Journal of Complexity 83 (2024) 101842

Contents lists available at ScienceDirect

Journal of Complexity

journal homepage: www.elsevier.com/locate/jco

Linear implicit approximations of invariant measures of semi-linear SDEs with non-globally Lipschitz coefficients ☆,☆☆

Chenxu Pang [a], Xiaojie Wang [a], Yue Wu [b,*]

[a] School of Mathematics and Statistics, HNP-LAMA, Central South University, Changsha, Hunan, PR China
[b] Department of Mathematics and Statistics, University of Strathclyde, Glasgow G1 1XH, UK

ARTICLE INFO

Article history:
Received 17 September 2023
Received in revised form 25 February 2024
Accepted 27 February 2024
Available online 13 March 2024

MSC:
60H35
37M25
65C30

Keywords:
Stochastic differential equations
Invariant measure
Weak convergence
Projected method
Kolmogorov equations

ABSTRACT

This article investigates the weak approximation towards the invariant measure of semi-linear stochastic differential equations (SDEs) under non-globally Lipschitz coefficients. For this purpose, we propose a linear-theta-projected Euler (LTPE) scheme, which also admits an invariant measure, to handle the potential influence of the linear stiffness. Under certain assumptions, both the SDE and the corresponding LTPE method are shown to converge exponentially to the underlying invariant measures, respectively. Moreover, with time-independent regularity estimates for the corresponding Kolmogorov equation, the weak error between the numerical invariant measure and the original one can be guaranteed with convergence of order one. In terms of computational complexity, the proposed ergodicity preserving scheme with the nonlinearity explicitly treated has a significant advantage over the ergodicity preserving implicit Euler method in the literature. Numerical experiments are provided to verify our theoretical findings.

☆ This work was supported by Natural Science Foundation of China (12071488, 12371417, 11971488). All three authors contributed equally to this work and the last two authors (X.W. and Y.W.) both serve as corresponding authors for the publication.
☆☆ Communicated by A. Jentzen.

* Corresponding author.
E-mail addresses: c.x.pang@csu.edu.cn (C. Pang), x.j.wang7@csu.edu.cn (X. Wang), yue.wu@strath.ac.uk (Y. Wu).

https://doi.org/10.1016/j.jco.2024.101842

Contents

1. Introduction

The primary objective of this paper is to study the invariant measures of semi-linear stochastic differential equations (SDEs) with multiplicative noise and their weak approximations. Given the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we consider the following $\mathbb{R}^d$-valued semi-linear SDEs of Itô type:

$$\begin{cases} \mathrm{d}X_t = AX_t + f(X_t)\,\mathrm{d}t + g(X_t)\,\mathrm{d}W_t, \quad t \in [0,\infty), \\ X_0 = x_0, \end{cases} \tag{1.1}$$

where $A \in \mathbb{R}^{d\times d}$ represents a negative definite matrix, $f\colon \mathbb{R}^d \to \mathbb{R}^d$ is the drift coefficient function, $g\colon \mathbb{R}^d \to \mathbb{R}^{d\times m}$ is the diffusion coefficient function, and $W_\cdot = \big(W_{1,\cdot}, \dots, W_{m,\cdot}\big)^T : [0,\infty) \times \Omega \to \mathbb{R}^m$ denotes the $\mathbb{R}^m$-valued standard Brownian motion with respect to $\{\mathcal{F}_t\}_{t\in[0,\infty)}$. Moreover, the initial data $x_0 : \Omega \to \mathbb{R}^d$ is assumed to be $\mathcal{F}_0$-measurable. This form covers a broad class of SDEs which are used to model real applications, for instance, the stochastic Ginzburg–Landau equation (see (6.1)), the mean-reverting model (see (6.2) or [19,12]) and space discretization of stochastic partial differential equations (SPDEs) (see (6.4) or [27,20]).

In this paper, we pay particular attention to a class of SDEs that, under certain conditions, converge exponentially to a unique invariant measure π. Evaluating the expectation of some function φ with respect to that invariant measure π is of great interest in mathematical biology, physics and Bayesian statistics:

$$\pi(\varphi) := \int_{\mathbb{R}^d} \varphi(x)\pi(\mathrm{d}x) \tag{1.2}$$

Generally speaking, it is not easy to obtain either the analytical solutions of SDEs or the explicit expression of the invariant measure. The study of the numerical approximations of π therefore receives increased attention. Previous research in this field typically focuses on stochastic differential equations (SDEs) characterized by coefficients that exhibit global Lipschitz continuity [24]. Such a strong condition is however rarely satisfied by SDEs from applications. On the other hand, conventional numerical tools lose their powers when attempting to simulate SDEs under relaxed conditions. For example, as claimed in [14,23], for a large class of SDEs with super-linear growth coefficients, the widely-used

Euler-Maruyama scheme leads to divergent numerical approximations in both finite and infinite time intervals. A natural question thus arises as to how to design the numerical scheme of the SDE (1.1) under a stiff condition caused by the linear operator in order to well approximate its invariant measure π and perform the error analysis.

Recent years have seen a proper growth of the literature on this topic, and it is worth mentioning that a majority of existing works analyze numerical approximations of invariant measures from SDEs via strong approximation error bounds (see [18,19,10,23,21,25]). The direct study of weak approximation errors (see [4,5,7,8]), which hold particular relevance in fields like financial engineering and statistics, is still in its early stages. In [8], the authors analyzed the backward Euler method of SDEs with piecewise continuous arguments (PCAs), where the drift is dissipative and the diffusion is globally Lipschitz, and recovered a time-independent convergence of order one. The author in [5] studied the tamed Euler scheme for ergodic SDEs with one-sided Lipschitz continuous drift coefficient and additive noise, and gave a moment bound that still depends on terminal time. We also mention that the authors in [1] provided new sufficient conditions for a numerical method to approximate with high order accuracy of the invariant measure of an ergodic SDE, independently of the weak order of accuracy of the method.

Each method exhibits drawbacks when approximating (1.2) weakly. Implicit methods by their nature have better stability but at a price of escalated complexity; explicit methods such as the tamed methods (see [15,28]) on the other hand may not preserve the long time property numerically since the taming factor has no positive lower bound. Even though the explicit projected method [26] does keep the asymptotic stability, it usually faces a severe stepsize restriction due to stability issues from solving stiff linear systems; to apply the truncated methods [18] to approximate the invariant distribution, one has to construct a strictly increasing function to control the growth of both drift and diffusion and to find its inverse version. Besides, the weak error analysis of such schemes is, to the best of our knowledge, still an open problem. We, therefore, aim to propose a family of linear-implicit methods that not only address the challenges posed by stiff systems but also preserve ergodicity and achieve weak convergence towards the invariant measure admitted by SDEs (1.1).

More formally, our scheme, called the linear-theta-implicit-projected Euler (LTPE) method, with a method parameter $\theta \in [0, 1]$ on a uniform timestep size h is given as follows,

$$Y_{n+1} - \theta A Y_{n+1} h = \mathscr{P}(Y_n) + (1-\theta) A \mathscr{P}(Y_n) h + f\big(\mathscr{P}(Y_n)\big) h + g\big(\mathscr{P}(Y_n)\big) \Delta W_n, \quad Y_0 = X_0, \tag{1.3}$$

where $\Delta W_n := W_{t_{n+1}} - W_{t_n}$, $n \in \mathbb{N}_0$, and $\mathscr{P} : \mathbb{R}^d \to \mathbb{R}^d$ is the projection operator denoted as

$$\mathscr{P}(x) := \min\left\{1, h^{-\frac{1}{2\gamma}} \|x\|^{-1}\right\} x, \quad \forall x \in \mathbb{R}^d, \tag{1.4}$$

with γ being determined in Assumption 2.4 later.

We point out that the scheme above is derived from the stochastic theta methods [22,29] used to deal with different models. Also, note that the parameter θ is pre-determined. Where there is a stiff system, we are able to treat the linear operator A implicitly (i.e. $\theta = 1$) without sacrificing numerical efficiency. And if one is working with the non-stiff system, using the explicit numerical scheme (i.e. $\theta = 0$) would be more appropriate. In addition, we follow the projected technique, previously used in [2,3] for SDEs in finite time interval, to prevent the nonlinear drift and diffusion from producing extraordinary large values. Under certain conditions, for $\forall \zeta \in L^{8\gamma+2}(\Omega, \mathbb{R}^d)$, where γ is given by Assumption 2.4, the projected process $\mathscr{P}(x)$ converges strongly to the original random variable ζ of order 2 (see Lemma 5.7 or [3]), i.e.

$$\|\mathscr{P}(\zeta) - \zeta\|_{L^2(\Omega, \mathbb{R}^d)} \leq C h^2. \tag{1.5}$$

Compared with the truncated method in [18], the implementation of the LTPE method in (1.3) is more straightforward, where the projection operator we have chosen depends only on the growth of the drift and diffusion. Besides, when facing with linear-stiff systems, our method with $\theta = 1$ may not suffer from too strict stepsize restriction.

To show the main result in Theorem 2.5, the derivations of the whole paper are organized in the following way: under Assumption 2.1-2.4, which is regarded as a kind of dissipative condition, we follow [9] to present the existence and uniqueness of the invariant measures of both SDEs (1.1) and the LTPE scheme (1.3), respectively in Theorem 3.1 and Theorem 4.1, and further establish their strongly mixing property, as detailed in (5.1) and (5.2), which paves the way to the subsequent error analysis. The main result regarding weak error analysis, presented in Theorem 5.8, is derived based on the associated Kolmogorov equation (5.5) of SDE (1.1). However, one is confronted with two main challenges. The first one is to get a couple of priori estimates that are independent of time and stepsize, including the uniform moment bounds of the LTPE method (1.3) and the time-independent regularity estimates of the Kolmogorov equation. Another one is the implicitness and discontinuity of the proposed LTPE method (1.3), which results in further difficulties in handling the weak error via the Kolmogorov equation. Different techniques are used to circumvent these obstacles. The discretization strategy based on the binomial theorem is adopted to obtain the uniform moment bounds of the LTPE scheme (see Lemma 4.3). Moreover, we revisit the mean square differetiability of random function (see Definition 5.1), which facilitates the exponential estimates for derivatives of SDEs (1.1) with respect to the initial state (see Lemma 5.2). Consequently, this leads to the time-independent regularity estimates of the Kolmogorov equation (see Lemma 5.3 and Corollary 5.5). To deal with possible implicitness and discontinuity of the LTPE scheme (1.3), we introduce its continuous-version $\{\mathbb{Z}^n(t)\}_{t\in[t_n,t_{n+1}]}$ with $n\in\mathbb{N}_0$, as

$$\begin{cases} \mathbb{Z}^n(t)=\mathbb{Z}^n(t_n)+F\big(\mathscr{P}(Y_n)\big)(t-t_n)+g\big(\mathscr{P}(Y_n)\big)(W_t-W_{t_n}),\\ \mathbb{Z}^n(t_n):=\mathscr{P}(Y_n)-\theta A\mathscr{P}(Y_n)h, \end{cases} \tag{1.6}$$

where $F(x):=Ax+f(x), \forall x\in\mathbb{R}^d$. We observe that $\mathbb{Z}^n(t_{n+1})=Y_{n+1}-\theta AY_{n+1}h$. In order to estimate the numerical approximation error of invariant measure, we introduce the function $u:[0,\infty)\times\mathbb{R}^d\to\mathbb{R}$ as

$$u(t,x):=\mathbb{E}\left[\varphi\big(X_t^x\big)\right]$$

based on the associated Kolmogorov equation (see Section 5 or [6, Chapter 1] for details). Following this, given any terminal time $T\in(0,\infty)$ such that $Nh=T$, $N\in\mathbb{N}$, one separates the weak error $\big|\mathbb{E}[\varphi(Y_N^{x_0})]-\mathbb{E}[\varphi(X_T^{x_0})]\big|$, i.e., $\big|\mathbb{E}\left[u(T,x_0)\right]-\mathbb{E}[u(0,Y_N^{x_0})]\big|$, into three parts,

$$\begin{aligned} \big|\mathbb{E}[\varphi(Y_N^{x_0})]-\mathbb{E}[\varphi(X_T^{x_0})]\big| &= \big|\mathbb{E}\left[u(T,x_0)\right]-\mathbb{E}[u(0,Y_N^{x_0})]\big|\\ &\le \underbrace{\big|\mathbb{E}[u(0,Y_N^{x_0})]-\mathbb{E}\left[u(0,Z_N)\right]\big|}_{:=\text{Error}_1}+\underbrace{\big|\mathbb{E}\left[u(T,Z_0)\right]-\mathbb{E}\left[u(T,x_0)\right]\big|}_{:=\text{Error}_2}\\ &\quad+\underbrace{\big|\mathbb{E}\left[u(0,Z_N)\right]-\mathbb{E}\left[u(T,Z_0)\right]\big|}_{:=\text{Error}_3}, \end{aligned}$$

where, for short, we denote $Z_n:=Y_n-\theta AY_nh$. Thanks to the fact that $Z_{n+1}=\mathbb{Z}^n(t_{n+1})$ and the time-independent regularity estimates of the Kolmogorov equation, one treats Error_1 and Error_2 directly and get $\max\{\text{Error}_1,\text{Error}_2\}=\mathcal{O}(h)$. For Error_3, we take full advantage of (1.6) and show further decomposition as

$$\begin{aligned} \text{Error}_3 \le &\left|\sum_{n=0}^{N-1}\mathbb{E}\left[u\big(T-t_n,\mathbb{Z}^n(t_n)\big)\right]-\mathbb{E}\big[u(T-t_n,Z_n)\big]\right|\\ &+\left|\sum_{n=0}^{N-1}\mathbb{E}\left[u\big(T-t_{n+1},\mathbb{Z}^n(t_{n+1})\big)\right]-\mathbb{E}\left[u\big(T-t_n,\mathbb{Z}^n(t_n)\big)\right]\right|. \end{aligned} \tag{1.7}$$

The first term on the right hand side of (1.7) is $\mathcal{O}(h)$ due to the regularity estimates of $u(t,\cdot)$ and (1.5); the second one, based on the Kolmogorov equation and Itô's formula, is proved to be $\mathcal{O}(h)$ (see more details in the proof of Theorem 5.8). Hence, we obtain the uniform weak error between the invariant measures, admitted by SDE (1.1) and the LTPE method (1.3), of order one eventually.

We summarize our main contributions:

- A family of linear implicit numerical methods, capable of dealing with stiff linear systems and inheriting invariant measures, is presented.
- Time-independent weak convergence between two invariant measures inherited by SDE (1.1) and LTPE scheme (1.3), respectively, is established under non-globally Lipschitz coefficients.

Also, we would like to point out that our semi-linear SDE setting (1.1) covers nonlinear SDEs when A vanishes (see Assumption 2.1), for which an explicit scheme (1.3) with $A = 0$ is proved to preserve the ergodicity.

Some numerical tests illustrate our findings in Section 6. Finally, the Appendix contains the detailed proof of auxiliary lemmas.

2. Settings and main result

Throughout this paper, we use $\mathbb{N}$ to denote the set of all positive integers and denote $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$. Let $d, m \in \mathbb{N}$ be given. Let $\|\cdot\|$ and $\langle \cdot, \cdot \rangle$ denote the Euclidean norm and the inner product of vectors in $\mathbb{R}^d$, respectively. We use $\max\{a, b\}$ and $\min\{a, b\}$ for the maximum and minimum values of between a and b respectively, and sometimes we also use a simplified notation $a \wedge b$ for $\min\{a, b\}$. Adopting the same notation as the vector norm, we denote $\|M\| := \sqrt{\mathrm{trace}(M^T M)}$ as the trace norm of a matrix $M \in \mathbb{R}^{d\times m}$, where M^T represents the transpose of a matrix M. Given a filtered probability space $\big(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,\infty)}, \mathbb{P}\big)$, we use $\mathbb{E}$ to mean the expectation and $L^r(\Omega, \mathbb{R}^d), r \geq 1$, to denote the family of $\mathbb{R}^d$-valued random variables ξ satisfying $\mathbb{E}[\|\xi\|^r] < \infty$. In addition, let $L^p(\mathbb{R}^d, \pi)$ be the space of all functions defined on $\mathbb{R}^d$ which are p-th integrable with respect to measure π. The diffusion coefficient function $g\colon \mathbb{R}^d \to \mathbb{R}^{d\times m}$ is frequently written as $g = (g_{i,j})_{d\times m} = (g_1, g_2, ..., g_m)$ for $g_{i,j}\colon \mathbb{R}^d \to \mathbb{R}$ and $g_j\colon \mathbb{R}^d \to \mathbb{R}^d, i \in \{1, 2, ..., d\}, j \in \{1, 2, ..., m\}$. Moreover, we introduce a new notation X_t^x for $t \in [0, \infty)$ denoting the solution of SDE (1.1) satisfying the initial condition $X_0^x = X_0 = x$. Also, let Y_n^x, $n \in \mathbb{N}_0$, be an approximation of the solution of SDE (1.1) with the initial point $Y_0^x = x$. In addition, denote by $C_b(\mathbb{R}^d)$ the Banach space of all uniformly continuous and bounded mappings $\phi : \mathbb{R}^d \to \mathbb{R}$ endowed with the norm $\|\phi\|_0 = \sup_{x\in\mathbb{R}^d} |\phi(x)|$.

For the vector-valued function $\mathbf{u} : \mathbb{R}^d \to \mathbb{R}^\ell$, $\mathbf{u} = (u_{(1)}, \ldots, u_{(\ell)})$, its first order partial derivative is considered as the Jacobian matrix as

$$D\mathbf{u} = \begin{pmatrix} \frac{\partial u_{(1)}}{\partial x_1} & \cdots & \frac{\partial u_{(1)}}{\partial x_d} \\ \vdots & \ddots & \vdots \\ \frac{\partial u_{(\ell)}}{\partial x_1} & \cdots & \frac{\partial u_{(\ell)}}{\partial x_d} \end{pmatrix}_{\ell\times d}.$$

For any $v_1 \in \mathbb{R}^d$, we have $D(\mathbf{u})v_1 \in \mathbb{R}^\ell$ and $D^2\mathbf{u}(v_1, v_2)$ defined as

$$D^2\mathbf{u}(v_1, v_2) := D\big(D(\mathbf{u})v_1\big)v_2, \quad \forall v_1, v_2 \in \mathbb{R}^d.$$

In the same manner, one defines

$$D^3\mathbf{u}(v_1, v_2, v_3) := D\Big(D\big(D(\mathbf{u})v_1\big)v_2\Big)v_3, \quad \forall v_1, v_2, v_3 \in \mathbb{R}^d$$

and for any integer $k \geq 3$ the k-th order partial derivatives of the function $\mathbf{u}$ are defined recursively. Given the Banach spaces $\mathbb{X}$ and $\mathbb{Y}$, we denote by $L(\mathbb{X}, \mathbb{Y})$ the Banach space of bounded linear operators from $\mathbb{X}$ into $\mathbb{Y}$. Then the partial derivatives of the function $\mathbf{u}$ are also regarded as the operators

$$D\mathbf{u}(\cdot)(\cdot) : \mathbb{R}^d \to L(\mathbb{R}^d, \mathbb{R}^\ell),$$
$$D^2\mathbf{u}(\cdot)(\cdot,\cdot) : \mathbb{R}^d \to L(\mathbb{R}^d, L(\mathbb{R}^d, \mathbb{R}^d)) \cong L(\mathbb{R}^d \otimes \mathbb{R}^d, \mathbb{R}^\ell)$$

and

$$D^3\mathbf{u}(\cdot)(\cdot,\cdot,\cdot):\mathbb{R}^d \to L(\mathbb{R}^d, L(\mathbb{R}^d, L(\mathbb{R}^d,\mathbb{R}^d))) \cong L((\mathbb{R}^d)^{\otimes 3},\mathbb{R}^\ell).$$

We remark that the partial derivatives of the scalar valued function are covered by the special case $\ell = 1$. For any $k \in \mathbb{N}$, let $C_b^k(\mathbb{R}^d)$ be the subspace of $C_b(\mathbb{R}^d)$ consisting of all functions with bounded partial derivatives $D^i\phi(x)$, $1 \le i \le k$, and with the norm $\|\phi\|_k := \|\phi\|_0 + \sum_{i=1}^k \sup_{x\in\mathbb{R}^d}\|D^i\phi(x)\|$. Further, let $\mathbf{1}_{\mathbb{B}}$ be the indicative function of a set $\mathbb{B}$. Denote $\frac{1}{0} := \infty$. To close this part, we let both C and C_A be the generic constant which are independent of the terminal time T and the stepsize, but more specially, the notation C_A further depends on the matrix A.

We present the following assumptions required to establish our main result.

Assumption 2.1. Assume the matrix $A \in \mathbb{R}^{d\times d}$ is self-adjoint and semi negative definite.

Assumption 2.1 immediately implies that there exists a sequence of non-decreasing positive real numbers $\{\lambda_i\}_{i=1}^d$ with $0\le\lambda_1 \le \lambda_2 \le \cdots \le \lambda_d < \infty$ and an orthonormal basis $\{e_i\}_{i\in\{1,\dots,d\}}$ such that $Ae_i = -\lambda_i e_i$, $i \in \{1,\dots,d\}$. Moreover, one also obtains

$$\langle x-y, A(x-y)\rangle \le -\lambda_1\|x-y\|^2, \quad \forall x,y\in\mathbb{R}^d.$$

Setting $y = 0$ leads to

$$\langle x, Ax\rangle \le -\lambda_1\|x\|^2, \quad \forall x\in\mathbb{R}^d. \tag{2.1}$$

We mention that the matrix A can vanish in the above setting, i.e. $\lambda_1 = \lambda_2 = \cdots = \lambda_d = 0$, so that the semi-linear SDEs (1.1) reduce into general non-linear SDEs, for which an explicit scheme given by (1.3) with $A = 0$ is proved to preserve the ergodicity. In addition, we need a coercivity condition put on nonlinear coefficients of SDEs.

Assumption 2.2. (Coercivity condition) For some $p_0 \in [1,\infty)$, there exist some constants $L_1 \in \mathbb{R}$ and $C_\star \in [0,\infty)$ such that,

$$2\langle x, f(x)\rangle + (2p_0-1)\|g(x)\|^2 \le L_1\|x\|^2 + C_\star, \quad \forall x\in\mathbb{R}^d.$$

Assumption 2.3. (Coupled monotoncity condition) For some $p_1 \in (1,\infty)$, there exists a constant $L_2 \in \mathbb{R}$ such that,

$$2\langle x-y, f(x)-f(y)\rangle + (2p_1-1)\|g(x)-g(y)\|^2 \le L_2\|x-y\|^2, \quad \forall x,y\in\mathbb{R}^d.$$

We remark that Assumption 2.3 implies Assumption 2.2 with $L_1 = L_2 + \epsilon$, $p_0 = p_1 - \epsilon$ and $C_\star = C(\epsilon)$ depending on ϵ, $f(0)$ and $g(0)$, where $\epsilon > 0$ is a sufficiently small constant. Note that Assumption 2.3 is equivalent to the following expression

$$2\langle x-y, f(x)-f(y)\rangle + (2p_1-1)\sum_{j=1}^m \left\|g_j(x)-g_j(y)\right\|^2 \le L_2\|x-y\|^2, \quad \forall x,y\in\mathbb{R}^d.$$

Thanks to Assumptions 2.1-2.3, one obtains that SDE (1.1) possesses a unique solution with continuous sample paths. Before proceeding on, we denote a mapping $\mathcal{P}_\cdot(\cdot,\cdot):\mathbb{R}\times\mathbb{R}^d\times\mathbb{R}^d \to [1,\infty)$ as

$$\mathcal{P}_{\bar\gamma}(x,\tilde x) := \max\left\{1, (1+\|x\|+\|\tilde x\|)^{\bar\gamma}\right\}, \quad \forall\bar\gamma\in\mathbb{R},\ \forall x,\tilde x\in\mathbb{R}^d.$$

In particular, let $\mathcal{P}_\cdot(\cdot):\mathbb{R}\times\mathbb{R}^d \to [1,\infty)$ be defined as

$$\mathcal{P}_{\bar\gamma}(x) := \mathcal{P}_{\bar\gamma}(x,0) = \max\left\{1, (1+\|x\|)^{\bar\gamma}\right\}, \quad \forall\bar\gamma\in\mathbb{R},\ \forall x\in\mathbb{R}^d. \tag{2.2}$$

Obviously, these mappings are non-decreasing with respect to $\bar{\gamma}$. Moreover, we require that the coefficients f and g have continuous partial derivatives up to the third order. The corresponding assumption is presented as below.

Assumption 2.4. (Polynomial growth of drift and diffusion) Assume that $f:\mathbb{R}^d \to \mathbb{R}^d$ and $g_j:\mathbb{R}^d \to \mathbb{R}^d$, $j \in \{1,\dots,m\}$, have all continuous derivatives up to order 3. Then there exist some positive constant $\gamma \in [1,\infty)$ such that

$$\left\| D^3 f(x)(v_1,v_2,v_3) \right\| \le C\mathcal{P}_{\gamma-3}(x) \cdot \|v_1\| \cdot \|v_2\| \cdot \|v_3\|, \quad \forall x, v_1, v_2, v_3 \in \mathbb{R}^d,$$
$$\left\| D^3 g_j(x)(v_1,v_2,v_3) \right\|^2 \le C\mathcal{P}_{\gamma-5}(x) \cdot \|v_1\|^2 \cdot \|v_2\|^2 \cdot \|v_3\|^2, \quad \forall x, v_1, v_2, v_3 \in \mathbb{R}^d.$$

Assumption 2.4 is regarded as a kind of polynomial growth conditions and in proofs which follow we will need some implications of this assumption. It follows immediately that,

$$\left\| D^2 f(x)(v_1,v_2) - D^2 f(\tilde{x})(v_1,v_2) \right\| \le C\mathcal{P}_{\gamma-3}(x,\tilde{x}) \cdot \|x-\tilde{x}\| \cdot \|v_1\| \cdot \|v_2\|, \ \forall x, \tilde{x}, v_1, v_2 \in \mathbb{R}^d,$$

and

$$\left\| D^2 f(x)(v_1,v_2) \right\| \le C\mathcal{P}_{\gamma-2}(x) \cdot \|v_1\| \cdot \|v_2\|, \quad \forall x, v_1, v_2 \in \mathbb{R}^d,$$

which in turns gives

$$\left\| Df(x)v_1 - Df(\tilde{x})v_1 \right\| \le C\mathcal{P}_{\gamma-2}(x,\tilde{x}) \cdot \|x-\tilde{x}\| \cdot \|v_1\|, \quad \forall x, \tilde{x}, v_1 \in \mathbb{R}^d,$$
$$\|Df(x)v_1\| \le C(1+\|x\|)^{\gamma-1}\|v_1\|, \quad \forall x, v_1 \in \mathbb{R}^d,$$

and

$$\begin{aligned} \left\| f(x) - f(\tilde{x}) \right\| &\le C_1(1+\|x\|+\|\tilde{x}\|)^{\gamma-1}\|x-\tilde{x}\|, \quad \forall x, \tilde{x} \in \mathbb{R}^d, \\ \|f(x)\| &\le C_2(1+\|x\|)^{\gamma}, \quad \forall x \in \mathbb{R}^d. \end{aligned} \tag{2.3}$$

Following the same idea, Assumption 2.4 also ensures, for $j \in \{1,\dots,m\}$,

$$\begin{aligned} \left\| D^2 g_j(x)(v_1,v_2) - D^2 g_j(\tilde{x})(v_1,v_2) \right\|^2 &\le C\mathcal{P}_{\gamma-5}(x,\tilde{x}) \cdot \|x-\tilde{x}\|^2 \cdot \|v_1\|^2 \cdot \|v_2\|^2, \\ &\qquad \forall x, \tilde{x}, v_1, v_2 \in \mathbb{R}^d, \\ \left\| D^2 g_j(x)(v_1,v_2) \right\|^2 &\le C\mathcal{P}_{\gamma-3}(x) \cdot \|v_1\|^2 \cdot \|v_2\|^2, \quad \forall x, v_1, v_2 \in \mathbb{R}^d. \end{aligned} \tag{2.4}$$

This in turns gives, for $j \in \{1,\dots,m\}$,

$$\left\| Dg_j(x)v_1 - Dg_j(\tilde{x})v_1 \right\|^2 \le C\mathcal{P}_{\gamma-3}(x,\tilde{x}) \cdot \|x-\tilde{x}\|^2 \cdot \|v_1\|^2, \quad \forall x, \tilde{x}, v_1 \in \mathbb{R}^d,$$
$$\left\| Dg_j(x)v_1 \right\|^2 \le C(1+\|x\|)^{\gamma-1}\|v_1\|^2, \quad \forall x, v_1 \in \mathbb{R}^d,$$

and

$$\left\| g_j(x) - g_j(\tilde{x}) \right\|^2 \le C(1+\|x\|+\|\tilde{x}\|)^{\gamma-1}\|x-\tilde{x}\|^2, \quad \forall x, \tilde{x} \in \mathbb{R}^d,$$
$$\left\| g_j(x) \right\|^2 \le C(1+\|x\|)^{\gamma+1}, \quad \forall x \in \mathbb{R}^d.$$

We remark that Assumptions 2.1-2.4 enable us to cover a broad class of SDEs with non-globally Lipschitz coefficients, which do not have closed-form solutions in general.

Now we are fully prepared to state the main result of this article as follows,

Theorem 2.5. *(Main result) Let Assumptions 2.1-2.4 hold with $p_0 \geq \max\{4\gamma+1, 5\gamma-4\}$ and $2\lambda_1 > \max\{L_1, L_2\}$ and consider SDE (1.1). Given $p \in [1, p_0) \cap \mathbb{N}$ and method parameter $\theta \in [0,1]$, let h be the uniform timestep satisfying*

$$h \in \left(0, \min\left\{\frac{1}{2(1-\theta)\lambda_1}, \frac{p_0-p}{(1-\theta)(2p_0-p-1)\lambda_1}, \frac{1}{(1-\theta)\lambda_d}, \frac{2\lambda_1-L_2}{4(1-\theta)^2\lambda_d^2}, \frac{(2\lambda_1-L_2)^\gamma}{(2\lambda_f)^{2\gamma}}, 1\right\}\right),$$

where $\lambda_f := 3C_1$, and C_1 is a constant depending only on the drift f, determined by (2.3). Then the SDE (1.1) and the corresponding LTPE scheme (1.3) method converge exponentially to a unique invariant measure, denoted by π and π_h, respectively. Moreover, for every test function $\varphi \in C_b^3(\mathbb{R}^d)$,

$$\left| \int_{\mathbb{R}^d} \varphi(x)\pi(dx) - \int_{\mathbb{R}^d} \varphi(x)\pi_h(dx) \right| \leq C_A h.$$

This theorem is divided into three parts as follows:

- Existence and uniqueness of an invariant measure of SDE (1.1).
- Existence and uniqueness of an invariant measure of the LTPE scheme (1.3).
- Time-independent weak error analysis between SDE (1.1) and the LTPE scheme (1.3).

In the following, more details of each part will be shown.

3. Invariant measure of the semi-linear SDE

Indeed, we show the following result.

Theorem 3.1. *Let Assumptions 2.1-2.3 be fulfilled with $2\lambda_1 > \max\{L_1, L_2\}$, given $\varphi \in C_b^1(\mathbb{R}^d)$, then the semi-linear SDE $\{X_t^{x_0}\}_{t\in[0,\infty)}$ in (1.1), with the initial condition $X_0 = x_0$, admits a unique invariant measure π and there exists some positive constant $c_1 \in (0, 2\lambda_1 - L_2]$ such that*

$$\left| \mathbb{E}\left[\varphi\left(X_t^{x_0}\right)\right] - \int_{\mathbb{R}^d} \varphi(x)\pi(dx) \right| \leq \|\varphi\|_1 e^{-\frac{c_1}{2}t}\left(1 + \mathbb{E}\left[\|x_0\|^2\right]\right), \quad \forall t \in [0,\infty).$$

With the condition $2\lambda_1 > \max\{L_1, L_2\}$, SDE (1.1) is regarded as a dissipative system. We follow the standard way, as shown in [9], to prove the existence and uniqueness of the invariant measure inherited by such systems. For completeness, we outline the central idea in the proof of Theorem 3.1 while the detailed proof of the following lemmas is found in Appendix.

It is desirable to consider SDE (1.1) with a negative initial time, that is,

$$\begin{cases} dX_t = AX_t + f(X_t)\,dt + g(X_t)\,d\widetilde{W}_t, & t \geq -\iota, \\ X_{-\iota} = x_0, \end{cases} \tag{3.1}$$

where $\iota \geq 0$, $\widetilde{W}_t$ is specified in the following way. Let $\overline{W}_t$ be another Brownian motion independent of W_t defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and define

$$\widetilde{W}_t = \begin{cases} W_t, & t \geq 0 \\ \overline{W}_t, & t < 0 \end{cases}$$

with the filtration $\widetilde{\mathcal{F}}_t := \sigma\{\widetilde{W}_s, s \leq t\}$, $t \in \mathbb{R}$. In what follows, we write $X_t^{s,x}$ in lieu of X_t to highlight the initial value $X_s = x$.

Before moving on, we introduce a useful lemma, which is a slight generalization of Lemma 8.1 in [16], as below,

Lemma 3.2. *If functions $r(\cdot) : [\tau, \infty) \to [0, \infty)$ and $m(\cdot) : [\tau, \infty) \to \mathbb{R}$, where $\tau \in \mathbb{R}$, are continuous, and if*

$$r(t) \le r(s) - \widetilde{c} \int_s^t r(u)du + \int_s^t m(u)du, \quad \tau \le s \le t < \infty,$$

where $\widetilde{c}$ is a positive constant, then

$$r(t) \le r(\tau) + \int_\tau^t e^{-\widetilde{c}(t-u)} m(u)du.$$

The proof of Lemma 3.2 has been shown in [11]. It is time to present the uniform moment bounds of the SDE (3.1).

Lemma 3.3. *(Uniform moment bounds of semi-linear SDEs.) Let the semi-linear SDEs $\{X_t^{-\iota,x_0}\}_{t\ge-\iota}$ in (3.1) satisfy Assumptions 2.1, 2.2 with $2\lambda_1 > L_1$. Then, for any $p \in [1, p_0]$ and $t \in [0, \infty)$,*

$$\mathbb{E}\left[\left\| X_t^{-\iota,x_0} \right\|^{2p}\right] \le C < \infty.$$

The proof of Lemma 3.3 is found in Appendix A.1. Note that Lemma 3.3 covers the case $p \in [0, 1)$ due to the Hölder inequality. Following Lemma 3.3, we obtain the contractive property of SDE (1.1) as follows,

Lemma 3.4. *(Contractivity of semi-linear SDEs.) Consider the pair of solutions of the semi-linear SDE (3.1), $X_t^{-\iota,x_{0,(1)}}$ and $X_t^{-\iota,x_{0,(2)}}$, driven by the same Brownian motion but with different initial state $X_{-\iota}^{-\iota,x_{0,(1)}} = x_0^{(1)}$, $X_{-\iota}^{-\iota,x_{0,(2)}} = x_0^{(2)}$. Let Assumptions 2.1, 2.3 hold with $2\lambda_1 > L_2$, then, there exists a constant $c_1 \in (0, 2\lambda_1 - L_2]$ such that, for any $p \in [1, p_1]$, $t \ge -\iota$,*

$$\mathbb{E}\left[\left\| X_t^{-\iota,x_{0,(1)}} - X_t^{-\iota,x_{0,(2)}} \right\|^{2p}\right] \le e^{-c_1 p(t+\iota)} \mathbb{E}\left[\left\| x_0^{(1)} - x_0^{(2)} \right\|^{2p}\right].$$

The proof of Lemma 3.4 is provided in Appendix A.2. The next Lemma is a direct consequence of Lemma 3.3 and Lemma 3.4.

Lemma 3.5. *Consider the semi-linear SDE in (3.1) satisfying Assumptions 2.1, 2.3 hold with $2\lambda_1 > \max\{L_1, L_2\}$. Let $X_t^{-s_1,x_0}$ and $X_t^{-s_2,x_0}$ with $s_1, s_2 > 0$ satisfying $-s_1 < -s_2 \le t < \infty$, be the solutions of SDE (3.1) at time t starting from the same point x_0 but at different times. Then, for any $p \in [1, p_0]$, there exists some constant $c_2 \in (0, 2\lambda_1 - L_2]$ such that*

$$\mathbb{E}\left[\| X_t^{-s_1,x_0} - X_t^{-s_2,x_0} \|^{2p}\right] \le C e^{-c_2 p(t+s_2)} \mathbb{E}\left[\left(1 + \|x_0\|^2\right)^p\right].$$

The proof of Lemma 3.5 is postponed to Appendix A.3. Equipped with the previously derived lemmas, it is not hard to show Theorem 3.1. To be precise, recalling Lemma 3.5, by sending s_1 to infinity, one directly observes that $\{X_0^{-s,x_0}\}_{s>0}$ is a Cauchy sequence in $L^2(\Omega, \mathbb{R}^d)$ and there exists ϑ^{x_0} in $L^2(\Omega, \mathbb{R}^d)$ such that

$$\vartheta^{x_0} := \lim_{s_1\to\infty} X_0^{-s_1,x_0}.$$

Using Lemma 3.5 again yields

$$\mathbb{E}\left[\| X_0^{-s_2,x_0} - \vartheta^{x_0} \|^2\right] = \lim_{s_1\to\infty} \mathbb{E}\left[\| X_0^{-s_2,x_0} - X_0^{-s_1,x_0} \|^2\right] \le e^{-c_2 s_2} \mathbb{E}\left[1 + \|x_0\|^2\right].$$

By Lemma 3.4, we know ϑ^{x_0} is independent of x_0, i.e.

$$\mathbb{E}\left[\left\|\vartheta^{x_0}-\vartheta^{x_1}\right\|^2\right]=\lim_{s_1\to\infty}\mathbb{E}\left[\left\|X_0^{-s_1,x_0}-X_0^{-s_1,x_1}\right\|^2\right]\le\lim_{s_1\to\infty}e^{-c_1s_1}\mathbb{E}\left[\|x_0-x_1\|^2\right]=0,$$

and thus denoted by ϑ. Let π be the law of the random variable ϑ, then π is the unique invariant measure for SDE (1.1). Moreover, since $X_t^{x_0}$ and X_0^{-t,x_0} have the same distribution, for any function $\varphi\in C_b^1(\mathbb{R}^d)$, we get

$$\begin{aligned}\left|\mathbb{E}\left[\varphi\left(X_t^{x_0}\right)\right]-\int_{\mathbb{R}^d}\varphi(x)\pi(\mathrm{d}x)\right|&=\left|\mathbb{E}\left[\varphi\left(X_t^{x_0}\right)-\varphi\left(\vartheta\right)\right]\right|\\&\le\|\varphi\|_1\mathbb{E}\left[\left\|X_0^{-t,x_0}-\vartheta\right\|\right]\\&\le\|\varphi\|_1e^{-\frac{c_2}{2}t}\left(1+\mathbb{E}\left[\|x_0\|^2\right]\right),\quad\forall t\in[0,\infty).\end{aligned}$$

4. Invariant measure of the LTPE scheme

The main result of this Section is provided as below.

Theorem 4.1. *Let Assumptions 2.1-2.4 hold with $2\lambda_1>\max\{L_1,L_2\}$. For a method parameter $\theta\in[0,1]$, consider the LTPE method in (1.3) subject to a uniform timestep h satisfying*

$$h\in\left(0,\min\left\{\tfrac{2\lambda_1-L_2}{4(1-\theta)^2\lambda_d^2},\tfrac{(2\lambda_1-L_2)^\gamma}{(2\lambda_f)^{2\gamma}},1\right\}\right).$$

Then the numerical simulation from LTPE (1.3) method, denoted by $\{Y_n^{x_0}\}_{n\in\mathbb{N}_0}$ with the initial point x_0, admits a unique invariant measure π_h. Moreover, there exists some positive constant $\widetilde{C}_1$ such that, for every function $\varphi\in C_b^1(\mathbb{R}^d)$, $t_n=nh$, $n\in\mathbb{N}_0$,

$$\left|\mathbb{E}\left[\varphi\left(Y_n^{x_0}\right)\right]-\int_{\mathbb{R}^d}\varphi(x)\pi_h(dx)\right|\le\|\varphi\|_1e^{-\frac{\widetilde{C}_1}{2}t_n}\left(1+\mathbb{E}\left[\|x_0\|^2\right]\right).$$

The theorem above is proved in exactly the same way that Theorem 3.1 is proved, where the ergodicity of the LTPE (1.3) boils down to verifying the uniform moment bounds (see Lemma 4.3) and the contractive property (see Lemma 4.4). Before proceeding further, we first establish some preliminary estimates necessary for the proof of Theorem 4.1.

Lemma 4.2. *Let the projection $\mathscr{P}$ be defined by (1.4). Let Assumptions 2.2, 2.4 be fulfilled, then for any $x\in\mathbb{R}^d$ the following estimates*

$$\begin{aligned}&\|\mathscr{P}(x)\|\le\min\left\{\|x\|,h^{-\frac{1}{2\gamma}}\right\},\quad\|f\left(\mathscr{P}(x)\right)\|\le C_fh^{-\frac{1}{2}},\\&\|g(\mathscr{P}(x))\|^2\le\tfrac{L_1}{2p_0-1}(1+\|\mathscr{P}(x)\|^2)+2C_fh^{-\frac{1}{2}}\|\mathscr{P}(x)\|\end{aligned}\tag{4.1}$$

hold true, where $C_f:=2C_2$. Especially, for any integer $p\ge1$, we have, for $x\in\mathbb{R}^d$,

$$\|g(\mathscr{P}(x))\|^{2p}\le\left(\tfrac{L_1}{2p_0-1}\right)^p\left(1+\|\mathscr{P}(x)\|^2\right)^p+Ch^{-\frac{p}{2}}\left(1+\|\mathscr{P}(x)\|^2\right)^{p-1}.$$

Moreover, for any $x,y\in\mathbb{R}^d$, the following estimates hold true

$$\begin{aligned} \|\mathscr{P}(x)-\mathscr{P}(y)\| &\le \|x-y\|, \\ \left\| f\left(\mathscr{P}(x)\right) - f\left(\mathscr{P}(y)\right)\right\| &\le \lambda_f h^{-\frac{\gamma-1}{2\gamma}} \|x-y\|, \end{aligned} \tag{4.2}$$

where $\lambda_f := 3C_1$*, depending only on* f*.*

The proof of Lemma 4.2 is shown in Appendix B.1. The next lemma provides the uniform moment estimates for the LTPE scheme (1.3).

Lemma 4.3. *(Uniform moment bounds of the LTPE method) Let Assumptions 2.1, 2.2 and 2.4 hold with* $2\lambda_1 > L_1$*. For a method parameter* $\theta \in [0,1]$*, consider the numerical simulation* Y_n *from LTPE method in (1.3). Then, for any uniform stepsize* $h \in (0,1)$ *and* $n \in \mathbb{N}_0$*,*

$$\mathbb{E}\left[\|Y_n\|^2\right] \le C_A < \infty. \tag{4.3}$$

Moreover, for $p \in (1, p_0) \cap \mathbb{N}$*, if the timestep h further satisfies*

$$h \in \left(0, \min\left\{\tfrac{1}{2(1-\theta)\lambda_1}, \tfrac{p_0-p}{(1-\theta)(2p_0-p-1)\lambda_1}, \tfrac{1}{(1-\theta)\lambda_d}, 1\right\}\right),$$

then, for any $n \in \mathbb{N}_0$*,*

$$\mathbb{E}\left[\|Y_n\|^{2p}\right] \le C_A < \infty. \tag{4.4}$$

Proof of Lemma 4.3. We first take square of (1.3) on both sides and analyze the left and right hand sides individually. With Assumption 2.1 being used, the left hand side goes to

$$\begin{aligned} \|Y_{n+1} - \theta A Y_{n+1} h\|^2 &= \|Y_{n+1}\|^2 - 2\theta h \langle Y_{n+1}, AY_{n+1}\rangle + \theta^2 h^2 \|AY_{n+1}\|^2 \\ &\ge (1 + 2\theta\lambda_1 h)\|Y_{n+1}\|^2. \end{aligned} \tag{4.5}$$

On the other hand, the right hand side reads

$$\begin{aligned} &\left\|\mathscr{P}(Y_n) + (1-\theta)A\mathscr{P}(Y_n)h + f\big(\mathscr{P}(Y_n)\big)h + g\big(\mathscr{P}(Y_n)\big)\Delta W_n\right\|^2 \\ &= \|\mathscr{P}(Y_n)\|^2 + (1-\theta)^2 h^2 \|A\mathscr{P}(Y_n)\|^2 + h^2 \left\|f\left(\mathscr{P}(Y_n)\right)\right\|^2 + \left\|g\big(\mathscr{P}(Y_n)\big)\Delta W_n\right\|^2 \\ &\quad + 2(1-\theta)h\left\langle \mathscr{P}(Y_n), A\mathscr{P}(Y_n)\right\rangle + 2h\left\langle \mathscr{P}(Y_n), f\left(\mathscr{P}(Y_n)\right)\right\rangle + 2\left\langle B\mathscr{P}(Y_n), g\big(\mathscr{P}(Y_n)\big)\Delta W_n\right\rangle \\ &\quad + 2(1-\theta)h^2\left\langle A\mathscr{P}(Y_n), f\left(\mathscr{P}(Y_n)\right)\right\rangle + 2h\left\langle f\left(\mathscr{P}(Y_n)\right), g\big(\mathscr{P}(Y_n)\big)\Delta W_n\right\rangle, \end{aligned} \tag{4.6}$$

where $B := I + (1-\theta)Ah$. In the following, let us start by the estimate of (4.3).

Case I: estimate of $\mathbb{E}\left[\|Y_{n+1}\|^{2p}\right]$ **when** $p = 1$**.**

Using the Young inequality yields

$$2(1-\theta)h^2\left\langle A\mathscr{P}(Y_n), f\left(\mathscr{P}(Y_n)\right)\right\rangle \le (1-\theta)^2 h^2 \|A\mathscr{P}(Y_n)\|^2 + h^2 \left\|f\left(\mathscr{P}(Y_n)\right)\right\|^2. \tag{4.7}$$

Taking expectations of (4.5) and (4.6) respectively with Lemma 4.2 and the fact that $\mathbb{E}\left[\Delta W_n | \mathcal{F}_{t_n}\right] = 0$ shows

$$\begin{aligned} &(1 + 2\theta\lambda_1 h)\mathbb{E}\left[\|Y_{n+1}\|^2\right] \\ &\le [1 - 2(1-\theta)\lambda_1 h]\,\mathbb{E}\left[\|\mathscr{P}(Y_n)\|^2\right] + h\mathbb{E}\left[\|g\big(\mathscr{P}(Y_n)\big)\|^2\right] + 2h\mathbb{E}\left[\left\langle \mathscr{P}(Y_n), f\left(\mathscr{P}(Y_n)\right)\right\rangle\right] \\ &\quad + 2\big(C_f\big)^2 h + 2(1-\theta)^2 \lambda_d^2 h^{2-\frac{1}{\gamma}}. \end{aligned}$$

In conjunction with Assumption 2.2 with $2\lambda_1 > L_1$ for some positive constants $C_A = C(L_1, \lambda_d, C_f, \theta)$ and $\overline{C} := (2\lambda_1 - L_1)/(1 + 2\theta\lambda_1)$, such that

$$0 < \overline{C} \le \frac{2\lambda_1 - L_1}{1 + 2\theta\lambda_1 h},$$

one arrives at

$$\begin{aligned}\mathbb{E}\left[\|Y_{n+1}\|^2\right] &\le \frac{1-2(1-\theta)\lambda_1 h + L_1 h}{(1+2\theta\lambda_1 h)}\mathbb{E}\big[\|\mathscr{P}(Y_n)\|^2\big] + C_A h\\ &\le \left(1-\overline{C}h\right)\mathbb{E}\big[\|\mathscr{P}(Y_n)\|^2\big] + C_A h\\ &\le \left(1-\overline{C}h\right)^{n+1}\mathbb{E}\left[\|x_0\|^2\right] + \frac{C_A}{\overline{C}}\\ &\le e^{-\overline{C}t_{n+1}}\mathbb{E}\left[\|x_0\|^2\right] + \frac{C_A}{\overline{C}},\end{aligned}$$

where $1 - x \le e^{-x}$ for any $x > 0$.

Case II: estimate of $\mathbb{E}\left[\|Y_{n+1}\|^{2p}\right]$ when $p \in (1, p_0) \cap \mathbb{N}$.

Proceeding to the estimate of higher order moment of the LTPE method (1.3), some restrictions need to be imposed on the timestep h. Recalling $B = I + (1-\theta)Ah$, with $h \in (0, 1/[(1-\theta)\lambda_d])$, obviously, the matrix B is positive definite and $\max_{i=1,\dots,d}\lambda_{B,i} = 1-(1-\theta)\lambda_1 h$. By the Young inequality, we get, for some positive constant $\epsilon_1 \in (0, (2p_0-2p)/(2p-1)]$,

$$2h\big\langle f\big(\mathscr{P}(Y_n)\big), g\big(\mathscr{P}(Y_n)\big)\Delta W_n\big\rangle \le \frac{h^2}{\epsilon_1}\left\|f\left(\mathscr{P}(Y_n)\right)\right\|^2 + \epsilon_1\left\|g\big(\mathscr{P}(Y_n)\big)\Delta W_n\right\|^2. \tag{4.8}$$

Plugging estimates (4.7), (4.8), together with Assumption 2.1 and Lemma 4.2, into (4.6) to shows that

$$\begin{aligned}&(1+2\theta\lambda_1 h)(1+\|Y_{n+1}\|^2)\\ &\le [1-2(1-\theta)\lambda_1 h]\left(1+\|\mathscr{P}(Y_n)\|^2\right) + (1+\epsilon_1)\left\|g\big(\mathscr{P}(Y_n)\big)\Delta W_n\right\|^2\\ &\quad + 2h\big\langle\mathscr{P}(Y_n), f\big(\mathscr{P}(Y_n)\big)\big\rangle\\ &\quad + 2\big\langle B\mathscr{P}(Y_n), g\big(\mathscr{P}(Y_n)\big)\Delta W_n\big\rangle + \left[(2+\tfrac{1}{\epsilon_1})\big(C_f\big)^2 + (1-\theta)^2\lambda_d^2 + 2\lambda_1\right]h\\ &=: \widetilde{L}\left(1+\|\mathscr{P}(Y_n)\|^2\right)(1+\Xi_{n+1}) + C_{\epsilon_1}h,\end{aligned}$$

where $\widetilde{L} := 1-2(1-\theta)\lambda_1 h \ge 0$, $C_{\epsilon_1} = C(\lambda_1, \lambda_d, \theta, C_f) = (2+\frac{1}{\epsilon_1})\big(C_f\big)^2 + (1-\theta)^2\lambda_d^2 + 2\lambda_1$, and

$$\Xi_{n+1} = \underbrace{\frac{(1+\epsilon_1)\|g(\mathscr{P}(Y_n))\Delta W_n\|^2}{\widetilde{L}(1+\|\mathscr{P}(Y_n)\|^2)}}_{=:I_1} + \underbrace{\frac{2h\langle\mathscr{P}(Y_n), f(\mathscr{P}(Y_n))\rangle}{\widetilde{L}(1+\|\mathscr{P}(Y_n)\|^2)}}_{=:I_2} + \underbrace{\frac{2\langle B\mathscr{P}(Y_n), g(\mathscr{P}(Y_n))\Delta W_n\rangle}{\widetilde{L}(1+\|\mathscr{P}(Y_n)\|^2)}}_{=:I_3}.$$

Following the binomial expansion theorem and taking the conditional mathematical expectation with respect to $\mathcal{F}_{t_n}$ on both sides to show that,

$$\begin{aligned}&(1+2\theta\lambda_1 h)^p\mathbb{E}\left[(1+\|Y_{n+1}\|^2)^p\big|\mathcal{F}_{t_n}\right]\\ &\le \underbrace{(1+\|\mathscr{P}(Y_n)\|^2)^p\widetilde{L}^p\mathbb{E}\left[(1+\Xi_{n+1})^p\big|\mathcal{F}_{t_n}\right]}_{=:\mathbb{I}_1}\\ &\quad + \underbrace{C_{\epsilon_1}h(1+\|\mathscr{P}(Y_n)\|^2)^{p-1}\sum_{i=0}^{p-1}\widetilde{L}^i\mathbb{E}\left[(1+\Xi_{n+1})^i\big|\mathcal{F}_{t_n}\right]}_{=:\mathbb{I}_2}\end{aligned} \tag{4.9}$$

with $C_{\epsilon_1} = C(\lambda_1, \lambda_d, \theta, C_f, p)$. Hence, the analysis is divided into the following two parts.

The estimate of $\mathbb{I}_1$:

According to the binomial expansion theorem again, one has

$$\mathbb{E}\left[(1+\Xi_{n+1})^p\big|\mathcal{F}_{t_n}\right] = \sum_{i=0}^{p}\mathscr{C}_p^i\mathbb{E}\left[\Xi_{n+1}^i\big|\mathcal{F}_{t_n}\right],$$

where $\mathscr{C}_p^i := p!/(i!(p-i)!)$. Let us decompose the estimate of $\mathbb{I}_1$ further into four steps.

Step I: the estimate of $\mathbb{E}\left[\Xi_{n+1}\big|\mathcal{F}_{t_n}\right]$.

Based on the property of Brownian motion and the fact that ΔW_n is independent of $\mathcal{F}_{t_n}$, we deduce

$$\mathbb{E}\left[\Delta W_{j,n}\big|\mathcal{F}_{t_n}\right]=0, \quad \mathbb{E}\left[|\Delta W_{j,n}|^2\big|\mathcal{F}_{t_n}\right]=h, \quad j\in\{1,\dots,m\}, \tag{4.10}$$

leading to

$$\mathbb{E}\left[\Xi_{n+1}\big|\mathcal{F}_{t_n}\right]=\tfrac{(1+\epsilon_1)h\|g(\mathscr{P}(Y_n))\|^2+2h\langle\mathscr{P}(Y_n),f(\mathscr{P}(Y_n))\rangle}{\widetilde{L}(1+\|\mathscr{P}(Y_n)\|^2)}.$$

Step II: the estimate of $\mathbb{E}\left[\Xi_{n+1}^2\big|\mathcal{F}_{t_n}\right]$.

Recalling some power properties of Brownian motions, we derive that, for any $\ell\in\mathbb{N}$,

$$\mathbb{E}\left[\left(\Delta W_{j,n}\right)^{2\ell-1}\big|\mathcal{F}_{t_n}\right]=0, \quad \mathbb{E}\left[\left(\Delta W_{j,n}\right)^{2\ell}\big|\mathcal{F}_{t_n}\right]=(2\ell-1)!!\,h^\ell, \quad \forall\, n\in\mathbb{N},\ j\in\{1,\dots,m\}, \tag{4.11}$$

where $(2\ell-1)!! := \Pi_{i=1}^{\ell}(2i-1)$. Before moving on, we here introduce a series of useful estimates. For any $\ell\in[2,\infty)\cap\mathbb{N}$, by Lemma 4.2 and (4.11), one achieves with some constant $C=C(L_1,C_f,p)$,

$$\begin{aligned}\mathbb{E}\left[(I_1)^\ell\big|\mathcal{F}_{t_n}\right]&=\tfrac{(1+\epsilon_1)^\ell(2\ell-1)!!\,h^\ell\|g(\mathscr{P}(Y_n))\|^{2\ell}}{\widetilde{L}^\ell(1+\|\mathscr{P}(Y_n)\|^2)^\ell}\\&\le\frac{(1+\epsilon_1)^\ell(2\ell-1)!!\,h^\ell\left[L_1^\ell(1+\|\mathscr{P}(Y_n)\|^2)^\ell+Ch^{-\frac{\ell}{2}}(1+\|\mathscr{P}(Y_n)\|^2)^{\ell-1}\right]}{(2p_0-1)^\ell\widetilde{L}^\ell(1+\|\mathscr{P}(Y_n)\|^2)^\ell}\\&\le\tfrac{(1+\epsilon_1)^\ell L_1^\ell(2\ell-1)!!\,h^\ell}{(2p_0-1)^\ell\widetilde{L}^\ell}+\tfrac{C(1+\epsilon)^\ell(2\ell-1)!!h^{\frac{\ell}{2}}}{\widetilde{L}^\ell(1+\|\mathscr{P}(Y_n)\|^2)}.\end{aligned}\tag{4.12}$$

Similarly, with the Cauchy Schwarz inequality, one gets

$$\mathbb{E}\left[(I_2)^\ell\big|\mathcal{F}_{t_n}\right]=\tfrac{2^\ell h^\ell(\langle\mathscr{P}(Y_n),f(\mathscr{P}(Y_n))\rangle)^\ell}{\widetilde{L}^\ell(1+\|\mathscr{P}(Y_n)\|^2)^\ell}\le\tfrac{2^\ell h^\ell\|\mathscr{P}(Y_n)\|^\ell\|f(\mathscr{P}(Y_n))\|^\ell}{\widetilde{L}^\ell(1+\|\mathscr{P}(Y_n)\|^2)^\ell}.$$

For any $\ell\ge2$ and $x\ge0$, we know that $x^{\frac{\ell}{2}}\le(1+x^2)^{\ell-1}$. Therefore, with $C=C(C_f,\ell)$, we obtain

$$\mathbb{E}\left[(I_2)^\ell\big|\mathcal{F}_{t_n}\right]\le\tfrac{Ch^{\frac{\ell}{2}}}{\widetilde{L}^\ell(1+\|\mathscr{P}(Y_n)\|^2)}. \tag{4.13}$$

One needs to be careful about the estimate of term I_3. Equipping with (4.10) yields

$$\mathbb{E}\left[(I_3)^2\big|\mathcal{F}_{t_n}\right]=\tfrac{4h\left\|(B\mathscr{P}(Y_n))^Tg(\mathscr{P}(Y_n))\right\|^2}{\widetilde{L}^2(1+\|\mathscr{P}(Y_n)\|^2)^2}\le\tfrac{4h[1-(1-\theta)\lambda_1h]^2}{1-2(1-\theta)\lambda_1h}\tfrac{\|g(\mathscr{P}(Y_n))\|^2}{\widetilde{L}(1+\|\mathscr{P}(Y_n)\|^2)}. \tag{4.14}$$

It is time to move on to the estimate of $\mathbb{E}\left[\Xi_{n+1}^2\big|\mathcal{F}_{t_n}\right]$. We begin with the following expansion

$$\begin{aligned}\mathbb{E}\left[\Xi_{n+1}^2\big|\mathcal{F}_{t_n}\right]&=\mathbb{E}\left[(I_1+I_2+I_3)^2\big|\mathcal{F}_{t_n}\right]\\&=\mathbb{E}\left[(I_1)^2+(I_2)^2+(I_3)^2+2I_1I_2+2I_1I_3+2I_2I_3\big|\mathcal{F}_{t_n}\right].\end{aligned}\tag{4.15}$$

As claimed before, one observes

$$\mathbb{E}\left[I_1I_3\big|\mathcal{F}_{t_n}\right]=\mathbb{E}\left[I_2I_3\big|\mathcal{F}_{t_n}\right]=0,$$

and, for $C=C(L_1,C_f)$,

$$\mathbb{E}\left[I_1I_2\big|\mathcal{F}_{t_n}\right]\le\tfrac{Ch}{\widetilde{L}^2(1+\|\mathscr{P}(Y_n)\|^2)}$$

from (4.12), (4.13) and the Hölder inequality. Plugging these with (4.12), (4.13) and (4.14) into (4.15) shows that

$$\mathbb{E}\left[\Xi_{n+1}^2\middle|\mathcal{F}_{t_n}\right] \le \frac{3(1+\epsilon_1)^2 L_1^2 h^2}{(2p_0-1)^2\widetilde{L}^2} + \frac{4h[1-(1-\theta)\lambda_1 h]^2}{1-2(1-\theta)\lambda_1 h}\frac{\|g(\mathscr{P}(Y_n))\|^2}{\widetilde{L}(1+\|\mathscr{P}(Y_n)\|^2)} + \frac{Ch}{\widetilde{L}^2(1+\|\mathscr{P}(Y_n)\|^2)},$$

where $C=C(L_1,C_f)$. As we know, for any positive constant $\ell\in[2,p]\cap\mathbb{N}$, $p<p_0$ and $\epsilon_1\in(0,(2p_0-2p)/(2p-1)]$,

$$(2\ell-1)!!(1+\epsilon_1)^\ell < (2p-1)^\ell(1+\epsilon_1)^\ell < (2p_0-1)^\ell, \tag{4.16}$$

so that we obtain

$$\mathbb{E}\left[\Xi_{n+1}^2\middle|\mathcal{F}_{t_n}\right] \le \frac{L_1^2h^2}{\widetilde{L}^2} + \frac{4h[1-(1-\theta)\lambda_1 h]^2}{1-2(1-\theta)\lambda_1 h}\frac{\|g(\mathscr{P}(Y_n))\|^2}{\widetilde{L}(1+\|\mathscr{P}(Y_n)\|^2)} + \frac{Ch}{\widetilde{L}^2(1+\|\mathscr{P}(Y_n)\|^2)}.$$

Step III: the estimate of $\mathbb{E}\left[\Xi_{n+1}^3\middle|\mathcal{F}_{t_n}\right]$.

By the similar procedure, we acquire that

$$\mathbb{E}\left[\Xi_{n+1}^3\middle|\mathcal{F}_{t_n}\right] = \mathbb{E}\left[(I_1+I_2+I_3)^3\middle|\mathcal{F}_{t_n}\right],$$

where (4.10) and (4.11) are used to imply that

$$\mathbb{E}\left[(I_3)^3\middle|\mathcal{F}_{t_n}\right] = \mathbb{E}\left[(I_1)^2 I_3\middle|\mathcal{F}_{t_n}\right] = \mathbb{E}\left[(I_2)^2 I_3\middle|\mathcal{F}_{t_n}\right] = \mathbb{E}\left[I_1I_2I_3\middle|\mathcal{F}_{t_n}\right] = 0.$$

Obeying (4.12)-(4.14), (4.16) yields with $C=C(L_1,C_f)$,

$$\begin{aligned}\mathbb{E}\left[\Xi_{n+1}^3\middle|\mathcal{F}_{t_n}\right] &= \mathbb{E}\left[(I_1)^3\middle|\mathcal{F}_{t_n}\right] + \mathbb{E}\left[(I_2)^3\middle|\mathcal{F}_{t_n}\right] + 3\mathbb{E}\left[(I_1)^2\cdot I_2\middle|\mathcal{F}_{t_n}\right] + 3\mathbb{E}\left[I_1\cdot(I_2)^2\middle|\mathcal{F}_{t_n}\right] \\ &\quad + 3\mathbb{E}\left[I_1\cdot(I_3)^2\middle|\mathcal{F}_{t_n}\right] + 3\mathbb{E}\left[I_2\cdot(I_3)^2\middle|\mathcal{F}_{t_n}\right] \\ &\le \frac{15(1+\epsilon_1)^3L_1^3h^3}{(2p_0-1)^3\widetilde{L}^3} + \frac{Ch}{\widetilde{L}^3(1+\|\mathscr{P}(Y_n)\|^2)} \\ &\le \frac{L_1^3h^3}{\widetilde{L}^3} + \frac{Ch}{\widetilde{L}^3(1+\|\mathscr{P}(Y_n)\|^2)}.\end{aligned}$$

Step IV: the estimate of $\mathbb{E}\left[\Xi_{n+1}^\ell\middle|\mathcal{F}_{t_n}\right]$, $\ell\in[4,p]\cap\mathbb{N}$.

It follows from Lemma 4.2 that, for $\ell\in[4,p]\cap\mathbb{N}$,

$$\begin{aligned}\mathbb{E}\left[(I_3)^\ell\middle|\mathcal{F}_{t_n}\right] &\le \frac{2^\ell h^{\frac{\ell}{2}}\|\mathscr{P}(Y_n)^T g(\mathscr{P}(Y_n))\|^\ell}{\widetilde{L}^\ell(1+\|\mathscr{P}(Y_n)\|^2)^\ell} \\ &\le \frac{2^\ell h^{\frac{\ell}{2}}\|\mathscr{P}(Y_n)\|^\ell\left[L_1^{\frac{\ell}{2}}(1+\|\mathscr{P}(Y_n)\|^2)^{\frac{\ell}{2}} + Ch^{-\frac{\ell}{4}}(1+\|\mathscr{P}(Y_n)\|^2)^{\frac{\ell}{2}-1}\right]}{\widetilde{L}^\ell(1+\|\mathscr{P}(Y_n)\|^2)^\ell} \\ &= \frac{2^\ell h^{\frac{\ell}{2}}\left[L_1^{\frac{\ell}{2}}(1+\|\mathscr{P}(Y_n)\|^2)^{\frac{\ell}{2}}\|\mathscr{P}(Y_n)\|^{\ell-2}\|\mathscr{P}(Y_n)\|^2 + Ch^{-\frac{\ell}{4}}(1+\|\mathscr{P}(Y_n)\|^2)^{\frac{\ell}{2}-1}\|\mathscr{P}(Y_n)\|^\ell\right]}{\widetilde{L}^\ell(1+\|\mathscr{P}(Y_n)\|^2)^\ell}.\end{aligned} \tag{4.17}$$

Bearing the fact from Lemma 4.2 that $\|\mathscr{P}(Y_n)\|^2\le h^{-\frac{1}{\gamma}}$ in mind, we deduce, for some constant $C=C(L_1,C_f,\ell)$,

$$\mathbb{E}\left[(I_3)^\ell\middle|\mathcal{F}_{t_n}\right] \le \frac{Ch^{\frac{\ell}{2}-\frac{1}{\gamma}}+Ch^{\frac{\ell}{4}}}{\widetilde{L}^\ell(1+\|\mathscr{P}(Y_n)\|^2)} \le \frac{Ch^{\frac{\ell}{4}}}{\widetilde{L}^\ell(1+\|\mathscr{P}(Y_n)\|^2)}.$$

Using the Young inequality yields, for some positive constants $\epsilon_\ell\in(0,(2\ell-1)^\ell/(2\ell-1)!!]$, $\ell\in[4,p]\cap\mathbb{N}$,

$$\mathbb{E}\left[\Xi_{n+1}^{\ell}\big|\mathcal{F}_{t_n}\right]=\mathbb{E}\left[(I_1+I_2+I_3)^{\ell}\big|\mathcal{F}_{t_n}\right]\le(1+\epsilon_\ell)\mathbb{E}\left[|I_1|^{\ell}\big|\mathcal{F}_{t_n}\right]+(1+\tfrac{1}{\epsilon_\ell})\mathbb{E}\left[|I_2+I_3|^{\ell}\big|\mathcal{F}_{t_n}\right].$$

In light of the estimates (4.12)-(4.13) and (4.17) with the elementary inequality, we obtain that, for some constant $C_{\epsilon_\ell}=C(L_1,C_f,\ell)$,

$$\mathbb{E}\left[\Xi_{n+1}^{\ell}\big|\mathcal{F}_{t_n}\right]\le(1+\epsilon_\ell)\frac{(2\ell-1)!!\,(1+\epsilon_1)^{\ell}L_1^{\ell}h^{\ell}}{(2p_0-1)^{\ell}\widetilde{L}^{\ell}}+\frac{C_{\epsilon_\ell}h}{\widetilde{L}^{\ell}(1+\|\mathscr{P}(Y_n)\|^2)}. \tag{4.18}$$

We would like to mention that the following inequality holds for any $\ell\in[4,p]\cap\mathbb{N}$ and $\epsilon_1\in(0,(2p_0-2p)/(2p-1)]$,

$$(1+\epsilon_\ell)(2\ell-1)!!(1+\epsilon_1)^{\ell}\le(2\ell-1)^{\ell}(1+\epsilon_1)^{\ell}\le(2p_0-1)^{\ell}.$$

Therefore, the estimate (4.18) is rewritten as

$$\mathbb{E}\left[\Xi_{n+1}^{\ell}\big|\mathcal{F}_{t_n}\right]\le\frac{L_1^{\ell}h^{\ell}}{\widetilde{L}^{\ell}}+\frac{C_{\epsilon_\ell}h}{\widetilde{L}^{\ell}(1+\|\mathscr{P}(Y_n)\|^2)}.$$

Combining **Step I**∼**Step IV** to show that, for some constant $C_{\epsilon_\ell}=C(L_1,C_f,p)$,

$$\begin{aligned}&\mathbb{E}\left[(1+\Xi_{n+1})^{p}\big|\mathcal{F}_{t_n}\right]\\&\le1+ph\frac{2\langle\mathscr{P}(Y_n),f(\mathscr{P}(Y_n))\rangle+\left[(2p-2)\frac{[1-(1-\theta)\lambda_1h]^2}{1-2(1-\theta)\lambda_1h}+1+\epsilon_1\right]\|g(\mathscr{P}(Y_n))\|^2}{\widetilde{L}(1+\|\mathscr{P}(Y_n)\|^2)}\\&\quad+\sum_{\ell=2}^{p}\mathscr{C}_p^{\ell}\frac{L_1^{\ell}h^{\ell}}{\widetilde{L}^{\ell}}+\frac{C_{\epsilon_\ell}h}{\widetilde{L}^{p}(1+\|\mathscr{P}(Y_n)\|^2)}.\end{aligned}$$

Moreover, we can choose an appropriate h such that

$$h\in\left(0,\frac{p_0-p}{(1-\theta)\lambda_1(2p_0-p-1)}\right)$$

to make sure

$$2p_0-1>(2p-2)\frac{1-(1-\theta)\lambda_1h}{1-2(1-\theta)\lambda_1h}+1\ge(2p-2)\frac{[1-(1-\theta)\lambda_1h]^2}{1-2(1-\theta)\lambda_1h}+1,$$

which leads to the following estimate by Assumption 2.2,

$$\widetilde{L}^{p}\mathbb{E}\left[(1+\Xi_{n+1})^{p}\big|\mathcal{F}_{t_n}\right]\le1+\sum_{\ell=1}^{p}\mathscr{C}_p^{\ell}\widetilde{L}^{p-\ell}L_1^{\ell}h^{\ell}+\frac{C_{\epsilon_\ell}h}{1+\|\mathscr{P}(Y_n)\|^2}=\left(\widetilde{L}+L_1h\right)^{p}+\frac{C_{\epsilon_\ell}h}{1+\|\mathscr{P}(Y_n)\|^2}.$$

Hence, we deduce that

$$\mathbb{I}_1\le\left(\widetilde{L}+L_1h\right)^{p}(1+\|\mathscr{P}(Y_n)\|^2)^{p}+C_{\epsilon_\ell}h(1+\|\mathscr{P}(Y_n)\|^2)^{p-1}.$$

The estimate of $\mathbb{I}_2$:

For the estimate of $\mathbb{I}_2$, the key point is to get the estimate of $\widetilde{L}^{i}\mathbb{E}\left[(1+\Xi_{n+1})^{i}\big|\mathcal{F}_{t_n}\right]$, $i\in(1,p)\in\mathbb{N}$, which is uniformly bounded with the same analysis as the estimate of $\mathbb{I}_1$, i.e., there exists some positive constant $C=C(L_1,C_f,p)$ such that,

$$\sum_{i=0}^{p-1}\widetilde{L}^{i}\mathbb{E}\left[(1+\Xi_{n+1})^{i}\big|\mathcal{F}_{t_n}\right]\le C,$$

leading to

$$\mathbb{I}_2\le C_{\epsilon_1}h(1+\|\mathscr{P}(Y_n)\|^2)^{p-1}.$$

Combining the estimates of $\mathbb{I}_1$ and $\mathbb{I}_2$:

Taking the estimates of $\mathbb{I}_1$ and $\mathbb{I}_2$ into (4.9), for some constant $C_{\epsilon_1,\epsilon_\ell} = C(\lambda_1, \lambda_d, \theta, L_1, C_f, p)$, we recall $\widetilde{L} = 1 - 2(1-\theta)\lambda_1 h$ to show

$$(1 + 2\theta\lambda_1 h)^p \mathbb{E}\left[(1 + \|Y_{n+1}\|^2)^p \big| \mathcal{F}_{t_n}\right] \le [1 - 2(1-\theta)\lambda_1 h + L_1 h]^p (1 + \|\mathscr{P}(Y_n)\|^2)^p + C_{\epsilon_1,\epsilon_\ell} h (1 + \|\mathscr{P}(Y_n)\|^2)^{p-1}. \tag{4.19}$$

For $2\lambda_1 > L_1$, we take expectations on both sides of (4.19) with Lemma 4.2 and the Young inequality to show that, for some $\epsilon_2 > 0$,

$$\begin{aligned}
&\mathbb{E}\left[(1 + \|Y_{n+1}\|^2)^p\right] \\
&\le \left(\tfrac{1-2(1-\theta)\lambda_1 h + L_1 h}{1+2\theta\lambda_1 h}\right)^p \mathbb{E}\left[(1 + \|\mathscr{P}(Y_n)\|^2)^p\right] + C_{\epsilon_1,\epsilon_\ell} h \mathbb{E}\left[(1 + \|\mathscr{P}(Y_n)\|^2)^{p-1}\right] \\
&= \left(1 - \tfrac{2\lambda_1 - L_1}{1+2\theta\lambda_1 h} h\right) \mathbb{E}\left[(1 + \|\mathscr{P}(Y_n)\|^2)^p\right] + C_{\epsilon_1,\epsilon_\ell} h \mathbb{E}\left[(1 + \|\mathscr{P}(Y_n)\|^2)^{p-1}\right] \\
&\le \left(1 - \tfrac{2\lambda_1 - L_1}{1+2\theta\lambda_1 h} h + \tfrac{p-1}{p} \epsilon_2 C_{\epsilon_1,\epsilon_\ell} h\right) \mathbb{E}\left[(1 + \|Y_n\|^2)^p\right] + \tfrac{\epsilon_2^{1-p}}{p} C_{\epsilon_1,\epsilon_\ell} h.
\end{aligned}$$

Then we choose a suitable ϵ_2 to ensure that

$$0 < \widetilde{C} := \tfrac{2\lambda_1 - L_1}{1+2\theta\lambda_1} - \tfrac{p-1}{p} \epsilon_2 C_{\epsilon_1,\epsilon_\ell} \le \tfrac{2\lambda_1 - L_1}{1+2\theta\lambda_1 h} - \tfrac{p-1}{p} \epsilon_2 C_{\epsilon_1,\epsilon_\ell}.$$

Therefore, for some constant $C_{\epsilon_1,\epsilon_2,\epsilon_\ell} = C(\lambda_1, \lambda_d, \theta, L_1, C_f, p)$, we get

$$\begin{aligned}
\mathbb{E}\left[(1 + \|Y_{n+1}\|^2)^p\right] &\le \left(1 - \widetilde{C} h\right) \mathbb{E}\left[(1 + \|Y_n\|^2)^p\right] + C_{\epsilon_1,\epsilon_2,\epsilon_\ell} h \\
&\le \left(1 - \widetilde{C} h\right)^{n+1} \mathbb{E}\left[(1 + \|x_0\|^2)^p\right] + \sum_{i=0}^{n} \left(1 - \widetilde{C} h\right)^i C_{\epsilon_1,\epsilon_2,\epsilon_\ell} h \\
&\le e^{-\widetilde{C} t_{n+1}} \mathbb{E}\left[(1 + \|x_0\|^2)^p\right] + \tfrac{C_{\epsilon_1,\epsilon_2,\epsilon_\ell}}{\widetilde{C}},
\end{aligned}$$

where we have used the fact that for any $x > 0$, $1 - x \le e^{-x}$. The proof is complete. □

We remark that to verify the existence and uniqueness of the invariant measure of the LTPE method (1.3), the uniform estimate of the second order moment (i.e. (4.3) in Lemma 4.3) is enough. The estimate of the $2p$-th order moment (i.e. (4.4) in Lemma 4.3) of the LTPE method (1.3) is essential to the error analysis that follows.

The contractivity of the LTPE method (1.3) follows directly from Lemma 4.3 and Lemma 4.2.

Lemma 4.4. *(Contractivity of the theta-linear-projected Euler method.) Consider the following pair of solutions of LTPE method (1.3) with a parameter $\theta \in [0, 1]$ driven by the same Brownian motion:*

$$\begin{aligned}
&Y_{n+1}^{(1)} - \theta A Y_{n+1}^{(1)} h \\
&= \mathscr{P}(Y_n^{(1)}) + (1-\theta) A \mathscr{P}(Y_n^{(1)}) h + f\big(\mathscr{P}(Y_n^{(1)})\big) h + g\big(\mathscr{P}(Y_n^{(1)})\big) \Delta W_n, \ Y_0^{(1)} = x_0^{(1)}; \\
&Y_{n+1}^{(2)} - \theta A Y_{n+1}^{(2)} h \\
&= \mathscr{P}(Y_n^{(2)}) + (1-\theta) A \mathscr{P}(Y_n^{(2)}) h + f\big(\mathscr{P}(Y_n^{(2)})\big) h + g\big(\mathscr{P}(Y_n^{(2)})\big) \Delta W_n, \ Y_0^{(2)} = x_0^{(2)},
\end{aligned}$$

where h is the uniform timestep with

$$h \in \left(0, \min\left\{\tfrac{2\lambda_1 - L_2}{4(1-\theta)^2 \lambda_d^2}, \tfrac{(2\lambda_1 - L_2)^\gamma}{(2\lambda_f)^{2\gamma}}, 1\right\}\right), \quad \lambda_f := 3C_1.$$

Let Assumptions 2.1, 2.3 and 2.4 hold for $2\lambda_1 > L_2$*. Then there exists a positive constant* $\widetilde{C}_1$*, such that, for any* $n \in \mathbb{N}_0$ *and* $t_n = nh$*,*

$$\mathbb{E}\left[\|Y_n^{(1)} - Y_n^{(2)}\|^2\right] \le e^{-\widetilde{C}_1 t_n}\mathbb{E}\big[\|x_0^{(1)} - x_0^{(2)}\|^2\big].$$

The proof of Lemma 4.4 is deferred to Appendix B.2.

Proof of Theorem 4.1. With Lemma 4.3 in mind, the existence of the invariant measure π_h admitted by the LTPE scheme (1.3) is obtained by Krylov-Bogoliubov theorem [9]. Further, the proof of the uniqueness of such invariant measure π_h follows almost the same idea quoted from Theorem 7.9 in [18], which is a consequence of Lemma 4.4, so that we omit it here. Then, using Lemma 4.4 and the Chapman-Kolmogorov equation yields, for $\varphi \in C_b^1(\mathbb{R}^d)$, $t_n = nh$, $n \in \mathbb{N}_0$,

$$\begin{aligned}
\left|\mathbb{E}\left[\varphi\left(Y_n^{x_0}\right)\right] - \int_{\mathbb{R}^d} \varphi(x)\pi_h(\mathrm{d}x)\right| &= \left|\int_{\mathbb{R}^d} \mathbb{E}\left[\varphi\left(Y_n^{x_0}\right)\right]\pi_h(\mathrm{d}x) - \int_{\mathbb{R}^d} \mathbb{E}\left[\varphi\left(Y_n^{x}\right)\right]\pi_h(\mathrm{d}x)\right| \\
&\le \|\varphi\|_1 \int_{\mathbb{R}^d} \mathbb{E}\left[\left\|Y_n^{x_0} - Y_n^{x}\right\|\right]\pi_h(\mathrm{d}x) \\
&\le \|\varphi\|_1 e^{-\frac{\widetilde{C}_1}{2}t_n}\left(1 + \mathbb{E}\left[\|x_0\|^2\right]\right). \quad \square
\end{aligned} \tag{4.20}$$

5. Time-independent weak error analysis

Our aim is to estimate the error between the invariant measure π and π_h. i.e.

$$\left|\int_{\mathbb{R}^d} \varphi(x)\pi(\mathrm{d}x) - \int_{\mathbb{R}^d} \varphi(x)\pi_h(\mathrm{d}x)\right|.$$

By Theorem 3.1 and Theorem 4.1, we claim that both $\{X_{t_n}\}_{n\in\mathbb{N}_0}$, defined by (1.1), and $\{Y_n\}_{n\in\mathbb{N}_0}$, defined by (1.3), are strongly mixing [13,9], i.e.

$$\lim_{n\to\infty} \mathbb{E}\big[\varphi(X_{t_n}^x)\big] = \int_{\mathbb{R}^d} \varphi(x)\pi(\mathrm{d}x), \ \ \text{in } L^2(\mathbb{R}^d, \pi), \tag{5.1}$$

and

$$\lim_{n\to\infty} \mathbb{E}\big[\varphi(Y_n^x)\big] = \int_{\mathbb{R}^d} \varphi(x)\pi_h(\mathrm{d}x), \ \text{in } L^2(\mathbb{R}^d, \pi_h), \tag{5.2}$$

for all $\varphi \in L^2(\mathbb{R}^d, \pi) \cap L^2(\mathbb{R}^d, \pi_h)$. Hence, the error estimate boils down to the time-independent weak convergence analysis of the LTPE method (1.3) as follows,

$$\left|\int_{\mathbb{R}^d} \varphi(x)\pi(\mathrm{d}x) - \int_{\mathbb{R}^d} \varphi(x)\pi_h(\mathrm{d}x)\right| \le \lim_{n\to\infty} \left|\mathbb{E}\big[\varphi(X_{t_n}^x)\big] - \mathbb{E}\big[\varphi(Y_n^x)\big]\right|. \tag{5.3}$$

In order to carry out the error analysis, we need some priori estimates and lemmas. The key ingredient is to introduce $u : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}$ defined by

$$u(t, x) := \mathbb{E}\left[\varphi\big(X_t^x\big)\right], \tag{5.4}$$

where $\varphi \in C_b^3(\mathbb{R}^d)$. In what follows, we will show that $u(\cdot,\cdot)$ is the unique solution of the associated Kolmogorov equations as

$$\partial_t u(t,x) = Du(t,x)F(x) + \tfrac{1}{2}\sum_{j=1}^{m} D^2u(t,x)\big(g_j(x), g_j(x)\big), \tag{5.5}$$

with initial condition $u(0,\cdot) = \varphi(\cdot)$, where we denote that $F(x) := Ax + f(x)$. To examine the regularity of u, we need the following properties.

For the matrix $A \in \mathbb{R}^{d\times d}$, it is apparent that

$$D(Ax)v_1 = Av_1, \quad \forall x, v_1 \in \mathbb{R}^d,$$

and for $i \in [2,\infty) \cap \mathbb{N}$,

$$D^i(Ax)(v_1,\cdots,v_i) = 0, \quad \forall x, v_1,\cdots,v_i \in \mathbb{R}^d.$$

Then it follows from Assumption 2.4 and its consequences that, for $j \in \{1,\ldots,m\}$,

$$\begin{aligned} \Big\| D^2F(x)(v_1,v_2) - D^2F(\tilde{x})(v_1,v_2)\Big\| &\le C\mathcal{P}_{\gamma-3}(x,\tilde{x})\cdot\|x-\tilde{x}\|\cdot\|v_1\|\cdot\|v_2\|, \\ &\qquad \forall x,\tilde{x},v_1,v_2 \in \mathbb{R}^d, \\ \Big\| D^2F(x)(v_1,v_2)\Big\| &\le C\mathcal{P}_{\gamma-2}(x)\cdot\|v_1\|\cdot\|v_2\|, \quad \forall x,v_1,v_2 \in \mathbb{R}^d, \end{aligned} \tag{5.6}$$

which directly implies

$$\begin{aligned} \big\| DF(x)v_1 - DF(\tilde{x})v_1\big\| &\le C\mathcal{P}_{\gamma-2}(x,\tilde{x})\cdot\|x-\tilde{x}\|\cdot\|v_1\|, \quad \forall x,\tilde{x},v_1 \in \mathbb{R}^d, \\ \|DF(x)v_1\| &\le C_A(1+\|x\|)^{\gamma-1}\cdot\|v_1\|, \quad \forall x,v_1 \in \mathbb{R}^d, \end{aligned} \tag{5.7}$$

and

$$\begin{aligned} \big\| F(x) - F(\tilde{x})\big\| &\le C_A(1+\|x\|+\|\tilde{x}\|)^{\gamma-1}\|x-\tilde{x}\|, \quad \forall x,\tilde{x} \in \mathbb{R}^d, \\ \|F(x)\| &\le C_A(1+\|x\|)^{\gamma}, \quad \forall x \in \mathbb{R}^d. \end{aligned} \tag{5.8}$$

Besides, Assumptions 2.1, 2.3 lead to, for some $p_1 \ge 1$ with $2\lambda_1 > L_2$,

$$2\langle DF(x)y, y\rangle + (2p_1-1)\sum_{j=1}^{m}\|Dg_j(x)y\|^2 \le -\alpha\|y\|^2, \quad \forall x,y \in \mathbb{R}^d, \tag{5.9}$$

where $\alpha := 2\lambda_1 - L_2 > 0$. For random functions, let us introduce the mean-square differentiability, quoted from [30], as follows.

Definition 5.1. (Mean-square differentiable) Let $\Psi : \Omega\times\mathbb{R}^d \to \mathbb{R}$ and $\psi_i : \Omega\times\mathbb{R}^d \to \mathbb{R}$ be random functions satisfying

$$\lim_{\tau\to 0}\mathbb{E}\left[\Big|\tfrac{1}{\tau}\left[\Psi(x+\tau e_i) - \Psi(x)\right] - \psi_i(x)\Big|^2\right] = 0, \quad \forall i \in \{1,2,\cdots,d\},$$

where e_i is the unit vector in $\mathbb{R}^d$ with the i-th element being 1. Then Ψ is called to be mean-square differentiable, with $\psi = (\psi_1,\ldots,\psi_d)$ being the derivative (in the mean-square differentiable sense) of Ψ at x. Also denoting $\mathcal{D}_{(i)}\Psi = \psi_i$ and $\mathcal{D}\Psi(x) = \psi$.

The above definition is generalized to vector-valued functions in a component-wise manner. Now we are in the position to derive the uniform estimate of the derivatives of $\{X_t^x\}_{t\in[0,\infty)}$ of (1.1) in the mean-square differentiable sense. Here for each t we take the function $X_t^{\cdot} : \mathbb{R}^d \to \mathbb{R}^d$, and write its derivative as $\mathcal{D}X_t^x \in L(\mathbb{R}^d,\mathbb{R}^d)$. Higher order derivatives $\mathcal{D}^2X_t^x$ and $\mathcal{D}^3X_t^x$ are defined similarly. From this point on, the convention $\times$ is applied to denote multiplication operations in cases where we are dealing with long expressions that extend over multiple lines.

Lemma 5.2. *Consider the SDE (1.1) subject to Assumptions 2.1-2.4 with* $2\lambda_1 > L_2$. *Then the solution* $\{X_t\}_{t\in[0,\infty)}$ *of (1.1) is three times mean-square differentiable. Moreover, recall* p_1 *given in Assumption 2.3, for any* $q_1 \in [1, p_1]$, $q_2 \in [1, q_1)$, *and some random variables* $v_1 \in L^{2\max\{q_1,\rho_2 q_2\}}(\Omega, \mathbb{R}^d)$, $v_2 \in L^{2\max\{\rho_3 q_2, \rho_3\rho_5\}}(\Omega, \mathbb{R}^d)$, $v_3 \in L^{2\rho_3}(\Omega, \mathbb{R}^d)$, *where* $\rho_1, \rho_2, \rho_3, \rho_4, \rho_5, \rho_6 > 1$ *satisfies* $1/\rho_1 + 1/\rho_2 + 1/\rho_3 = 1$, $1/\rho_4 + 1/\rho_5 + 1/\rho_6 = 1$ *and* p_0 *in Assumption 2.2 fulfills*

$$p_0 \in \Big[\max\{\rho_1 q_2, \rho_2\} \times (\gamma - 2), \infty\Big) \cap [1, \infty), \tag{5.10}$$

such that

$$\begin{aligned}
&\left\|\mathcal{D}X_t^x v_1\right\|_{L^{2q_1}(\Omega,\mathbb{R}^d)} \le Ce^{-\alpha_1 t}\,\|v_1\|_{L^{2q_1}(\Omega,\mathbb{R}^d)},\\
&\left\|\mathcal{D}^2 X_t^x(v_1, v_2)\right\|_{L^{2q_2}(\Omega,\mathbb{R}^d)}\\
&\le Ce^{-\alpha_2 t}\sup_{r\in[0,\infty)}\left\|\mathcal{P}_{\gamma-2}(X_r^x)\right\|_{L^{2\rho_1 q_2}(\Omega,\mathbb{R})}\|v_1\|_{L^{2\rho_2 q_2}(\Omega,\mathbb{R}^d)}\|v_2\|_{L^{2\rho_3 q_2}(\Omega,\mathbb{R}^d)},\\
&\left\|\mathcal{D}^3 X_t^x(v_1, v_2, v_3)\right\|_{L^{2}(\Omega,\mathbb{R}^d)} \le Ce^{-\alpha_3 t}\sup_{r\in[0,\infty)}\left\|\mathcal{P}_{\gamma-2}(X_r^x)\right\|_{L^{2\max\{\rho_1,\rho_3\rho_4\}}(\Omega,\mathbb{R})}\times\\
&\qquad\qquad \|v_1\|_{L^{2\rho_2}(\Omega,\mathbb{R}^d)}\|v_2\|_{L^{2\rho_3\rho_5}(\Omega,\mathbb{R}^d)}\|v_3\|_{L^{2\rho_3\rho_6}(\Omega,\mathbb{R}^d)},
\end{aligned}$$

where $\alpha_1 = \alpha/2$, $\alpha_2 = (q_2\alpha - \tilde{\epsilon}_1)/q_2$, $\tilde{\epsilon}_1 \in (0, q_2\alpha)$ *and* $\alpha_3 = \alpha - \tilde{\epsilon}_4$, $\tilde{\epsilon}_4 \in (0, \alpha)$.

The proof of Lemma 5.2 will be presented in Appendix C.1. As a consequence of Lemma 5.2, the uniform estimate of the derivatives of $u(t.\cdot)$ is obtained by the following lemma.

Lemma 5.3. *For any* $x \in \mathbb{R}^d$ *and some random variables* $v_1 \in L^{2\rho_2}(\Omega, \mathbb{R}^d)$, $v_2 \in L^{2\rho_3\rho_5}(\Omega, \mathbb{R}^d)$, $v_3 \in L^{2\rho_3\rho_6}(\Omega, \mathbb{R}^d)$, *where* $\rho_1, \rho_2, \rho_3, \rho_4, \rho_5, \rho_6 > 1$ *satisfying* $1/\rho_1 + 1/\rho_2 + 1/\rho_3 = 1$ *and* $1/\rho_4 + 1/\rho_5 + 1/\rho_6 = 1$, *let Assumptions 2.1-2.4 be fulfilled with* $2\lambda_1 > \max\{L_1, L_2\}$ *and*

$$p_0 \in \Big[\max\{\rho_1, \rho_3\rho_4\} \times (\gamma - 2), \infty\Big) \cap [1, \infty). \tag{5.11}$$

Then the function u defined by (5.4) satisfies

$$\begin{aligned}
&\|Du(t,x)v_1\|_{L^1(\Omega,\mathbb{R})} \le Ce^{-\alpha_1 t}\|v_1\|_{L^2(\Omega,\mathbb{R}^d)},\\
&\left\|D^2 u(t,x)(v_1, v_2)\right\|_{L^1(\Omega,\mathbb{R})}\\
&\le Ce^{-\tilde{\alpha}_2 t}\sup_{r\in[0,\infty)}\left\|\mathcal{P}_{\gamma-2}(X_r^x)\right\|_{L^{2\rho_1}(\Omega,\mathbb{R})}\|v_1\|_{L^{2\rho_2}(\Omega,\mathbb{R}^d)}\|v_2\|_{L^{2\rho_3}(\Omega,\mathbb{R}^d)},
\end{aligned}$$

and

$$\begin{aligned}
&\left\|D^3 u(t,x)(v_1, v_2, v_3)\right\|_{L^1(\Omega,\mathbb{R})} \le Ce^{-\tilde{\alpha}_3 t}\sup_{r\in[0,\infty)}\left\|\mathcal{P}_{\gamma-2}(X_r^x)\right\|_{L^{2\max\{\rho_1,\rho_3\rho_4\}}(\Omega,\mathbb{R})}\times\\
&\qquad\qquad \|v_1\|_{L^{2\rho_2}(\Omega,\mathbb{R}^d)}\|v_2\|_{L^{2\rho_3\rho_5}(\Omega,\mathbb{R}^d)}\|v_3\|_{L^{2\rho_3\rho_6}(\Omega,\mathbb{R}^d)},
\end{aligned}$$

where α_1, $\tilde{\alpha}_2$ *and* $\tilde{\alpha}_3$ *are positive constants, with the latter two depending on* α_1, α_2 *and* α_3 *defined as Lemma 5.2, i.e.*

$$\tilde{\alpha}_2 := \min\{2\alpha_1, \alpha_2\}, \quad \tilde{\alpha}_3 := \min\{3\alpha_1, \alpha_1 + \alpha_2, \alpha_3\}.$$

Remark 5.4. Bearing Lemma 5.3 in mind, we obtain that given the test function $\varphi \in C_b^3(\mathbb{R}^d)$ and $t > 0$, the function $u(t, \cdot) \in C_b^3(\mathbb{R}^d)$. Then $u(t, x)$ is the unique solution of (5.5) with initial data $x \in \mathbb{R}^d$ (see Theorem 1.6.2 in [6]).

The proof of Lemma 5.3 is given in Appendix C.2. Moreover, Lemma 5.3 apparently yields the contractivity of $u(t,\cdot)$, which is also derived by Lemma 3.4. Thus, one has the following result.

Corollary 5.5. *Let Assumptions 2.1-2.4 hold with $2\lambda_1 > \max\{L_1, L_2\}$, and recall that $\alpha_1 = \alpha/2$, then*

$$\|u(t,\zeta_1) - u(t,\zeta_2)\|_{L^1(\Omega,\mathbb{R})} \le Ce^{-\alpha_1 t}\|\zeta_1 - \zeta_2\|_{L^2(\Omega,\mathbb{R}^d)}.$$

Before proceeding further, there is no guarantee that the LTPE method (1.3) is continuous in the whole time interval since the numerical solutions are prevented from leaving a ball, whose radius depends on the timestep size, in each iteration. To address this issue and fully exploit the Kolmogorov equations, we recall the continuous version of the LTPE scheme (1.3) defined in (1.6), i.e., $\{\mathbb{Z}^n(t)\}_{t\in[t_n,t_{n+1}]}$, $n \in \mathbb{N}_0$. It is time to show the next lemma concerning some regular estimates of this process.

Lemma 5.6. *Let Assumptions 2.1, 2.2, 2.4 hold with $2\lambda_1 > L_1$. For $p \in [1, p_0)$ and $t \in [t_n, t_{n+1}]$, $n \in \mathbb{N}_0$,*

$$\mathbb{E}\left[\left\|\mathbb{Z}^n(t)\right\|^{2p}\right] \le C_A. \tag{5.12}$$

Moreover, for $p \in [1, p_0/\gamma]$ and $s, t \in [t_n, t_{n+1}]$, then

$$\mathbb{E}\left[\left\|\mathbb{Z}^n(t) - \mathbb{Z}^n(s)\right\|^{2p}\right] \le C_A|t-s|^p. \tag{5.13}$$

The proof of Lemma 5.6 is presented in Appendix C.3. At this time, we would like to present the error estimate between the random variable $\zeta \in \mathbb{R}^d$ and its projection $\mathscr{P}(\zeta) \in \mathbb{R}^d$, which is defined by (1.3).

Lemma 5.7. *Recall γ given in Assumption 2.4, let $\zeta \in L^{8\gamma+2}(\Omega,\mathbb{R}^d)$ and let $\mathscr{P}(\zeta)$ be defined as (1.3), then*

$$\mathbb{E}\left[\|\zeta - \mathscr{P}(\zeta)\|^2\right] \le Ch^4\mathbb{E}\left[\|\zeta\|^{8\gamma+2}\right].$$

The proof of Lemma 5.7 is found in Appendix C.4. Up to this point, we have developed sufficient machinery to obtain the uniform weak error estimate of the SDE (1.1) and the LTPE scheme (1.3) as below.

Theorem 5.8. *Let Assumptions 2.1-2.4 hold with $p_0 \ge \max\{4\gamma+1, 5\gamma-4\}$ and $2\lambda_1 > \max\{L_1, L_2\}$. Also, let h satisfy*

$$h \in \left(0, \min\left\{\frac{1}{2(1-\theta)\lambda_1}, \frac{p_0-p}{(1-\theta)(2p_0-p-1)\lambda_1}, \frac{1}{(1-\theta)\lambda_d}, 1\right\}\right),$$

where $p \in (1, p_0) \cap \mathbb{N}$. Moreover, denote by $\{X_t^{x_0}\}_{t\in[0,\infty)}$ and $\{Y_n^{x_0}\}_{n\in\mathbb{N}_0}$, the solutions to SDE (1.1) and the LTPE numerical scheme (1.3) with the initial state x_0, respectively. Then, given any terminal time $T \in (0,\infty)$ such that $Nh = T$, $N \in \mathbb{N}$, for every test function $\varphi \in C_b^3(\mathbb{R}^d)$,

$$\left|\mathbb{E}\left[\varphi(Y_N^{x_0})\right] - \mathbb{E}\left[\varphi(X_T^{x_0})\right]\right| \le C_A h.$$

Proof of Theorem 5.8. We begin with the following notation, for $n \in \{0, 1, \dots, N-1\}$, $N \in \mathbb{N}$,

$$Z_n := Y_n^{x_0} - \theta A Y_n^{x_0} h.$$

Due to the fact that

$$\mathbb{E}\left[\varphi(X_T^{x_0})\right] = u(T, x_0), \quad \mathbb{E}\left[\varphi(Y_N^{x_0})\right] = u(0, Y_N^{x_0}),$$

the weak error analysis is divided into several parts as

$$
\begin{aligned}
\left|\mathbb{E}\left[\varphi(Y_N^{x_0})\right]-\mathbb{E}\left[\varphi(X_T^{x_0})\right]\right| &= \left|\mathbb{E}\left[u(T,x_0)\right]-\mathbb{E}\left[u(0,Y_N^{x_0})\right]\right| \\
&\le \left|\mathbb{E}\left[u(0,Y_N^{x_0})\right]-\mathbb{E}\left[u(0,Z_N)\right]\right|+\left|\mathbb{E}\left[u(T,Z_0)\right]-\mathbb{E}\left[u(T,x_0)\right]\right| \\
&\quad+\left|\mathbb{E}\left[u(0,Z_N)\right]-\mathbb{E}\left[u(T,Z_0)\right]\right| \\
&=: J_1+J_2+J_3.
\end{aligned} \tag{5.14}
$$

For the estimate of J_1, one observes by the construction of Z_N and Lemma 4.3 that

$$
J_1 \le h\mathbb{E}\left[\|AY_N\|\right] \le C_A h. \tag{5.15}
$$

For the estimate of J_2, due to the fact $Z_0=x_0-\theta Ax_0$, Lemma 4.3 and Corollary 5.5, we derive, for some positive constant α_1 defined in Corollary 5.5,

$$
J_2 \le Ce^{-\alpha_1 T}\|Z_0-x_0\|_{L^2(\Omega,\mathbb{R}^d)} \le C_A e^{-\alpha_1 T}h. \tag{5.16}
$$

About J_3, by (1.6), it is easy to see $Z_{n+1}=\mathbb{Z}^n(t_{n+1})$. Then using a telescoping sum argument shows that

$$
\begin{aligned}
J_3 &= \left|\sum_{n=0}^{N-1}\mathbb{E}\big[u(T-t_{n+1},Z_{n+1})\big]-\mathbb{E}\big[u(T-t_n,Z_n)\big]\right| \\
&\le \left|\sum_{n=0}^{N-1}\mathbb{E}\left[u\big(T-t_n,\mathbb{Z}^n(t_n)\big)\right]-\mathbb{E}\big[u(T-t_n,Z_n)\big]\right| \\
&\quad+\left|\sum_{n=0}^{N-1}\mathbb{E}\left[u\big(T-t_{n+1},\mathbb{Z}^n(t_{n+1})\big)\right]-\mathbb{E}\left[u\big(T-t_n,\mathbb{Z}^n(t_n)\big)\right]\right| \\
&=: J_{3,1}+J_{3,2}.
\end{aligned}
$$

Together with the same analysis as (5.16), applying Lemma 5.7, Corollary 5.5 and the construction of $\mathbb{Z}^n(t_n)$ yields

$$
\begin{aligned}
J_{3,1} &\le C\sum_{n=0}^{N-1}e^{-\alpha_1(T-t_n)}\|Z_n-\mathbb{Z}^n(t_n)\|_{L^2(\Omega,\mathbb{R}^d)} \\
&\le C_A\sum_{n=0}^{N-1}he^{-\alpha_1(T-t_n)}\left(1+\sup_{0\le r\le N}\|Y_r\|_{L^{8\gamma+2}(\Omega,\mathbb{R}^d)}^{4\gamma}\right)h \\
&\le C_A\left(1+\sup_{0\le r\le N}\|Y_r\|_{L^{8\gamma+2}(\Omega,\mathbb{R}^d)}^{4\gamma+1}\right)h,
\end{aligned}
$$

where $\sum_{n=0}^{N-1}he^{-\alpha_1(T-t_n)}$ is uniformly bounded.

For the remaining term $J_{3,2}$, recalling the associated Kolmogorov equation (5.5), the Itô formula and (1.6), we obtain that, for every $n\in\{0,1,\dots,N-1\}$, $N\in\mathbb{N}$,

$$
\begin{aligned}
&\mathbb{E}\left[u\big(T-t_{n+1},\mathbb{Z}^n(t_{n+1})\big)\right]-\mathbb{E}\left[u\big(T-t_n,\mathbb{Z}^n(t_n)\big)\right] \\
&=\mathbb{E}\left[\int_{t_n}^{t_{n+1}}Du\big(T-s,\mathbb{Z}^n(s)\big)\Big(F\big(\mathscr{P}(Y_n)\big)-F\big(\mathbb{Z}^n(s)\big)\Big)\mathrm{d}s\right] \\
&\quad+\tfrac{1}{2}\sum_{j=1}^{m}\mathbb{E}\Bigg[\int_{t_n}^{t_{n+1}}D^2u\big(T-s,\mathbb{Z}^n(s)\big)\Big(g_j\big(\mathscr{P}(Y_n)\big),g_j\big(\mathscr{P}(Y_n)\big)\Big)
\end{aligned}
$$

$$
\begin{aligned}
&\quad - D^2 u\big(T-s, \mathbb{Z}^n(s)\big)\Big(g_j\big(\mathbb{Z}^n(s)\big), g_j\big(\mathbb{Z}^n(s)\big)\Big)\mathrm{d}s\Bigg] \\
&=: \mathbb{J}_1 + \mathbb{J}_2.
\end{aligned}
$$

A further decomposition is introduced for $\mathbb{J}_1$

$$
\begin{aligned}
\mathbb{J}_1 &= \mathbb{E}\left[\int_{t_n}^{t_{n+1}} Du\big(T-s, \mathbb{Z}^n(s)\big)\Big(F\big(\mathscr{P}(Y_n)\big) - F\big(\mathbb{Z}^n(t_n)\big)\Big)\mathrm{d}s\right] \\
&\quad + \mathbb{E}\left[\int_{t_n}^{t_{n+1}} Du\big(T-s, \mathbb{Z}^n(t_n)\big)\Big(F\big(\mathbb{Z}^n(t_n)\big) - F\big(\mathbb{Z}^n(s)\big)\Big)\mathrm{d}s\right] \\
&\quad + \mathbb{E}\left[\int_{t_n}^{t_{n+1}} \Big(Du\big(T-s, \mathbb{Z}^n(s)\big) - Du\big(T-s, \mathbb{Z}^n(t_n)\big)\Big)\Big(F\big(\mathbb{Z}^n(t_n)\big) - F\big(\mathbb{Z}^n(s)\big)\Big)\mathrm{d}s\right] \\
&=: \mathbb{J}_{1,1} + \mathbb{J}_{1,2} + \mathbb{J}_{1,3}.
\end{aligned}
$$

Now we are in a position to estimate $\mathbb{J}_{1,1}$. By Lemma 4.2, Lemma 5.3, (5.8) and the Hölder inequality, we get

$$
\begin{aligned}
\mathbb{J}_{1,1} &\le C \int_{t_n}^{t_{n+1}} e^{-\alpha_1(T-s)} \left\| F\big(\mathscr{P}(Y_n)\big) - F\big(\mathbb{Z}^n(t_n)\big)\right\|_{L^2(\Omega,\mathbb{R}^d)} \mathrm{d}s \\
&\le C_A \int_{t_n}^{t_{n+1}} e^{-\alpha_1(T-s)}\mathrm{d}s \left(1 + \sup_{0\le r\le N} \|Y_r\|^{\gamma}_{L^{2\gamma}(\Omega,\mathbb{R}^d)}\right) h.
\end{aligned}
$$

For the estimate of $\mathbb{J}_{1,2}$, the Taylor expansion and a conditional expectation argument gives

$$
-\mathbb{J}_{1,2} = \mathbb{E}\left[\int_{t_n}^{t_{n+1}} \Big\langle Du\big(T-s, \mathbb{Z}^n(t_n)\big), DF\big(\mathbb{Z}^n(t_n)\big)F\big(\mathscr{P}(Y_n)\big)(s-t_n) + \mathcal{R}_F\big(\mathbb{Z}^n(s), \mathbb{Z}^n(t_n)\big)\Big\rangle \mathrm{d}s\right],
$$

where

$$
\begin{aligned}
&\mathcal{R}_F\big(\mathbb{Z}^n(s), \mathbb{Z}^n(t_n)\big) \\
&:= \int_0^1 \Big(DF\Big(\mathbb{Z}^n(t_n) + r\big(\mathbb{Z}^n(s) - \mathbb{Z}^n(t_n)\big)\Big) - DF\big(\mathbb{Z}^n(t_n)\big)\Big)\Big(\mathbb{Z}^n(s) - \mathbb{Z}^n(t_n)\Big)\mathrm{d}r.
\end{aligned}
$$

Keeping (5.7) and (5.8) in mind, we obtain that

$$
\begin{aligned}
&(s-t_n)\left\| DF\big(\mathbb{Z}^n(t_n)\big)F\big(\mathscr{P}(Y_n)\big)\right\|_{L^2(\Omega,\mathbb{R}^d)} \\
&\le C_A \left\| \big(1 + \|\mathbb{Z}^n(t_n)\|\big)^{\gamma-1}\big(1 + \|\mathscr{P}(Y_n)\|\big)^{\gamma}\right\|_{L^2(\Omega,\mathbb{R})} (s-t_n).
\end{aligned}
$$

If $\gamma = 1$, by Lemma 4.3, one directly arrives at

$$
(s-t_n)\left\| DF\big(\mathbb{Z}^n(t_n)\big)F\big(\mathscr{P}(Y_n)\big)\right\|_{L^2(\Omega,\mathbb{R}^d)} \le C_A\left(1 + \sup_{0\le r\le N}\|Y_r\|_{L^2(\Omega,\mathbb{R}^d)}\right) h.
$$

If $\gamma > 1$, using the Hölder inequality yields

$$\begin{aligned}&\Big\|\big(1+\|\mathbb{Z}^n(t_n)\|\big)^{\gamma-1}\big(1+\|\mathscr{P}(Y_n)\|\big)^{\gamma}\Big\|_{L^2(\Omega,\mathbb{R})}\\&\le\Big\|\big(1+\|\mathbb{Z}^n(t_n)\|\big)^{\gamma-1}\Big\|_{L^{2k_1}(\Omega,\mathbb{R})}\Big\|\big(1+\|\mathscr{P}(Y_n)\|\big)^{\gamma}\Big\|_{L^{2k_2}(\Omega,\mathbb{R})},\end{aligned}$$

where we take $k_1=(2\gamma-1)/(\gamma-1)$ and $k_2=(2\gamma-1)/\gamma$ with Lemma 4.2, Lemma 3.3 and Lemma 5.6 to get

$$\begin{aligned}&(s-t_n)\left\|DF\big(\mathbb{Z}^n(t_n)\big)F\big(\mathscr{P}(Y_n)\big)\right\|_{L^2(\Omega,\mathbb{R}^d)}\\&\le C_A\left(1+\|\mathbb{Z}^n(t_n)\|^{\gamma-1}_{L^{4\gamma-2}(\Omega,\mathbb{R}^d)}\right)\left(1+\|Y_n\|^{\gamma}_{L^{4\gamma-2}(\Omega,\mathbb{R}^d)}\right)h\\&\le C_A\left(1+\sup_{0\le r\le N}\|Y_r\|^{2\gamma-1}_{L^{4\gamma-2}(\Omega,\mathbb{R}^d)}\right)h.\end{aligned}$$

Similarly, we also attain

$$\begin{aligned}&\left\|\mathcal{R}_F\big(\mathbb{Z}^n(s),\mathbb{Z}^n(t_n)\big)\right\|_{L^2(\Omega,\mathbb{R}^d)}\\&\le C\left\|\mathcal{P}_{\gamma-2}\Big(\mathbb{Z}^n(t_n)+r\big(\mathbb{Z}^n(s)-\mathbb{Z}^n(t_n)\big),\mathbb{Z}^n(t_n)\Big)\left\|\mathbb{Z}^n(s)-\mathbb{Z}^n(t_n)\right\|^2\right\|_{L^2(\Omega,\mathbb{R})}\\&\le C_A\left(1+\sup_{0\le r\le N}\|Y_r\|^{\max\{2\gamma,3\gamma-2\}}_{L^{\max\{4\gamma,6\gamma-4\}}(\Omega,\mathbb{R}^d)}\right)h.\end{aligned}$$

Then it follows from Lemma 5.3 that

$$\mathbb{J}_{1,2}\le C_A\int_{t_n}^{t_{n+1}}e^{-\alpha_1(T-s)}\mathrm{d}s\left(1+\sup_{0\le r\le N}\|Y_r\|^{\max\{2\gamma,3\gamma-2\}}_{L^{\max\{4\gamma,6\gamma-4\}}(\Omega,\mathbb{R}^d)}\right)h.$$

For the estimate of $\mathbb{J}_{1,3}$, the Taylor expansion to $u(t,\cdot)$ shows, denoting $\widetilde{\upsilon}(\bar{r}):=\mathbb{Z}^n(t_n)+\bar{r}\big(\mathbb{Z}^n(s)-\mathbb{Z}^n(t_n)\big)$,

$$\mathbb{J}_{1,3}=\mathbb{E}\left[\int_{t_n}^{t_{n+1}}\int_0^1 D^2u(T-s,\widetilde{\upsilon}(\bar{r}))\Big(\mathbb{Z}^n(s)-\mathbb{Z}^n(t_n),F\big(\mathbb{Z}^n(t_n)\big)-F\big(\mathbb{Z}^n(s)\big)\Big)\mathrm{d}\bar{r}\mathrm{d}s\right].$$

Applying Lemma 3.3, Lemma 5.3 and the Hölder inequality yields,

$$\begin{aligned}\mathbb{J}_{1,3}\le C\int_{t_n}^{t_{n+1}}e^{-\tilde{\alpha}_2(T-s)}&\sup_{r\in[0,T]}\left\|\mathcal{P}_{\gamma-2}(X_r^{\widetilde{\upsilon}})\right\|_{L^{2\rho_1}(\Omega,\mathbb{R})}\left\|\mathbb{Z}^n(s)-\mathbb{Z}^n(t_n)\right\|_{L^{2\rho_2}(\Omega,\mathbb{R}^d)}\times\\&\left\|F\big(\mathbb{Z}^n(t_n)\big)-F\big(\mathbb{Z}^n(s)\big)\right\|_{L^{2\rho_3}(\Omega,\mathbb{R}^d)}\mathrm{d}s.\end{aligned}$$

We need to discuss the estimation of $\mathbb{J}_{1,3}$ through the range of γ. For the case that $\gamma>2$, taking $\rho_1=(4\gamma-3)/(\gamma-2)$, $\rho_2=(4\gamma-3)/\gamma$ and $\rho_3=(4\gamma-3)/(2\gamma-1)$ with Assumption 2.4 and Lemma 5.6 gives

$$\mathbb{J}_{1,3}\le C_A\int_{t_n}^{t_{n+1}}e^{-\tilde{\alpha}_2(T-s)}\mathrm{d}s\left(1+\sup_{0\le r\le N}\|Y_r\|^{4\gamma-3}_{L^{8\gamma-6}(\Omega,\mathbb{R}^d)}\right)h.$$

For the case that $1\le\gamma\le 2$, choosing $\rho_1=\infty$, $\rho_2=(3\gamma-1)/\gamma$ and $\rho_3=(3\gamma-1)/(2\gamma-1)$ shows

$$\mathbb{J}_{1,3}\le C_A\int_{t_n}^{t_{n+1}}e^{-\tilde{\alpha}_2(T-s)}\mathrm{d}s\left(1+\sup_{0\le r\le N}\|Y_r\|^{3\gamma-1}_{L^{6\gamma-2}(\Omega,\mathbb{R}^d)}\right)h.$$

Consequently, combining the estimations of $\mathbb{J}_{1,1}$-$\mathbb{J}_{1,3}$ leads to

$$\mathbb{J}_1 \le C_A \int_{t_n}^{t_{n+1}} e^{-\min(\alpha_1,\tilde{\alpha}_2)(T-s)} \mathrm{d}s \left(1 + \sup_{0\le r\le N} \|Y_r\|^{\max\{3\gamma-1,4\gamma-3\}}_{L^{\max\{6\gamma-2,8\gamma-6\}}(\Omega,\mathbb{R}^d)}\right) h. \tag{5.17}$$

For the estimate of $\mathbb{J}_2$, we here use the following equality, for any matrix $U \in \mathbb{R}^{d\times d}$ and any $a, b \in \mathbb{R}^d$,

$$a^T U a - b^T U b = (a-b)^T U (a-b) + (a-b)^T U b + b^T U (a-b).$$

As a result, one shows a further decomposition of $\mathbb{J}_2$ as follows,

$$\begin{aligned}
\mathbb{J}_2 &= \tfrac{1}{2}\sum_{j=1}^m \mathbb{E}\left[\int_{t_n}^{t_{n+1}} D^2 u\big(T-s, \mathbb{Z}^n(s)\big)\Big(g_j\big(\mathbb{Z}^n(s)\big) - g_j\big(\mathscr{P}(Y_n)\big), g_j\big(\mathbb{Z}^n(s)\big) - g_j\big(\mathscr{P}(Y_n)\big)\Big)\mathrm{d}s\right] \\
&\quad \tfrac{1}{2}\sum_{j=1}^m \mathbb{E}\left[\int_{t_n}^{t_{n+1}} D^2 u\big(T-s, \mathbb{Z}^n(s)\big)\Big(g_j\big(\mathbb{Z}^n(s)\big) - g_j\big(\mathscr{P}(Y_n)\big), g_j\big(\mathscr{P}(Y_n)\big)\Big)\mathrm{d}s\right] \\
&\quad \tfrac{1}{2}\sum_{j=1}^m \mathbb{E}\left[\int_{t_n}^{t_{n+1}} D^2 u\big(T-s, \mathbb{Z}^n(s)\big)\Big(g_j\big(\mathscr{P}(Y_n)\big), g_j\big(\mathbb{Z}^n(s)\big) - g_j\big(\mathscr{P}(Y_n)\big)\Big)\mathrm{d}s\right] \\
&=: \mathbb{J}_{2,1} + \mathbb{J}_{2,2} + \mathbb{J}_{2,3}.
\end{aligned}$$

For $\mathbb{J}_{2,1}$, combining Lemma 5.2, Lemma 5.3 with $\rho_2 = \rho_3$ implies

$$\begin{aligned}
\mathbb{J}_{2,1} \le C \sum_{j=1}^m \int_{t_n}^{t_{n+1}} e^{-\tilde{\alpha}_2(T-s)} \sup_{r\in[0,T]} \left\|\mathcal{P}_{\gamma-2}\left(X_r^{\mathbb{Z}^n(s)}\right)\right\|_{L^{2\rho_1}(\Omega,\mathbb{R})} \big\|g_j\big(\mathbb{Z}^n(s)\big) \\
- g_j\big(\mathscr{P}(Y_n)\big)\big\|^2_{L^{2\rho_2}(\Omega,\mathbb{R}^d)} \mathrm{d}s,
\end{aligned}$$

where using Lemma 5.6, Assumption 2.4 and the Hölder inequality gives that, for some $\rho_2 \ge 1$, $j \in \{1,\dots,m\}$,

$$\begin{aligned}
&\big\|g_j\big(\mathbb{Z}^n(s)\big) - g_j\big(\mathscr{P}(Y_n)\big)\big\|_{L^{2\rho_2}(\Omega,\mathbb{R}^d)} \\
&\le \big\|g_j(\mathbb{Z}^n(s)) - g_j\big(\mathbb{Z}^n(t_n)\big)\big\|_{L^{2\rho_2}(\Omega,\mathbb{R}^d)} + \big\|g_j\big(\mathbb{Z}^n(t_n)\big) - g_j\big(\mathscr{P}(Y_n)\big)\big\|_{L^{2\rho_2}(\Omega,\mathbb{R}^d)} \\
&\le C_A h^{\frac{1}{2}} \left(1 + \sup_{0\le r\le N} \|Y_r\|^{\frac{3\gamma-1}{2}}_{L^{\rho_2(3\gamma-1)}(\Omega,\mathbb{R}^d)}\right) + C_A h \left(1 + \sup_{0\le r\le N} \|Y_r\|^{\frac{\gamma+1}{2}}_{L^{\rho_2(\gamma+1)}(\Omega,\mathbb{R}^d)}\right) \\
&\le C_A h^{\frac{1}{2}} \left(1 + \sup_{0\le r\le N} \|Y_r\|^{\frac{3\gamma-1}{2}}_{L^{\rho_2(3\gamma-1)}(\Omega,\mathbb{R}^d)}\right).
\end{aligned}$$

For $\gamma > 2$, choosing $\rho_1 = (4\gamma-3)/(\gamma-2)$, $\rho_2 = \rho_3 = (8\gamma-6)/(3\gamma-1)$ yields

$$\mathbb{J}_{2,1} \le C_A \int_{t_n}^{t_{n+1}} e^{-\tilde{\alpha}_2(T-s)} \mathrm{d}s \left(1 + \sup_{0\le r\le N} \|Y_r\|^{4\gamma-3}_{L^{8\gamma-6}(\Omega,\mathbb{R}^d)}\right) h.$$

For $1 \le \gamma \le 2$, taking $\rho_1 = \infty$, $\rho_2 = \rho_3 = 2$ leads to

$$\mathbb{J}_{2,1} \le C_A \int_{t_n}^{t_{n+1}} e^{-\tilde{\alpha}_2(T-s)} \mathrm{d}s \left(1 + \sup_{0\le r\le N} \|Y_r\|^{3\gamma-1}_{L^{6\gamma-2}(\Omega,\mathbb{R}^d)}\right) h.$$

The estimates of $\mathbb{J}_{2,2}$ and $\mathbb{J}_{2,3}$ are in the same way. As a consequence, we take $\mathbb{J}_{2,2}$ as an example. Then an application of the Taylor expansion with a conditional expectation argument yields that

$$
\begin{aligned}
\mathbb{J}_{2,2} &= \tfrac{1}{2}\sum_{j=1}^{m}\mathbb{E}\left[\int_{t_n}^{t_{n+1}} D^2u\big(T-s,\mathbb{Z}^n(s)\big)\Big(g_j\big(\mathbb{Z}^n(s)\big)-g_j\big(\mathbb{Z}^n(t_n)\big),g_j\big(\mathscr{P}(Y_n)\big)\Big)\mathrm{d}s\right] \\
&\quad + \tfrac{1}{2}\sum_{j=1}^{m}\mathbb{E}\left[\int_{t_n}^{t_{n+1}} D^2u\big(T-s,\mathbb{Z}^n(s)\big)\Big(g_j\big(\mathbb{Z}^n(t_n)\big)-g_j\big(\mathscr{P}(Y_n)\big),g_j\big(\mathscr{P}(Y_n)\big)\Big)\mathrm{d}s\right] \\
&= \tfrac{1}{2}\sum_{j=1}^{m}\mathbb{E}\left[\int_{t_n}^{t_{n+1}} (s-t_n)D^2u\big(T-s,\mathbb{Z}^n(t_n)\big)\Big(Dg_j\big(\mathbb{Z}^n(t_n)\big)F\big(\mathscr{P}(Y_n)\big),g_j\big(\mathscr{P}(Y_n)\big)\Big)\mathrm{d}s\right] \\
&\quad + \tfrac{1}{2}\sum_{j=1}^{m}\mathbb{E}\left[\int_{t_n}^{t_{n+1}} D^2u\big(T-s,\mathbb{Z}^n(t_n)\big)\Big(\mathcal{R}_{g_j}\big(\mathbb{Z}^n(s),\mathbb{Z}^n(t_n)\big),g_j\big(\mathscr{P}(Y_n)\big)\Big)\mathrm{d}s\right] \\
&\quad + \tfrac{1}{2}\sum_{j=1}^{m}\mathbb{E}\left[\int_{t_n}^{t_{n+1}} \Big(D^2u\big(T-s,\mathbb{Z}^n(s)\big)-D^2u\big(T-s,\mathbb{Z}^n(t_n)\big)\Big)\Big(g_j(\mathbb{Z}^n(s))\right. \\
&\quad \left. -\, g_j\big(\mathbb{Z}^n(t_n)\big),g_j\big(\mathscr{P}(Y_n)\big)\Big)\mathrm{d}s\right] \\
&\quad + \tfrac{1}{2}\sum_{j=1}^{m}\mathbb{E}\left[\int_{t_n}^{t_{n+1}} D^2u\big(T-s,\mathbb{Z}^n(s)\big)\Big(g_j\big(\mathbb{Z}^n(t_n)\big)-g_j\big(\mathscr{P}(Y_n)\big),g_j\big(\mathscr{P}(Y_n)\big)\Big)\mathrm{d}s\right] \\
&=: \mathbb{J}_{2,2,1}+\mathbb{J}_{2,2,2}+\mathbb{J}_{2,2,3}+\mathbb{J}_{2,2,4},
\end{aligned}
$$

where we denote that, for $j\in\{1,\dots,m\}$,

$$
\begin{aligned}
&\mathcal{R}_{g_j}\Big(\mathbb{Z}^n(s),\mathbb{Z}^n(t_n)\Big) \\
&:= \int_0^1 \left[Dg_j\left(\mathbb{Z}^n(t_n)+r\big(\mathbb{Z}^n(s)-\mathbb{Z}^n(t_n)\big)\right)-Dg_j\big(\mathbb{Z}^n(t_n)\big)\right]\left(\mathbb{Z}^n(s)-\mathbb{Z}^n(t_n)\right)\mathrm{d}r.
\end{aligned}
$$

Using Lemma 3.3 and Lemma 5.3 implies that

$$
\begin{aligned}
\mathbb{J}_{2,2,1} \le C_A \int_{t_n}^{t_{n+1}} (s-t_n)e^{-\tilde{\alpha}_2(T-s)}\mathrm{d}s \sup_{r\in[0,T]}\left\|\mathcal{P}_{\gamma-2}\left(X_r^{\mathbb{Z}^n(t_n)}\right)\right\|_{L^{2\rho_1}(\Omega,\mathbb{R})}\times \\
\sum_{j=1}^{m}\left\|Dg_j\big(\mathbb{Z}^n(t_n)\big)F\big(\mathscr{P}(Y_n)\big)\right\|_{L^{2\rho_2}(\Omega,\mathbb{R}^d)}\left\|g_j\big(\mathscr{P}(Y_n)\big)\right\|_{L^{2\rho_3}(\Omega,\mathbb{R}^d)}.
\end{aligned}
$$

For the case that $\gamma\ge 2$, it follows from Assumption 2.4 and Lemma 4.3 that

$$
\mathbb{J}_{2,2,1} \le C_A \int_{t_n}^{t_{n+1}} e^{-\tilde{\alpha}_2(T-s)}\mathrm{d}s\left(1+\sup_{0\le r\le N}\|Y_r\|^{3\gamma-2}_{L^{6\gamma-4}(\Omega,\mathbb{R}^d)}\right)h,
$$

where we let $\rho_1=(3\gamma-2)/(\gamma-2)$, $\rho_2=(6\gamma-4)/(3\gamma-1)$ and $\rho_3=(6\gamma-4)/(\gamma-1)$. For $1\le\gamma\le 2$, taking $\rho_1=\infty$, $\rho_2=4\gamma/(3\gamma-1)$ and $\rho_3=4\gamma/(\gamma+1)$ leads to

$$\mathbb{J}_{2,2,1} \le C_A \int_{t_n}^{t_{n+1}} e^{-\tilde{\alpha}_2(T-s)} \mathrm{d}s \left(1 + \sup_{0\le r\le N} \|Y_r\|^{2\gamma}_{L^{4\gamma}(\Omega,\mathbb{R}^d)}\right) h.$$

Similarly, one gets

$$\begin{aligned}\mathbb{J}_{2,2,2} \le C \int_{t_n}^{t_{n+1}} e^{-\tilde{\alpha}_2(T-s)} \sup_{r\in[0,T]} \left\| \mathcal{P}_{\gamma-2}\left(X_r^{\mathbb{Z}^n(t_n)}\right)\right\|_{L^{2\rho_1}(\Omega,\mathbb{R})} \times \\ \sum_{j=1}^m \left\| \mathcal{R}_{g_j}\left(\mathbb{Z}^n(s), \mathbb{Z}^n(t_n)\right)\right\|_{L^{2\rho_2}(\Omega,\mathbb{R}^d)} \left\| g_j\big(\mathscr{P}(Y_n)\big)\right\|_{L^{2\rho_3}(\Omega,\mathbb{R}^d)} \mathrm{d}s,\end{aligned} \tag{5.18}$$

where one obtains easily from Assumption 2.4 and Lemma 5.6 that, for some $\rho_2 \ge 1$, $j \in \{1,\dots,m\}$,

$$\begin{aligned}&\left\| \mathcal{R}_{g_j}\left(\mathbb{Z}^n(s), \mathbb{Z}^n(t_n)\right)\right\|_{L^{2\rho_2}(\Omega,\mathbb{R}^d)} \\ &\le \int_0^1 \left\| \Big[Dg_j\Big(\mathbb{Z}^n(t_n) + r\big(\mathbb{Z}^n(s) - \mathbb{Z}^n(t_n)\big)\Big) - Dg_j\big(\mathbb{Z}^n(t_n)\big)\Big]\Big(\mathbb{Z}^n(s) - \mathbb{Z}^n(t_n)\Big)\right\|_{L^{2\rho_2}(\Omega,\mathbb{R}^d)} \mathrm{d}r \\ &\le C \int_0^1 \left\| \Big(1 + \|r\mathbb{Z}^n(s) + (1-r)\mathbb{Z}^n(t_n)\| + \|\mathbb{Z}^n(t_n)\|\Big)^{\frac{\gamma-3}{2}} \left\|\mathbb{Z}^n(s) - \mathbb{Z}^n(t_n)\right\|^2 \right\|_{L^{2\rho_2}(\Omega;\mathbb{R})} \mathrm{d}r \\ &\le C_A \left(1 + \sup_{0\le r\le N} \|Y_r\|^{\max\{2\gamma, \frac{5\gamma-3}{2}\}}_{L^{\max\{4\rho_2\gamma, \rho_2(5\gamma-3)\}}(\Omega,\mathbb{R}^d)}\right) h.\end{aligned}$$

Putting this estimate into (5.18) with the Hölder inequality and using the same distinction of cases for γ as above yield

$$\mathbb{J}_{2,2,2} \le C_A \int_{t_n}^{t_{n+1}} e^{-\tilde{\alpha}_2(T-s)} \mathrm{d}s \left(1 + \sup_{0\le r\le N} \|Y_r\|^{\max\{\frac{5\gamma+1}{2}, \frac{7\gamma-3}{2}, 4\gamma-3\}}_{L^{\max\{5\gamma+1, 7\gamma-3, 8\gamma-6\}}(\Omega,\mathbb{R}^d)}\right) h.$$

Then, the Taylor expansion and Lemma 5.3 are used to give that, setting $\widetilde{\upsilon}_1(\bar{r}) := \mathbb{Z}^n(t_n) + \bar{r}\big(\mathbb{Z}^n(s) - \mathbb{Z}^n(t_n)\big)$,

$$\begin{aligned}\mathbb{J}_{2,2,3} &= \tfrac{1}{2} \sum_{j=1}^m \mathbb{E}\Bigg[\int_{t_n}^{t_{n+1}} \int_0^1 D^3 u(T-s, \widetilde{\upsilon}_1(\bar{r}))\Big(\mathbb{Z}^n(s) - \mathbb{Z}^n(t_n), g_j(\mathbb{Z}^n(s)) \\ &\qquad - g_j\big(\mathbb{Z}^n(t_n)\big), g_j\big(\mathscr{P}(Y_n)\big)\Big) \mathrm{d}\bar{r}\mathrm{d}s\Bigg] \\ &\le C \sum_{j=1}^m \int_{t_n}^{t_{n+1}} e^{-\tilde{\alpha}_3(T-s)} \sup_{r\in[0,T]} \left\| \mathcal{P}_{\gamma-2}\left(X_r^{\widetilde{\upsilon}_1}\right)\right\|_{L^{2\max\{\rho_1,\rho_3\rho_4\}}(\Omega,\mathbb{R})} \times \\ &\qquad \left\| \mathbb{Z}^n(s) - \mathbb{Z}^n(t_n)\right\|_{L^{2\rho_2}(\Omega,\mathbb{R}^d)} \left\| g_j(\mathbb{Z}^n(s)) - g_j\big(\mathbb{Z}^n(t_n)\big)\right\|_{L^{2\rho_3\rho_5}(\Omega,\mathbb{R}^d)} \times \\ &\qquad \left\| g_j\big(\mathscr{P}(Y_n)\big)\right\|_{L^{2\rho_3\rho_6}(\Omega,\mathbb{R}^d)} \mathrm{d}s.\end{aligned}$$

Equipped with Lemma 5.6, Assumption 2.4 and the Hölder inequality, for $\gamma > 2$, one chooses $\rho_1 = (5\gamma-4)/(\gamma-2)$, $\rho_2 = (5\gamma-4)/\gamma$, $\rho_3 = (5\gamma-4)/(3\gamma-2)$, $\rho_4 = (3\gamma-2)/(\gamma-2)$, $\rho_5 = (6\gamma-4)/(3\gamma-1)$ and $\rho_6 = (6\gamma-4)/(\gamma+1)$ to get,

$$\mathbb{J}_{2,2,3} \leq C_A \int_{t_n}^{t_{n+1}} e^{-\tilde{\alpha}_3(T-s)} \mathrm{d}s \left(1 + \sup_{0 \leq r \leq N} \|Y_r\|^{5\gamma-4}_{L^{10\gamma-8}(\Omega,\mathbb{R}^d)}\right) h.$$

For $1 \leq \gamma \leq 2$, taking $\rho_1 = \rho_4 = \infty$, $\rho_2 = 3$, $\rho_3 = 3/2$, $\rho_5 = 4\gamma/(3\gamma-1)$ and $\rho_6 = 4\gamma/(\gamma+1)$ yields

$$\mathbb{J}_{2,2,3} \leq C_A \int_{t_n}^{t_{n+1}} e^{-\tilde{\alpha}_3(T-s)} \mathrm{d}s \left(1 + \sup_{0 \leq r \leq N} \|Y_r\|^{3\gamma}_{L^{6\gamma}(\Omega,\mathbb{R}^d)}\right) h.$$

By Lemma 5.3 with $q_2 = 1$, it is quite obvious that

$$\begin{aligned}\mathbb{J}_{2,2,4} \leq C \sum_{j=1}^{m} \int_{t_n}^{t_{n+1}} e^{-\tilde{\alpha}_2(T-s)} \sup_{r \in [0,T]} \left\| \mathcal{P}_{\gamma-2}\left(X_r^{\mathbb{Z}^n(s)}\right)\right\|_{L^{2\rho_1}(\Omega,\mathbb{R})} \big\| g_j\big(\mathbb{Z}^n(t_n)\big) \\ - g_j\big(\mathscr{P}(Y_n)\big)\big\|_{L^{2\rho_2}(\Omega,\mathbb{R}^d)} \times \\ \big\|g_j\big(\mathscr{P}(Y_n)\big)\big\|_{L^{2\rho_3}(\Omega,\mathbb{R}^d)} \mathrm{d}s.\end{aligned}$$

Following the same argument, we show that

$$\mathbb{J}_{2,2,4} \leq C_A \int_{t_n}^{t_{n+1}} e^{-\tilde{\alpha}_2(T-s)} \mathrm{d}s \left(1 + \sup_{0 \leq r \leq N} \|Y_r\|^{\max\{\gamma+1,2\gamma-1\}}_{L^{\max\{2\gamma+2,4\gamma-2\}}(\Omega,\mathbb{R}^d)}\right) h.$$

Hence, by the estimates of $\mathbb{J}_{2,2,1} - \mathbb{J}_{2,2,4}$, we deduce that

$$\max\left\{\mathbb{J}_{2,2}, \mathbb{J}_{2,3}\right\} \leq C_A \int_{t_n}^{t_{n+1}} e^{-\min(\tilde{\alpha}_2,\tilde{\alpha}_3)(T-s)} \mathrm{d}s \left(1 + \sup_{0 \leq r \leq N} \|Y_r\|^{\max\{3\gamma,5\gamma-4\}}_{L^{\max\{6\gamma,10\gamma-8\}}(\Omega,\mathbb{R}^d)}\right) h,$$

resulting in

$$\mathbb{J}_2 \leq C_A \int_{t_n}^{t_{n+1}} e^{-\min(\tilde{\alpha}_2,\tilde{\alpha}_3)(T-s)} \mathrm{d}s \left(1 + \sup_{0 \leq r \leq N} \|Y_r\|^{\max\{3\gamma,5\gamma-4\}}_{L^{\max\{6\gamma,10\gamma-8\}}(\Omega,\mathbb{R}^d)}\right) h.$$

Combining this with (5.17) leads to

$$\begin{aligned}J_{3,2} &\leq C_A h \left(1 + \sup_{0 \leq r \leq N} \|Y_r\|^{\max\{3\gamma,5\gamma-4\}}_{L^{\max\{6\gamma,10\gamma-8\}}(\Omega,\mathbb{R}^d)}\right) \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} e^{-\min(\alpha_1,\tilde{\alpha}_2,\tilde{\alpha}_3)(T-s)} \mathrm{d}s \\ &= C_A h \left(1 + \sup_{0 \leq r \leq N} \|Y_r\|^{\max\{3\gamma,5\gamma-4\}}_{L^{\max\{6\gamma,10\gamma-8\}}(\Omega,\mathbb{R}^d)}\right) \int_0^T e^{-\min(\alpha_1,\tilde{\alpha}_2,\tilde{\alpha}_3)(T-s)} \mathrm{d}s.\end{aligned}$$

It is known that

$$\int_0^T e^{-\min(\alpha_1,\tilde{\alpha}_2,\tilde{\alpha}_3)(T-s)} \mathrm{d}s = \tfrac{1-e^{-\min(\alpha_1,\tilde{\alpha}_2,\tilde{\alpha}_3)T}}{\min(\alpha_1,\tilde{\alpha}_2,\tilde{\alpha}_3)}$$

is uniformly bounded. All in all, we are in a position to derive the estimate of J_3 as

$$J_3 \leq C_A \left(1 + \sup_{0 \leq r \leq N} \|Y_r\|^{\max\{4\gamma+1,5\gamma-4\}}_{L^{\max\{8\gamma+2,10\gamma-8\}}(\Omega,\mathbb{R}^d)}\right) h. \tag{5.19}$$

Plugging (5.15), (5.16) and (5.19) into (5.14) gives

$$\left|\mathbb{E}\left[\varphi(Y_N^{X_0})\right]-\mathbb{E}\left[\varphi(X_T^{X_0})\right]\right|\le C_A\left(1+\sup_{0\le r\le N}\|Y_r\|_{L^{\max\{8\gamma+2,10\gamma-8\}}(\Omega,\mathbb{R}^d)}^{\max\{4\gamma+1,5\gamma-4\}}\right)h,$$

which completes the proof. □

To conclude, we deduce from Theorem 5.8 that the weak convergence order of the π and π_h is 1, i.e.

$$\left|\int_{\mathbb{R}^d}\varphi(x)\pi(\mathrm{d}x)-\int_{\mathbb{R}^d}\varphi(x)\pi_h(\mathrm{d}x)\right|\le C_A h,$$

since the constant C_A is independent of n in (5.3).

6. Numerical experiments

In this section, we illustrate the previous theoretical findings through three numerical examples: the scalar stochastic Ginzburg-Landau equation [17] in **Example 1**, the mean-reverting type model with super-linear coefficients [19,12] in **Example 2** and the third is the semi-linear stochastic partial equation (SPDE) [20,27] in **Example 3**.

For all three numerical experiments, we consider a terminal time $T=5$, the timesteps $h=2^{-6},2^{-7},2^{-8},2^{-9}$ and four different choices for test function $\varphi(\cdot)$,

$$\varphi(x)\in\{\arctan(\|x\|),\ e^{-\|x\|^2},\ \cos(\|x\|),\ \sin(\|x\|^2)\}.$$

The empirical mean of $\mathbb{E}[\varphi(X_T)]$ is estimated by a Monte Carlo approximation, involving 10,000 independent trajectories. It is worth noting that in **Example 2** we will test that the terminal time $T=5$ what we have chosen is appropriate.

Example 1. Consider the stochastic Ginzburg-Landau equation [17] from the theory of superconductivity as follows,

$$dX_t=\left(-X_t^3+\left(\alpha+\tfrac{1}{2}\sigma^2\right)X_t\right)\mathrm{d}t+\sigma X_t\,\mathrm{d}W_t,\quad \alpha,\sigma\in\mathbb{R}. \tag{6.1}$$

Let $\alpha=-2$, $\sigma=0.5$ and $X_0=1$. Then, all conditions in Assumptions 2.1-2.4 are meet with $\gamma=3$ and for any $p_0\ge 13$. We compute the equation (6.1) numerically using the explicit projected Euler method, i.e. $\theta=0$ in (1.3), and the *exact* solutions are identified with the corresponding numerical approximations at a fine stepsize $h_{exact}=2^{-14}$. Also, the reference lines of slope 0.5 and 1 are given here. It turns out in Fig. 1 that the weak convergence rate of the approximation errors of the projected Euler method decrease at a slope close to 1.

Example 2. Consider a scalar mean-reverting type model with super-linear coefficients in financial and energy markets as follows,

$$\mathrm{d}X_t=\left(b-\alpha X_t-\beta X_t^3\right)\mathrm{d}t+\sigma X_t^2\,\mathrm{d}W_t,\quad b,\alpha,\beta,\sigma\in\mathbb{R}. \tag{6.2}$$

Setting $b=0.3$, $\alpha=1$, $\beta=0.6$, $\sigma=0.2$ and $X_0=1$. The requirements from Assumptions 2.1-2.4 are verified with $\gamma=3$ and for any $p_0\in[13,31/2]$.

To assess the probability density of the LTPE scheme (1.3), we run 10,000 individual sample trajectories and discretize model (6.2) with three different θ, $\theta=0,0.5,1$, at the terminal time $T=5$ using a stepsize $h=2^{-14}$. Then the output data are divided into 20 uniformly spaced bins, enabling a detailed quantification of the data in each bin. As shown in Fig. 2, each discrete point on the histogram

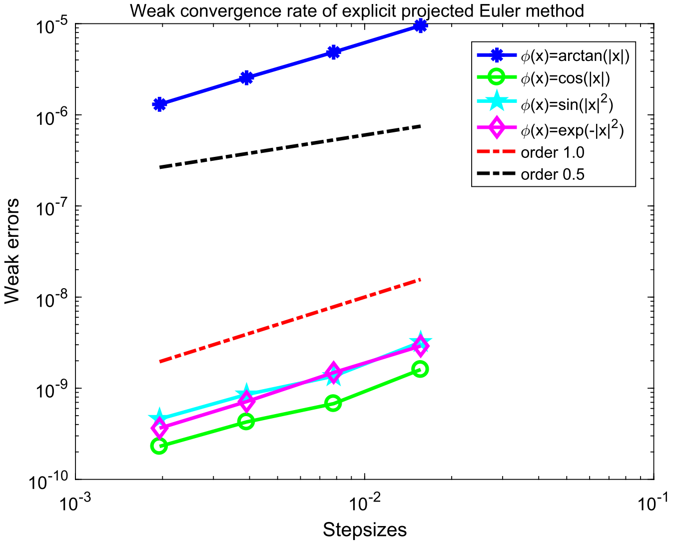

Fig. 1. Weak convergence rates of the explicit projected Euler method for stochastic Ginzburg-Landau model (6.1).

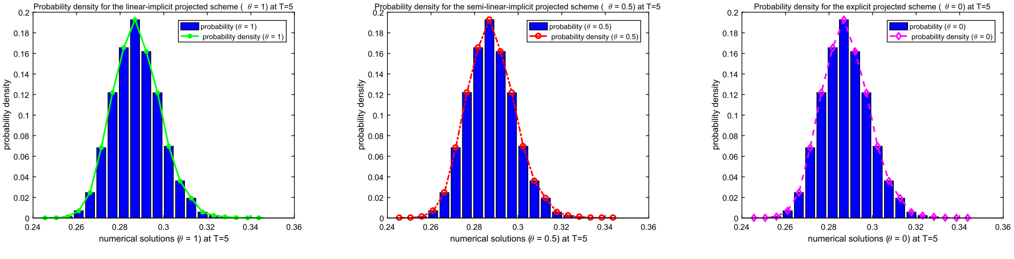

Fig. 2. Probability density of LTPE scheme method for discretizing the mean reverting model (6.2) with different θ.

is representative of the approximate probability of data falling into each designated bin with different θ. Following this, the linear interpolation technique is used to construct the approximate curve of the probability density with different choice of θ. The difference among approximated distributions has been measured, where the maximum absolute error is about 10^{-4}. Such a negligible difference indicates that the choice of time $T = 5$ is appropriate and explains why the distribution curves in Fig. 2 are almost identical.

Moving on to the convergence test, we discretize this model (6.2) by the semi-linear-implicit projected Euler method (i.e. $\theta = 0.5$ in (1.3)). To find the *exact* solutions, we discrete this model by the linear-implicit projected Euler method ($\theta = 1$ in (1.3)) at a fine stepsize $h_{exact} = 2^{-14}$. In Fig. 3, the weak error lines have slopes close to 1 for all cases.

Example 3. Consider the following semi-linear stochastic partial differential equation (SPDE),

$$\begin{cases} \mathrm{d}u(t,x) = \left[\frac{\partial^2}{\partial x^2} u(t,x) + u(t,x) - u^3(t,x)\right]\mathrm{d}t + g(u(t,x))\mathrm{d}W_t, & t \in (0,T],\ x \in (0,1), \\ u(t,0) = u(t,1) = 0, \\ u(0,x) = u_0(x), \end{cases} \tag{6.3}$$

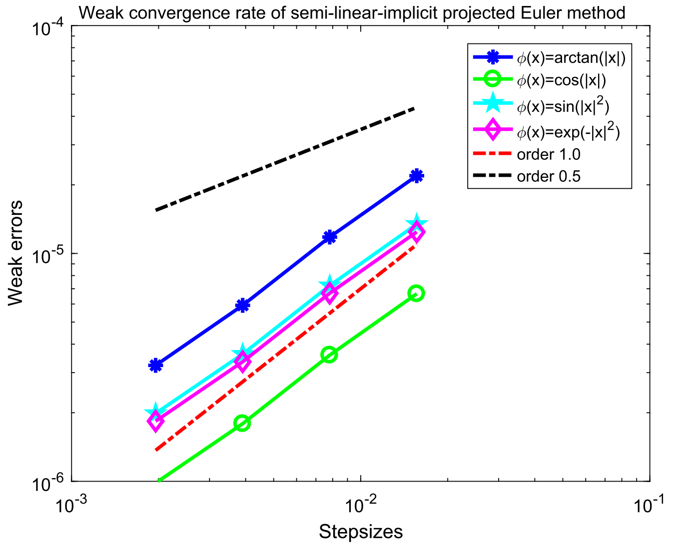

Fig. 3. Weak convergence rates of semi-linear-implicit projected Euler method for the mean reverting model (6.2).

where $g : \mathbb{R} \to \mathbb{R}$ and $W_\cdot : [0, T] \times \Omega \to \mathbb{R}$ is the real-valued standard Brownian motions. Such an SPDE is usually termed as the stochastic Allen-Cahn equation. Discretizing such SPDE (6.3) spatially by a finite difference method yields a system of SDE as below,

$$\mathrm{d}X_t = [\mathbb{A}X_t + \mathbb{F}(X_t)]\,\mathrm{d}t + \mathbb{G}(X_t)\,\mathrm{d}W_t, \quad t \in (0, T], \quad X_0 = x_0, \tag{6.4}$$

where $X_t = (X_{1,t}, X_{2,t}, \cdots, X_{K-1,t})^T := (u(t, x_1), u(t, x_2), \cdots, u(t, x_{K-1}))^T$, $\mathbb{A} \in \mathbb{R}^{(K-1)\times(K-1)}$, $x_0 = (u_0(x_1), u_0(x_2), ..., u_0(x_{K-1}))^T$ and

$$\mathbb{A} = K^2 \begin{bmatrix} -2 & 1 & 0 & \cdots & 0 & 0 \\ 1 & -2 & 1 & \cdots & 0 & 0 \\ 0 & 1 & -2 & \cdots & 0 & 0 \\ & \cdots & & \cdots & & \\ 0 & 0 & 0 & \cdots & -2 & 1 \\ 0 & 0 & 0 & \cdots & 1 & -2 \end{bmatrix}, \quad \mathbb{F}(X) = \begin{bmatrix} X_1 - (X_1)^3 \\ X_2 - (X_2)^3 \\ \vdots \\ X_{K-1} - (X_{K-1})^3 \end{bmatrix},$$

$$\mathbb{G}(X) = \begin{bmatrix} g(X_1) \\ g(X_2) \\ \vdots \\ g(X_{K-1}) \end{bmatrix}.$$

Here we only focus on the temporal discretization of the SDE system (6.4). In what follows we set $g(u) = \sin(u) + 1$ and $u_0(x) \equiv 1$. The eigenvalues $\{\lambda_i\}_{i=1}^{K-1}$ of the matrix $\mathbb{A}$ are $\lambda_i = -4K^2 \sin^2(i\pi/2K) < 0$ [27], resulting in a very stiff system (6.4). Further, it is easy to check all conditions in Assumptions 2.1-2.4 are fulfilled with $\gamma = 3$ and for any $p_0 \geq 13$.

Here we take the case $K = 4$ as an example. To deal with the stiffness, we take the linear-implicit projected Euler method, i.e., $\theta = 1$ in (1.3), to discretize (6.4) in time and the *exact* solutions are given numerically by using a fine stepsize $h_{\text{exact}} = 2^{-14}$. As is evident from Fig. 4, the weak convergence rate of the linear-implicit Euler method is 1.

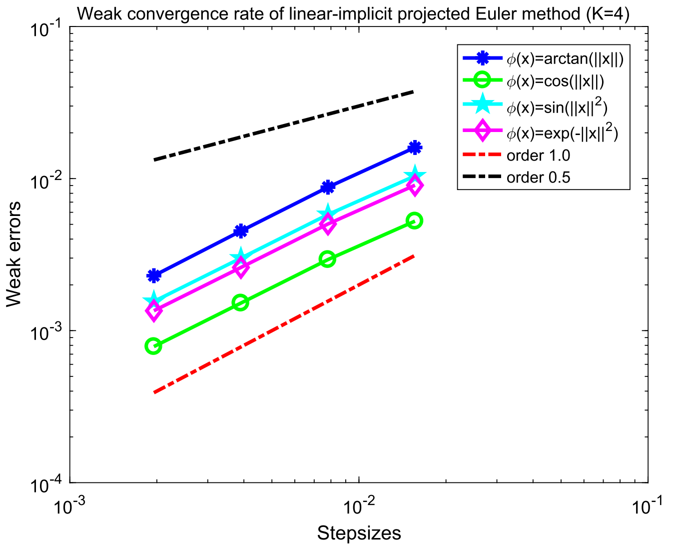

Fig. 4. Weak convergence rates of linear-implicit projected Euler method for model (6.4) (K=4).

Appendix A. Proof of Lemmas in Section 3

A.1. Proof of Lemma 3.3

Proof of Lemma 3.3. By the Itô formula and the Cauchy-Schwarz inequality, for any $p \in [1, \infty)$, we get

$$
\begin{aligned}
&\left(1 + \|X_t^{-\iota,x_0}\|^2\right)^p \\
&\le \left(1 + \|x_0\|^2\right)^p + 2p \int_{-\iota}^{t} \left(1 + \left\|X_s^{-\iota,x_0}\right\|^2\right)^{p-1} \left\langle X_s^{-\iota,x_0}, AX_s^{-\iota,x_0}\right\rangle \mathrm{d}s \\
&\quad + 2p \int_{-\iota}^{t} \left(1 + \left\|X_s^{-\iota,x_0}\right\|^2\right)^{p-1} \left\langle X_s^{-\iota,x_0}, f\left(X_s^{-\iota,x_0}\right)\right\rangle \mathrm{d}s \\
&\quad + 2p \int_{-\iota}^{t} \left(1 + \left\|X_s^{-\iota,x_0}\right\|^2\right)^{p-1} \left\langle X_s^{-\iota,x_0}, g\left(X_s^{-\iota,x_0}\right) \mathrm{d}\widetilde{W}_s\right\rangle \\
&\quad + p(2p-1) \int_{-\iota}^{t} \left(1 + \left\|X_s^{-\iota,x_0}\right\|^2\right)^{p-1} \|g(X_s^{-\iota,x_0})\|^2 \mathrm{d}s.
\end{aligned}
$$

Here we define a stopping time as

$$
\tau_n = \inf\{s \ge -\iota : \|X_s^{-\iota,x_0}\| > n\}.
$$

Taking expectations on both sides with (2.1) and Assumption 2.2 shows that

$$
\mathbb{E}\left[\left(1 + \|X_{t\wedge\tau_n}^{-\iota,x_0}\|^2\right)^p\right]
$$

$$
\begin{aligned}
&\leq \mathbb{E}\left[\left(1+\|x_0\|^2\right)^p\right]-p\left(2\lambda_1-L_1\right)\mathbb{E}\left[\int_{-\iota}^{t\wedge\tau_n}\left(1+\left\|X_s^{-\iota,x_0}\right\|^2\right)^{p-1}\left\|X_s^{-\iota,x_0}\right\|^2\mathrm{d}s\right]\\
&\quad+C_\star p\mathbb{E}\left[\int_{-\iota}^{t\wedge\tau_n}\left(1+\left\|X_s^{-\iota,x_0}\right\|^2\right)^{p-1}\mathrm{d}s\right]\\
&=\mathbb{E}\left[\left(1+\|x_0\|^2\right)^p\right]-p\left(2\lambda_1-L_1\right)\mathbb{E}\left[\int_{-\iota}^{t\wedge\tau_n}\left(1+\left\|X_s^{-\iota,x_0}\right\|^2\right)^{p}\mathrm{d}s\right]\\
&\quad+p\left(2\lambda_1-L_1+C_\star\right)\mathbb{E}\left[\int_{-\iota}^{t\wedge\tau_n}\left(1+\left\|X_s^{-\iota,x_0}\right\|^2\right)^{p-1}\mathrm{d}s\right].
\end{aligned}
$$

For $p\in[1,p_0]$, using the Young inequality

$$
a^{p-1}b\leq\epsilon\frac{p-1}{p}a^p+\epsilon^{1-p}\frac{b^p}{p},\quad\forall a,b\geq1\quad\text{with}\quad\epsilon\in\left(0,\frac{p(2\lambda_1-L_1)}{(p-1)(2\lambda_1-L_1+C_\star)}\right),
$$

indicates that,

$$
\begin{aligned}
&p\left(2\lambda_1-L_1+C_\star\right)\mathbb{E}\left[\int_{-\iota}^{t\wedge\tau_n}\left(1+\left\|X_s^{-\iota,x_0}\right\|^2\right)^{p-1}\mathrm{d}s\right]\\
&\leq(p-1)\left(2\lambda_1-L_1+C_\star\right)\epsilon\mathbb{E}\left[\int_{-\iota}^{t\wedge\tau_n}\left(1+\left\|X_s^{-\iota,x_0}\right\|^2\right)^{p}\mathrm{d}s\right]+\int_{-\iota}^{t\wedge\tau_n}\left(2\lambda_1-L_1+C_\star\right)\epsilon^{1-p}\mathrm{d}s.
\end{aligned}
$$

Then one achieves that

$$
\begin{aligned}
&\mathbb{E}\left[\left(1+\left\|X_{t\wedge\tau_n}^{-\iota,x_0}\right\|^2\right)^p\right]\\
&+\left[p\left(2\lambda_1-L_1\right)-(p-1)\left(2\lambda_1-L_1+C_\star\right)\epsilon\right]\mathbb{E}\left[\int_{-\iota}^{t\wedge\tau_n}\left(1+\left\|X_s^{-\iota,x_0}\right\|^2\right)^{p}\mathrm{d}s\right]\\
&\leq\mathbb{E}\left[\left(1+\|x_0\|^2\right)^p\right]+\int_{-\iota}^{t\wedge\tau_n}\left(2\lambda_1-L_1+C_\star\right)\epsilon^{1-p}\mathrm{d}s.
\end{aligned}
$$

Due to the Fatou Lemma, let $n\to\infty$, we obtain that

$$
\begin{aligned}
&\mathbb{E}\left[\left(1+\left\|X_{t}^{-\iota,x_0}\right\|^2\right)^p\right]\\
&+\left[p\left(2\lambda_1-L_1\right)-(p-1)\left(2\lambda_1-L_1+C_\star\right)\epsilon\right]\mathbb{E}\left[\int_{-\iota}^{t}\left(1+\left\|X_s^{-\iota,x_0}\right\|^2\right)^{p}\mathrm{d}s\right]\\
&\leq\mathbb{E}\left[\left(1+\|x_0\|^2\right)^p\right]+\int_{-\iota}^{t}\left(2\lambda_1-L_1+C_\star\right)\epsilon^{1-p}\mathrm{d}s.
\end{aligned}
$$

As $p\left(2\lambda_1-L_1\right)-(p-1)\left(2\lambda_1-L_1+C_\star\right)\epsilon>0$, the proof is done by Lemma 3.2. □

A.2. Proof of Lemma 3.4

Proof of Lemma 3.4. For brevity, we define

$$\Delta x_t = X_t^{-\iota,x_{0,(1)}} - X_t^{-\iota,x_{0,(2)}}, \quad \Delta f_t = f\left(X_t^{-\iota,x_{0,(1)}}\right) - f\left(X_t^{-\iota,x_{0,(2)}}\right),$$
$$\Delta g_t = g\left(X_t^{-\iota,x_{0,(1)}}\right) - g\left(X_t^{-\iota,x_{0,(2)}}\right).$$

With the stopping time defined as follows,

$$\overline{\tau}_n^{(1)} = \inf\{s \geq -\iota : \|X_s^{-\iota,x_0^{(1)}}\| > n \quad \text{or} \quad \|X_s^{-\iota,x_0^{(2)}}\| > n\},$$

one obtains by using the Itô formula,

$$\begin{aligned}
&e^{c_1 p(t\wedge\overline{\tau}_n^{(1)})}\left\|\Delta x_{t\wedge\overline{\tau}_n^{(1)}}\right\|^{2p} \\
&\leq \|\Delta x_0\|^{2p} + c_1 p \int_{-\iota}^{t\wedge\overline{\tau}_n^{(1)}} e^{c_1 ps}\|\Delta x_s\|^{2p} \mathrm{d}s + 2p \int_{-\iota}^{t\wedge\overline{\tau}_n^{(1)}} e^{c_1 ps}\|\Delta x_s\|^{2p-2} \langle \Delta x_s, A\Delta x_s\rangle \, \mathrm{d}s \\
&\quad + 2p \int_{-\iota}^{t\wedge\overline{\tau}_n^{(1)}} e^{c_1 ps}\|\Delta x_s\|^{2p-2} \langle \Delta x_s, \Delta f_s\rangle \, \mathrm{d}s + 2p \int_{-\iota}^{t\wedge\overline{\tau}_n^{(1)}} e^{c_1 ps}\|\Delta x_s\|^{2p-2} \langle \Delta x_s, \Delta g_s \mathrm{d}W_s\rangle \\
&\quad + p(2p-1) \int_{-\iota}^{t\wedge\overline{\tau}_n^{(1)}} e^{c_1 ps}\|\Delta x_s\|^{2p-2} \|\Delta g_s\|^2 \, \mathrm{d}s.
\end{aligned}$$

Hence, by taking expectations on both sides with Assumptions 2.1, 2.3 and the Fatou lemma, we reach that, for some positive constant $c_1 \in (0, 2\lambda_1 - L_2]$,

$$\mathbb{E}\left[e^{c_1 p(t\wedge\overline{\tau}_n^{(1)})}\left\|\Delta x_{t\wedge\overline{\tau}_n^{(1)}}\right\|^{2p}\right] \leq \mathbb{E}\left[\|\Delta x_0\|^{2p}\right] + \underbrace{p\left[c_1 - (2\lambda_1 - L_2)\right] \int_{-\iota}^{t\wedge\overline{\tau}_n^{(1)}} e^{c_1 ps}\|\Delta x_s\|^{2p} \mathrm{d}s}_{\leq 0},$$

leading to

$$\mathbb{E}\left[\|\Delta x_t\|^{2p}\right] \leq e^{-c_1 pt}\mathbb{E}\left[\|\Delta x_0\|^{2p}\right].$$

The proof is complete. □

A.3. Proof of Lemma 3.5

Proof of Lemma 3.5. Let

$$\Delta X_t = X_t^{-s_1,x_0} - X_t^{-s_2,x_0}, \; \Delta \bar{f}_t = f\left(X_t^{-s_1,x_0}\right) - f\left(X_t^{-s_2,x_0}\right),$$
$$\Delta \bar{g} = g\left(X_t^{-s_1,x_0}\right) - g\left(X_t^{-s_2,x_0}\right).$$

With reference to the proof of Lemma 3.4, setting the stopping time as

$$\overline{\tau}_n^{(2)} = \inf\{s \geq -s_2 : \|X_s^{-s_2,x_0}\| > n\},$$

by the Itô formula, we deduce that

$$
\begin{aligned}
&\mathbb{E}\left[e^{c_2 p(t\wedge\overline{\tau}_n^{(2)}+s_2)}\left\|\Delta X_{t\wedge\overline{\tau}_n^{(2)}}\right\|^{2p}\right]\\
&\leq \mathbb{E}\left[\|\Delta X_{-s_2}\|^{2p}\right]+c_2 p\mathbb{E}\left[\int_{-s_2}^{t\wedge\overline{\tau}_n^{(2)}} e^{c_2 p(s+s_2)}\|\Delta X_s\|^{2p}\mathrm{d}s\right]\\
&\quad+2p\mathbb{E}\left[\int_{-s_2}^{t\wedge\overline{\tau}_n^{(2)}} e^{c_2 p(s+s_2)}\|\Delta X_s\|^{2p-2}\left\langle\Delta X_s, A\Delta X_s\right\rangle\mathrm{d}s\right]\\
&\quad+2p\mathbb{E}\left[\int_{-s_2}^{t\wedge\overline{\tau}_n^{(2)}} e^{c_2 p(s+s_2)}\|\Delta X_s\|^{2p-2}\left\langle\Delta X_s, \Delta\bar{f}_s\right\rangle\mathrm{d}s\right]\\
&\quad+p(2p-1)\mathbb{E}\left[\int_{-s_2}^{t\wedge\overline{\tau}_n^{(2)}} e^{c_2 p(s+s_2)}\|\Delta X_s\|^{2p-2}\left\|\Delta\bar{g}_s\right\|^2\mathrm{d}s\right].
\end{aligned}
$$

According to Lemma 3.3, one directly obtains

$$
\mathbb{E}\left[\|\Delta X_{-s_2}\|^{2p}\right]=\mathbb{E}\left[\|X_{-s_2}^{-s_1,x_0}-x_0\|^{2p}\right]\leq C\mathbb{E}\left[\left(1+\|x_0\|^2\right)^p\right].
$$

Taking this with Assumption 2.1, Assumption 2.3 and the Fatou lemma into account yields

$$
\begin{aligned}
&e^{c_2 p(t\wedge\overline{\tau}_n^{(2)}+s_2)}\mathbb{E}\left[\left\|\Delta X_{t\wedge\overline{\tau}_n^{(2)}}\right\|^{2p}\right]\\
&\leq C\mathbb{E}\left[\left(1+\|x_0\|^2\right)^p\right]+\underbrace{p\left[c_2-(2\lambda_1-L_2)\right]\int_{-s_2}^{t\wedge\overline{\tau}_n^{(2)}} e^{c_2 p(s+s_2)}\|\Delta X_s\|^{2p}\mathrm{d}s}_{\leq 0},
\end{aligned}
$$

resulting in

$$
\mathbb{E}\left[\|\Delta X_t\|^{2p}\right]\leq Ce^{-c_2 p(t+s_2)}\mathbb{E}\left[\left(1+\|x_0\|^2\right)^p\right].
$$

The proof is complete. □

Appendix B. Proof of Lemmas in Section 4

B.1. *Proof of Lemma 4.2*

Proof of Lemma 4.2. The first and the second estimates are obvious from (2.3) and (1.4). Equipped with these above, by Assumption 2.2 and the Cauchy-Schwarz inequality, one obtains

$$
\begin{aligned}
(2p_0-1)\|g(\mathscr{P}(x))\|^2&\leq L_1(1+\|\mathscr{P}(x)\|^2)-2\left\langle\mathscr{P}(x), f\left(\mathscr{P}(x)\right)\right\rangle\\
&\leq L_1(1+\|\mathscr{P}(x)\|^2)+2\left\|\mathscr{P}(x)\right\|\cdot\left\|f\left(\mathscr{P}(x)\right)\right\|\\
&\leq L_1(1+\|\mathscr{P}(x)\|^2)+2C_f h^{-\frac{1}{2}}\|\mathscr{P}(x)\|,
\end{aligned}
$$

where $C_f := 2C_2$. Owing to the fact that $p_0 \in [1, \infty)$, the proof of the third estimate in (4.1) is complete. Then taking p-th square on both sides yields

$$(2p_0-1)^p \|g(\mathscr{P}(x))\|^{2p} \le L_1(1+\|\mathscr{P}(x)\|^2)^p + 2pC_f L_1^{p-1} h^{-\frac{1}{2}}(1+\|\mathscr{P}(x)\|^2)^{p-1}\|\mathscr{P}(x)\| + \sum_{i=2}^{p} \mathscr{C}_p^i (2C_f)^i L_1^{p-i} h^{-\frac{i}{2}} (1+\|\mathscr{P}(x)\|^2)^{p-i}\|\mathscr{P}(x)\|^i,$$

where $\mathscr{C}_p^i := p!/(i!(p-i)!)$. As we have claimed, $\|\mathscr{P}(x)\| \le h^{-\frac{1}{2\gamma}}$ and $\|\mathscr{P}(x)\|^i \le (1+\|\mathscr{P}(x)\|^2)^{\frac{i}{2}}$ for any $i \ge 2$, so that

$$\begin{aligned}
&(2p_0-1)^p \|g(\mathscr{P}(x))\|^{2p} \\
&\le L_1^p(1+\|\mathscr{P}(x)\|^2)^p + 2C_f p L_1^{p-1} h^{-\left(\frac{1}{2}+\frac{1}{2\gamma}\right)}(1+\|\mathscr{P}(x)\|^2)^{p-1} \\
&\quad + h^{-\frac{p}{2}}(1+\|\mathscr{P}(x)\|^2)^{p-1}\sum_{i=2}^{p}\mathscr{C}_p^i (2C_f)^i L_1^{p-i} \\
&\le L_1^p(1+\|\mathscr{P}(x)\|^2)^p + Ch^{-\frac{p}{2}}(1+\|\mathscr{P}(x)\|^2)^{p-1},
\end{aligned}$$

where $C = C(L_1, C_f, p) = \sum_{i=1}^{p}\mathscr{C}_p^i (2C_f)^i L_1^{p-i}$.

Turning now on to the estimate (4.2), the proof of the first estimate in (4.2) is found from Lemma 6.2 in [2]. For the second estimate, we know from (1.3), Assumption 2.4 and Lemma 4.2 that,

$$\begin{aligned}
\left\| f\big(\mathscr{P}(x)\big) - f\big(\mathscr{P}(y)\big)\right\| &\le C_1\big(1+\|\mathscr{P}(x)\|^{\gamma-1}+\|\mathscr{P}(y)\|^{\gamma-1}\big)\|\mathscr{P}(x)-\mathscr{P}(x)\| \\
&\le C_1(1+2h^{-\frac{\gamma-1}{2\gamma}})\|\mathscr{P}(x)-\mathscr{P}(x)\| \\
&\le \lambda_f h^{-\frac{\gamma-1}{2\gamma}}\|\mathscr{P}(x)-\mathscr{P}(x)\|,
\end{aligned}$$

where we recall $\lambda_f := 3C_1$, and one follows the first estimate to complete the proof. □

B.2. *Proof of Lemma 4.4*

Proof of Lemma 4.4. Shortly, we denote

$$\Delta Y_n = Y_n^{(1)} - Y_n^{(2)}, \quad \Delta\mathscr{P}(Y_n) = \mathscr{P}(Y_n^{(1)}) - \mathscr{P}(Y_n^{(2)}), \quad \Delta\widetilde{f}_n = f\big(\mathscr{P}(Y_n^{(1)})\big) - f\big(\mathscr{P}(Y_n^{(2)})\big),$$
$$\Delta\widetilde{g}_n = g\big(\mathscr{P}(Y_n^{(1)})\big) - g\big(\mathscr{P}(Y_n^{(2)})\big).$$

It is apparent to show that

$$\Delta Y_{n+1} - \theta A\Delta Y_{n+1}h = \Delta\mathscr{P}(Y_n) + (1-\theta)A\Delta\mathscr{P}(Y_n)h + \Delta\widetilde{f}_n h + \Delta\widetilde{g}_n\Delta W_n.$$

Taking square on both sides, we then take expectations and follow Assumption 2.1 and Assumption 2.3 to imply

$$\begin{aligned}
&(1+2\theta\lambda_1 h+\theta^2h^2)\mathbb{E}\left[\|\Delta Y_{n+1}\|^2\right] \\
&\le [1-2(1-\theta)\lambda_1 h]\,\mathbb{E}\left[\|\Delta\mathscr{P}(Y_n)\|^2\right] + (1-\theta)^2h^2\mathbb{E}\left[\|A\Delta\mathscr{P}(Y_n)\|^2\right] + h^2\mathbb{E}\left[\|\Delta\widetilde{f}_n\|^2\right] \\
&\quad + h\mathbb{E}\left[\|\Delta\widetilde{g}_n\|^2\right] + 2h\mathbb{E}\big[\langle\Delta\mathscr{P}(Y_n), \Delta\widetilde{f}_n\rangle\big] + 2(1-\theta)h^2\mathbb{E}\big[\langle A\Delta\mathscr{P}(Y_n), \Delta\widetilde{f}_n\rangle\big].
\end{aligned}$$

Using the Cauchy-Schwarz inequality leads to

$$2(1-\theta)h^2\mathbb{E}\big[\langle A\Delta\mathscr{P}(Y_n), \Delta\widetilde{f}_n\rangle\big] \le 2(1-\theta)h^2\mathbb{E}\left[\left\|A\Delta\mathscr{P}(Y_n)\right\|\cdot\left\|\Delta\widetilde{f}_n\right\|\right].$$

Recalling Assumption 2.1, Assumption 2.3, Lemma 4.2, we obtain that

$$
\begin{aligned}
&(1+2\theta\lambda_1 h)\mathbb{E}\left[\|\Delta Y_{n+1}\|^2\right] \\
&\le [1-2(1-\theta)\lambda_1 h]\,\mathbb{E}\left[\|\Delta\mathscr{P}(Y_n)\|^2\right] + 2h\mathbb{E}\big[\langle\Delta\mathscr{P}(Y_n),\Delta\widetilde{f}_n\rangle\big] + (2p_1-1)h\mathbb{E}\left[\|\Delta\widetilde{g}_n\|^2\right] \\
&\quad + (1-\theta)^2\lambda_d^2 h^2\mathbb{E}\left[\|\Delta\mathscr{P}(Y_n)\|^2\right] + 2(1-\theta)\lambda_d\lambda_f h^{1+\frac{\gamma+1}{2\gamma}}\,\mathbb{E}\left[\|\Delta\mathscr{P}(Y_n)\|^2\right] \\
&\quad + \lambda_f^2 h^{1+\frac{1}{\gamma}}\,\mathbb{E}\left[\|\Delta\mathscr{P}(Y_n)\|^2\right] \\
&\le \left\{1-2(1-\theta)\lambda_1 h + L_2 h + \left[(1-\theta)\lambda_d h^{\frac12} + \lambda_f h^{\frac{1}{2\gamma}}\right]^2 h\right\}\mathbb{E}\left[\|\Delta Y_n\|^2\right].
\end{aligned}
$$

Here we select an appropriate stepsize h such that

$$
(1-\theta)\lambda_d h^{\frac12} < \tfrac12\sqrt{2\lambda_1 - L_2}, \quad \lambda_f h^{\frac{1}{2\gamma}} < \tfrac12\sqrt{2\lambda_1 - L_2},
$$

which leads to

$$
h \in \left(0, \min\left\{\frac{2\lambda_1 - L_2}{4(1-\theta)^2\lambda_d^2}, \frac{(2\lambda_1-L_2)^\gamma}{(2\lambda_f)^{2\gamma}}, 1\right\}\right),
$$

to ensure

$$
2\lambda_1 - L_2 - \left[(1-\theta)\lambda_d h^{\frac12} + \lambda_f h^{\frac{1}{2\gamma}}\right]^2 > 0.
$$

As a result, there exists some positive constant $\widetilde{C}_1$ satisfying

$$
0 < \widetilde{C}_1 \le \frac{2\lambda_1 - L_2 - \big[(1-\theta)\lambda_d h^{\frac12} + \lambda_f h^{\frac{1}{2\gamma}}\big]^2}{1+2\theta L_1 h},
$$

such that

$$
\mathbb{E}\left[\|\Delta Y_{n+1}\|^2\right] \le (1-\widetilde{C}_1 h)\mathbb{E}\left[\|\Delta Y_n\|^2\right] \le e^{-\widetilde{C}_1 t_{n+1}}\mathbb{E}\left[\|x_0^{(1)} - x_0^{(2)}\|^2\right].
$$

The proof is complete. □

Appendix C. Proof of Lemmas in Section 5

C.1. Proof of Lemma 5.2

Proof of Lemma 5.2. The existence of the mean-square derivatives up to the third order is proved in a similar way as shown in [6], which is also found in [31, Appendix C]. Based on our assumptions, we would like to obtain the time-independent estimate of the derivatives of solutions $\{X_t^x\}_{t\in[0,\infty)}$ given by (1.1) with respect to the initial condition x.

For simplicity, we denote that

$$
\eta^{v_1}(t,x) := \mathcal{D}X_t^x v_1, \quad \xi^{v_1,v_2}(t,x) := \mathcal{D}^2 X_t^x(v_1,v_2), \quad \zeta^{v_1,v_2,v_3}(t,x) := \mathcal{D}^3 X_t^x(v_1,v_2,v_3).
$$

Part I: estimate of the first variation process

For the first variation process of SDE (1.1), we have

$$
\mathrm{d}\eta^{v_1}(t,x) = DF(X_t^x)\eta^{v_1}(t,x)\,\mathrm{d}t + \sum_{j=1}^m Dg_j(X_t^x)\eta^{v_1}(t,x)\,\mathrm{d}W_{j,t}, \quad \eta^{v_1}(0,x) = v_1.
$$

Define a stopping time as

$$\widetilde{\tau}_n^{(1)} = \inf\left\{s \geq 0 : \|\eta^{v_1}(s,x)\| > n \quad \text{or} \quad \left\|X_s^x\right\| > n\right\}.$$

Using the Itô formula, the Cauchy-Schwarz inequality and (5.9) to attain that, for some $q_1 \in [1, p_1]$ and $\delta > 0$,

$$\begin{aligned}
&\mathbb{E}\left[e^{(\alpha q_1 t)\wedge \widetilde{\tau}_n^{(1)}} \|\eta^{v_1}(t \wedge \widetilde{\tau}_n^{(1)}, x)\|^{2q_1}\right] \\
&\leq \mathbb{E}\left[\|v_1\|^{2q_1}\right] + \alpha q_1 \mathbb{E}\left[\int_0^{t\wedge\widetilde{\tau}_n^{(1)}} e^{\alpha q_1 s}\|\eta^{v_1}(s,x)\|^{2q_1} \mathrm{d}s\right] \\
&\quad + 2q_1\mathbb{E}\left[\int_0^{t\wedge\widetilde{\tau}_n^{(1)}} e^{\alpha q_1 s}\|\eta^{v_1}(s,x)\|^{2q_1-2}\left\langle \eta^{v_1}(s,x), DF(X_s^x)\eta^{v_1}(s,x)\right\rangle \mathrm{d}s\right] \\
&\quad + q_1(2q_1-1)\sum_{j=1}^m \mathbb{E}\left[\int_0^{t\wedge\widetilde{\tau}_n^{(1)}} e^{\alpha q_1 s}\|\eta^{v_1}(s,x)\|^{2q_1-2}\left\|Dg_j(X_s^x)\eta^{v_1}(s,x)\right\|^2 \mathrm{d}s\right] \\
&\leq \mathbb{E}\left[\|v_1\|^{2q_1}\right].
\end{aligned} \tag{C.1}$$

Hence, by Fatou lemma and taking $\alpha_1 = \alpha/2$, the estimate above leads to

$$\mathbb{E}\left[\|\eta^{v_1}(t,x)\|^{2q_1}\right] \leq e^{-2\alpha_1 q_1 t}\mathbb{E}\left[\|v_1\|^{2q_1}\right]. \tag{C.2}$$

Part II: estimate of the second variation process

For the second variation process of SDE (1.1), we then acquire that,

$$\begin{aligned}
&\mathrm{d}\xi^{v_1,v_2}(t,x) \\
&= \Big(DF(X_t^x)\xi^{v_1,v_2}(t,x) + D^2F(X_t^x)\big(\eta^{v_1}(t,x), \eta^{v_2}(t,x)\big)\Big)\mathrm{d}t \\
&\quad + \sum_{j=1}^m \Big(Dg_j(X_t^x)\xi^{v_1,v_2}(t,x) + D^2g_j(X_t^x)\big(\eta^{v_1}(t,x), \eta^{v_2}(t,x)\big)\Big)\mathrm{d}W_{j,t}, \quad \xi^{v_1,v_2}(0,x_0) = 0.
\end{aligned}$$

Following the same idea as (C.1), we begin with the definition of the stopping time as follows,

$$\widetilde{\tau}_n^{(2)} = \inf\left\{s \geq 0 : \|\xi^{v_1,v_2}(s,x)\| > n \quad \text{or} \quad \left\|X_s^x\right\| > n\right\}.$$

Then, for some $q_2 \in [1, q_1)$ and $\delta > 0$, by taking the Itô formula, one will arrive at

$$\begin{aligned}
&\mathbb{E}\left[\left(\delta + \|\xi^{v_1,v_2}(t\wedge\widetilde{\tau}_n^{(2)}, x)\|^2\right)^{q_2}\right] \\
&\leq \delta^{q_2} + 2q_2\mathbb{E}\left[\int_0^{t\wedge\widetilde{\tau}_n^{(2)}} \left(\delta + \|\xi^{v_1,v_2}(s,x)\|^2\right)^{q_2-1}\left\langle \xi^{v_1,v_2}(s,x), DF(X_s^x)\xi^{v_1,v_2}(s,x)\right\rangle \mathrm{d}s\right] \\
&\quad + 2q_2\mathbb{E}\left[\int_0^{t\wedge\widetilde{\tau}_n^{(2)}} \underbrace{\left(\delta + \|\xi^{v_1,v_2}(s,x)\|^2\right)^{q_2-1}\left\langle \xi^{v_1,v_2}(s,x), D^2F(X_s^x)\big(\eta^{v_1}(s,x), \eta^{v_2}(s,x)\big)\right\rangle}_{=:\mathbb{T}_1} \mathrm{d}s\right] \\
&\quad + q_2(2q_2-1)\sum_{j=1}^m \mathbb{E}\left[\int_0^{t\wedge\widetilde{\tau}_n^{(2)}} \underbrace{\left(\delta + \|\xi^{v_1,v_2}(s,x)\|^2\right)^{q_2-1}\left\|Dg_j(X_s^x)\xi^{v_1,v_2}(s,x) + D^2g_j(X_s^x)\big(\eta^{v_1}(s,x), \eta^{v_2}(s,x)\big)\right\|^2}_{=:\mathbb{T}_2} \mathrm{d}s\right].
\end{aligned}$$

The Cauchy-Schwarz inequality and the Young inequality are used several times to indicate that, for two positive constants $\tilde{\epsilon}_1, \tilde{\epsilon}_2$ with $\tilde{\epsilon}_1 \in (0, q_2\alpha)$ and $\tilde{\epsilon}_2 \in (0, (p_1 - q_2)/q_2]$,

$$\mathbb{T}_1 \leq \tfrac{\tilde{\epsilon}_1}{2} \left(\delta + \|\xi^{v_1,v_2}(s,x)\|^2\right)^{q_2} + C_{\tilde{\epsilon}_1} \left\| D^2 F(X_s^x)\big(\eta^{v_1}(s,x), \eta^{v_2}(s,x)\big) \right\|^{2q_2},$$

and

$$\begin{aligned}
\mathbb{T}_2 &\leq (1+\tilde{\epsilon}_2) \left(\delta + \|\xi^{v_1,v_2}(s,x)\|^2\right)^{q_2-1} \left\| Dg_j(X_s^x)\xi^{v_1,v_2}(s,x)\right\|^2 \\
&\quad + C_{\tilde{\epsilon}_2} \left(\delta + \|\xi^{v_1,v_2}(s,x)\|^2\right)^{q_2-1} \left\| D^2 g_j(X_s^x)\big(\eta^{v_1}(s,x), \eta^{v_2}(s,x)\big)\right\|^2 \\
&\leq (1+\tilde{\epsilon}_2) \left(\delta + \|\xi^{v_1,v_2}(s,x)\|^2\right)^{q_2-1} \left\| Dg_j(X_s^x)\xi^{v_1,v_2}(s,x)\right\|^2 + \tfrac{\tilde{\epsilon}_1}{2} \left(\delta + \|\xi^{v_1,v_2}(s,x)\|^2\right)^{q_2} \\
&\quad + C_{\tilde{\epsilon}_1,\tilde{\epsilon}_2} \left\| D^2 g_j(X_s^x)\big(\eta^{v_1}(s,x), \eta^{v_2}(s,x)\big)\right\|^{2q_2}.
\end{aligned}$$

With these estimates above, we obtain that

$$\begin{aligned}
&\mathbb{E}\left[\left(\delta + \|\xi^{v_1,v_2}(t\wedge\widetilde{\tau}_n^{(2)}, x)\|^2\right)^{q_2}\right] \\
&\leq \delta^{q_2} + 2q_2\mathbb{E}\left[\int_0^{t\wedge\widetilde{\tau}_n^{(2)}} \left(\delta + \|\xi^{v_1,v_2}(s,x)\|^2\right)^{q_2-1} \left(\left\langle \xi^{v_1,v_2}(s,x), DF(X_s^x)\xi^{v_1,v_2}(s,x)\right\rangle\right) \mathrm{d}s\right] \\
&\quad + q_2(2q_2-1)(1+\tilde{\epsilon}_2)\sum_{j=1}^m \mathbb{E}\left[\int_0^{t\wedge\widetilde{\tau}_n^{(2)}} \left(\delta + \|\xi^{v_1,v_2}(s,x)\|^2\right)^{q_2-1} \left\| Dg_j(X_s^x)\xi^{v_1,v_2}(s,x)\right\|^2 \mathrm{d}s\right] \\
&\quad + 2\tilde{\epsilon}_1\mathbb{E}\left[\int_0^{t\wedge\widetilde{\tau}_n^{(2)}} \left(\delta + \|\xi^{v_1,v_2}(s,x)\|^2\right)^{q_2} \mathrm{d}s\right] \\
&\quad + C_{\tilde{\epsilon}_1}\mathbb{E}\left[\int_0^{t\wedge\widetilde{\tau}_n^{(2)}} \left\| D^2 F(X_s^x)\big(\eta^{v_1}(s,x), \eta^{v_2}(s,x)\big)\right\|^{2q_2} \mathrm{d}s\right] \\
&\quad + C_{\tilde{\epsilon}_1,\tilde{\epsilon}_2}\sum_{j=1}^m \mathbb{E}\left[\int_0^{t\wedge\widetilde{\tau}_n^{(2)}} \left\| D^2 g_j(X_s^x)\big(\eta^{v_1}(s,x), \eta^{v_2}(s,x)\big)\right\|^{2q_2} \mathrm{d}s\right].
\end{aligned}$$

With Assumption 2.4, Lemma 3.3, the Hölder inequality and (C.2) in mind, and recall the definition $\mathcal{P}_\cdot(\cdot)$ in (2.2) its property in (5.6), we are able to show that, for some positive constants ρ_1, ρ_2, ρ_3 satisfying $1/\rho_1 + 1/\rho_2 + 1/\rho_3 = 1$ and (5.10),

$$\begin{aligned}
&\left\| D^2 F(X_s^x)\big(\eta^{v_1}(s,x), \eta^{v_2}(s,x)\big)\right\|_{L^{2q_2}(\Omega,\mathbb{R}^d)} \\
&\leq C\left\| \mathcal{P}_{\gamma-2}(X_s^x)\cdot\|\eta^{v_1}(s,x)\|\cdot\|\eta^{v_2}(t,x)\|\right\|_{L^{2q_2}(\Omega,\mathbb{R})} \\
&\leq C\left\| \mathcal{P}_{\gamma-2}(X_s^x)\right\|_{L^{2\rho_1 q_2}(\Omega,\mathbb{R})} \left\|\eta^{v_1}(s,x)\right\|_{L^{2\rho_2 q_2}(\Omega,\mathbb{R}^d)} \left\|\eta^{v_2}(s,x)\right\|_{L^{2\rho_3 q_2}(\Omega,\mathbb{R}^d)} \\
&\leq Ce^{-2\alpha_1 s}\left\| \mathcal{P}_{\gamma-2}(X_s^x)\right\|_{L^{2\rho_1 q_2}(\Omega,\mathbb{R})} \|v_1\|_{L^{2\rho_2 q_2}(\Omega,\mathbb{R}^d)} \|v_2\|_{L^{2\rho_3 q_2}(\Omega,\mathbb{R}^d)}.
\end{aligned}$$

Following the same idea and taking into account (2.4), one gets, for $j \in \{1, \ldots, m\}$ and $\rho_1 q_2(\gamma-3) \leq 2p_0$,

$$
\begin{aligned}
&\left\| D^2 g_j(X_s^x)\big(\eta^{v_1}(s,x),\eta^{v_2}(s,x)\big)\right\|_{L^{2q_2}(\Omega,\mathbb{R}^d)}\\
&\le C\left\|\mathcal{P}_{(\gamma-3)/2}(X_s^x)\right\|_{L^{2\rho_1 q_2}(\Omega,\mathbb{R})}\left\|\eta^{v_1}(s,x)\right\|_{L^{2\rho_2 q_2}(\Omega,\mathbb{R}^d)}\left\|\eta^{v_2}(s,x)\right\|_{L^{2\rho_3 q_2}(\Omega,\mathbb{R}^d)}\\
&\le Ce^{-2\alpha_1 s}\left\|\mathcal{P}_{(\gamma-3)/2}(X_s^x)\right\|_{L^{2\rho_1 q_2}(\Omega,\mathbb{R})}\left\|v_1\right\|_{L^{2\rho_2 q_2}(\Omega,\mathbb{R}^d)}\left\|v_2\right\|_{L^{2\rho_3 q_2}(\Omega,\mathbb{R}^d)}.
\end{aligned}
$$

Combining these estimates with (5.9), (5.10), Lemma 3.3 and the Young inequality and the monotonicity of $\mathcal{P}_{\cdot}(X_s^x)$ yields,

$$
\begin{aligned}
&\mathbb{E}\left[\left(\delta+\|\xi^{v_1,v_2}(t\wedge\widetilde{\tau}_n^{(2)},x)\|^2\right)^{q_2}\right]\\
&\le \delta^{q_2}-(2q_2\alpha-\tilde{\epsilon}_1)\mathbb{E}\left[\int_0^{t\wedge\widetilde{\tau}_n^{(2)}}\left(\delta+\|\xi^{v_1,v_2}(s,x)\|^2\right)^{q_2}\mathrm{d}s\right]\\
&\quad+2q_2\alpha\delta\mathbb{E}\left[\int_0^{t\wedge\widetilde{\tau}_n^{(2)}}\left(\delta+\|\xi^{v_1,v_2}(s,x)\|^2\right)^{q_2-1}\mathrm{d}s\right]\\
&\quad+C_{\tilde{\epsilon}_1,\tilde{\epsilon}_2}\sup_{r\in[0,\infty)}\left\|\mathcal{P}_{\gamma-2}(X_r^x)\right\|^{2q_2}_{L^{2\rho_1 q_2}(\Omega,\mathbb{R})}\left\|v_1\right\|^{2q_2}_{L^{2\rho_2 q_2}(\Omega,\mathbb{R}^d)}\left\|v_2\right\|^{2q_2}_{L^{2\rho_3 q_2}(\Omega,\mathbb{R}^d)}\int_0^{t\wedge\widetilde{\tau}_n^{(2)}}e^{-2q_2\alpha_1 s}\mathrm{d}s\\
&\le \delta^{q_2}-\left(2q_2\alpha-2\tilde{\epsilon}_1\right)\mathbb{E}\left[\int_0^{t\wedge\widetilde{\tau}_n^{(2)}}\left(\delta+\|\xi^{v_1,v_2}(s,x)\|^2\right)^{q_2}\mathrm{d}s\right]+C_{\tilde{\epsilon}_2}\left(2q_2\delta\right)^{q_2}\\
&\quad+C_{\tilde{\epsilon}_1,\tilde{\epsilon}_2}\sup_{r\in[0,\infty)}\left\|\mathcal{P}_{\gamma-2}(X_r^x)\right\|^{2q_2}_{L^{2\rho_1 q_2}(\Omega,\mathbb{R})}\left\|v_1\right\|^{2q_2}_{L^{2\rho_2 q_2}(\Omega,\mathbb{R}^d)}\left\|v_2\right\|^{2q_2}_{L^{2\rho_3 q_2}(\Omega,\mathbb{R}^d)}\int_0^{t\wedge\widetilde{\tau}_n^{(2)}}e^{-2q_2\alpha_1 s}\mathrm{d}s.
\end{aligned}
$$

Setting $\delta\to 0^+$, one observes

$$
\begin{aligned}
&\mathbb{E}\left[\left(\delta+\|\xi^{v_1,v_2}(t\wedge\widetilde{\tau}_n^{(2)},x_0)\|^2\right)^{q_2}\right]+\left(2q_2\alpha-2\tilde{\epsilon}_1\right)\mathbb{E}\left[\int_0^{t\wedge\widetilde{\tau}_n^{(2)}}\left(\delta+\|\xi^{v_1,v_2}(s,x)\|^2\right)^{q_2}\mathrm{d}s\right]\\
&\le C_{\tilde{\epsilon}_1,\tilde{\epsilon}_2}\sup_{r\in[0,\infty)}\left\|\mathcal{P}_{\gamma-2}(X_r^x)\right\|^{2q_2}_{L^{2\rho_1 q_2}(\Omega,\mathbb{R})}\left\|v_1\right\|^{2q_2}_{L^{2\rho_2 q_2}(\Omega,\mathbb{R}^d)}\left\|v_2\right\|^{2q_2}_{L^{2\rho_3 q_2}(\Omega,\mathbb{R}^d)}\int_0^{t\wedge\widetilde{\tau}_n^{(2)}}e^{-2q_2\alpha_1 s}\mathrm{d}s.
\end{aligned}
$$

By virtue of (5.10), Lemma 3.2, Lemma 3.3 and the Fatou lemma, one will arrive at,

$$
\begin{aligned}
&\mathbb{E}\left[\|\xi^{v_1,v_2}(t,x)\|^{2q_2}\right]\\
&\le C_{\tilde{\epsilon}_1,\tilde{\epsilon}_2}\sup_{r\in[0,\infty)}\left\|\mathcal{P}_{\gamma-2}(X_r^x)\right\|^{2q_2}_{L^{2\rho_1 q_2}(\Omega,\mathbb{R})}\left\|v_1\right\|^{2q_2}_{L^{2\rho_2 q_2}(\Omega,\mathbb{R}^d)}\left\|v_2\right\|^{2q_2}_{L^{2\rho_3 q_2}(\Omega,\mathbb{R}^d)}\times\\
&\quad\int_0^t e^{-2(q_2\alpha_1-\tilde{\epsilon}_1)(t-s)}e^{-2q_2\alpha_1 s}\mathrm{d}s\\
&\le C_{\tilde{\epsilon}_1,\tilde{\epsilon}_2}e^{-\alpha_2 q_2 t}\sup_{r\in[0,\infty)}\left\|\mathcal{P}_{\gamma-2}(X_r^x)\right\|^{2q_2}_{L^{2\rho_1 q_2}(\Omega,\mathbb{R})}\left\|v_1\right\|^{2q_2}_{L^{2\rho_2 q_2}(\Omega,\mathbb{R}^d)}\left\|v_2\right\|^{2q_2}_{L^{2\rho_3 q_2}(\Omega,\mathbb{R}^d)},
\end{aligned}
\tag{C.3}
$$

where $\alpha_2 := (q_2\alpha_1 - \tilde{\epsilon}_1)/q_2$.

Part III: estimate of the third variation process

For the third variation process of the SDE (1.1), we get

$$
\begin{aligned}
&\mathrm{d}\zeta^{v_1,v_2,v_3}(t,x)\\
&= \Big(DF(X_t^x)\zeta^{v_1,v_2,v_3}(t,x) + D^2F(X_t^x)\big(\eta^{v_1}(t,x), \xi^{v_2,v_3}(t,x)\big)\\
&\quad + D^2F(X_t^x)\big(\xi^{v_1,v_3}(t,x), \eta^{v_2}(t,x)\big)\\
&\quad + D^2F(X_t^x)\big(\xi^{v_1,v_2}(t,x), \eta^{v_3}(t,x)\big) + D^3F(X_t^x)\big(\eta^{v_1}(t,x), \eta^{v_2}(t,x), \eta^{v_3}(t,x)\big)\Big)\mathrm{d}t\\
&\quad + \sum_{j=1}^{m}\Big(Dg_j(X_t^x)\zeta^{v_1,v_2,v_3}(t,x) + D^2g_j(X_t^x)\big(\eta^{v_1}(t,x), \xi^{v_2,v_3}(t,x)\big)\\
&\quad + D^2g_j(X_t^x)\big(\xi^{v_1,v_3}(t,x), \eta^{v_2}(t,x)\big)\\
&\quad + D^2g_j(X_t^x)\big(\xi^{v_2,v_3}(t,x), \eta^{v_1}(t,x)\big) + D^3g_j(X_t^x)\big(\eta^{v_1}(t,x), \eta^{v_2}(t,x), \eta^{v_3}(t,x)\big)\Big)\mathrm{d}W_{j,t}\\
&=: \Big(DF(X_t^x)\zeta^{v_1,v_2,v_3}(t,x) + H(X_t^x)\Big)\mathrm{d}t + \sum_{j=1}^{m}\Big(Dg_j(X_t^x)\zeta^{v_1,v_2,v_3}(t,x) + G_j(X_t^x)\Big)\mathrm{d}W_{j,t}.
\end{aligned}
$$

Similarly, given the stopping time as below,

$$
\widetilde{\tau}_n^{(3)} = \inf\left\{s \geq 0 : \|\zeta^{v_1,v_2,v_3}(s,x)\| > n \quad \text{or} \quad \left\|X_s^x\right\| > n\right\},
$$

due to the Itô formula and the Young inequality, we obtain that, for positive constants $\tilde{\epsilon}_3 \in (0, 2p_1 - 2)$ and $\tilde{\epsilon}_4 \in (0, \alpha)$,

$$
\begin{aligned}
&\mathbb{E}\left[\|\zeta^{v_1,v_2,v_3}(t\wedge\widetilde{\tau}_n^{(3)}, x)\|^2\right]\\
&\leq 2\mathbb{E}\left[\int_0^{t\wedge\widetilde{\tau}_n^{(3)}} \big\langle\zeta^{v_1,v_2,v_3}(s,x), DF(X_s^x)\zeta^{v_1,v_2,v_3}(s,x)\big\rangle\mathrm{d}s\right]\\
&\quad + 2\mathbb{E}\left[\int_0^{t\wedge\widetilde{\tau}_n^{(3)}} \big\langle\zeta^{v_1,v_2,v_3}(s,x), H(X_s^x)\big\rangle\mathrm{d}s\right]\\
&\quad + \sum_{j=1}^{m}\mathbb{E}\left[\int_0^{t\wedge\widetilde{\tau}_n^{(3)}} \left\|Dg_j(X_s^x)\zeta^{v_1,v_2,v_3}(s,x) + G_j(X_s^x)\right\|^2\mathrm{d}s\right]\\
&\leq \mathbb{E}\left[\int_0^{t\wedge\widetilde{\tau}_n^{(3)}} 2\big\langle\zeta^{v_1,v_2,v_3}(s,x), DF(X_s^x)\zeta^{v_1,v_2,v_3}(s,x)\big\rangle\right.\\
&\quad \left. + (1+\tilde{\epsilon}_3)\sum_{j=1}^{m}\left\|Dg_j(X_s^x)\zeta^{v_1,v_2,v_3}(s,x)\right\|^2\mathrm{d}s\right]\\
&\quad + \tilde{\epsilon}_4\mathbb{E}\left[\int_0^{t\wedge\widetilde{\tau}_n^{(3)}} \|\zeta^{v_1,v_2,v_3}(s,x)\|^2\mathrm{d}s\right] + C_{\tilde{\epsilon}_4}\mathbb{E}\left[\int_0^{t\wedge\widetilde{\tau}_n^{(3)}} \|H(X_s^x)\|^2\mathrm{d}s\right]
\end{aligned}
\tag{C.4}
$$

$$
\begin{aligned}
&+ C_{\tilde{\epsilon}_3} \sum_{j=1}^{m} \mathbb{E}\left[\int_0^{t\wedge\tilde{\tau}_n^{(3)}} \|G_j(X_s^x)\|^2 \mathrm{d}s \right] \\
&\le -(\alpha - \tilde{\epsilon}_4)\mathbb{E}\left[\int_0^{t\wedge\tilde{\tau}_n^{(3)}} \|\zeta^{v_1,v_2,v_3}(s,x)\|^2 \mathrm{d}s \right] + C_{\tilde{\epsilon}_4}\mathbb{E}\left[\int_0^{t\wedge\tilde{\tau}_n^{(3)}} \|H(X_s^x)\|^2 \mathrm{d}s \right] \\
&+ C_{\tilde{\epsilon}_3} \sum_{j=1}^{m} \mathbb{E}\left[\int_0^{t\wedge\tilde{\tau}_n^{(3)}} \|G_j(X_s^x)\|^2 \mathrm{d}s \right].
\end{aligned}
$$

The elementary inequality is used to imply that

$$
\begin{aligned}
&\|H(X_s^x)\|_{L^2(\Omega,\mathbb{R}^d)} \\
&\le \left\| D^2F(X_s^x)\big(\eta^{v_1}(s,x), \xi^{v_2,v_3}(s,x)\big) \right\|_{L^2(\Omega,\mathbb{R}^d)} + \left\| D^2F(X_s^x)\big(\xi^{v_1,v_3}(s,x), \eta^{v_2}(s,x)\big) \right\|_{L^2(\Omega,\mathbb{R}^d)} \\
&+ \left\| D^2F(X_s^x)\big(\xi^{v_1,v_2}(s,x), \eta^{v_3}(s,x)\big) \right\|_{L^2(\Omega,\mathbb{R}^d)} \\
&+ \left\| D^3F(X_s^x)\big(\eta^{v_1}(s,x), \eta^{v_2}(s,x), \eta^{v_3}(s,x)\big) \right\|_{L^2(\Omega,\mathbb{R}^d)}.
\end{aligned} \tag{C.5}
$$

In the following, we first show that the analysis of the first term to the third term on the right hand of (C.5) is equivalent. Taking the first term and the second term as examples, by Assumption 2.4, (C.2), (C.3), Lemma 3.3 and the Hölder inequality, for some positive constants ρ_1, ρ_2, ρ_3, ρ_4, ρ_5 and ρ_6 with $1/\rho_1 + 1/\rho_2 + 1/\rho_3 = 1$, $1/\rho_4 + 1/\rho_5 + 1/\rho_6 = 1$ and (5.10), one derives,

$$
\begin{aligned}
&\left\| D^2F(X_s^x)\big(\eta^{v_1}(s,x), \xi^{v_2,v_3}(s,x)\big) \right\|_{L^2(\Omega,\mathbb{R}^d)} \\
&\le C\left\| \mathcal{P}_{\gamma-2}(X_s^x)\cdot\|\eta^{v_1}(s,x)\|\cdot\|\xi^{v_2,v_3}(s,x)\| \right\|_{L^2(\Omega,\mathbb{R})} \\
&\le C\left\| \mathcal{P}_{\gamma-2}(X_s^x)\right\|_{L^{2\rho_1}(\Omega,\mathbb{R})} \|\eta^{v_1}(s,x)\|_{L^{2\rho_2}(\Omega,\mathbb{R}^d)} \|\xi^{v_2,v_3}(s,x)\|_{L^{2\rho_3}(\Omega,\mathbb{R}^d)} \\
&\le Ce^{-(\alpha_1+\alpha_2)s} \sup_{r\in[0,\infty)} \left\| \mathcal{P}_{\gamma-2}(X_r^x)\right\|_{L^{2\max\{\rho_1,\rho_3\rho_4\}}(\Omega,\mathbb{R})} \times \\
&\qquad \|v_1\|_{L^{2\rho_2}(\Omega,\mathbb{R}^d)} \|v_2\|_{L^{2\rho_3\rho_5}(\Omega,\mathbb{R}^d)} \|v_3\|_{L^{2\rho_3\rho_6}(\Omega,\mathbb{R}^d)}.
\end{aligned} \tag{C.6}
$$

For the second term, we choose another series of constants κ_1, κ_2, κ_3, κ_4, κ_5 and κ_6 with $1/\kappa_1 + 1/\kappa_2 + 1/\kappa_3 = 1$, $1/\kappa_4 + 1/\kappa_5 + 1/\kappa_6 = 1$ and (5.10) to show,

$$
\begin{aligned}
&\left\| D^2F(X_s^x)\big(\xi^{v_1,v_3}(s,x), \eta^{v_2}(s,x)\big) \right\|_{L^2(\Omega,\mathbb{R}^d)} \\
&\le C\left\| \mathcal{P}_{\gamma-2}(X_s^x)\cdot\|\xi^{v_1,v_3}(s,x)\|\cdot\|\eta^{v_2}(s,x)\| \right\|_{L^2(\Omega,\mathbb{R})} \\
&\le C\left\| \mathcal{P}_{\gamma-2}(X_s^x)\right\|_{L^{2\kappa_1}(\Omega,\mathbb{R})} \|\xi^{v_1,v_3}(s,x)\|_{L^{2\kappa_2}(\Omega,\mathbb{R}^d)} \|\eta^{v_2}(s,x)\|_{L^{2\kappa_3}(\Omega,\mathbb{R}^d)} \\
&\le Ce^{-(\alpha_1+\alpha_2)s} \sup_{r\in[0,\infty)} \left\| \mathcal{P}_{\gamma-2}(X_r^x)\right\|_{L^{2\max\{\kappa_1,\kappa_2\kappa_4\}}(\Omega,\mathbb{R})} \times \\
&\qquad \|v_1\|_{L^{2\kappa_2\kappa_5}(\Omega,\mathbb{R}^d)} \|v_2\|_{L^{2\kappa_3}(\Omega,\mathbb{R}^d)} \|v_3\|_{L^{2\kappa_2\kappa_6}(\Omega,\mathbb{R}^d)}.
\end{aligned} \tag{C.7}
$$

Then, we take $\kappa_1 = \rho_1$, $\kappa_2\kappa_5 = \rho_2$, $\kappa_3 = \rho_3\rho_5$, $\kappa_2\kappa_6 = \rho_3\rho_6$, $\kappa_2\kappa_4 = \rho_3\rho_4$. It is obvious that

$$
\tfrac{1}{\rho_3} = \tfrac{1}{\kappa_3} + \tfrac{1}{\kappa_2\kappa_4} + \tfrac{1}{\kappa_2\kappa_6} = 1 - \tfrac{1}{\rho_1} - \tfrac{1}{\rho_2} = 1 - \tfrac{1}{\kappa_1} - \tfrac{1}{\kappa_2\kappa_5},
$$

which also leads to

$$\frac{1}{\kappa_1}+\frac{1}{\kappa_3}+\frac{1}{\kappa_2\kappa_4}+\frac{1}{\kappa_2\kappa_5}+\frac{1}{\kappa_2\kappa_6}=\frac{1}{\kappa_1}+\frac{1}{\kappa_3}+\frac{1}{\kappa_2}\left(\frac{1}{\kappa_4}+\frac{1}{\kappa_5}+\frac{1}{\kappa_6}\right)=1.$$

This implies that there is a one-to-one correspondence between ρ_i and κ_i, $i\in\{1,2,\cdots,6\}$. About the fourth item in (C.5), we deduce by Assumption 2.4 and the Hölder inequality, for some positive constants ρ_1', ρ_2', ρ_3' and ρ_4' with $1/\rho_1'+1/\rho_2'+1/\rho_3'+1/\rho_4'=1$,

$$\begin{aligned}
&\left\|D^3F(X_s^x)\big(\eta^{v_1}(s,x),\eta^{v_2}(s,x),\eta^{v_3}(s,x)\big)\right\|_{L^2(\Omega,\mathbb{R}^d)}\\
&\le C\left\|\mathcal{P}_{\gamma-3}(X_s^x)\cdot\|\eta^{v_1}(s,x)\|\cdot\|\eta^{v_2}(s,x)\|\cdot\|\eta^{v_3}(s,x)\|\right\|_{L^2(\Omega,\mathbb{R})}\\
&\le C\left\|\mathcal{P}_{\gamma-3}(X_s^x)\right\|_{L^{2\rho_1'}(\Omega,\mathbb{R})}\|\eta^{v_1}(s,x)\|_{L^{2\rho_2'}(\Omega,\mathbb{R}^d)}\|\eta^{v_2}(s,x)\|_{L^{2\rho_3'}(\Omega,\mathbb{R}^d)}\|\eta^{v_3}(s,x)\|_{L^{2\rho_4'}(\Omega,\mathbb{R}^d)}\\
&\le Ce^{-3\alpha_1 s}\sup_{r\in[0,\infty)}\left\|\mathcal{P}_{\gamma-3}(X_r^x)\right\|_{L^{2\rho_1'}(\Omega,\mathbb{R})}\|v_1\|_{L^{2\rho_2'}(\Omega,\mathbb{R}^d)}\|v_2\|_{L^{2\rho_3'}(\Omega,\mathbb{R}^d)}\|v_3\|_{L^{2\rho_4'}(\Omega,\mathbb{R}^d)}.
\end{aligned}$$

Assuming, for example, $\rho_1'=\rho_1$ and $\rho_2'=\rho_2$, it is obvious for us to choose ρ_3' and ρ_4' satisfying $\rho_3\rho_5\ge\rho_3'$ and $\rho_3\rho_6\ge\rho_4'$. That is to say, the fourth term in the right hand side of (C.5) is controlled by the first three terms. Hence, by (5.10), we get,

$$\begin{aligned}
&\max\left\{\|H(X_s^x)\|_{L^2(\Omega,\mathbb{R}^d)},\ \|G_j(X_s^x)\|_{L^2(\Omega,\mathbb{R}^d)}\right\}\\
&\le Ce^{-\min\{\alpha_1+\alpha_2,3\alpha_1\}s}\sup_{r\in[0,\infty)}\left\|\mathcal{P}_{\gamma-2}(X_r^x)\right\|_{L^{2\max\{\rho_1,\rho_3\rho_4\}}(\Omega,\mathbb{R})}\times\\
&\qquad\|v_1\|_{L^{2\rho_2}(\Omega,\mathbb{R}^d)}\|v_2\|_{L^{2\rho_3\rho_5}(\Omega,\mathbb{R}^d)}\|v_3\|_{L^{2\rho_3\rho_6}(\Omega,\mathbb{R}^d)},
\end{aligned}$$

where the analysis of $\|G_j(X_s^x)\|_{L^2(\Omega,\mathbb{R}^d)}$, $j\in\{1,\dots,m\}$, is virtually identical to the estimate of $\|H(X_s^x)\|_{L^2(\Omega,\mathbb{R}^d)}$ so that we omit it here. Plugging these estimates with (5.9) into (C.4) yields,

$$\begin{aligned}
&\mathbb{E}\left[\|\zeta^{v_1,v_2,v_3}(t\wedge\tilde{\tau}_n^{(3)},x)\|^2\right]+(\alpha-\tilde{\epsilon}_4)\mathbb{E}\left[\int_0^{t\wedge\tilde{\tau}_n^{(3)}}\|\zeta^{v_1,v_2,v_3}(s,x)\|^2\mathrm{d}s\right]\\
&\le C_{\tilde{\epsilon}_3,\tilde{\epsilon}_4}\int_0^{t\wedge\tilde{\tau}_n^{(3)}}e^{-2\min\{\alpha_1+\alpha_2,3\alpha_1\}s}\mathrm{d}s\sup_{r\in[0,\infty)}\left\|\mathcal{P}_{\gamma-2}(X_r^x)\right\|_{L^{2\max\{\rho_1,\rho_3\rho_4\}}(\Omega,\mathbb{R})}\times\\
&\qquad\|v_1\|_{L^{2\rho_2}(\Omega,\mathbb{R}^d)}\|v_2\|_{L^{2\rho_3\rho_5}(\Omega,\mathbb{R}^d)}\|v_3\|_{L^{2\rho_3\rho_6}(\Omega,\mathbb{R}^d)}.
\end{aligned}$$

As a direct consequence of the Fatou lemma, Lemma 3.2, Lemma 3.3 and (5.10), we have with $\alpha_3:=\alpha-\tilde{\epsilon}_4$,

$$\begin{aligned}
&\|\zeta^{v_1,v_2,v_3}(t,x)\|_{L^2(\Omega,\mathbb{R}^d)}\\
&\le C_{\tilde{\epsilon}_3,\tilde{\epsilon}_4}e^{-\alpha_3 t}\sup_{r\in[0,\infty)}\left\|\mathcal{P}_{\gamma-2}(X_r^x)\right\|_{L^{2\max\{\rho_1,\rho_3\rho_4\}}(\Omega,\mathbb{R})}\times\\
&\qquad\|v_1\|_{L^{2\rho_2}(\Omega,\mathbb{R}^d)}\|v_2\|_{L^{2\rho_3\rho_5}(\Omega,\mathbb{R}^d)}\|v_3\|_{L^{2\rho_3\rho_6}(\Omega,\mathbb{R}^d)}.
\end{aligned}$$

The proof is complete. □

C.2. Proof of Lemma 5.3

Proof of Lemma 5.3. As we know, the first-order derivatives of $u(t,x)$ are

$$Du(t,x)v_1=\mathbb{E}\left[D\varphi\left(X_t^x\right)\eta^{v_1}(t,x)\right].$$

Hence, for $\varphi \in C_b^3(\mathbb{R}^d)$, we obtain from Lemma 3.3, Lemma 5.2 and the Hölder inequality that

$$\|Du(t,x)v_1\|_{L^1(\Omega,\mathbb{R})} \le \|\varphi\|_1 \cdot \left\|\eta^{v_1}(t,x)\right\|_{L^2(\Omega,\mathbb{R}^d)} \le Ce^{-\alpha_1 t}\|v_1\|_{L^2(\Omega,\mathbb{R}^d)}.$$

And the second-order derivatives of $u(t,x)$ go to

$$D^2u(t,x)(v_1,v_2) = \mathbb{E}\left[D\varphi\left(X_t^x\right)\xi^{v_1,v_2}(t,x)\right] + \mathbb{E}\left[D^2\varphi\left(X_t^x\right)\left(\eta^{v_1}(t,x),\eta^{v_2}(t,x)\right)\right].$$

In a similar way, by (5.11), we get,

$$\begin{aligned}
&\left\|D^2u(t,x)(v_1,v_2)\right\|_{L^1(\Omega,\mathbb{R})} \\
&\le \|\varphi\|_1 \cdot \left\|\xi^{v_1,v_2}(t,x)\right\|_{L^2(\Omega,\mathbb{R}^d)} + \|\varphi\|_2 \cdot \left\|\eta^{v_1}(t,x)\right\|_{L^2(\Omega,\mathbb{R}^d)}\left\|\eta^{v_2}(t,x)\right\|_{L^2(\Omega,\mathbb{R}^d)} \\
&\le Ce^{-\tilde{\alpha}_2 t}\left\|\mathcal{P}_{\gamma-2}(X_r^x)\right\|_{L^{2\rho_1}(\Omega,\mathbb{R})}\|v_1\|_{L^{2\rho_2}(\Omega,\mathbb{R}^d)}\|v_2\|_{L^{2\rho_3}(\Omega,\mathbb{R}^d)},
\end{aligned}$$

where $\tilde{\alpha}_2 := \min\{2\alpha_1, \alpha_2\}$. In addition, it follows

$$\begin{aligned}
&D^3u(t,x)(v_1,v_2,v_3) \\
&= \mathbb{E}\left[D\varphi\left(X_t^x\right)\zeta^{v_1,v_2,v_3}(t,x)\right] + \mathbb{E}\left[D^2\varphi\left(X_t^x\right)\left(\xi^{v_1,v_2}(t,x),\eta^{v_3}(t,x)\right)\right] \\
&\quad + \mathbb{E}\left[D^2\varphi\left(X_t^x\right)\left(\xi^{v_1,v_3}(t),\eta^{v_2}(t,x)\right)\right] + \mathbb{E}\left[D^2\varphi\left(X_t^x\right)\left(\eta^{v_1}(t,x),\xi^{v_2,v_3}(t,x)\right)\right] \\
&\quad + \mathbb{E}\left[D^3\varphi\left(X_t^x\right)\left(\eta^{v_1}(t,x),\eta^{v_2}(t,x),\eta^{v_3}(t,x)\right)\right].
\end{aligned} \tag{C.8}$$

As shown in the proof of Lemma 5.2, the analysis from the second term to the fourth term in (C.8) is equivalent and the estimate of the last term in the right hand side of (C.8) is bounded by the other terms. With Lemma 3.3, Lemma 5.2 and the Hölder inequality, we get, for $\bar{\rho}_1, \bar{\rho}_2, \bar{\rho}_3, \rho_1, \rho_2, \rho_3, \rho_4, \rho_5, \rho_6 > 1$ satisfying $1/\bar{\rho}_1 + 1/\bar{\rho}_2 + 1/\bar{\rho}_3 = 1$, $1/\rho_1 + 1/\rho_2 + 1/\rho_3 = 1$, $1/\rho_4 + 1/\rho_5 + 1/\rho_6 = 1$ and (5.11),

$$\begin{aligned}
&\left\|D^3u(t,x)(v_1,v_2,v_3)\right\|_{L^1(\Omega,\mathbb{R})} \\
&\le \|\varphi\|_1 \cdot \left\|\zeta^{v_1,v_2,v_3}(t,x)\right\|_{L^2(\Omega,\mathbb{R}^d)} + \|\varphi\|_2 \cdot \left\|\xi^{v_1,v_2}(t,x)\right\|_{L^2(\Omega,\mathbb{R}^d)}\left\|\eta^{v_3}(t,x)\right\|_{L^2(\Omega,\mathbb{R}^d)} \\
&\quad + \|\varphi\|_2 \cdot \left\|\xi^{v_1,v_3}(t,x)\right\|_{L^2(\Omega,\mathbb{R}^d)}\left\|\eta^{v_2}(t,x)\right\|_{L^2(\Omega,\mathbb{R}^d)} \\
&\quad + \|\varphi\|_2 \cdot \left\|\xi^{v_2,v_3}(t,x)\right\|_{L^2(\Omega,\mathbb{R}^d)}\left\|\eta^{v_1}(t,x)\right\|_{L^2(\Omega,\mathbb{R}^d)} \\
&\quad + \|\varphi\|_3 \cdot \left\|\eta^{v_1}(t,x)\right\|_{L^{2\bar{\rho}_1}(\Omega,\mathbb{R}^d)}\left\|\eta^{v_2}(t,x)\right\|_{L^{2\bar{\rho}_2}(\Omega,\mathbb{R}^d)}\left\|\eta^{v_3}(t,x)\right\|_{L^{2\bar{\rho}_3}(\Omega,\mathbb{R}^d)}.
\end{aligned} \tag{C.9}$$

If we take $\bar{\rho}_2 = \rho_3\rho_5$, $\bar{\rho}_3 = \rho_3\rho_6$, then

$$\tfrac{1}{\bar{\rho}_1} = 1 - \tfrac{1}{\rho_3\rho_5} - \tfrac{1}{\rho_3\rho_6} = 1 - \tfrac{1}{\rho_3}(1 - \tfrac{1}{\rho_4}) = \tfrac{1}{\rho_1} + \tfrac{1}{\rho_2} + \tfrac{1}{\rho_3\rho_4} \ge \tfrac{1}{\rho_2},$$

leading to $\bar{\rho}_1 \le \rho_2$. Combining this with (C.9) and Lemma 5.2 yields

$$\begin{aligned}
&\left\|D^3u(t,x)(v_1,v_2,v_3)\right\|_{L^1(\Omega,\mathbb{R})} \\
&\le Ce^{-\tilde{\alpha}_3 t}\sup_{r\in[0,\infty)}\left\|\mathcal{P}_{\gamma-2}(X_r^x)\right\|_{L^{2\max\{\rho_1,\rho_3\rho_4\}}(\Omega,\mathbb{R})}\|v_1\|_{L^{2\rho_2}(\Omega,\mathbb{R}^d)}\|v_2\|_{L^{2\rho_3\rho_5}(\Omega,\mathbb{R}^d)}\|v_3\|_{L^{2\rho_3\rho_6}(\Omega,\mathbb{R}^d)},
\end{aligned}$$

where $\tilde{\alpha}_3 := \min\{3\alpha_1, \alpha_1+\alpha_2, \alpha_3\}$. The proof is complete. □

C.3. Proof of Lemma 5.6

Proof of Lemma 5.6. In view of (4.11), (1.6), and the triangle inequality, we obtain that, for $t \in [t_n, t_{n+1}]$, $n \in \mathbb{N}_0$,

$$\|\mathbb{Z}^n(t)\|_{L^{2p}(\Omega,\mathbb{R}^d)} \leq \|\mathbb{Z}^n(t_n)\|_{L^{2p}(\Omega,\mathbb{R}^d)} + (t-t_n)\|F(\mathscr{P}(Y_n))\|_{L^{2p}(\Omega,\mathbb{R}^d)} + (t-t_n)^{\frac{1}{2}}\|g(\mathscr{P}(Y_n))\|_{L^{2p}(\Omega,\mathbb{R}^{d\times m})}.$$

According to Lemma 4.2, it suffices to show that

$$(t-t_n)\|F(\mathscr{P}(Y_n))\|_{L^{2p}(\Omega,\mathbb{R}^d)} \leq C_A h^{\frac{1}{2}},$$

and

$$(t-t_n)^{\frac{1}{2}}\|g(\mathscr{P}(Y_n))\|_{L^{2p}(\Omega,\mathbb{R}^{d\times m})} \leq C h^{\frac{1}{4}}(1 + \|\mathscr{P}(Y_n)\|_{L^{2p}(\Omega,\mathbb{R}^d)}).$$

Hence, one obtains from Lemma 4.3 that

$$\|\mathbb{Z}^n(t)\|_{L^{2p}(\Omega,\mathbb{R}^d)} \leq C_A(1 + \|x_0\|_{L^{2p}(\Omega,\mathbb{R}^d)}).$$

For the estimate of (5.13), the proof is obvious due to (5.12) and Assumption 2.4 with $p \in [1, p_0/\gamma]$. The proof is complete. □

C.4. Proof of Lemma 5.7

Proof of Lemma 5.7. Consider these two measurable sets

$$\mathcal{A}_h := \left\{\omega \in \Omega : \|\zeta(\omega)\| \leq h^{-\frac{1}{2\gamma}}\right\}, \quad \mathcal{A}_h^c := \Omega \backslash \mathcal{A}_h.$$

Therefore, owing to the Hölder inequality, for $1/q + 1/q' = 1$, we obtain

$$\mathbb{E}\left[\|\zeta - \mathscr{P}(\zeta)\|^2\right] = \mathbb{E}\left[\|\zeta - \mathscr{P}(\zeta)\|^2 \mathbf{1}_{\mathcal{A}_h^c}\right] \leq \|\zeta - \mathscr{P}(\zeta)\|^2_{L^{2q}(\Omega,\mathbb{R}^d)} \|\mathbf{1}_{\mathcal{A}_h^c}\|_{L^{q'}(\Omega,\mathbb{R})}.$$

Here, using Lemma 4.2 with the triangular inequality yields

$$\|\zeta - \mathscr{P}(\zeta)\|^2_{L^{2q}(\Omega,\mathbb{R}^d)} \leq \|\zeta\|^2_{L^{2q}(\Omega,\mathbb{R}^d)} + \|\mathscr{P}(\zeta)\|^2_{L^{2q}(\Omega,\mathbb{R}^d)} \leq 2\|\zeta\|^2_{L^{2q}(\Omega,\mathbb{R}^d)}.$$

In addition, it follows from the Markov inequality that,

$$\|\mathbf{1}_{\mathcal{A}_h^c}\|_{L^{q'}(\Omega,\mathbb{R})} = \left(\mathbb{P}(\mathcal{A}_h^c)\right)^{\frac{1}{q'}} \leq h^{\frac{\beta}{2\gamma q'}} \|\zeta\|^{\frac{\beta}{q'}}_{L^{\beta}(\Omega,\mathbb{R}^d)}.$$

We choose $q = 4\gamma + 1$, $q' = 1 + 1/4\gamma$ and $\beta = 8\gamma + 2$, then the proof is complete. □

References

[1] A. Abdulle, G. Vilmart, K.C. Zygalakis, High order numerical approximation of the invariant measure of ergodic SDEs, SIAM J. Numer. Anal. 52 (4) (2014) 1600–1622.
[2] W.-J. Beyn, E. Isaak, R. Kruse, Stochastic C-stability and B-consistency of explicit and implicit Euler-type schemes, J. Sci. Comput. 67 (3) (2016) 955–987.
[3] W.-J. Beyn, E. Isaak, R. Kruse, Stochastic C-stability and B-consistency of explicit and implicit Milstein-type schemes, J. Sci. Comput. 70 (3) (2017) 1042–1077.
[4] C.-E. Bréhier, Approximation of the invariant measure with an Euler scheme for stochastic PDEs driven by space-time white noise, Potential Anal. 40 (1) (2014) 1–40.
[5] C.-E. Brehier, Approximation of the invariant distribution for a class of ergodic SDEs with one-sided Lipschitz continuous drift coefficient using an explicit tamed Euler scheme, ESAIM Probab. Stat. 27 (2023) 841–866.
[6] S. Cerrai, Second Order PDE's in Finite and Infinite Dimension: a Probabilistic Approach, Springer, 2001.
[7] C. Chen, J. Hong, X. Wang, Approximation of invariant measure for damped stochastic nonlinear Schrödinger equation via an ergodic numerical scheme, Potential Anal. 46 (2017) 323–367.

[8] C. Chen, J. Hong, Y. Lu, Stochastic differential equation with piecewise continuous arguments: Markov property, invariant measure and numerical approximation, Discrete Contin. Dyn. Syst., Ser. B 28 (1) (2023) 765–807.
[9] G. Da Prato, An Introduction to Infinite-Dimensional Analysis, Springer Science & Business Media, 2006.
[10] W. Fang, M.B. Giles, Adaptive Euler–Maruyama method for SDEs with non-globally Lipschitz drift, in: Monte Carlo and Quasi-Monte Carlo Methods, MCQMC 2016, Stanford, CA, August 14-19 12, Springer, 2018, pp. 217–234.
[11] Y. Guo, X. Wang, Y. Wu, Order-one convergence of the backward Euler method for random periodic solutions of semilinear SDEs, arXiv preprint arXiv:2306.06689.
[12] D.J. Higham, X. Mao, L. Szpruch, Convergence, nonnegativity and stability of a new Milstein scheme with applications to finance, Discrete Contin. Dyn. Syst., Ser. B 8 (8) (2013) 2083–2100.
[13] J. Hong, X. Wang, Invariant Measures for Stochastic Nonlinear Schrödinger Equations, Springer, 2019.
[14] M. Hutzenthaler, A. Jentzen, 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. A, Math. Phys. Eng. Sci. 467 (2130) (2011) 1563–1576.
[15] M. Hutzenthaler, A. Jentzen, P.E. Kloeden, Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients, Ann. Appl. Probab. 22 (4) (2012) 1611–1641.
[16] K. Ito, M. Nisio, On stationary solutions of a stochastic differential equation, Matematika 11 (5) (1967) 117–174.
[17] P.E. Kloeden, E. Platen, Numerical Solution of Stochastic Differential Equations, Springer-Verlag, 1992.
[18] X. Li, X. Mao, 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, X. Mao, Y. Wu, The backward Euler-Maruyama method for invariant measures of stochastic differential equations with super-linear coefficients, Appl. Numer. Math. 184 (2023) 137–150.
[20] Z. Liu, Z. Qiao, Strong approximation of monotone stochastic partial differential equations driven by multiplicative noise, Stoch. Partial Differ. Equ., Anal. Computat. 9 (2021) 559–602.
[21] M.B. Majka, A. Mijatović, L. Szpruch, Nonasymptotic bounds for sampling algorithms without log-concavity, Ann. Appl. Probab. 30 (4) (2020) 1534–1581.
[22] X. Mao, L. Szpruch, Strong convergence and stability of implicit numerical methods for stochastic differential equations with non-globally Lipschitz continuous coefficients, J. Comput. Appl. Math. 238 (2013) 14–28.
[23] J.C. Mattingly, A.M. Stuart, D.J. Higham, Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise, Stoch. Process. Appl. 101 (2) (2002) 185–232.
[24] G.N. Milstein, M.V. Tretyakov, Stochastic Numerics for Mathematical Physics, Springer Science & Business Media, 2013.
[25] A. Neufeld, M.N.C. En, Y. Zhang, Non-asymptotic convergence bounds for modified tamed unadjusted Langevin algorithm in non-convex setting, arXiv preprint arXiv:2207.02600.
[26] Ł. Szpruch, X. Zhāng, V-integrability, asymptotic stability and comparison property of explicit numerical schemes for non-linear SDEs, Math. Comput. 87 (310) (2018) 755–783.
[27] X. Wang, Mean-square convergence rates of implicit Milstein type methods for SDEs with non-Lipschitz coefficients, Adv. Comput. Math. 49 (2023) 37.
[28] X. Wang, S. Gan, The tamed Milstein method for commutative stochastic differential equations with non-globally Lipschitz continuous coefficients, J. Differ. Equ. Appl. 19 (3) (2013) 466–490.
[29] X. Wang, J. Wu, B. Dong, Mean-square convergence rates of stochastic theta methods for SDEs under a coupled monotonicity condition, BIT Numer. Math. 60 (3) (2020) 759–790.
[30] X. Wang, Y. Zhao, Z. Zhang, Weak error analysis for strong approximation schemes of SDEs with super-linear coefficients, IMA J. Numer. Anal. (2023) drad083.
[31] Y. Zhao, X. Wang, Z. Zhang, Second-order numerical methods of weak convergence for SDEs with super-linear coefficients, preprint.