}_k^x\big)\Delta W_k, \sigma\big(\bar{X}_k^x\big)\Delta W_k\Big)\mathrm{d}\theta.$$

Similar to the derivation of (4.35), one can use (4.24), (4.36) and Theorem 4.1 to get that for any $p \geq 1$ and $\tau < 1$,

$$\mathbf{E}\big\|\nabla^3\varphi\big(\bar{X}_k^x+\theta\Delta\bar{X}_k^x\big)\big\|_{\otimes}^p \leq K\big(1+|x|^{2pq'q}\big), \quad \theta\in[0,1]. \tag{4.38}$$

By (4.33), (4.38) and the Hölder inequality, one has

$$\mathbf{E}\big|\mathcal{R}_{\tau,6}^1\big| \leq K(x)\tau^{\frac{3}{2}}, \quad \mathbf{E}\big|\mathcal{R}_{\tau,6}^2\big| \leq K(x)\tau, \quad \mathbf{E}\big|\mathcal{R}_{\tau,6}^3\big| \leq K(x)\tau^{\frac{1}{2}}. \tag{4.39}$$

Further, applying the Taylor expansion for $\nabla^3\varphi$, we write

$$\mathcal{R}_{\tau,6}^4 = \mathcal{R}_{\tau,6}^{4,1} + \mathcal{R}_{\tau,6}^{4,2},$$

where

$$\mathcal{R}_{\tau,6}^{4,1} := -\frac{\tau^{\frac{1}{2}}}{6}\sum_{k=0}^{m-1}\nabla^3\varphi\big(\bar{X}_k^x\big)\Big(\sigma\big(\bar{X}_k^x\big)\Delta W_k, \sigma\big(\bar{X}_k^x\big)\Delta W_k, \sigma\big(\bar{X}_k^x\big)\Delta W_k\Big),$$

$$\mathcal{R}_{\tau,6}^{4,2} := -\frac{\tau^{\frac{1}{2}}}{2}\sum_{k=0}^{m-1}\int_0^1\int_0^1 \nabla^4\varphi\big(\bar{X}_k^x+r\theta\Delta\bar{X}_k^x\big) \times\Big(\sigma\big(\bar{X}_k^x\big)\Delta W_k, \sigma\big(\bar{X}_k^x\big)\Delta W_k, \sigma\big(\bar{X}_k^x\big)\Delta W_k, \Delta\bar{X}_k^x\Big)\mathrm{d}r\theta(1-\theta)^2\mathrm{d}\theta.$$

It can be shown that

$$\mathbf{E}\big|\mathcal{R}_{\tau,6}^{4,1}\big|^2 \leq K(x)\tau^2, \quad \mathbf{E}\big|\mathcal{R}_{\tau,6}^{4,2}\big| \leq K(x)\tau^{\frac{1}{2}}.$$

Thus, $\mathbf{E}|\mathcal{R}_{\tau,6}^4| \leq K(x)\tau^{1/2}$ for $\tau < 1$, which combined with (4.39) yields $\mathbf{E}|\mathcal{R}_{\tau,6}| \leq K(x)\tau^{1/2}$.

Combining the above estimates for $\mathcal{R}_{\tau,i}$, $i = 1,\ldots,6$, we obtain $\lim_{\tau\to 0}\mathbf{E}|\mathcal{R}_\tau| = 0$. This gives the desired conclusion. □

5. Numerical Experiments

In this section, we perform numerical experiments to verify our theoretical results. First, for a given test function h, we obtain the approximation of the ergodic limit $\pi(h)$ numerically by virtue of the fact $\lim_{t\to\infty}\mathbf{E}(h(X(t))) = \pi(h)$ (see (2.8)). Here, $\lim_{t\to\infty}\mathbf{E}(h(X(t)))$ is simulated by using the numerical solution $\{\bar{X}_n\}_{n\geq 0}$ of the BEM method. More precisely, let the step-size τ be small enough, n sufficiently large, and we use the Monte-Carlo method to simulate the expectation. Then we have

$$\lim_{t\to\infty}\mathbf{E}\big(h\big(X(t)\big)\big) \approx \frac{1}{M}\sum_{i=1}^{M} h\big(\bar{X}_n^i\big)$$

with $\{\bar{X}_n^i\}_{i=1}^M$ being M samplings of $\bar{X}_n$. Second, we verify the CLT for $\Pi_{\tau,\alpha}(h), \alpha \in (1,2]$. Denote

$$Z_{\tau,\alpha}(h) = \frac{1}{\tau^{\frac{\alpha-1}{2}}}\left(\frac{1}{\tau^{-\alpha}}\sum_{k=0}^{\tau^{-\alpha}-1} h\big(\bar{X}_k\big) - \pi(h)\right).$$

Then, the CLT shows that for any $f \in \mathbf{C}_b(\mathbb{R}^d)$,

$$\lim_{\tau\to 0}\mathbf{E}f\big(Z_{\tau,\alpha}(h)\big) = \int_{\mathbb{R}^d} f(x)\mathcal{N}\left(0, \pi\big(\big|\sigma^\top\nabla\varphi\big|^2\big)\right)(\mathrm{d}x).$$

We will numerically verify that $\mathbf{E}f(Z_{\tau,\alpha}(h))$ tends to some constant as τ decreases.

Example 5.1. Consider the following SODE:

$$\begin{cases}\mathrm{d}X(t) = -\big(X^3(t) + 8X(t)\big)\mathrm{d}t + \sin\big(X(t)\big)\mathrm{d}W(t), \quad t > 0,\\ X(0) = x_0 \in \mathbb{R}.\end{cases}$$

It is not difficult to verify that the coefficients of the above equation satisfy Assumptions 2.1-2.3. Thus, both Theorems 3.1 and 3.2 apply to the above SODE.

First, we numerically simulate the ergodic limit $\pi(h)$ using the aforementioned method. The expectation is realized by 10000 sample paths, using the Monte-Carlo method. Fig. 5.1 displays the evolution of the Monte-Carlo approximation of $\mathbf{E}h(\bar{X}_n)$ w.r.t. n with $h(x) = \sin(x) + 1$ starting from different initial values. It is observed that the ergodic limit is 1.

Second, we numerically verify

$$\lim_{\tau\to 0}\mathbf{E}f\big(Z_{\tau,2}(h)\big) = \int_{\mathbb{R}^d} f(x)\mathcal{N}\left(0, \pi\big(\big|\sigma^\top\nabla\varphi\big|^2\big)\right)(\mathrm{d}x).$$

For this end, we present the error in Table 5.1 between the Monte-Carlo approximation of $\mathbf{E}f(Z_{\tau,2}(h))$ and its numerical limit. As is shown in Table 5.1, for different test functions f, the Monte-Carlo approximation of $\mathbf{E}f(Z_{\tau,2}(h))$ will tend to its limit as τ decreases, which numerically verifies Theorem 3.2.

Finally, we numerically verify Theorem 3.1 and observe the impact of α on the convergence speed of $\mathbf{E}f(Z_{\tau,\alpha}(h))$. It is observed in Table 5.2 that for different α, the Monte-Carlo approximation of $\mathbf{E}f(Z_{\tau,\alpha}(h))$ will tend to its numerical limit, which validates Theorem 3.1. Fixing the step-size and comparing the errors for different α, we find that the bigger α is, the smaller the error is. This means a faster convergence speed of $Z_{\tau,\alpha}(h)$ in distribution if provided a larger α.

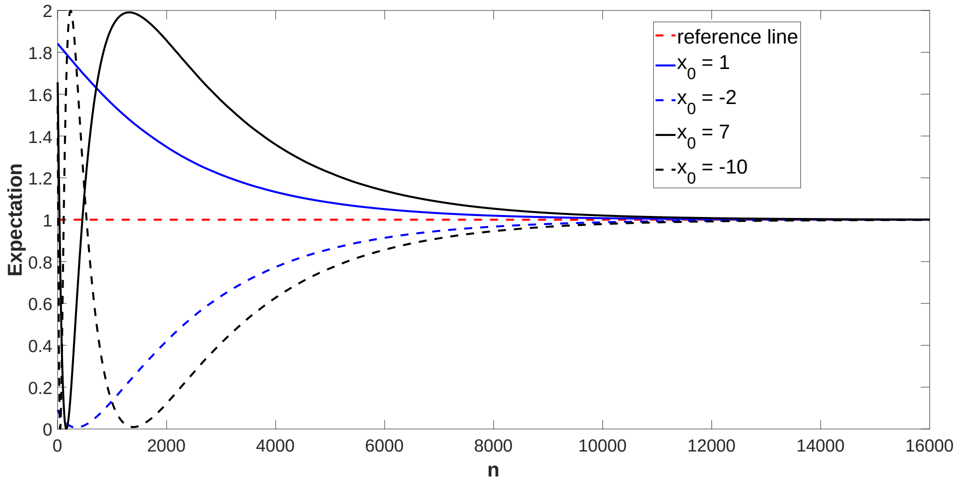

Fig. 5.1. Monte-Carlo approximation of $\mathbf{E}(\sin(\bar{X}_n)+1)$ versus n starting from different initial values with $\tau = 2^{-14}, M = 10000$.

Table 5.1: Error between the Monte-Carlo approximation of $\mathbf{E}f(Z_{\tau,2}(h))$ and its numerical limit for different step-sizes and f ($x_0 = 1$, $M = 10000$, $h(x) = \sin(x) + 1$).

error τ / f	0.05	0.045	0.04	0.035	0.03
$\cos(x)$	6.524181E-4	5.543409E-4	4.602158E-4	3.822945E-4	3.075336E-4
e^{-x^2+1}	3.544669E-3	3.012076E-3	2.500865E-3	2.077585E-3	1.67142E-3
error τ / f	0.025	0.02	0.015	0.01	0.005
$\cos(x)$	2.383274E-4	1.791496E-4	1.254459E-4	7.710036E-5	3.614681E-5
e^{-x^2+1}	1.29538E-3	9.737865E-4	6.819094E-4	4.191288E-4	1.965074E-4

Table 5.2: Error between the Monte-Carlo approximation of $\mathbf{E}f(Z_{\tau,\alpha}(h))$ and its numerical limit for different step-sizes and α ($x_0 = 1$, $M = 10000$, $f(x) = \cos(x^2)$ and $h(x) = \sin(x) + 1$).

error τ / α	0.01	0.0095	0.009	0.0085	0.008
$\alpha = 1.2$	1.239016E-4	1.054936E-4	9.069994E-5	7.741894E-5	6.201456E-5
$\alpha = 1.5$	2.001556E-5	1.664316E-5	1.233981E-5	1.003665E-5	7.46804E-6
$\alpha = 1.8$	3.294311E-6	2.481035E-6	1.83967E-6	1.290746E-6	9.055742E-7
$\alpha = 2$	9.909509E-7	7.218078E-7	4.981431E-7	3.457368E-7	2.232158E-7
τ / α	0.0075	0.007	0.0065	0.006	0.0055
$\alpha = 1.2$	5.100249E-5	4.004329E-5	3.126639E-5	2.322351E-5	1.504971E-5
$\alpha = 1.5$	5.478683E-6	3.875865E-6	2.483616E-6	1.455308E-6	6.260837E-7
$\alpha = 1.8$	5.952795E-7	3.556935E-7	2.045761E-7	9.201485E-8	2.59693E-8
$\alpha = 2$	1.334745E-7	7.590217E-8	3.755847E-8	1.42301E-8	3.12967E-9

Example 5.2. Consider the following two-dimensional SODE:

$$\begin{cases} \mathrm{d}X(t) = -\big(8X(t) + Y(t)\big)\mathrm{d}t + \mathrm{d}W_1(t), & t > 0, \\ \mathrm{d}Y(t) = -\big(X(t) + 8Y(t) + Y(t)^3\big)\mathrm{d}t + \mathrm{d}W_2(t), & t > 0, \\ \big(X(0), Y(0)\big) = (x_0, y_0) \in \mathbb{R}^2, \end{cases}$$

where W_1 and W_2 are two independent standard Brownian motions. It is not hard to check that the coefficients of the above equation satisfy Assumptions 2.1-2.3. Thus, it admits a unique invariant measure and $\Pi_{\tau,2}(h)$ $(h \in \mathbf{C}_b^4(\mathbb{R}^2))$ satisfies the CLT. In this experiment, we will numerically verify Theorem 6.1. Compared with Theorem 3.2, Theorem 6.1 relaxes the boundedness condition on $\nabla^i h$ $(i = 0, 1, \ldots, 4)$, allowing $\nabla^i h$ $(i = 0, 1, \ldots, 4)$ to grow polynomially.

First, we numerically compute the ergodic limit $\pi(h)$ with $h(x, y) = x^2 + y^2 + 1$. Fig. 5.2 displays the evolution of the Monte-Carlo approximation of $\mathbf{E}(\bar{X}_n^2 + \bar{Y}_n^2 + 1)$ with respect to n. We observe that the Monte-Carlo approximation of $\mathbf{E}(\bar{X}_n^2 + \bar{Y}_n^2 + 1)$, starting from different initial values, tends to the ergodic limit which approximately equals to 1.1243. This validates the unique ergodicity.

Second, we perform numerical experiments to present the convergence of the Monte-Carlo approximation of $\mathbf{E}f(Z_{\tau,2}(h))$ with $h(x, y) = x^2 + y^2 + 1$. We choose four test functions f to compute the error between the Monte-Carlo approximation of $\mathbf{E}f(Z_{\tau,2}(h))$ and its numerical limit. As is seen from Table 5.3, the Monte-Carlo approximation of $\mathbf{E}f(Z_{\tau,2}(h))$ will tend to

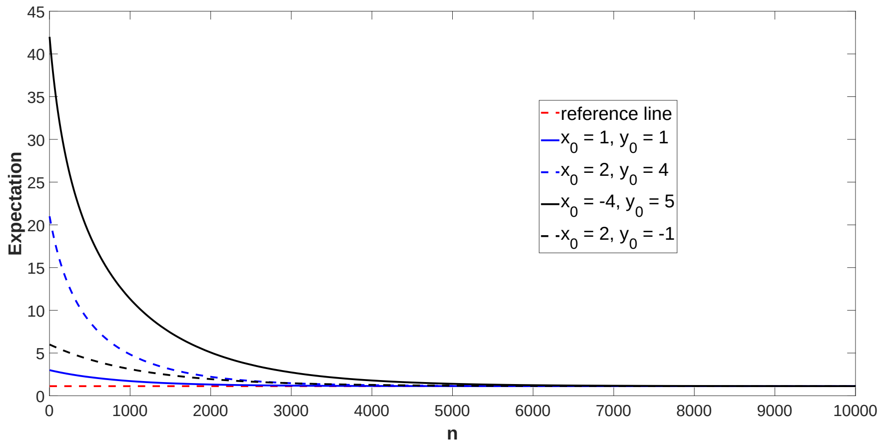

Fig. 5.2. Monte-Carlo approximation of $\mathbf{E}(\bar{X}_n^2 + \bar{Y}_n^2 + 1)$ versus n starting from different initial values with $\tau = 2^{-14}$, $M = 10000$.

Table 5.3: Error between the Monte–Carlo approximation of $\mathbf{E}f(Z_{\tau,2}(h))$ and its numerical limit for different step-sizes and f ($(x_0, y_0) = (1, 1)$, $M = 10000$ and $h(x, y) = x^2 + y^2 + 1$).

error τ / f	0.01	0.009	0.008	0.007	0.006
$cos(x)$	1.847026E-3	1.783466E-3	1.79837E-3	1.694887E-3	1.641087E-3
$cos(x^2)$	1.660341E-5	1.574741E-5	1.614185E-5	1.490238E-5	1.402271E-5
$cos(x^3)$	2.205107E-7	2.096067E-7	2.186593E-7	2.002097E-7	1.819951E-7
$cos(x^4)$	3.785467E-9	3.667431E-9	3.905721E-9	3.546581E-9	3.058579E-9

its limit as τ decreases. In addition, it seems that the faster $f(x)$ changes around $x = 0$, the faster $\mathbf{E}f(Z_{\tau,2}(h))$ converges.

6. Conclusions and Future Work

In this work, we prove the CLT for the temporal average of the BEM method, which characterizes the asymptotics of the BEM method in distribution. The drift coefficients of underlying SODEs are allowed to grow super-linearly. Different proof strategies are used for different deviation orders, which relies on the relationship between the deviation order and optimal strong order of the BEM method.

- Deviation order $\alpha \in (1, 2)$. In this case, we obtain the CLT for a class of numerical methods provided numerical methods has mean-square order $1/2$ in the infinite time horizon. As a byproduct, both the BEM method and the truncated EM method studied in [18] satisfy the CLT for $\alpha \in (1, 2)$.

- Deviation order $\alpha = 2$. In this case, we obtain the CLT of the temporal average of the BEM method. The arguments mainly depend on the Poisson equation and also require the p-th ($p > 2$) moment boundedness of the numerical solution (Theorem 4.1) and the exponential ergodicity of the BEM method (see (4.9)). We remark that the above arguments for the CLT of the numerical temporal average may apply to other numerical methods, if the corresponding p-th ($p > 2$) moment boundedness and the exponential ergodicity hold. For example, the truncated EM method in [18] will satisfy the CLT, once the exponential ergodicity is justified (Note that the p-th ($p > 2$) moment boundedness of the truncated EM method has been established in [18, Theorem 5.5]).

In fact, it is possible to weaken the conditions of Theorem 3.2, and we have the following result.

Theorem 6.1. *Let Assumption* 2.2 *hold with* $c_1 > 15L_1^2/2$ *replaced by* c_1 *being sufficiently large, and Assumption* 2.3 *hold. Assume that* σ *is globally Lipschitz. Let* $h \in \mathbf{C}^4(\mathbb{R}^d)$ *with* $\nabla^i h \in Poly(q'', \mathbb{R}^d)$, $i = 0, 1, \ldots, 4$. *Then for any* $x \in \mathbb{R}^d$,

$$\frac{1}{\sqrt{\tau}}\big(\Pi_{\tau,2}(h) - \pi(h)\big) \ \xrightarrow{d}\ \mathcal{N}\big(0, \pi(|\sigma^\top \nabla\varphi|^2)\big) \quad as \quad \tau \ \to\ 0.$$

Proof. The proof is similar to that of Theorem 3.2 and we only give its sketch. Note that the main difference lies in the assumptions on σ and h, compared with conditions of Theorem 3.2.

First, the assumptions of σ mainly make a difference on the proof of Theorem 4.1. Fortunately, we can still follow the same argument in Theorem 4.1 to give the p-th moment boundedness for the BEM method. Roughly speaking, in this case, (4.3) still holds. Similar to (4.7), we obtain

$$(1 + pc_1\tau)|\bar{X}_{n+1}^x|^{2p} \le \Big(\big|\bar{X}_n^x\big|^2 + 2\big\langle \bar{X}_n^x, \sigma\big(\bar{X}_n^x\big)\Delta W_n\big\rangle + K(\tau + |\Delta W_n|^2) + K\big|\bar{X}_n^x\big|^2|\Delta W_n|^2\Big)^p$$

due to the linear growth of σ. By a similar analysis for (4.8), one can show that

$$\mathbf{E}|\bar{X}_{n+1}^x|^{2p} \le \frac{(1 + A(p, D)\tau)}{(1 + pc_1\tau)}\mathbf{E}\big|\bar{X}_n^x\big|^{2p} + K(p)\frac{(1 + |x|^{2p-2})\tau}{(1 + pc_1\tau)}$$

for some $A(p, D) > 0$ dependent on p and D. Using the condition that c_1 is sufficiently large, one can finally obtain

$$\sup_{n\geq 0} \mathbf{E}\left|\bar{X}_n^x\right|^r \leq K(1+|x|^r)$$

for some r large enough.

Second, the assumptions on h mainly impact the regularity of φ as the solution to the Poisson equation. Following the arguments in the proof of Lemmas 4.1-4.2, we have that $\nabla^i \varphi \in Poly(L_0, \mathbb{R}^d)$, $i = 0, 1, \ldots, 4$ for some integer L_0 dependent on q' and q''.

Finally, other conclusions still hold on basis of the moment boundedness of $\{\bar{X}_n\}_{n\geq 0}$ and the regularity of φ. Thus, one can establish the CLT for $\Pi_{\tau,2}(h)$. □

When σ is Lipschitz or of super-linear growth, it is interesting to study how to prove the p-th ($p > 2$) moment boundedness of the BEM method in the infinite time horizon for a relatively small c_1. We will study this problem in the future.

Acknowledgements. We would like to thank the editors for their help and the referees for their constructive and insightful comments.

This work is supported by the National Natural Science Foundation of China (Grant Nos. 12201228, 12171047), and by the Fundamental Research Funds for the Central Universities (Grant No. 3004011142).

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**:4 (2014), 1600–1622.

[2] R.N. Bhattacharya, On the functional central limit theorem and the law of the iterated logarithm for Markov processes, *Z. für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, **60**:2 (1982), 185–201.

[3] C.E. Bréhier, D. Cohen, and T. Jahnke, Splitting integrators for stochastic Lie-Poisson systems, *Math. Comp.*, **92**:343 (2023), 2167–2216.

[4] M. Cai, S. Gan, and Y. Hu, Weak convergence of the backward Euler method for stochastic Cahn-Hilliard equation with additive noise, *Appl. Numer. Math.*, **188** (2023), 1–20.

[5] S. Cerrai, *Second Order PDE's in Finite and Infinite Dimension: A Probabilistic Approach*, in: *Lecture Notes in Mathematics*, Vol. 1762, Springer, 2001.

[6] C. Chen, T. Dang, J. Hong, and T. Zhou, CLT for approximating ergodic limit of SPDEs via a full discretization, *Stochastic Process. Appl.*, **157** (2023), 1–41.

[7] Z. Chen, S. Gan, and X. Wang, Mean-square approximations of Lévy noise driven SDEs with super-linearly growing diffusion and jump coefficients, *Discrete Contin. Dyn. Syst. Ser. B*, **24**:8 (2019), 4513–4545.

[8] I.P. Cornfeld, S.V. Fomin, and Ya. G. Sinaĭ, *Ergodic Theory*, in: *Grundlehren der Mathematischen Wissenschaften*, Vol. 245, Springer, 1982.

[9] Y. Geng, M. Song, and M. Liu, The convergence of truncated Euler-Maruyama method for stochastic differential equations with piecewise continuous arguments under generalized one-sided Lipschitz condition, *J. Comput. Math.*, **41**:4 (2023), 639–658.

[10] D.J. Higham, Stochastic ordinary differential equations in applied and computational mathematics, *IMA J. Appl. Math.*, **76**:3 (2011), 449–474.

[11] J. Hong, C. Huang, M. Kamrani, and X. Wang, Optimal strong convergence rate of a backward Euler type scheme for the Cox-Ingersoll-Ross model driven by fractional Brownian motion, *Stochastic Process. Appl.*, **130**:5 (2020), 2675–2692.

[12] 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**:6 (2017), 3006–3029.

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

[14] J. Hong, X. Wang, and L. Zhang, Numerical analysis on ergodic limit of approximations for stochastic NLS equation via multi-symplectic scheme, *SIAM J. Numer. Anal.*, **55**:1 (2017), 305–327.

[15] M. Hutzenthaler, A. Jentzen, and P.E. Kloeden, Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients, *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, **467**:2130 (2011), 1563–1576.

[16] J. Jacod and P. Protter, *Probability Essentials*, in: *Universitext*, Springer, 2004.

[17] L. Li and Y. Wang, A sharp uniform-in-time error estimate for stochastic gradient langevin dynamics, *arXiv:2207.09304v2*, 2022.

[18] X. Li, X. Mao, and G. Yin, Explicit numerical approximations for stochastic differential equations in finite and infinite horizons: Truncation methods, convergence in pth moment and stability, *IMA J. Numer. Anal.*, **39**:2 (2019), 847–892.

[19] W. Liu and X. Mao, Numerical stationary distribution and its convergence for nonlinear stochastic differential equations, *J. Comput. Appl. Math.*, **276** (2015), 16–29.

[20] W. Liu, X. Mao, and Y. Wu, The backward Euler-Maruyama method for invariant measures of stochastic differential equations with super-linear coefficients, *Appl. Numer. Math.*, **184** (2023), 137–150.

[21] Z. Liu, L^p-convergence rate of backward Euler schemes for monotone SDEs, *BIT*, **62**:4 (2022), 1573–1590.

[22] Z. Liu and Z. Qiao, Strong approximation of monotone stochastic partial differential equations driven by white noise, *IMA J. Numer. Anal.*, **40**:2 (2020), 1074–1093.

[23] J. Lu, Y. Tan, and L. Xu, Central limit theorem and self-normalized Cramér-type moderate deviation for Euler-Maruyama scheme, *Bernoulli*, **28**:2 (2022), 937–964.

[24] 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**:2 (2010), 552–577.

[25] D.L. McLeish, Dependent central limit theorems and invariance principles, *Ann. Probab.*, **2**:4 (1974), 620–628.

[26] G. Pagès and F. Panloup, Ergodic approximation of the distribution of a stationary diffusion: Rate of convergence, *Ann. Appl. Probab.*, **22**:3 (2012), 1059–1100.

[27] E. Pardoux and A. Yu. Veretennikov, On the Poisson equation and diffusion approximation 3, *Ann. Probab.*, **33**:3 (2005), 1111–1133.

[28] G.A. Pavliotis, *Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations*, in: *Texts in Applied Mathematics*, Vol. 60, Springer, 2014.

[29] N. Sandrić, Stability of the overdamped Langevin equation in double-well potential, *J. Math. Anal. Appl.*, **467**:1 (2018), 734–750.

[30] J. Schenker, Diffusion in the mean for an ergodic Schrödinger equation perturbed by a fluctuating potential, *Comm. Math. Phys.*, **339**:3 (2015), 859–901.

[31] D. Talay, Efficient numerical schemes for the approximation of expectations of functionals of the solution of a S.D.E., and applications, in: *Filtering and Control of Random Processes. Lecture Notes in Control and Information Sciences*, Vol. 61, 294–313, Springer, 1984.

[32] X. Wang, Y. Zhao, and Z. Zhang, Weak error analysis for strong approximation schemes of SDEs with super-linear coefficients, *arXiv:2303.14748v1*, 2023.

[33] Y. Yang and D. Jiang, Long-time behavior of a perturbed enzymatic reaction model under negative feedback process by white noise, *J. Math. Chem.*, **54**:4 (2016), 854–865.