[Communications on Stochastic Analysis](https://digitalcommons.lsu.edu/cosa)

Volume 13 | Number 3 Article 4

9-2019

Euler-Maruyama Method for Regime Switching Stochastic Differential Equations with Hölder Coefficients

Dung T. Nguyen
Department of Applied Mathematics, Faculty of Applied Science, HCMC University of Technology, Vietnam, dungnt@hcmut.edu.vn

Son L. Nguyen
Department of Mathematics, University of Puerto Rico-Rio Piedras, San Juan, PR 00936, sonluu.nguyen@upr.edu

Communications on Stochastic Analysis
Vol. 13, No. 3 (2019) Article 4 (34 pages)
DOI: 10.31390/cosa.13.3.04

EULER-MARUYAMA METHOD FOR REGIME SWITCHING STOCHASTIC DIFFERENTIAL EQUATIONS WITH HÖLDER COEFFICIENTS

DUNG T. NGUYEN* AND SON L. NGUYEN

Abstract. In this paper, we develop Euler-Maruyama scheme for a wide-ranging class of stochastic differential equations with regime switching under such conditions that allow drift and diffusion coefficients being Hölder continuous. The strong convergence of the numerical method is proved. In addition, the rate of convergence is obtained under similar conditions to the case of usual diffusions. Some numerical examples are provided to illustrate the results.

1. Introduction

This article is devoted to Euler-Maruyama scheme for a broad class of hybrid stochastic differential equations (SDEs) known as regime switching diffusions that may not satisfy common Lipschitz conditions. Because of their ability to demonstrate both continuous dynamics and discrete events, regime switching diffusions have gained a great deal of attention during the past three decades. They have been used to model a wide variety of uncertain complex systems arising from real world applications such as automatic control and differential games, economics, financial engineering, manufacturing systems, mathematical biology, among others. For complete lists of references, applications, together with comprehensive and systematic treatments for switching diffusion systems we refer the interested readers to the monographs [18, 29, 30].

Due to the nonlinearity and complexity of these equations, finding their analytic solutions is a virtually impossible task, and developing approximation methods for these solutions turns into a natural alternative [14, 20]. Numerical schemes for regime switching SDEs therefore have become an active area since the pioneer work by Yuan and Mao [31] with numerous results on various aspects [11,12,15,16,19,22, 23,24,25,27,28,32,33]. See, for instance, [31] for Euler-Maruyama method, [24] for weak Euler-Maruyama method, [22] for tamed-Euler method, [23] for Milstein-type algorithm, [11, 32] for stability of numerical approximations, [33] for stabilization of numerical solutions, [15] for approximation of invariant measures, [25, 27] for numerical scheme for state-dependent switching systems, [28] for scheme for hybrid

Received 2019-7-28; Accepted 2019-8-30; Communicated by guest editor George Yin.
2010 *Mathematics Subject Classification.* Primary 60J60; Secondary 60J27, 60J22, 65C30.
Key words and phrases. Euler-Maruyama scheme, switching diffusion, stochastic differential equation, convergence, rate of convergence, Holder continuous.
* Corresponding author.

systems with jumps, [16] for approximation of delayed hybrid systems (see also [12]). However, most of the aforementioned works (except [19], to the best of our knowledge, which focuses on somewhat specific models) require the global or local Lipschitz conditions for the drift and diffusion coefficients despite of a vital fact that many models in reality violate these conditions. To illustrate this and motivate our study, we present below three examples arising in applications, including the constant elasticity of variance model, the Cox-Ingersoll-Ross model, and the Longstaff interest rate model with regime-switchings, in which the common assumptions of linear growth and Lipschitz continuity are violated.

Example 1.1. The first model we want to discuss is constant elasticity of variance (CEV) model with Markov switching [6]. Suppose that the price process $S(t)$ of the risky asset S evolves over time according to the following stochastic differential equation with regime-switching

$$dS(t) = \mu(\beta(t))S(t)dt + \sigma(\beta(t))S^{\theta}(t)dW(t), \quad S(0) = s > 0, \tag{1.1}$$

where $W(\cdot)$ is a standard Brownian motion, $\beta(\cdot)$ is a Markov chain with finite state space $\mathcal{M}$, $\mu(i_0)$ and $\sigma(i_0) \geq 0$, $i_0 \in \mathcal{M}$, are the appreciation rate and the volatility of the risky asset S at time t, respectively, and $0 < \theta < 1$ is the constant elasticity parameter.

This regime switching CEV process generalizes the standard CEV process by replacing the constant market parameters with the corresponding market parameters depending on a finite-state continuous-time Markov chain known as the regime switching process. In [26], the authors mentioned that it is of practical importance to allow the market parameters to respond to the movements of the general market levels since the trend of general market levels is a key factor which governs the price movements of individual risky assets. Regime switching models provide a more realistic way to model the situation where the market parameters depend on a market mode which switches among a finite number of states and reflects the state of the underlying economy, the macro-economic condition, the general mood of the investors in the market, business cycles and other economic factors.

Example 1.2. In this second example, we consider the Cox-Ingersoll-Ross (CIR) model with regime switching [4,5]. The regime switching CIR model specifies that the instantaneous interest rate follows the Markov-switching stochastic differential equation

$$dr(t) = a(\beta(t))[b(\beta(t)) - r(t)]dt + \sigma(\beta(t))\sqrt{r(t)}dW(t), \tag{1.2}$$

where $W(\cdot)$ is a Brownian motion modelling the random market risk factor, the regime-switching process $\beta(\cdot)$ is Markov chain with finite state space $\mathcal{M}$ modelling a random environment, and $a(i_0), b(i_0)$, and $\sigma(i_0)$, $i_0 \in \mathcal{M}$, are the parameters. The parameter $a(i_0)$ corresponds to the speed of adjustment, $b(i_0)$ to the mean and $\sigma(i_0)$ to the volatility.

It should be noted that the standard Cox-Ingersoll-Ross model does not contain the changes in regime. However, after the pioneer work in [7, 10] using regime-switching models in economics, regime-switching behavior of interest rate models have been widely used in interest rates modelling. Empirical evidence provided in the finance literatures Aug and Bekaert [1,2] suggests that the switching of regimes in interest rates matches well with business cycles. In addition to the statistical evidence, there are economic reasons as well to believe that the regime shifts are important to understand the behavior of entire process. As shown in [3], based on the empirical data of US. treasure yields, the poor empirical performance of the standard CIR model without switching may well suggest the existence of regime shifts. Furthermore, a generalisation of the CIR model, in which $\sqrt{r(t)}$ is replaced by $[r(t)]^\gamma$ with $\gamma > 0$, was investigated in [8].

While the drift terms of the CEV model and CIR model are still Lipschitz continuous, this property is no longer valid in the following Longstaff model.

Example 1.3. In this example, let us consider the Longstaff interest rate model with regime-switching [17], which is also known as the double square root model of stochastic interest rate, described by the following equation

$$dr(t) = \kappa(\beta(t))\Big[\theta(\beta(t)) - \sqrt{r(t)}\Big]dt + \sigma(\beta(t))\sqrt{r(t)}dW(t), \tag{1.3}$$

where $W(\cdot)$ is a Brownian motion, $\beta(\cdot)$ is a continuous time Markov chain, and $\kappa(i_0), \theta(i_0)$, and $\sigma(i_0)$, $i_0 \in \mathcal{M}$, are constant parameters.

This interest rate process is similar to the CIR model with regime switching since the interest rate is elastically drawn toward central value. However, the restoring force in (1.3) is proportional to $\theta(\beta(t)) - \sqrt{r(t)}$, rather than $\theta(\beta(t)) - r(t)$. This nonlinear restoring force has many implications in the behavior of the interest rate. Longstaff reported in [17] that this so-called nonlinear term structure outperforms the CIR model in describing actual Treasury Bill yields for the period of 1964-1986. In that paper, he used the GMM to estimate the parameter. Adopting his double square root process to specify the interest rates, we obtain the risk-neutral process described by equation (1.3).

As seen from the above examples, Lipschitz and local Lipchitz conditions are not satisfied in many applications. Nevertheless, relaxing these conditions in approximating solutions of regime-switching SDEs brings a lot of challenges. In existing literature, these conditions play an important role in proving key estimates for the differences of the drift and diffusion coefficients to establish the convergence as well as obtain the rate of convergence of numerical methods. Lacking of these properties naturally requires new approaches. However, different from the cases of usual diffusions, many powerful tools have not been well-studied for hybrid systems such as Malliavin calculus and heat kernel estimates. In this paper, in order to overcome these difficulties, we shall use the Yamada-Watanabe approximation to approximate the function $\phi(x) = |x|$ by differentiable functions (see [9, 21]) and then study Euler-Maruyama scheme for regime-switching SDEs without assuming Lipchitz conditions. Under reasonable conditions on Hölder continuity of the coefficients, the convergence for this scheme is proved and the convergence rate is asserted. The appearance of the switching process modulated in our present system on one hand makes the problem more complicated since continuous dynamics and discrete events coexist and are intertwined. On the other hand, it leads to several additional terms which require substantial improvements to estimate.

The rest of the paper is organized as follows. In the next section, we will formulate the problem and introduce the assumptions. The numerical method and some preliminary estimates are also included in this part. The rates of convergence of the numerical method are provided and proved in Section 3. Section 4 is devoted to the simulation results. Finally, Section 5 includes some further discussions. In order to keep the presentation more transparent, proofs of some technical lemmas are relegated in an appendix at the end of the paper.

2. Numerical Methods and Main Results

2.1. Formulation. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. On this probability space, let $W(\cdot)$ be a real-valued, standard Brownian motion, $\{\beta(t), t \geq 0\}$ be a Markov chain that takes values in the finite set $\mathcal{M} = \{1, 2, \ldots, m_0\}$, and Y_0 be an $\mathbb{R}$-valued random variable. We assume that $W(\cdot)$, $\beta(\cdot)$, and X_0 are independent throughout the paper. We consider the following SDEs

$$\begin{aligned} dX(t) &= b(t, X(t), \beta(t))dt + \sigma(t, X(t), \beta(t))dW(t), \quad t \in [0, T], \\ X(0) &= Y_0, \end{aligned} \tag{2.1}$$

for some $T > 0$, where $b(\cdot, \cdot, \cdot) : [0, T] \times \mathbb{R} \times \mathcal{M} \to \mathbb{R}$ and $\sigma(\cdot, \cdot, \cdot) : [0, T] \times \mathbb{R} \times \mathcal{M} \to \mathbb{R}$ are given functions and $\mathbb{E}\,|X_0|^p < \infty$.

Assumption A. We assume that $b = u + v$ where v is monotone decreasing, and there exists constants $K > 0$, $\gamma \in (0, 1]$, and $\theta \in [0, 1/2]$ such that for any $t \in [0, T]$, $x \in \mathbb{R}$, $y \in \mathbb{R}$, and $i_0 \in M$,

A1. $|u(t, x, i_0) - u(t, y, i_0)| \leq K|x - y|$.
A2. $|v(t, x, i_0) - v(t, y, i_0)| \leq K|x - y|^\gamma$.
A3. $|\sigma(t, x, i_0) - \sigma(t, y, i_0)| \leq K|x - y|^{\theta+1/2}$.
A4. $|u(t, 0, i_0)| + |v(t, 0, i_0)| + |\sigma(t, 0, i_0)| \leq K$.

Combining Assumptions A1, A2 and A4, we derive that for all $t \in [0, T]$, for all $x \in \mathbb{R}$, and for all $q \geq p > 0$,

$$|b(t, x, i_0)|^p \leq K(1 + |x|^q). \tag{2.2}$$

Also, it follows from Assumptions A3 and A4 that for all $t \in [0, T]$, for all $x \in \mathbb{R}$, and for all $p > 0$ and $q \geq (\theta + 1/2)p$,

$$|\sigma(t, x, i_0)|^p \leq K(1 + |x|^q). \tag{2.3}$$

In the next remark we will point out that Assumption (A) is valid for all examples mentioned in Section 1.

Remark 2.1. i) In Example 1.1, $b(t, x, i_0) = \mu(i_0)x$ and $\sigma(t, x, i_0) = \sigma(i_0)x^\theta$. It is easily seen that the drift coefficient is Lipschitz continuous and the diffusion coefficient is θ-Hölder continuous.
ii) Similarly, in Example 1.2, the drift $b(t, x, i_0) = a(i_0)(b(i_0) - x)$ is Lipschitz continuous and the diffusion $\sigma(t, x, i_0) = \sigma(i_0)\sqrt{x}$ is Hölder continuous of order 1/2.
iii) In Example 1.3, the drift coefficient can be decomposed into two terms: $u(t, x, i_0) = \kappa(i_0)\theta(i_0)$ is Lipschitz continuous and $v(t, x, i_0) = \kappa(i_0)\sqrt{x}$ is 1/2-Hölder continuous. The diffusion coefficient $\sigma(t, x, i_0) = \sigma(i_0)\sqrt{x}$ is Hölder continuous of order 1/2.

To proceed, we present a result concerning the existence and uniqueness of solutions of the hybrid SDEs under Assumption (A). For convenience, a proof of the following proposition is provided in the Appendix.

Proposition 2.2. *Under Assumption* (A)*, equation* (2.1) *has a unique solution.*

2.2. Numerical Methods. Let $h > 0$ be the step size. Put $N_t = \lfloor t/h \rfloor$. For any $k \geq 1$, we define $\zeta^h : [0, T] \to [0, T]$ by

$$\zeta^h(t) = kh := t_k^h \quad \text{if} \quad t \in [kh, (k+1)h).$$

Construction of the Markov chain $\beta(t)$**.** To generate a continuous-time Markov chain with a given generator $Q = (q_{i_0 j_0}) \in \mathbb{R}^{m_0 \times m_0}$, we quote the method of constructing the Markov chain from [29], Section 2.4. Constructing the sample paths of $\beta(t)$ requires determining its sojourn time at each state and its subsequent moves. The chain sojourns in any given state i_0, $i_0 \in \mathcal{M}$, for a random length of time, ζ_{i_0}, which has an exponential distribution with parameter $-q_{i_0 i_0}$. Subsequently, the process will enter another state. Each state j_0 (with $j_0 \in \mathcal{M},\ j_0 \neq i_0$) has a probability $q_{i_0 j_0}/(-q_{i_0 i_0})$ of being the chain's next residence. The post-jump location is determined by a discrete random variable Z_{i_0} taking values in $\{1, 2, \ldots, i_0 - 1, i_0 + 1, \ldots, m_0\}$. Its value is specified by

$$Z_{i_0} = \begin{cases} 1 & \text{if } \ \widehat{Z} \leq q_{i_0 1}/(-q_{i_0 i_0}), \\ 2 & \text{if } \ q_{i_0 1}/(-q_{i_0 i_0}) < \widehat{Z} \leq (q_{i_0 1} + q_{i_0 2})/(-q_{i_0 i_0}), \\ \vdots & \vdots \\ m_0 & \text{if } \ \sum_{j_0 \neq i_0, j_0 < m_0} q_{i_0 j_0}/(-q_{i_0 i_0}) \leq \widehat{Z}, \end{cases} \tag{2.4}$$

where $\widehat{Z}$ is a random variable uniformly distributed in $(0, 1)$. Thus, the sample path of $\beta(t)$ is constructed by sampling from exponential and uniform random variables alternately. With the $\beta(t)$ generated above, set $\beta(\zeta^h(t)) = \beta(kh) = \beta(t_k^h)$ if $t \in [kh, (k+1)h)$ for $h > 0$ and $k = 0, 1, \ldots$, which is the h-skeleton of the Markov chain.

Euler-Maruyama scheme. We consider the following Euler-Maruyama scheme for the solution of equation (2.1)

$$\begin{aligned} Y_0^h &= X_0, \\ Y_{k+1}^h &= Y_k^h + hb(t_k^h, Y_k^h, \beta_k^h) + \sigma(t_k^h, Y_k^h, \beta_k^h)\Delta_k^h W \quad \text{for all } k \geq 0, \end{aligned}$$

where $\beta_k^h = \beta(t_k^h)$ and $\Delta_k^h W = W(t_{k+1}^h) - W(t_k^h)$. If $b(\cdot, \cdot, \cdot)$ and $\sigma(\cdot, \cdot, \cdot)$ satisfy Assumption (A), we define the continuous-time interpolation of the approximation as

$$Y^h(t) = X_0 + \int_0^t b(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s)))ds + \int_0^t \sigma(s, Y^h(\zeta^h(t)), \beta(\zeta^h(t)))dW(s). \tag{2.5}$$

It is clear that $Y^h(t_k^h) = Y_k^h$. In the next section, for convenience, we will use $Y^h(t)$ to approximate solution $X(t)$ of (2.1).

2.3. Preliminary Estimates. Denote

$$\widetilde{Y}^h(t) = Y^h(t) - Y^h(\zeta^h(t)), \quad \widehat{Y}^h(t) = X(t) - Y^h(t). \tag{2.6}$$

We have the following estimate on the bound of the moment of $Y^h(t)$. (See Appendix A.2 for the proof)

Lemma 2.3. *Suppose that Assumptions* A1 *to* A4 *hold and* $p > 0$. *If* $\mathbb{E}\,|X_0|^p < \infty$ *then there exists a constant* C *such that*

$$\sup_{h>0} \mathbb{E}\left[\sup_{t\in[0,T]} \left|Y^h(t)\right|^p\right] < C. \tag{2.7}$$

The bound of the moment of $\widetilde{Y}^h(t)$ can be estimated as follow. (See Appendix A.3 for the proof).

Lemma 2.4. *Let* $p > 0$. *If* $\mathbb{E}\,|X_0|^{\max\{p,1+2\theta\}} < \infty$ *then under Assumptions* A1 *to* A4 *there exists a constant* C *such that*

$$\sup_{t\in[0,T]} \mathbb{E}\,|\widetilde{Y}^h(t)|^p \le Ch^{p/2}, \tag{2.8}$$

As a result, for any $\varrho > 0$,

$$\mathbb{E}\left(\int_0^t |\widetilde{Y}^h(t)|^{\varrho} ds\right)^p \le Ch^{\varrho p/2}, \tag{2.9}$$

where $C > 0$ *is a constant only depending on* p *and* $\mathbb{E}\,|X_0|^{\max\{\varrho,\varrho p,1+2\theta\}} < \infty$.

In the proofs of the main theorems and the Theorem 2.7 below, we will use the Yamada-Watanabe method to approximate the function $\varphi(x) = |x|$. Let $\delta > 1$ and $\epsilon > 0$. Then $\int_{\epsilon/\delta}^{\epsilon} \frac{dx}{x} = \ln\delta$ and therefore there is a continuous nonnegative function $\psi(x), x \in [0,\infty)$, which is zero outside $[\epsilon/\delta, \epsilon]$, has integral 1 and satisfies $\psi(x) \le \frac{2}{x\ln\delta}$ (see [13, p. 168]). Define

$$\phi(x) = \int_0^{|x|}\int_0^y \psi(z)dzdy, \quad x \in \mathbb{R}. \tag{2.10}$$

We have the following remark on the properties of $\phi(x)$.

Remark 2.5. For any $x \in \mathbb{R}$, the following properties hold true:

(a) $|x| \le \epsilon + \phi(x)$.
(b) $\phi'(x) = \frac{x}{|x|}\phi'(|x|)$.
(c) $0 \le \phi'(|x|) \le \min\left\{1, \frac{\delta|x|}{\epsilon}\right\}$.
(d) $\phi''(x) = \psi(|x|) \le \frac{2}{|x|\ln\delta}\mathbb{1}_{[\epsilon/\delta,\epsilon]}(|x|)$.

More details on these properties can be found in [13, p. 168].

Denote

$$U^h(s) = \left|b(s, Y^h(\zeta^h(s)), \beta(s)) - b(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s)))\right|, \tag{2.11}$$

$$V^h(s) = \left|\sigma(s, Y^h(\zeta^h(s)), \beta(s)) - \sigma(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s)))\right|^2, \tag{2.12}$$

and

$$L^h(t) = \int_0^t \phi'(\widehat{Y}^h(s)) \left[\sigma(s, X(s), \beta(s)) - \sigma(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s)))\right] dW(s). \tag{2.13}$$

To proceed, we need the following estimates for $U^h(t)$ and $V^h(t)$. (See Appendix A.4 for the proof)

Lemma 2.6. *Suppose that Assumptions* A1 *to* A4 *hold and let ϱ be a positive constant. If $\mathbb{E}\,|X_0|^\varrho < \infty$, then there exists a constant C such that*

$$\mathbb{E} \int_0^t \left[U^h(s)\right]^\varrho ds \le Ch. \tag{2.14}$$

Similarly, if $\mathbb{E}\,|X_0|^{(1+2\theta)\varrho} < \infty$, then there exists a constant C such that

$$\mathbb{E} \int_0^t \left[V^h(s)\right]^\varrho ds \le Ch. \tag{2.15}$$

By using the properties of the function ϕ introduced in Theorem 2.5 and Itô's formula, the next theorem allows us to estimate $|\widehat{Y}^h(t)|$ and provides a key step to prove the main theorems. (See Appendix A.5 for the proof)

Proposition 2.7. *Suppose that Assumptions* A1 *to* A3 *hold. Then*

$$\begin{aligned} |\widehat{Y}^h(t)| \le \left[\epsilon + \frac{Ct\epsilon^{2\theta}}{\ln\delta}\right] + K\int_0^t |\widehat{Y}^h(s)| ds + K\int_0^t \left[\left|\widetilde{Y}^h(s)\right| + \left|\widetilde{Y}^h(s)\right|^\gamma + U^h(s)\right] ds \\ + \frac{C\delta}{\epsilon\ln\delta}\int_0^t \left[\left|\widetilde{Y}^h(s)\right|^{1+2\theta} + V^h(s)\right] ds + L^h(t). \end{aligned} \tag{2.16}$$

It is worth mentioning that, different from the case of common diffusion, the estimate in Theorem 2.7 is not enough to prove the convergence or convergence rate when $p \ge 2$. Therefore, in order to deal with this case, we need another result. To proceed, for each $t \in [0, T]$, we define

$$Z^h(t) = \sup_{s\in[0,t]} \left|\widehat{Y}^h(s)\right|. \tag{2.17}$$

We have the following proposition.

Proposition 2.8. *Suppose that Assumptions* A1 *to* A3 *hold. If $\mathbb{E}\,|X_0|^p < \infty$ then for all $p \ge 2$,*

$$\mathbb{E}\left|Z^h(t)\right|^p \le C\,\mathbb{E}\int_0^t \left|\widehat{Y}^h(s)\right|^{p-1+2\theta} ds + C\left[h^{p\gamma/2} + h + h^{(p-1+2\theta)/2}\right].$$

In particular, if $\theta = 1/2$ then

$$\mathbb{E}\left|Z^h(t)\right|^p \le C\left(h^{p\gamma/2} + h\right).$$

3. Rates of Convergence

In order to prove Theorem 3.4, we need the following theorem whose proof is put in Appendix A.7.

Lemma 3.1. *Let $(\widehat{Y}(t))_{t\geq 0}$ be a nonnegative stochastic process and put $Z(t) = \sup_{0\leq s\leq t}\widehat{Y}(s)$. Assume that for some $p>0, q\geq 1, \rho\in[1,q]$ and constants K_1, K_2 and $\delta\geq 0$*

$$\mathbb{E}[Z(t)]^p \leq K_1\,\mathbb{E}\int_0^t [Z(s)]^p ds + K_2\,\mathbb{E}\left[\int_0^t Y^\rho(s)ds\right]^{p/q} + \delta < \infty, \tag{3.1}$$

for all $t\in[0,T]$. Then the following statements hold
(1) If $\rho = q$ then there exists a constant C depending on K_1, K_2, ρ, q and p, such that for all $t\in[0,T]$,

$$\mathbb{E}[Z(t)]^p \leq \delta\exp\left(Ct + Ct^{\max\{p/q,1\}}\right).$$

In particular, there is a constant C such that

$$\mathbb{E}[Z(T)]^p \leq C_1\delta. \tag{3.2}$$

The constant C_1 depends only on C, p, q and ρ, and increases in T.
(2) If $\rho\in[1,q)$ and, in addition, if $p\geq q$ or $p>q+1-\rho$, then there exists a constant C, depending on K_1, K_2, ρ, q and p, such that for all $t\in[0,T]$,

$$\mathbb{E}[Z(t)]^p \leq e^{Ct}\left[Ct^{\frac{(p-1)\max\{p-q,0\}}{p(q-\rho)}}\int_0^t \mathbb{E}[\widehat{Y}(s)]ds + \delta\right]. \tag{3.3}$$

In particular, there exist constants C_1 and C_2 such that

$$\mathbb{E}[Z(T)]^p \leq C_1\delta + C_2\int_0^T \mathbb{E}[\widehat{Y}(s)]ds. \tag{3.4}$$

The constants C_1 and C_2 depend only on C, p, q and ρ, and increases in T.
The same results can also be achieved if inequality (3.1) *is replaced by*

$$\mathbb{E}[Z(t)]^p \leq K_1\,\mathbb{E}\left[\int_0^t Z(s)ds\right]^p + K_2\,\mathbb{E}\left[\int_0^t [\widehat{Y}(s)]^\rho ds\right]^{p/q} + \delta < \infty. \tag{3.5}$$

The main results of this section are presented in the following theorems and corollaries.

Theorem 3.2. *Suppose that Assumption* (A) *holds. If $\mathbb{E}\,|X_0|^{1+2\theta}<\infty$ then there exists a constant C such that for all h,*

$$\sup_{\tau\in\mathcal{T}}\mathbb{E}\left|X(\tau)-Y^h(\tau)\right| \leq \begin{cases}\frac{C}{|\ln(h)|} & \text{if } \theta = 0,\\ Ch^{\min\{\theta,\gamma/2\}} & \text{if } 0<\theta\leq 1/2,\end{cases}$$

where $\mathcal{T}$ is the set of all stopping times $\tau\leq T$.

Proof. Let $U^h(s), V^h(s)$, and $M^h(s)$ be the notations introduced in Theorem 2.7. We can derive from Theorem 2.6 that

$$\mathbb{E}\int_0^t U^h(s)ds \leq Ch, \tag{3.6}$$

and

$$\mathbb{E}\int_0^t V^h(s)ds \le Ch. \tag{3.7}$$

Besides, owing to Theorems 2.4 and 2.6, it follows from Theorem 2.7 that

$$\begin{aligned}
&\mathbb{E}\,|\widehat{Y}^h(t)| \\
&\le \left[\epsilon + \frac{Ct\epsilon^{2\theta}}{\ln\delta}\right] + K\int_0^t \mathbb{E}\,|\widehat{Y}^h(s)|ds + K\int_0^t \mathbb{E}\left[\left|\widetilde{Y}^h(s)\right| + \left|\widetilde{Y}^h(s)\right|^{\gamma}\right]ds \\
&\quad + \mathbb{E}\int_0^t U^h(s)ds + \frac{C\delta}{\epsilon\ln\delta}\int_0^t \mathbb{E}\left[\left|\widetilde{Y}^h(s)\right|^{1+2\theta} + V^h(s)\right]ds \\
&\le K\int_0^t \mathbb{E}\,|\widehat{Y}^h(s)|ds + \epsilon + C\Big(h^{1/2} + h^{\gamma/2}\Big) + \frac{C\epsilon^{2\theta}}{\ln\delta} + \frac{C\delta}{\epsilon\ln\delta}\Big(h^{\theta+1/2} + h\Big).
\end{aligned} \tag{3.8}$$

Let $\overline{Y}^h(t) = \left|\widehat{Y}^h(t\wedge\tau)\right|$ for any stopping time $\tau \le T$. Then inequality (3.8) becomes

$$\begin{aligned}
\mathbb{E}[\overline{Y}^h(t)] \le K\int_0^t \mathbb{E}[\overline{Y}^h(s)]ds \\
+ C\left[\epsilon + h^{1/2} + h^{\gamma/2} + \frac{\epsilon^{2\theta}}{\ln\delta} + \frac{\delta}{\epsilon\ln\delta}\Big(h^{\theta+1/2} + h\Big)\right].
\end{aligned}$$

By Gronwall's inequality

$$\mathbb{E}[\overline{Y}^h(t)] \le C\left[\epsilon + h^{1/2} + h^{\gamma/2} + \frac{\epsilon^{2\theta}}{\ln\delta} + \frac{\delta}{\epsilon\ln\delta}\Big(h^{\theta+1/2} + h\Big)\right].$$

If $0 < \theta \le 1/2$ then, by choosing $\epsilon = h^{1/2}$ and $\delta = 2$, we obtain

$$\mathbb{E}[\overline{Y}^h(t)] \le Ch^{\min\{\theta,\gamma/2\}}.$$

Letting $t \to \infty$ yields

$$\sup_{\tau\in\mathcal{T}} \mathbb{E}\left|X(\tau) - Y^h(\tau)\right| \le Ch^{\min\{\theta,\gamma/2\}}.$$

If $\theta = 0$ then, by choosing $\epsilon = \frac{1}{|\ln(h)|}$ and $\delta = h^{-1/3}$, we arrive at

$$\mathbb{E}[\overline{Y}^h(t)] \le \frac{C}{|\ln(h)|}.$$

Letting $t \to \infty$ yields

$$\sup_{\tau\in\mathcal{T}} \mathbb{E}\left|X(\tau) - Y^h(\tau)\right| \le \frac{C}{|\ln(h)|}.$$

□

Corollary 3.3. *Under Assumption* (A), *if* $\mathbb{E}\,|X_0|^{1+2\theta} < \infty$, *then there exists a constant* C *such that for all* $h > 0$,

$$\mathbb{E}\left[\sup_{0\le t\le T}\left|X(t) - Y^h(t)\right|\right] \le \begin{cases} \frac{C}{\sqrt{|\ln(h)|}} & \text{if } \theta = 0, \\ Ch^{\min\left\{2\theta^2,\theta\gamma\right\}} & \text{if } 0 < \theta \le 1/2. \end{cases} \tag{3.9}$$

Proof. Recall that $Z^h(t) = \sup_{s\in[0,t]} \left|\widehat{Y}^h(s)\right|$. In order to estimate $Z^h(t)$, we can use Theorem 2.7, in light of Theorems 2.4 and 2.6, to derive

$$\begin{aligned}\mathbb{E}\left[Z^h(t)\right] \le K\int_0^t \mathbb{E}\left[Z^h(s)\right] ds + \left[\epsilon + \frac{C\epsilon^{2\theta}}{\ln\delta}\right] + C(h^{1/2} + h^{\gamma/2} + h) \\ + \frac{C\delta}{\epsilon\ln\delta}(h^{\theta+1/2} + h) + \mathbb{E}\sup_{s\in[0,t]}\left|L^h(s)\right|,\end{aligned} \tag{3.10}$$

where $L^h(s)$ is defined by equation (2.13). It remains to estimate $L^h(s)$. To do so, put

$$\begin{aligned}&L_1^h(t) \\ &= C\,\mathbb{E}\left(\int_0^t \left|\sigma(s, Y^h(\zeta^h(s)), \beta(s)) - \sigma(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s))\right|^2 ds\right)^{1/2}.\end{aligned} \tag{3.11}$$

An application of Burkholder-Davis-Gundy inequality, by virtue of Assumption A3, leads to

$$\begin{aligned}&\mathbb{E}\sup_{s\in[0,t]}\left|M^h(s)\right| \\ &\le C\,\mathbb{E}\left(\int_0^t \left|\phi'(\widehat{Y}^h(s))\right|^2 |\sigma(s, X(s), \beta(s)) \right. \\ &\qquad\qquad\qquad \left. - \sigma(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s)))|^2 ds\right)^{1/2} \\ &\le C\,\mathbb{E}\left(\int_0^t \left|\sigma(s, X(s), \beta(s)) - \sigma(s, Y^h(s), \beta(s))\right|^2 ds\right)^{1/2} \\ &\quad + C\,\mathbb{E}\left(\int_0^t \left|\sigma(s, Y^h(s), \beta(s)) - \sigma(s, Y^h(\zeta^h(s)), \beta(s))\right|^2 ds\right)^{1/2} + L_1^h(t) \\ &\le C\,\mathbb{E}\left(\int_0^t \left|\widehat{Y}^h(s)\right|^{1+2\theta} ds\right)^{1/2} + C\,\mathbb{E}\left(\int_0^t \left|\widetilde{Y}^h(s)\right|^{1+2\theta} ds\right)^{1/2} + L_1^h(t),\end{aligned} \tag{3.12}$$

where $L_1^h(t)$ is defined by equation (3.11). It follows from Theorem 2.4 that

$$\mathbb{E}\left[\int_0^t \left|\widetilde{Y}^h(s)\right|^{1+2\theta} ds\right]^{1/2} \le Ch^{(1+2\theta)/4}. \tag{3.13}$$

Observe that for all $s \in (t_k^h, t_{k+1}^h)$ and $i_0 \in \mathcal{M}$ the following inequality holds

$$\begin{aligned}\mathbb{P}\left(\beta(s) \ne i_0, \beta(t_k^h) = i_0 \middle| \mathcal{F}_{t_k^h}\right) &= \mathbb{P}\left(\beta(s) \ne i_0 \middle| \beta(t_k^h) = i_0\right) \mathbb{1}_{\{\beta(t_k^h)=i_0\}} \\ &\le Ch.\end{aligned} \tag{3.14}$$

Besides, $L_1^h(t)$ can be written as

$$\begin{aligned}&L_1^h(t) = \\ &C\left[\sum_{k=0}^{N_t}\int_{t_k^h}^{t_{k+1}^h\wedge t} \mathbb{E}\left|\sigma(s, Y^h(t_k^h), \beta(s)) - \sigma(s, Y^h(t_k^h), \beta(t_k^h))\right|^2 \mathbb{1}_{\{\beta(s)\ne\beta(t_k^h)\}} ds\right]^{1/2}.\end{aligned}$$

Thus, we can estimate $L_1^h(t)$, in view of inequality (2.3), as follows

$$
\begin{aligned}
&L_1^h(t)\\
&= C\Bigg\{\sum_{k=0}^{N_t}\int_{t_k^h}^{t_{k+1}^h\wedge t}\mathbb{E}\Big[\mathbb{E}\Big(\big|\sigma(s,Y^h(t_k^h),\beta(s))-\sigma(s,Y^h(t_k^h),\beta(t_k^h))\big|^2\\
&\qquad\qquad\qquad\qquad\qquad\qquad \times \mathbb{1}_{\{\beta(s)\neq\beta(t_k^h)\}}\Big|\mathcal{F}_{t_k^h}\Big)\Big]ds\Bigg\}^{1/2}\\
&= C\Bigg\{\sum_{k=0}^{N_t}\int_{t_k^h}^{t_{k+1}^h\wedge t}\sum_{i_0=1}^{m_0}\mathbb{E}\Big[\mathbb{E}\Big(\big|\sigma(s,Y^h(t_k^h),\beta(s))-\sigma(s,Y^h(t_k^h),i_0)\big|^2\\
&\qquad\qquad\qquad\qquad\qquad\qquad \times \mathbb{1}_{\{\beta(s)\neq i_0,\beta(t_k^h)=i_0\}}\Big|\mathcal{F}_{t_k^h}\Big)\Big]ds\Bigg\}^{1/2}\\
&\le C\Bigg\{\sum_{k=0}^{N_t}\int_{t_k^h}^{t_{k+1}^h\wedge t}\sum_{1\le i_0\neq j_0\le m_0}\mathbb{E}\Big[\big|\sigma(s,Y^h(t_k^h),j_0)-\sigma(s,Y^h(t_k^h),i_0)\big|^2\\
&\qquad\qquad\qquad\qquad\qquad\qquad \times \mathbb{E}\Big(\mathbb{1}_{\{\beta(s)=j_0,\beta(t_k^h)=i_0\}}\Big|\mathcal{F}_{t_k^h}\Big)\Big]ds\Bigg\}^{1/2}\\
&= C\Bigg[\sum_{k=0}^{N_t}\int_{t_k^h}^{t_{k+1}^h\wedge t}\sum_{1\le i_0\neq j_0\le m_0}\mathbb{E}\Big[\big|\sigma(s,Y^h(t_k^h),j_0)-\sigma(s,Y^h(t_k^h),i_0)\big|^2\\
&\qquad\qquad\qquad\qquad\qquad\qquad \times \mathbb{P}\Big(\beta(s)=j_0,\beta(t_k^h)=i_0\Big|\mathcal{F}_{t_k^h}\Big)\Big]ds\Bigg]^{1/2}\\
&\le C\left[h\sum_{k=0}^{N_t}\int_{t_k^h}^{t_{k+1}^h\wedge t}\Big(1+\mathbb{E}\big|Y^h(t_k^h)\big|^2\Big)ds\right]^{1/2}\le Ch^{1/2}.
\end{aligned}
\tag{3.15}
$$

Note that we have used the estimate (3.14) in the last line of inequality (3.15).

As a result, substituting inequalities (3.12), (3.13) and (3.15) into inequality (3.10) leads to

$$
\begin{aligned}
\mathbb{E}\left[Z^h(t)\right] &\le K\int_0^t\mathbb{E}\left[Z^h(s)\right]ds+\left[\epsilon+\frac{C\epsilon^{2\theta}}{2\ln\delta}\right]+C(h^{1/2}+h^{\gamma/2})\\
&\quad+\frac{C\delta}{2\epsilon\ln\delta}(h^{\theta+1/2}+h)+C\,\mathbb{E}\left(\int_0^t\left|\widehat{Y}^h(s)\right|^{1+2\theta}ds\right)^{1/2}+Ch^{(1+2\theta)/4}+Ch^{1/2},
\end{aligned}
$$

which can be simplified as follows

$$
\begin{aligned}
\mathbb{E}\left[Z^h(t)\right] &\le K\int_0^t\mathbb{E}\left[Z^h(s)\right]ds+\left[\epsilon+\frac{C\epsilon^{2\theta}}{2\ln\delta}\right]+C(h^{1/2}+h^{\gamma/2}+h^{(1+2\theta)/4})\\
&\qquad+\frac{C\delta}{2\epsilon\ln\delta}(h^{\theta+1/2}+h)+C\,\mathbb{E}\left(\int_0^t\left|\widehat{Y}^h(s)\right|^{1+2\theta}ds\right)^{1/2}.
\end{aligned}
\tag{3.16}
$$

In order to estimate the first integral $\mathbb{E}\left(\int_0^t \left|\widehat{Y}^h(s)\right|^{1+2\theta} ds\right)^{1/2}$ on the right-hand side of inequality (3.16), we consider two cases:
Case 1: $\theta = 0$. By using Jensen's inequality and Theorem 3.2, we get

$$\mathbb{E}\left(\int_0^t \left|\widehat{Y}^h(s)\right| ds\right)^{1/2} \leq \left(\int_0^t \mathbb{E}\left|\widehat{Y}^h(s)\right| ds\right)^{1/2} \leq \frac{C}{\sqrt{|\ln(h)|}},$$

which, in view of inequality (3.16) with $\epsilon = \frac{1}{|\ln(h)|}$ and $\delta = h^{-1/3}$, implies that

$$\mathbb{E}\left[Z^h(t)\right] \leq C\int_0^t \mathbb{E}\left[Z^h(s)\right] ds + \frac{C}{\sqrt{|\ln(h)|}},$$

This inequality, in light of Gronwall's inequality, verifies inequality (3.9) for $\theta = 0$.
Case 2: $\theta \in (0, 1/2]$. By using Jensen's inequality, Cauchy's inequality, and Theorem 3.2, we deduce that for any $\epsilon_1 > 0$,

$$\begin{aligned}\mathbb{E}\left(\int_0^t \left|\widehat{Y}^h(s)\right|^{1+2\theta} ds\right)^{1/2} &\leq \mathbb{E}\left(Z^h(t)\int_0^t \left|\widehat{Y}^h(s)\right|^{2\theta} ds\right)^{1/2}\\ &\leq \epsilon_1\,\mathbb{E}\left[Z^h(t)\right] + \frac{1}{4\epsilon_1}\int_0^t \mathbb{E}\left|\widehat{Y}^h(s)\right|^{2\theta} ds\\ &\leq \epsilon_1\,\mathbb{E}\left[Z^h(t)\right] + \frac{1}{4\epsilon_1}\int_0^t \left(\mathbb{E}\left|\widehat{Y}^h(s)\right|\right)^{2\theta} ds\\ &\leq \epsilon_1\,\mathbb{E}\left[Z^h(t)\right] + \frac{C}{4\epsilon_1}h^{\min\left\{2\theta^2,\gamma\theta\right\}},\end{aligned}$$

which, taking into account inequality (3.16) with $\epsilon = h^{1/2}$, $\delta = 2$, and $\epsilon_1 = \frac{1}{2C}$, implies that

$$\mathbb{E}\left[Z^h(t)\right] \leq C\int_0^t \mathbb{E}\left[Z^h(s)\right] ds + Ch^{\min\left\{2\theta^2,\gamma\theta\right\}}.$$

This inequality together with Gronwall's inequality verifies inequality (3.9) for $\theta \in (0, 1/2]$. □

The above corollary gives an estimate for the 1st moment of the error. Regarding to an estimate for its p-th moment, we have the following theorem.

Theorem 3.4. *Under Assumption* (A)*, if* $p \geq 2$ *and* $\mathbb{E}|X_0|^{(1+2\theta)p} < \infty$*, then there exists a constant* C *such that for all* $h > 0$*,*

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left|X(t) - Y^h(t)\right|^p\right] \leq \begin{cases} \frac{C}{|\ln(h)|} & \text{if } \theta = 0,\\ Ch^{\min\{\theta,\gamma/2\}} & \text{if } \theta \in (0,1/2),\\ Ch^{\min\{1,p\gamma/2\}} & \text{if } \theta = 1/2.\end{cases} \tag{3.17}$$

Proof. Bearing in mind Theorem 2.7, similar arguments used in the proof of Theorem 3.5 show that

$$\begin{aligned}|\widehat{Y}^h(t)|^p &\le C\left[\epsilon+\frac{t\epsilon^{2\theta}}{\ln\delta}\right]^p + C\left(\int_0^t\left[\left|\widetilde{Y}^h(s)\right|+\left|\widetilde{Y}^h(s)\right|^{\gamma}\right]ds\right)^p\\ &\quad + C\left(\frac{\delta}{\epsilon\ln\delta}\right)^p\left[\left(\int_0^t\left|\widetilde{Y}^h(s)\right|^{1+2\theta}ds\right)^p+\left(\int_0^t V^h(s)ds\right)^p\right] \qquad (3.18)\\ &\quad + C\left(\int_0^t|\widehat{Y}^h(s)|ds\right)^p + C\left(\int_0^t U^h(s)ds\right)^p + C\left|L^h(t)\right|^p.\end{aligned}$$

We will evaluate all terms on the right side of inequality (3.18) and then use Theorem 3.1 to get the desired estimate as follows. Taking into consideration Theorem 2.6, we obtain

$$\mathbb{E}\left(\int_0^t U^h(s)ds\right)^p \le C\int_0^t \mathbb{E}\left[U^h(s)\right]^p ds \le Ch,$$

and

$$\mathbb{E}\left(\int_0^t V^h(s)ds\right)^p \le C\int_0^t \mathbb{E}\left[V^h(s)\right]^p ds \le Ch.$$

Likewise, Theorem 2.4 implies

$$\mathbb{E}\left(\int_0^t\left|\widetilde{Y}^h(s)\right|ds\right)^p \le h^{p/2},\quad \mathbb{E}\left(\int_0^t\left|\widetilde{Y}^h(s)\right|^{\gamma}ds\right)^p \le h^{p\gamma/2},$$

and

$$\mathbb{E}\left(\int_0^t\left|\widetilde{Y}^h(s)\right|^{1+2\theta}ds\right)^p \le h^{(\theta+1/2)p}.$$

Consequently, inequality (3.18) leads to

$$\begin{aligned}\mathbb{E}\left[Z^h(t)\right]^p &\le C\,\mathbb{E}\left[\int_0^t Z^h(s)ds\right]^p + C\left[\epsilon+\frac{\epsilon^{2\theta}}{\ln\delta}\right]^p + C(h^{p/2}+h^{p\gamma/2})\\ &\quad + Ch + \left(\frac{\delta}{\epsilon\ln\delta}\right)^p C\left(h^{(\theta+1/2)p}+h\right) + \mathbb{E}\sup_{s\in[0,t]}\left|L^h(s)\right|^p.\end{aligned} \qquad (3.19)$$

It is remaining to estimate $L^h(t)$. We first put

$$L_2^h(t) = \mathbb{E}\left(\int_0^t\left|\sigma(s,Y^h(\zeta^h(s)),\beta(s)) - \sigma(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))\right|^2 ds\right)^{p/2}.$$

We will estimate $L_2^h(t)$ first.

On one hand, Hölder's inequality, Theorem 2.3, and inequality (2.3) in that order yield

$$\begin{aligned}
L_2^h(t) &\le t^{p/2-1}\,\mathbb{E}\int_0^t \left|\sigma(s,Y^h(\zeta^h(s)),\beta(s)) - \sigma(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s))\right|^p ds \\
&\le Ct^{p/2-1}\sum_{k=0}^{N_t}\int_{t_k^h}^{t_{k+1}^h\wedge t}\sum_{i_0=1}^{m_0}\mathbb{E}\left|\sigma(s,Y^h(t_k^h),\beta(s)) - \sigma(s,Y^h(t_k^h),i_0)\right|^p \\
&\qquad\qquad \times \mathbb{1}_{\left\{\beta(s)\neq i_0,\beta(t_k^h)=i_0\right\}}ds \\
&\le Ct^{p/2-1}\sum_{k=0}^{N_t}\int_{t_k^h}^{t_{k+1}^h\wedge t}\sum_{1\le i_0\neq j_0\le m_0}\mathbb{E}\Bigg[\left|\sigma(s,Y^h(t_k^h),j_0) - \sigma(s,Y^h(t_k^h),i_0)\right|^p \\
&\qquad\qquad \times \mathbb{P}\left(\beta(s)=j_0,\beta(t_k^h)=i_0\Big|\mathcal{F}_{t_k^h}\right)\Bigg]ds \\
&\le Ch\sum_{k=0}^{N_t}\int_{t_k^h}^{t_{k+1}^h\wedge t}\left(1+\mathbb{E}\left|Y^h(t_k^h)\right|^p\right)ds \le Ch.
\end{aligned} \tag{3.20}$$

On the other hand, applying Burkholder-Davis-Gundy inequality to

$$L^h(t) = \int_0^t \phi'(\widehat{Y}^h(s))\Big[\sigma(s,X(s),\beta(s)) - \sigma(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))\Big]dW(s),$$

and taking note of Assumption A3, one can arrive at the following inequality

$$\begin{aligned}
&\mathbb{E}\sup_{s\in[0,t]}\left|L^h(s)\right|^p \\
&\le C\,\mathbb{E}\left(\int_0^t \left|\phi'(\widehat{Y}^h(s))\right|^2\left|\sigma(s,X(s),\beta(s)) - \sigma(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s))\right|^2 ds\right)^{p/2} \\
&\le C\,\mathbb{E}\left(\int_0^t \left|\sigma(s,X(s),\beta(s)) - \sigma(s,Y^h(s),\beta(s))\right|^2 ds\right)^{p/2} \\
&\quad + C\,\mathbb{E}\left(\int_0^t \left|\sigma(s,Y^h(s),\beta(s)) - \sigma(s,Y^h(\zeta^h(s)),\beta(s))\right|^2 ds\right)^{p/2} + CL_2^h(t) \\
&\le C\,\mathbb{E}\left(\int_0^t \left|\widehat{Y}^h(s)\right|^{1+2\theta} ds\right)^{p/2} + C\,\mathbb{E}\left(\int_0^t \left|\widetilde{Y}^h(s)\right|^{1+2\theta} ds\right)^{p/2} + CL_2^h(t).
\end{aligned} \tag{3.21}$$

Besides, Theorem 2.4 yields

$$\mathbb{E}\left(\int_0^t \left|\widetilde{Y}^h(s)\right|^{1+2\theta} ds\right)^{p/2} \le Ch^{(1+2\theta)p/4}. \tag{3.22}$$

It follows from inequalities (3.20) to (3.22) that

$$\mathbb{E}\sup_{s\in[0,t]}\left|L^h(s)\right|^p \le C\,\mathbb{E}\left(\int_0^t \left|\widehat{Y}^h(s)\right|^{1+2\theta} ds\right)^{p/2} + Ch^{(1+2\theta)p/4} + Ch. \tag{3.23}$$

Combining inequalities (3.19) and (3.23), we get

$$\begin{aligned}\mathbb{E}\left[Z^h(t)\right]^p \le C\,\mathbb{E}\left[\int_0^t Z^h(s)ds\right]^p + C\left[\epsilon + \frac{\epsilon^{2\theta}}{\ln\delta}\right]^p + C\left(h^{p\gamma/2} + h + h^{(1+2\theta)p/4}\right)\\ + \left(\frac{\delta}{\epsilon\ln\delta}\right)^p Ch + C\,\mathbb{E}\left(\int_0^t \left|\widehat{Y}^h(s)\right|^{1+2\theta} ds\right)^{p/2}.\end{aligned} \tag{3.24}$$

In (3.24), the term $\left(\frac{\delta}{\epsilon\ln\delta}\right)^p Ch^{(\theta+1/2)p}$ is eliminated because $(\theta+1/2)p \ge 1$. Since Theorem 2.8 showed that

$$\mathbb{E}\left[Z^h(t)\right]^p \le Ch^{\min\{1,p\gamma/2\}}$$

when $\theta = 1/2$, it is sufficient to prove this theorem for $\theta \in [0, 1/2)$. We consider two cases:

Case 1: $\theta = 0$. Choosing $\delta = h^{-\frac{1}{p+1}}$ and $\epsilon = \frac{1}{|\ln(h)|}$, we derive from inequality (3.24) that

$$\mathbb{E}\left[Z^h(t)\right]^p \le C\,\mathbb{E}\left[\int_0^t Z^h(s)ds\right]^p + C\,\mathbb{E}\left(\int_0^t \left|\widehat{Y}^h(s)\right| ds\right)^{p/2} + \frac{C}{|\ln(h)|^p}. \tag{3.25}$$

By virtue of Theorem 3.2 and Theorem 3.1 with $p \ge q = 2$ and $\rho = 1$, inequality (3.25) yields

$$\mathbb{E}\left[Z^h(t)\right]^p \le \frac{C}{|\ln(h)|^p} + C\int_0^t \mathbb{E}\left|\widehat{Y}^h(s)\right| ds \le \frac{C}{|\ln h|}.$$

Case 2: $\theta \in (0, 1/2)$. We choose $\delta = 2$ and $\epsilon = h^{\frac{1}{2p}}$. Inequality (3.24) becomes

$$\begin{aligned}\mathbb{E}\left[Z^h(t)\right]^p \le C\,\mathbb{E}\left[\int_0^t Z^h(s)ds\right]^p + C\left(h^{p\gamma/2} + h^{\theta}\right)\\ + C\,\mathbb{E}\left(\int_0^t \left|\widehat{Y}^h(s)\right|^{1+2\theta} ds\right)^{p/2}.\end{aligned} \tag{3.26}$$

It can be derived from inequality (3.26) and Theorem 3.1 (with $p \ge q = 2$ and $\rho = 1 + 2\theta < 2$) that

$$\mathbb{E}\left[Z^h(t)\right]^p \le C\left(h^{p\gamma/2} + h^{\theta}\right) + \int_0^t \mathbb{E}\left|\widehat{Y}^h(s)\right| ds.$$

According to Theorem 3.2,

$$\int_0^t \mathbb{E}\left|\widehat{Y}^h(s)\right| ds \le Ch^{\min\{\theta,\gamma/2\}}.$$

Therefore,

$$\mathbb{E}\left[Z^h(t)\right]^p \le Ch^{\min\{\theta,\gamma/2\}}.$$

These completes the proof of the theorem. □

Corollary 3.5. *Under Assumption* (A)*, if* $0 < p < 2$ *and* $\mathbb{E}\,|X_0|^{2(1+2\theta)} < \infty$*, then there exists a constant* C *such that for all* $h > 0$*,*

$$\mathbb{E}\left[\sup_{0\le t\le T}\left|X(t)-Y^h(t)\right|^p\right] \le \begin{cases} \frac{C}{|\ln(h)|^{p/2}} & \text{if } \theta = 0,\\ Ch^{\min\{p\theta/2,p\gamma/4\}} & \text{if } \theta\in(0,1/2),\\ Ch^{p\gamma/2} & \text{if } \theta = 1/2.\end{cases}$$

Proof. If $0 < p < 2$ then Lyapunov inequality implies that

$$\mathbb{E}\left(\sup_{t\in[0,T]}\left|Y^h(t)-X(t)\right|^p\right) \le \left[\mathbb{E}\left(\sup_{t\in[0,T]}\left|Y^h(t)-X(t)\right|^2\right)\right]^{p/2}. \tag{3.27}$$

For our purpose here, we also consider three cases: $\theta = 0$, $\theta = 1/2$, and $\theta \in (0, 1)$. Using inequality (3.27) and the corresponding results obtained in Theorem 3.4, we can get the desired estimates. For instance, if $\theta = 0$ then we deduce from inequality (3.27) that

$$\mathbb{E}\left(\sup_{t\in[0,T]}\left|Y^h(t)-X(t)\right|^p\right) \le \left[\frac{C}{|\ln(h)|}\right]^{p/2} \le \frac{C}{|\ln(h)|^{p/2}}.$$

Similar arguments can be used for other cases. This completes the proof of the theorem. □

Remark 3.6. Theorems 3.2 and 3.5 extend the results in [9] for the case Markovian switching diffusions. In fact, we obtain the same rates of convergence. Regarding to the strong convergence in L^p sense for the Euler-Maruyama approximation with and without Markovian switching, we also achieve the same rate of convergence provided that $\theta \ne 1/2$.

If $\theta = 1/2$, that is σ is Lipchitz continuous, then it follows from Theorem 3.4 and Theorem 3.5 that

$$\mathbb{E}\left[\sup_{0\le t\le T}\left|X(t)-Y^h(t)\right|^p\right]^{1/p} \le Ch^{\min\{1/p,\gamma/2\}}.$$

In particular, the rate of convergence of EM algorithm in L^p sense is $\gamma/2$ for $0 < p \le 2$. This actually generalizes the standard L^1 and L^2 convergence rate of 1/2 for Euler-Maruyama method for SDE with Markovian switching under Lipschitz condition (i.e., $\gamma = 1$). Under certain non-Lipschitz conditions, only L^1 and L^2-convergence (without convergence rate) are discussed in [19]. Note that different from numerical methods for common SDEs, the L^p convergence rates for methods of regime switching SDEs would reduce for $p > 2$ because of the effect of estimations relating the switching process. All the existing results on numerical methods for SDE with Markovian switching therefore mainly focus on L^1 and L^2 convergences (see [18,19,22,23,24,30,31]). To the best of our knowledge, there are no results with better rates of convergence in L^p $(0 < p \le 2)$ sense for numerical method for regime switching diffusions when the diffusion coefficients are Lipschitz continuous and the drift coefficients are Hölder continuous.

4. Simulation Results

In this section, we consider several examples to demonstrate the theoretical findings.

Example 4.1. In the first example, we consider the following equation

$$\begin{aligned} &dX(t) = b(X(t),\beta(t))dt + \sigma(X(t),\beta(t))dW(t), \quad 0 \le t \le 1, \\ &X(0) = 5, \end{aligned} \tag{4.1}$$

where $W(\cdot)$ is a Brownian motion, $\beta(\cdot)$ is a two-state Markov chain with the state space $\mathcal{M} = \{1,2\}$ and the generator Q given by

$$Q = \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix}. \tag{4.2}$$

Furthermore, we suppose that the drift $b(\cdot,\cdot)$ and the diffusion $\sigma(\cdot,\cdot)$ are given as follows

$$\begin{aligned} b(x,1) &= 2 + x - 2\sqrt{x}, \qquad & b(x,2) &= 1 + 2x - \sqrt{x} \\ \sigma(x,1) &= 0.1x, & \sigma(x,2) &= 0.3x. \end{aligned}$$

That is equation (4.1) becomes

$$dX(t) = \Big(2 + X(t) - 2\sqrt{X(t)}\Big)dt + 0.1X(t)dW(t)$$

provided that the Markov chain $\beta(t)$ resides at the state 1; otherwise, if it resides at the state 2, equation (4.1) becomes

$$dX(t) = \Big(1 + 2X(t) - \sqrt{X(t)}\Big)dt + 0.3X(t)dW(t).$$

It is easy to verify that Assumption (A) holds for equation (4.1) with

$$\begin{aligned} &\theta = 0.5, \qquad && u(x,1) = 2 + x, \qquad && v(x,1) = -2\sqrt{x}, \qquad && \sigma(x,1) = 0.1x, \\ &\gamma = 0.5, && u(x,2) = 1 + 2x, && v(x,2) = -\sqrt{x}, && \sigma(x,2) = 0.3x. \end{aligned}$$

We use the Euler scheme with $h = \delta = 2^{-15}$ to be a good approximation of the exact solution and compare this with the Euler approximations using $h = 2^{-4}, h = 2^{-5}, h = 2^{-6}$, and $h = 2^{-7}$. The resulting log-log error plots are shown

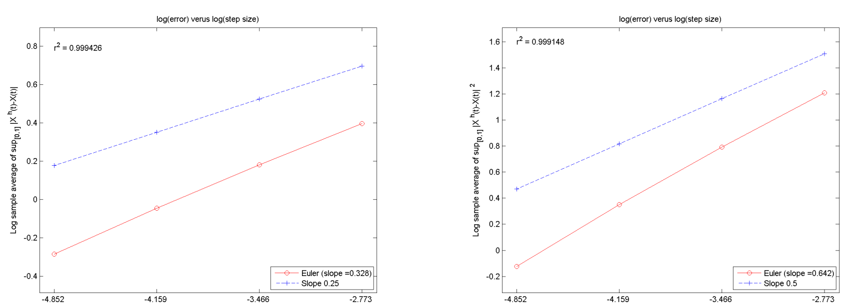

FIGURE 1. log-log plot for Theorem 4.1

with a dashed reference line of slope 0.25 and 0.5 in Fig. 1. The empirical rates of convergence of the Euler schemes, which are the slope of the first regression line and 1/2 the slope of the second regression line, can be estimated by 0.328 and $\frac{0.642}{2} = 0.321$ respectively. In this example, the empirical convergence rates are close to the verified rates.

Example 4.2. This example concerns a CEV model with Markov switching.

$$\begin{aligned} &dS(t) = \mu(\beta(t))S(t)dt + \sigma(\beta(t))[S(t)]^{4/5}dW(t), \quad 0 \le t \le 1,\\ &S(0) = 0.1, \end{aligned}$$

where $W(\cdot)$ is a Brownian motion modeling the random market risk factor; $\mu(\cdot)$ and $\sigma(\cdot)$ are the parameters of the model; $\beta(\cdot)$ is a two-state Markov chain with the state space $\mathcal{M} = \{1, 2\}$ and the generator Q given by

$$Q = \begin{pmatrix} -4 & 4 \\ 6 & -6 \end{pmatrix}.$$

Assume further that

$$\mu(1) = 5, \quad \mu(2) = 3, \quad \sigma(1) = 0.3, \quad \sigma(2) = 0.4.$$

It is easy to verify that Assumption (A) holds for this equation with $\theta = 0.3$, $g \equiv 0$. We take the Euler scheme with $h = 2^{-15}$ to be a good approximation of the exact solution and compare this with the Euler approximations using $h = 2^{-3}, h = 2^{-4}$, and $h = 2^{-5}$. The resulting log-log error plots are shown with a dashed reference

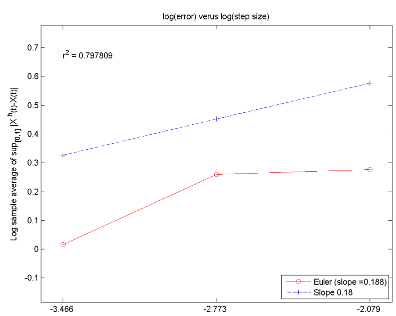

Figure 2. log-log plot for Theorem 4.2

line of slope 0.18 in Fig. 2. The empirical rates of convergence of the Euler schemes, which is the slope of the regression line, can be estimated by 0.188. In this example, the empirical convergence rate is close to 0.18.

Example 4.3. In Theorem 1.2, we introduced the regime switching Cox-Ingersoll-Ross (CIR) model for the instantaneous interest rate that follows the Markov-switching stochastic differential equation:

$$\begin{aligned} &dr(t) = a(\beta(t))\,[b(\beta(t)) - r(t)]\,dt + \sigma(\beta(t))\sqrt{r(t)}dW(t) \quad 0 \le t \le 1,\\ &r(0) = 0.05, \end{aligned} \tag{4.3}$$

where $W(\cdot)$ is a Brownian motion modeling the random market risk factor; $a(\cdot)$, $b(\cdot)$, and $\sigma(\cdot)$ are the parameters of the model; $\beta(\cdot)$ is a two-state Markov chain with the state space $\mathcal{M} = \{1, 2\}$ and the generator Q given by

$$Q = \begin{pmatrix} -3 & 3 \\ 1 & -1 \end{pmatrix}.$$

Assume further that

$$\begin{aligned} a(1) &= 0.8, & a(2) &= 1.0, \\ b(1) &= 0.05, & b(2) &= 0.03, \\ \sigma(1) &= 0.05, & \sigma(2) &= 0.10. \end{aligned}$$

It is easy to verify that Assumption (A) holds for this equation with $\theta = 0$, $g \equiv 0$. We take the Euler scheme with $h = \delta = 2^{-15}$ to be a good approximation of the exact solution and compare this with the Euler approximations using $h = 2^{-3}, h = 2^{-4}$, and $h = 2^{-5}$. The resulting log-log error plots are shown with a dashed

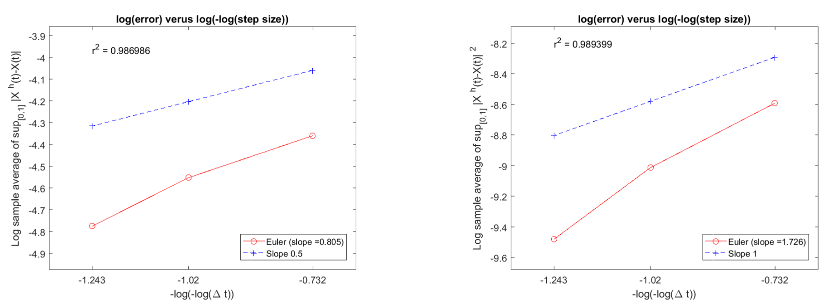

FIGURE 3. log-log plot for Theorem 4.3

reference line of slope 0.5 and 1.0 in Fig. 3. The empirical rates of convergence of the Euler schemes (as powers of $\frac{1}{|\ln(h)|}$), which are the slope of the first regression line and 1/2 the slope of the second regression line, can be estimated by 0.805 and $\frac{1.726}{2} = 0.863$ respectively.

5. Further discussion

In this paper, we study Euler-Maruyama approximation for one-dimensional stochastic differential equations with Hölder coefficients using the method of Yamada and Watanabe approximating the function $\phi(x) = |x|$ for $x \in \mathbb{R}$, which is discussed at the beginning of Section 2.3. Extending the results for multi-dimensional stochastic differential equations is a natural and interesting problem. Dealing with multi-dimensional stochastic differential equations would require a totally different approach. This open problem will be one of our future work.

Appendix A. Appendix

A.1. Proof of Theorem 2.2. We first verify the existence of equation (2.1) under Assumption (A). In fact, let $\tau_0 = 0$ and $0 < \tau_1 < \tau_2 < \dots$ be all the jump times of the switching process such that $\beta(t)$ is a constant in each interval $[\tau_i, \tau_{i+1})$ for each $i \geq 0$. Since equation (2.1) always has a unique solution within $[\tau_i, \tau_{i+1})$ under Assumption (A), the existence of solution of equation (2.1) is then confirmed.

Next, we will prove the uniqueness of equation (2.1) under Assumption (A). Suppose that $Y_1(t)$ and $Y_2(t)$ are two solutions of equation (2.1). Put $\overline{Y}(t) = Y_1(t) - Y_2(t)$. Fix $T > 0$. Then

$$\begin{aligned} d\overline{Y}(t) &= \Big[b(t, Y_1(t), \beta(t)) - b(t, Y_2(t), \beta(t))\Big]dt \\ &\quad + \Big[\sigma(t, Y_1(t), \beta(t)) - \sigma(t, Y_2(t), \beta(t))\Big]dW(t), \quad t \in [0, T], \\ \overline{Y}(0) &= 0. \end{aligned}$$

Applying Itô's formula for $\phi(\widehat{Y}^h(t))$ and using property (a) of ϕ in Theorem 2.5, we obtain

$$\begin{aligned} \big|\overline{Y}(t)\big| &\leq \epsilon + \phi(\overline{Y}(t)) \\ &= \epsilon + \int_0^t \phi'\big(\overline{Y}(s)\big)\left[u(s, Y_1(s), \beta(s)\big) - f\big(s, Y_2(s), \beta(s)\big)\right] ds \\ &\quad + \int_0^t \phi'\big(\overline{Y}(s)\big)\left[v(s, Y_1(s), \beta(s)\big) - g\big(s, Y_2(s), \beta(s)\big)\right] ds \\ &\quad + \frac{1}{2}\int_0^t \phi''\big(\overline{Y}(s)\big)\left[\sigma(s, Y_1(s), \beta(s)\big) - \sigma\big(s, Y_2(s), \beta(s)\big)\right]^2 ds \\ &\quad + \int_0^t \phi'(\overline{Y}(s))\left[\sigma(s, Y_1(s), \beta(s)) - \sigma(s, Y_2(s), \beta(s))\right] dW(s). \end{aligned} \tag{A.1}$$

We are going to estimate the integrals on the right hand side of inequality (A.1). First of all, using properties (b) and (c), and Assumption A1 yields

$$\begin{aligned} &\int_0^t \phi'\big(\overline{Y}(s)\big)\left[u(s, Y_1(s), \beta(s)\big) - f\big(s, Y_2(s), \beta(s)\big)\right] ds \\ &\leq \int_0^t \big|\phi'\big(\overline{Y}(s)\big)\big|\,\big|u(s, Y_1(s), \beta(s)\big) - f\big(s, Y_2(s), \beta(s)\big)\big| ds \\ &\leq K\int_0^t \big|\overline{Y}(s)\big| ds. \end{aligned} \tag{A.2}$$

Since g is monotone decreasing, we can use properties (b) and (c) to obtain

$$\begin{aligned} &\int_0^t \phi'\left(\overline{Y}(s)\right)\left[v(s, Y_1(s), \beta(s)\big) - g\big(s, Y_2(s), \beta(s)\big)\right] ds \\ &= \int_0^t \frac{\phi\left(|\overline{Y}(s)|\right)}{|\overline{Y}(s)|}\overline{Y}(s)\left[v(s, Y_1(s), \beta(s)\big) - g\big(s, Y_2(s), \beta(s)\big)\right] ds \leq 0. \end{aligned} \tag{A.3}$$

Thanks to property (d) and Assumption A3, the third integral in inequality (A.1) can be estimated as follows

$$\begin{aligned}
&\int_0^t \phi''\big(\overline{Y}(s)\big)\left[\sigma\big(s, Y_1(s), \beta(s)\big) - \sigma\big(s, Y_2(s), \beta(s)\big)\right]^2 ds \\
&\le \int_0^t \left|\overline{Y}(s)\right|^{2\theta+1} \frac{2}{|\overline{Y}(s)| \ln \delta} \mathbb{1}_{[\epsilon/\delta, \epsilon]}(|\overline{Y}(s)|) ds \\
&\le \int_0^t \epsilon^{2\theta} \frac{2}{\ln \delta} ds \le \frac{2T\epsilon^{2\theta}}{\ln \delta}.
\end{aligned} \tag{A.4}$$

As a result, inequalities (A.1) to (A.4) imply

$$\mathbb{E}\left|\overline{Y}(t)\right| \le \epsilon + K \int_0^t \mathbb{E}\left|\overline{Y}(s)\right| ds + \frac{2T\epsilon^{2\theta}}{\ln \delta},$$

from which, by using Gronwall's inequality, we can derive that

$$\mathbb{E}\left|\overline{Y}(t)\right| \le \left(\epsilon + \frac{2T\epsilon^{2\theta}}{\ln \delta}\right) e^{Kt}.$$

If we choose $\delta = 2$ and let $\epsilon \to 0$, then this inequality leads to $\mathbb{E}\left|\overline{Y}(t)\right| = 0$; it means that $\overline{Y}(t) = 0$ almost surely for all $t \in [0, T]$. Hence, equation (2.1) has a unique solution. □

A.2. Proof of Theorem 2.3. For each $k = 1, 2, \ldots$, put

$$\tau_k = \inf\left\{t \ge 0 : \left|Y^h(t)\right| \ge k\right\}.$$

We first note that

$$\begin{aligned}
|Y^h(t \wedge \tau_k)|^p \le C|X_0|^p &+ C\left|\int_0^{t\wedge\tau_k} b(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s)))ds\right|^p \\
&+ C\left|\int_0^{t\wedge\tau_k} \sigma(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s)))dW(s)\right|^p.
\end{aligned} \tag{A.5}$$

Using inequality (2.2) and the elementary inequality $(a_1 + b_1)^p \le C(a_1^p + b_1^p)$ for positive numbers a_1 and b_1 (constant C only depends on p), we have the following estimate for the first integral on the right hand side of inequality (A.5)

$$\begin{aligned}
\left|\int_0^{t\wedge\tau_k} b(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s)))ds\right|^p &\le \left[\int_0^{t\wedge\tau_k} \left(1 + \left|Y^h(\zeta^h(s))\right|\right) ds\right]^p \\
&\le C + C\left(\int_0^{t\wedge\tau_k} \sup_{r\in[0,s]} \left|Y^h(r)\right| ds\right)^p \\
&\le C + C\left(\int_0^t \sup_{r\in[0,s]} \left|Y^h(r \wedge \tau_k)\right| ds\right)^p.
\end{aligned} \tag{A.6}$$

We can also estimate the second integral on the right hand side of inequality (A.5) using inequality (2.3) and Burkholder-Davis-Gundy inequality as follows

$$\begin{aligned}
&\mathbb{E}\left[\sup_{r\in[0,t]}\left|\int_0^{r\wedge\tau_k}\sigma(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))dW(s)\right|^p\right]\\
&\le C\,\mathbb{E}\left(\int_0^{t\wedge\tau_k}\left|\sigma(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))\right|^2 ds\right)^{p/2}\\
&\le C\,\mathbb{E}\left(\int_0^{t\wedge\tau_k}\left(1+\left|Y^h(s)\right|^2\right)ds\right)^{p/2}\\
&\le C + C\,\mathbb{E}\left(\int_0^{t}\left|Y^h(s\wedge\tau_k)\right|^2 ds\right)^{p/2}.
\end{aligned}\tag{A.7}$$

As a result, a combination of inequalities (A.5) to (A.7) gives

$$\begin{aligned}
\mathbb{E}\sup_{r\in[0,t]}\left|Y^h(r\wedge\tau_k)\right|^p &\le C + C\,\mathbb{E}\left(\int_0^t\sup_{r\in[0,s]}\left|Y^h(r\wedge\tau_k)\right|ds\right)^p\\
&\quad + C\,\mathbb{E}\left(\int_0^t\left|Y^h(s\wedge\tau_k)\right|^2ds\right)^{p/2},
\end{aligned}$$

and thus an application of Theorem 3.1 with $\rho = q = 2$ yields that

$$\mathbb{E}\sup_{r\in[0,T]}|Y^h(r\wedge\tau_k)|^p \le C,$$

where C is a constant depending only on T, h, λ, K, p and $\mathbb{E}\,|X_0|^p$. Therefore, $\lim_{k\to\infty}\tau_k > T$ a.s. By letting $k\to\infty$ and using Fatou's lemma we obtain the desired estimate $\mathbb{E}\sup_{r\in[0,T]}|Y^h(r)|^p \le C$. □

A.3. Proof of Theorem 2.4. For any $p > 0$,

$$\begin{aligned}
\mathbb{E}\,|\widetilde{Y}^h(t)|^p &\le 2^{p-1}\,\mathbb{E}\left|\int_{\zeta^h(t)}^t b(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))ds\right|^p\\
&\quad + 2^{p-1}\,\mathbb{E}\left|\int_{\zeta^h(t)}^t \sigma(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))dW(s)\right|^p.
\end{aligned}\tag{A.8}$$

First of all, we want to estimate the first integral on the right-hand side of inequality (A.8) by considering two cases: $p \ge 1$ or $0 < p < 1$. If $p \ge 1$ then Hölder's inequality, inequality (2.2) and Theorem 2.3 imply

$$\begin{aligned}
&\mathbb{E}\left|\int_{\zeta^h(t)}^t b(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))ds\right|^p\\
&\le h^{p-1}\,\mathbb{E}\int_{\zeta^h(t)}^t\left|b(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))\right|^p ds\\
&\le Ch^{p-1}\,\mathbb{E}\int_{\zeta^h(t)}^t\left[1+\left|Y^h(\zeta^h(s))\right|^p\right]ds \le Ch^p.
\end{aligned}\tag{A.9}$$

If $0 < p < 1$, Jensen's inequality, inequality (2.2) and Theorem 2.3 also imply

$$\begin{aligned}
&\mathbb{E}\left|\int_{\zeta^h(t)}^t b(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s)))ds\right|^p \\
&\le \left[\mathbb{E}\int_{\zeta^h(t)}^t \left|b(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s)))\right| ds\right]^p \qquad \text{(A.10)}\\
&\le \left[\mathbb{E}\int_{\zeta^h(t)}^t \left(1 + \left|Y^h(\zeta^h(s))\right|\right) ds\right]^p \le Ch^p.
\end{aligned}$$

Thus, it follows from inequalities (A.9) and (A.10) that for all $p > 0$,

$$\mathbb{E}\left|\int_{\zeta^h(t)}^t b(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s)))ds\right|^p \le Ch^p. \tag{A.11}$$

Now we will estimate the second integral on the right-hand side of inequality (A.8) by considering three cases: $p \ge 2$, $1 \le p < 2$, or $0 < p < 1$. If $p \ge 2$ then the Burkholder-Davis-Gundy inequality, inequality (2.3), Hölder's inequality, Theorem 2.3 in that order derive that

$$\begin{aligned}
&\mathbb{E}\left|\int_{\zeta^h(t)}^t \sigma(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s)))dW(s)\right|^p \\
&\le C\,\mathbb{E}\left[\int_{\zeta^h(t)}^t \left|\sigma(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s)))\right|^2 ds\right]^{p/2} \\
&\le C\,\mathbb{E}\left[\int_{\zeta^h(t)}^t \left(1 + \left|Y^h(\zeta^h(s))\right|^2\right) ds\right]^{p/2} \\
&\le Ch^{p/2-1}\,\mathbb{E}\int_{\zeta^h(t)}^t \left(1 + \left|Y^h(\zeta^h(s))\right|^p\right) ds \le Ch^{p/2}.
\end{aligned}$$

If $1 \le p < 2$ then Burkholder-Davis-Gundy inequality, Jensen's inequality, inequality (2.3) and Theorem 2.3 in that order yield

$$\begin{aligned}
&\mathbb{E}\left|\int_{\zeta^h(t)}^t \sigma(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s)))dW(s)\right|^p \\
&\le C\,\mathbb{E}\left[\int_{\zeta^h(t)}^t \left|\sigma(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s)))\right|^2 ds\right]^{p/2} \\
&\le C\left[\mathbb{E}\int_{\zeta^h(t)}^t \left|\sigma(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s)))\right|^2 ds\right]^{p/2} \\
&\le C\left[\mathbb{E}\int_{\zeta^h(t)}^t \left(1 + \left|Y^h(\zeta^h(s))\right|^{1+2\theta}\right) ds\right]^{p/2} \le Ch^{p/2}.
\end{aligned}$$

If $0 < p < 1$, then Jensen's inequality, Burkholder-Davis-Gundy inequality, and inequality (2.3) give

$$\begin{aligned}
&\mathbb{E}\left|\int_{\zeta^h(t)}^t \sigma(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s)))dW(s)\right|^p \\
&\leq \left[\mathbb{E}\left|\int_{\zeta^h(t)}^t \sigma(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s)))dW(s)\right|\right]^p \\
&\leq C\left[\mathbb{E}\left|\int_{\zeta^h(t)}^t \left|\sigma(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s)))\right|^2 ds\right|^{1/2}\right]^p \\
&\leq C\left[\mathbb{E}\left|\int_{\zeta^h(t)}^t \left(1 + \left|Y^h(\zeta^h(s))\right|^{1+2\theta}\right) ds\right|^{1/2}\right]^p,
\end{aligned}$$

while Jensen's inequality and Theorem 2.3 imply

$$\begin{aligned}
\mathbb{E}\left|\int_{\zeta^h(t)}^t \left(1 + \left|Y^h(\zeta^h(s))\right|^{1+2\theta}\right) ds\right|^{1/2} &\leq \left[\mathbb{E}\int_{\zeta^h(t)}^t \left(1 + \left|Y^h(\zeta^h(s))\right|^{1+2\theta}\right) ds\right]^{1/2} \\
&\leq h^{1/2}.
\end{aligned}$$

Thus,

$$\mathbb{E}\left|\int_{\zeta^h(t)}^t \sigma(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s)))dW(s)\right|^p \leq Ch^{p/2}.$$

for all $0 < p < 1$. Hence for all $p > 0$,

$$\mathbb{E}\left|\int_{\zeta^h(t)}^t \sigma(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s)))dW(s)\right|^p \leq Ch^{p/2}. \tag{A.12}$$

Combining inequalities (A.8), (A.11) and (A.12), we obtain inequality (2.8). Next, we will verify inequality (2.9). In fact, if $0 < p < 1$ then we can use Jensen's inequality and inequality (2.8) to deduce

$$\mathbb{E}\left(\int_0^t |\widetilde{Y}^h(t)|^\varrho ds\right)^p \leq \left(\int_0^t \mathbb{E}\,|\widetilde{Y}^h(t)|^\varrho ds\right)^p \leq Ch^{\varrho p/2}.$$

Besides, if $p \geq 1$ then Hölder's inequality and inequality (2.8) yield

$$\mathbb{E}\left(\int_0^t |\widetilde{Y}^h(t)|^\varrho ds\right)^p \leq C\,\mathbb{E}\int_0^t |\widetilde{Y}^h(t)|^{\varrho p} ds \leq Ch^{\varrho p/2}.$$

Therefore, inequality (2.9) is valid for all $\varrho > 0$ and $p > 0$. $\square$

A.4. Proof of Theorem 2.6. By Assumptions A1, A2 and A4, it follows that

$$
\begin{aligned}
&\mathbb{E}\int_0^t \big[U^h(s)\big]^{\varrho}\,ds\\
&\le \sum_{k=0}^{N_t}\int_{t_k^h}^{t_{k+1}^h\wedge t}\mathbb{E}\left|b(s,Y^h(t_k^h),\beta(s))-b(s,Y^h(t_k^h),\beta(t_k^h))\right|^{\varrho}ds\\
&\le \sum_{k=0}^{N_t}\int_{t_k^h}^{t_{k+1}^h\wedge t}\sum_{1\le i_0\ne j_0\le m_0}\mathbb{E}\bigg[\left|b(s,Y^h(t_k^h),j_0)-b(s,Y^h(t_k^h),i_0)\right|^{\varrho} \qquad (A.13)\\
&\qquad\qquad\qquad\qquad \times\mathbb{P}\Big(\beta(s)=j_0,\beta(t_k^h)=i_0\Big|\mathcal{F}_{t_k^h}\Big)\bigg]ds\\
&\le Ch\sum_{k=0}^{N_t}\int_{t_k^h}^{t_{k+1}^h\wedge t}\Big(1+\mathbb{E}\left|Y^h(t_k^h)\right|^{\varrho}+\mathbb{E}\left|Y^h(t_k^h)\right|^{\gamma\varrho}\Big)ds\le Ch.
\end{aligned}
$$

The last inequality is valid due to Theorem 2.3 with $p=\varrho$ and $p=\gamma\varrho$ respectively.

A similar computation as for the estimation of $\mathbb{E}\int_0^t\big[U^h(s)\big]^{\varrho}ds$ above, on account of Assumptions A3 and A4, also gives

$$
\begin{aligned}
&\mathbb{E}\int_0^t \big[V^h(s)\big]^{\varrho}\,ds\\
&\le \sum_{k=0}^{N_t}\int_{t_k^h}^{t_{k+1}^h\wedge t}\mathbb{E}\left|\sigma(s,Y^h(t_k^h),\beta(s))-\sigma(s,Y^h(t_k^h),\beta(t_k^h))\right|^{2\varrho}ds\\
&\le \sum_{k=0}^{N_t}\int_{t_k^h}^{t_{k+1}^h\wedge t}\sum_{1\le i_0\ne j_0\le m_0}\mathbb{E}\left|\sigma(s,Y^h(t_k^h),j_0)-\sigma(s,Y^h(t_k^h),i_0)\right|^{2\varrho} \qquad (A.14)\\
&\qquad\qquad\qquad\qquad \times\mathbb{P}\Big(\beta(s)=j_0,\beta(t_k^h)=i_0\Big|\mathcal{F}_{t_k^h}\Big)\bigg]ds\\
&\le Ch\sum_{k=0}^{N_t}\int_{t_k^h}^{t_{k+1}^h\wedge t}\Big(1+\mathbb{E}\left|Y^h(t_k^h)\right|^{(1+2\theta)\varrho}\Big)ds\le Ch.
\end{aligned}
$$

This completes the proof. □

A.5. Proof of Theorem 2.7. Notice that

$$
\begin{aligned}
\widehat{Y}^h(t)=&\int_0^t\big[b(t,X(t),\beta(t))-b(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))\big]ds\\
&+\int_0^t\big[\sigma(t,X(t),\beta(t))-\sigma(s,Y^h(\zeta^h(t)),\beta(\zeta^h(t)))\big]dW(s).
\end{aligned}
$$

Using property (a) of ϕ in Theorem 2.5 and applying Itô's formula for $\phi(\widehat{Y}^h(t))$, we obtain

$$
\left|\widehat{Y}^h(t)\right|\le\epsilon+\phi(\widehat{Y}^h(t))=\epsilon+\int_0^t U_1^h(s)ds+\frac{1}{2}\int_0^t V_1^h(s)ds+L^h(t), \qquad (A.15)
$$

where $L^h(t)$ is defined as in equation (2.13) and

$$U_1^h(s) = \phi'\big(\widehat{Y}^h(s)\big)\left[b\big(s, X(s), \beta(s)\big) - b\big(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s))\big)\right],$$
$$V_1^h(s) = \phi''\big(\widehat{Y}^h(s)\big)\left[\sigma\big(s, X(s), \beta(s)\big) - \sigma\big(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s))\big)\right]^2.$$

For our purposes here, we need to estimate $U_1^h(s)$ and $V_1^h(s)$. Let $U^h(s)$ be defined as in equation (2.11). Then,

$$\begin{aligned}
U_1^h(s) = {} & \phi'(\widehat{Y}^h(s))\left[u(s, X(s), \beta(s)) - u(s, Y^h(s), \beta(s))\right] \\
& + \frac{\phi'(|\widehat{Y}^h(s)|)}{|\widehat{Y}^h(s)|}\widehat{Y}^h(s)\left[v(s, X(s), \beta(s)) - v(s, Y^h(s), \beta(s))\right] \\
& + \phi'(\widehat{Y}^h(s))\left[u(s, Y^h(s), \beta(s)) - u(s, Y^h(\zeta^h(s)), \beta(s))\right] \\
& + \phi'(\widehat{Y}^h(s))\left[v(s, Y^h(s), \beta(s)) - v(s, Y^h(\zeta^h(s)), \beta(s))\right] \\
& + \phi'(\widehat{Y}^h(s))\left[b(s, Y^h(\zeta^h(s)), \beta(s)) - b\big(s, Y^h(\zeta^h(s)), \beta(\zeta^h(s))\big)\right].
\end{aligned} \tag{A.16}$$

On the right-hand side of equation (A.16), the first and the third term can be estimated by using Theorem 2.5 and Assumption A1, the second term is negative due to the decreasing monotonicity of function g, the fourth term can be estimated by using Theorem 2.5 and Assumption A2. Therefore,

$$U_1^h(s) \le K|\widehat{Y}^h(s)| + K\left|\widetilde{Y}^h(s)\right| + K\left|\widetilde{Y}^h(s)\right|^\gamma + U^h(s). \tag{A.17}$$

It remains to estimate $V_1^h(s)$. Let $V^h(s)$ be defined as in equation (2.12). Then,

$$V_1^h(s) \le 2\phi''(\widehat{Y}^h(s))\left[\left|\sigma(s, X(s), \beta(s)) - \sigma(s, Y^h(\zeta^h(s)), \beta(s))\right|^2 + V^h(s)\right].$$

On the other hand, Assumption A3 implies

$$\begin{aligned}
\left|\sigma(s, X(s), \beta(s)) - \sigma(s, Y^h(\zeta^h(s)), \beta(s))\right|^2 & \le C\left|X(s) - Y^h(\zeta^h(s))\right|^{1+2\theta} \\
& \le C\left[\left|\widehat{Y}^h(s)\right|^{1+2\theta} + \left|\widetilde{Y}^h(s)\right|^{1+2\theta}\right].
\end{aligned}$$

As a result, Property (d) of ϕ in Theorem 2.5 gives

$$\begin{aligned}
V_1^h(s) & \le \frac{C}{|\widehat{Y}^h(s)|\ln\delta}\mathbb{1}_{[\epsilon/\delta,\epsilon]}(|\widehat{Y}^h(s)|)\left[\left|\widehat{Y}^h(s)\right|^{1+2\theta} + \left|\widetilde{Y}^h(s)\right|^{1+2\theta} + V^h(s)\right] \\
& \le \frac{C\epsilon^{2\theta}}{\ln\delta} + \frac{C\delta}{\epsilon\ln\delta}\left[\left|\widetilde{Y}^h(s)\right|^{1+2\theta} + V^h(s)\right].
\end{aligned} \tag{A.18}$$

The combination of inequalities (A.15) and (A.17) and equation (A.18) yields inequality (2.16). □

A.6. Proof of Theorem 2.8. An application of Itô's formula for $\left|\widehat{Y}^h(t)\right|^p$ gives

$$\begin{aligned}
&\left|\widehat{Y}^h(t)\right|^p \\
&= \int_0^t p\left|\widehat{Y}^h(s)\right|^{p-2}\widehat{Y}^h(s)\left[b(s,X(s),\beta(s)) - b(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))\right]ds \\
&\quad + \frac{1}{2}\int_0^t p(p-1)\left|\widehat{Y}^h(s)\right|^{p-2}\left[\sigma(s,X(s),\beta(s)) - \sigma(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))\right]^2 ds \\
&\quad + \int_0^t p\left|\widehat{Y}^h(s)\right|^{p-2}\widehat{Y}^h(s)\left[\sigma(s,X(s),\beta(s)) - \sigma(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))\right]dW(s).
\end{aligned} \tag{A.19}$$

Recall that $Z^h(t) = \sup_{s\in[0,t]}\left|\widehat{Y}^h(s)\right|$. We will estimate all integrals on the right-hand side of inequality (A.19) in three corresponding steps.

Step 1: For the first integral, we observe that

$$\begin{aligned}
&\mathbb{E}\int_0^t p\left|\widehat{Y}^h(s)\right|^{p-2}\widehat{Y}^h(s)\left[b(s,X(s),\beta(s)) - b(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))\right]ds \\
&\le \mathbb{E}\int_0^t p\left|\widehat{Y}^h(s)\right|^{p-2}\widehat{Y}^h(s)\left[u(s,X(s),\beta(s)) - u(s,Y^h(s),\beta(s))\right]ds \\
&\quad + \mathbb{E}\int_0^t p\left|\widehat{Y}^h(s)\right|^{p-2}\widehat{Y}^h(s)\left[v(s,X(s),\beta(s)) - v(s,Y^h(s),\beta(s))\right]ds \\
&\quad + \mathbb{E}\int_0^t p\left|\widehat{Y}^h(s)\right|^{p-2}\widehat{Y}^h(s)\left[b(s,Y^h(s),\beta(s)) - b(s,Y^h(\zeta^h(s)),\beta(s))\right]ds \\
&+ \mathbb{E}\int_0^t p\left|\widehat{Y}^h(s)\right|^{p-2}\widehat{Y}^h(s)\left[b(s,Y^h(\zeta^h(s)),\beta(s)) - b(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))\right]ds.
\end{aligned} \tag{A.20}$$

The first term on the right-hand side of inequality (A.20) can be estimated by using Assumption A1, while the second term is negative due to the decreasing monotonicity of function g and the third term can be estimated by using Assumptions A1 and A2. That is, inequality (A.20) becomes

$$\begin{aligned}
&\mathbb{E}\int_0^t p\left|\widehat{Y}^h(s)\right|^{p-2}\widehat{Y}^h(s)\left[b(s,X(s),\beta(s)) - b(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))\right]ds \\
&\quad \le C\,\mathbb{E}\int_0^t p\left|\widehat{Y}^h(s)\right|^p ds + C\,\mathbb{E}\int_0^t p\left|\widehat{Y}^h(s)\right|^{p-1}\left[\left|\widetilde{Y}^h(s)\right| + \left|\widetilde{Y}^h(s)\right|^\gamma + U^h(s)\right]ds \\
&\quad \le C\,\mathbb{E}\int_0^t \left|\widehat{Y}^h(s)\right|^p ds + C\,\mathbb{E}\int_0^t \left[\left|\widehat{Y}^h(s)\right|^p + \left|\widetilde{Y}^h(s)\right|^{p\gamma} + \left[U^h(s)\right]^p\right]ds.
\end{aligned}$$

On the other hand, applications of Theorem 2.4 and Theorem 2.6 give

$$\mathbb{E}\int_0^t \left[\left|\widetilde{Y}^h(s)\right|^p + \left|\widetilde{Y}^h(s)\right|^{p\gamma} + \left[U^h(s)\right]^p\right]ds \le C\left(h^{p/2} + h^{p\gamma/2} + h\right).$$

Therefore,

$$\begin{aligned}
&\mathbb{E}\int_0^t p\left|\widehat{Y}^h(s)\right|^{p-2}\widehat{Y}^h(s)\left[b(s,X(s),\beta(s))-b(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))\right]ds\\
&\quad\le C\,\mathbb{E}\int_0^t\left|Z^h(s)\right|^p ds+C\left(h^{p/2}+h^{p\gamma/2}+h\right).
\end{aligned}\tag{A.21}$$

Step 2: The second integral on the right hand side of inequality (A.19) can also be estimated by using Assumptions A1 and A2 as follows

$$\begin{aligned}
&\mathbb{E}\int_0^t\left|\widehat{Y}^h(s)\right|^{p-2}\left[\sigma(s,X(s),\beta(s))-\sigma(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))\right]^2 ds\\
&\quad\le C\,\mathbb{E}\int_0^t\left|\widehat{Y}^h(s)\right|^{p-2}\left[\sigma(s,X(s),\beta(s))-\sigma(s,Y^h(s),\beta(s))\right]^2 ds\\
&\qquad+C\,\mathbb{E}\int_0^t\left|\widehat{Y}^h(s)\right|^{p-2}\left[\sigma(s,Y^h(s),\beta(s))-\sigma(s,Y^h(\zeta^h(s)),\beta(s))\right]^2 ds\\
&\qquad+C\,\mathbb{E}\int_0^t\left|\widehat{Y}^h(s)\right|^{p-2}\left[\sigma(s,Y^h(\zeta^h(s)),\beta(s))-\sigma(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))\right]^2 ds\\
&\quad\le C\,\mathbb{E}\int_0^t\left|\widehat{Y}^h(s)\right|^{p-1+2\theta}ds+C\,\mathbb{E}\int_0^t\left|\widehat{Y}^h(s)\right|^{p-2}\left[\left|\widetilde{Y}^h(s)\right|^{1+2\theta}+V^h(s)\right]ds\\
&\quad\le C\,\mathbb{E}\int_0^t\left|\widehat{Y}^h(s)\right|^{p-1+2\theta}ds+C\,\mathbb{E}\int_0^t\left[\left|\widetilde{Y}^h(s)\right|^{p-1+2\theta}+\left[V^h(s)\right]^{\frac{p-1+2\theta}{1+2\theta}}\right]ds.
\end{aligned}$$

Thanks to Theorem 2.4 and Theorem 2.6, we can verify that

$$\mathbb{E}\int_0^t\left[\left|\widetilde{Y}^h(s)\right|^{p-1+2\theta}+\left[V^h(s)\right]^{\frac{p-1+2\theta}{1+2\theta}}\right]ds\le C\left(h^{(p-1+2\theta)/2}+h\right).$$

Hence,

$$\begin{aligned}
&\mathbb{E}\int_0^t\left|\widehat{Y}^h(s)\right|^{p-2}\left[\sigma(s,X(s),\beta(s))-\sigma(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))\right]^2 ds\\
&\quad\le C\,\mathbb{E}\int_0^t\left|\widehat{Y}^h(s)\right|^{p-1+2\theta}ds+C\left(h^{(p-1+2\theta)/2}+h\right).
\end{aligned}\tag{A.22}$$

Step 3:
Finally, we will estimate the third integral on the right hand side of inequality (A.19) using Burkholder-Davis-Gundy inequality as follows

$$\begin{aligned}
&\mathbb{E}\sup_{r\in[0,t]}\left|\int_0^r\left|\widehat{Y}^h(s)\right|^{p-2}\widehat{Y}^h(s)[\sigma(s,X(s),\beta(s))-\sigma(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))]dW(s)\right|\\
&\le C\,\mathbb{E}\left[\left(\int_0^t\left|\widehat{Y}^h(s)\right|^{2(p-1)}\left|\sigma(s,X(s),\beta(s))-\sigma(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))\right|^2 ds\right)^{1/2}\right]
\end{aligned}$$

which is bounded by

$$\begin{aligned}
&C\,\mathbb{E}\left[\left(\int_0^t \left|\widehat{Y}^h(s)\right|^{2(p-1)} \left|\sigma(s,X(s),\beta(s))-\sigma(s,Y^h(s),\beta(s))\right|^2 ds\right)^{1/2}\right]\\
&\quad+C\,\mathbb{E}\left[\left(\int_0^t \left|\widehat{Y}^h(s)\right|^{2(p-1)} \left|\sigma(s,Y^h(s),\beta(s))-\sigma(s,Y^h(\zeta^h(s)),\beta(s))\right|^2 ds\right)^{1/2}\right]\\
&\quad+C\,\mathbb{E}\left[\left(\int_0^t \left|\widehat{Y}^h(s)\right|^{2(p-1)} |\sigma(s,Y^h(\zeta^h(s)),\beta(s))\right.\right.\\
&\qquad\qquad\qquad\left.\left.-\,\sigma(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))|^2 ds\right)^{1/2}\right].
\end{aligned}$$

Similar arguments as in Step 1 and 2 yield

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_0^t \left|\widehat{Y}^h(s)\right|^{2(p-1)} \left|\sigma(s,X(s),\beta(s))-\sigma(s,Y^h(s),\beta(s))\right|^2 ds\right)^{1/2}\right]\\
&\quad\le C\,\mathbb{E}\left[\left(\int_0^t \left|\widehat{Y}^h(s)\right|^{2p-1+2\theta} ds\right)^{1/2}\right]\\
&\quad\le C\,\mathbb{E}\left[\left|Z^h(t)\right|^{p/2}\left(\int_0^t \left|\widehat{Y}^h(s)\right|^{p-1+2\theta} ds\right)^{1/2}\right]\\
&\quad\le C\,\mathbb{E}\int_0^t \left|\widehat{Y}^h(s)\right|^{p-1+2\theta} ds+\frac{1}{6}\,\mathbb{E}\left|Z^h(t)\right|^p.
\end{aligned}$$

Likewise,

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_0^t \left|\widehat{Y}^h(s)\right|^{2(p-1)} \left|\sigma(s,Y^h(s),\beta(s))-\sigma(s,Y^h(\zeta^h(s)),\beta(s))\right|^2 ds\right)^{1/2}\right]\\
&\quad\le C\,\mathbb{E}\left[\left(\int_0^t \left|\widehat{Y}^h(s)\right|^{2(p-1)} \left|\widetilde{Y}^h(s)\right|^{1+2\theta} ds\right)^{1/2}\right]\\
&\quad\le C\,\mathbb{E}\left[\left|Z^h(t)\right|^{p-1}\left(\int_0^t \left|\widetilde{Y}^h(s)\right|^{1+2\theta} ds\right)^{1/2}\right]\\
&\quad\le \frac{1}{6}\,\mathbb{E}\left|Z^h(t)\right|^p+C\,\mathbb{E}\left[\left(\int_0^t \left|\widetilde{Y}^h(s)\right|^{1+2\theta} ds\right)^{p/2}\right],
\end{aligned}$$

and

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_0^t \left|\widehat{Y}^h(s)\right|^{2(p-1)} \left|\sigma(s,Y^h(\zeta^h(s)),\beta(s))-\sigma(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))\right|^2 ds\right)^{1/2}\right]\\
&\quad\le C\,\mathbb{E}\left[\left|Z^h(t)\right|^{p-1}\left(\int_0^t V^h(s)ds\right)^{1/2}\right]\\
&\quad\le \frac{1}{6}\,\mathbb{E}\left|Z^h(t)\right|^p+C\,\mathbb{E}\left[\left(\int_0^t V^h(s)ds\right)^{p/2}\right].
\end{aligned}$$

Thus,

$$\begin{aligned}&\mathbb{E}\sup_{r\in[0,t]}\left|\int_0^r\left|\widehat{Y}^h(s)\right|^{p-2}\widehat{Y}^h(s)[\sigma(s,X(s),\beta(s))-\sigma(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))]dW(s)\right|\\&\quad\le C\,\mathbb{E}\int_0^t\left|\widehat{Y}^h(s)\right|^{p-1+2\theta}ds+\frac{1}{2}\,\mathbb{E}\left|Z^h(t)\right|^p+C\,\mathbb{E}\left[\left(\int_0^t\left|\widetilde{Y}^h(s)\right|^{1+2\theta}ds\right)^{p/2}\right]\\&\qquad+C\,\mathbb{E}\left[\left(\int_0^t V^h(s)ds\right)^{p/2}\right],\end{aligned}$$

whence we can use Theorem 2.4 and Theorem 2.6 to derive

$$\begin{aligned}&\mathbb{E}\sup_{r\in[0,t]}\left|\int_0^r\left|\widehat{Y}^h(s)\right|^{p-2}\widehat{Y}^h(s)[\sigma(s,X(s),\beta(s))-\sigma(s,Y^h(\zeta^h(s)),\beta(\zeta^h(s)))]dW(s)\right|\\&\quad\le C\,\mathbb{E}\int_0^t\left|\widehat{Y}^h(s)\right|^{p-1+2\theta}ds+\frac{1}{2}\,\mathbb{E}\left|Z^h(t)\right|^p+C\left[h^{(1+2\theta)p/2}+h\right].\end{aligned}\tag{A.23}$$

In consequence, a combination of inequalities (A.19) and (A.21) to (A.23) leads to

$$\begin{aligned}\mathbb{E}\left|Z^h(t)\right|^p&\le C\,\mathbb{E}\int_0^t\left|Z^h(t)\right|^p ds+C\,\mathbb{E}\int_0^t\left|\widehat{Y}^h(s)\right|^{p-1+2\theta}ds\\&\quad+C\left[h^{p\gamma/2}+h+h^{(p-1+2\theta)/2}\right].\end{aligned}$$

Hence, using Gronwall's inequality, we get

$$\mathbb{E}\left|Z^h(t)\right|^p\le C\,\mathbb{E}\int_0^t\left|\widehat{Y}^h(s)\right|^{p-1+2\theta}ds+C\left[h^{p\gamma/2}+h+h^{(p-1+2\theta)/2}\right].$$

In particular, if $\theta=1/2$ then

$$\begin{aligned}\mathbb{E}\left|Z^h(t)\right|^p&\le C\,\mathbb{E}\int_0^t\left|\widehat{Y}^h(s)\right|^p ds+C\left[h^{p\gamma/2}+h+h^{p/2}\right]\\&\le C\,\mathbb{E}\int_0^t\left|Z^h(s)\right|^p ds+C\left[h^{p\gamma/2}+h\right].\end{aligned}$$

Again, an application of Gronwall's inequality also yields

$$\mathbb{E}\left|Z^h(t)\right|^p\le C\left[h^{p\gamma/2}+h\right].$$

□

A.7. Proof of Theorem 3.1. Let $p>0$. We claim that: For any $C_1>0$, there exists a constant $C_2>0$ such that

$$\left(\int_0^t\widehat{Y}(s)ds\right)^p\le C_1[Z(t)]^p+C_2t^{\max\{p-1,0\}}\int_0^t[\widehat{Y}(s)]^p ds,\tag{A.24}$$

for all $t\in[0,T]$, where C_2 depends only on p,C_1, and T. Indeed, if $p\ge 1$ then Hölder's inequality implies that

$$\left(\int_0^t\widehat{Y}(s)ds\right)^p\le t^{p-1}\int_0^t[\widehat{Y}(s)]^p ds,$$

and if $0 < p < 1$ then Young's inequality implies that

$$\begin{aligned}\left(\int_0^t \widehat{Y}(s)ds\right)^p &\le [Z(t)]^{p(1-p)}\left(\int_0^t [\widehat{Y}(s)]^p ds\right)^p\\ &\le C_1[Z(t)]^p + C_2\int_0^t [\widehat{Y}(s)]^p ds.\end{aligned}$$

Thus inequality (A.24) holds for all $p > 0$ and $t \in [0, T]$. Using inequality (A.24) with V in place of Y and $C_1 = 1/2$, we obtain

$$\left(\int_0^t Z(s)ds\right)^p \le \frac{1}{2}[Z(t)]^p + C_2 T^{\max\{p-1,0\}}\int_0^t [Z(s)]^p ds$$

for all $t \in [0, T]$. As a result, inequality (3.1) can be derived from inequality (3.5). Therefore, it it only necessary to prove this theorem under assumption (3.1) instead of (3.5). Now we assume that inequality (3.1) holds. We consider the following cases.

i. If $\rho = q$ then using inequality (A.24) with Y^q in place of Y, p/q in place of p, and $C_1 = \frac{1}{2K_2}$ gives

$$\left(\int_0^t [\widehat{Y}(s)]^q ds\right)^{p/q} \le \frac{1}{2K_2}[Z(t)]^p + C_2 t^{\max\{(p-q)/q,0\}}\int_0^t [\widehat{Y}(s)]^p ds.$$

Thus, by inequality (3.1),

$$\mathbb{E}[Z(t)]^p \le \left(C + Ct^{\max\{(p-q)/q,0\}}\right)\mathbb{E}\int_0^t [Z(s)]^p ds + \delta,$$

and hence, for any $r \in [0, t]$

$$\mathbb{E}[Z(r)]^p \le \left(C + Ct^{\max\{(p-q)/q,0\}}\right)\mathbb{E}\int_0^r [Z(s)]^p ds + \delta,$$

which, in view of Gronwall's inequality, leads to

$$\mathbb{E}[Z(t)]^p \le \delta\exp\left(Ct + Ct^{\max\{p/q,1\}}\right).$$

By choosing $t = T$ in this estimate, we obtain inequality (3.2).

ii. If $p \ge q$ with $q \ge 1$ and $\rho \in [1, q)$ then Hölder's inequality and Young's inequality imply

$$\begin{aligned}\left(\int_0^t [\widehat{Y}(s)]^\rho ds\right)^{p/q} &\le t^{\frac{p-q}{q}}\int_0^t [\widehat{Y}(s)]^{\rho p/q} ds \qquad (A.25)\\ &\le t^{\frac{p-q}{q}}\int_0^t [\widehat{Y}(s)]^{\frac{p}{p-1}\left(1-\frac{\rho}{q}\right)}[Z(s)]^{\frac{p}{p-1}\left(\frac{p\rho}{q}-1\right)} ds\\ &\le Ct^{\frac{(p-q)(p-1)}{p(q-\rho)}}\int_0^t \widehat{Y}(s)ds + C\int_0^t [Z(s)]^p ds. \qquad (A.26)\end{aligned}$$

Here inequality (A.25) is valid due to Hölder's inequality and inequality (A.26) is valid due to Young's inequality. In view of inequality (3.1),

$$\mathbb{E}[Z(t)]^p \le C\,\mathbb{E}\int_0^t [Z(s)]^p ds + Ct^{\frac{(p-q)(p-1)}{p(q-\rho)}}\int_0^t \widehat{Y}(s)ds + \delta,$$

which, in view of Gronwall's inequality, leads to

$$\mathbb{E}[Z(t)]^p \leq e^{Ct}\left[Ct^{\frac{(p-q)(p-1)}{p(q-\rho)}}\int_0^t \mathbb{E}[\widehat{Y}(s)]ds+\delta\right].$$

Thus inequality (3.3) holds if $p \geq q$. Then choosing $t = T$ in inequality (3.3) verifies inequality (3.4).

iii. If $q + 1 - \rho < p < q$ with $q \geq 1$ and $\rho \in [1, q)$ then Young's inequality implies

$$\begin{aligned}\left(\int_0^t [\widehat{Y}(s)]^\rho ds\right)^{p/q} &\leq [Z(t)]^{\frac{(q-p)p}{q}}\left(\int_0^t [\widehat{Y}(s)]^{\rho-(q-p)}ds\right)^{p/q}\\ &\leq \frac{1}{2K_2}[Z(t)]^p + C\int_0^t [\widehat{Y}(s)]^{p-(q-\rho)}ds,\end{aligned}$$

and

$$\begin{aligned}\int_0^t [\widehat{Y}(s)]^{p-(q-\rho)}ds &\leq \int_0^t [\widehat{Y}(s)]^{\frac{q-\rho}{p-1}}[Z(s)]^{p\left(1-\frac{q-\rho}{p-1}\right)}\\ &\leq C\int_0^t \widehat{Y}(s)ds + C\int_0^t [Z(s)]^p ds.\end{aligned}$$

Again, by inequality (3.1),

$$\mathbb{E}[Z(t)]^p \leq C\int_0^t \mathbb{E}[Z(s)]^p ds + C\int_0^t \mathbb{E}[\widehat{Y}(s)]ds + \delta,$$

which, in view of Gronwall's inequality, leads to

$$\mathbb{E}[Z(t)]^p \leq e^{Ct}\left[C\int_0^t \mathbb{E}[\widehat{Y}(s)]ds + \delta\right].$$

Thus inequality (3.3) holds if $p \geq q+1-\rho$. Then choosing $t = T$ in inequality (3.3) verifies inequality (3.4).

The proof is complete. □

Acknowledgment. The first author's research is funded by Ho Chi Minh City University of Technology – VNU-HCM under grand number T-KHUD-2018-18. The second author's research was supported by a seed fund of Department of Mathematics at University of Puerto Rico, Rio Piedras.

References

1. Ang, A. and Bekaert, G.: International Asset Allocation With Regime Shifts. *The Review of Financial Studies* **15** (2002), 1137–1187.
2. Ang, A. and Bekaert, G.: Short Rate Nonlinearities and Regime Switches. *Journal of Economic Dynamics and Control* **26** (2002), 1243–1274.
3. Brown, S. J. and Dybvig, P. H.: The Empirical Implications of The Cox, Ingersoll, Ross Theory of The Term Structure of Interest Rates. *The Journal of Finance* **41** (1986), 617–630.
4. Choi, S.: Regime-Switching Univariate Diffusion Models of The Short-Term Interest Rate. *Studies in Nonlinear Dynamics & Econometrics* **13** (2009).
5. Driffill, J., Kenc, T., and Sola, M.: An Empirical Examination of Term Structure Models with Regime Shifts. *EFMA 2003 Helsinki Meetings*. 2003.

6. Elliott, R. J., Chan, L., and Siu, T. K.: Option Valuation Under a Regime-Switching Constant Elasticity of Variance Process. *Applied Mathematics and Computation* **219** (2013), 4434–4443.
7. Garcia, R. and Perron, P.: An Analysis of The Real Interest Rate Under Regime Shifts. *The Review of Economics and Statistics* **78** (1996), 111–125.
8. Gray, S. F.: Modeling The Conditional Distribution of Interest Rates As a Regime-Switching Process. *Journal of Financial Economics* **42** (1996), 27–62.
9. Gyöngy, I. and Rásonyi, M.: A Note on Euler Approximations for SDEs with Hölder Continuous Diffusion Coefficients. *Stochastic Processes and Their Applications* **121** (2011), 2189–2200.
10. Hamilton, J. D.: A New Approach to The Economic Analysis of Nonstationary Time Series and The Business Cycle. *Econometrica* **57** (1989), 357–384.
11. Higham, D. J., Mao, X., and Yuan, C.: Preserving Exponential Mean-Square Stability in The Simulation of Hybrid Stochastic Differential Equations. *Numerische Mathematik* **108** (2007), 295–325.
12. Hu, Y., Chen, H., and Yuan, C.: Numerical Solutions of Neutral Stochastic Functional Differential Equations with Markovian Switching. *Advances in Difference Equations* **2019** (2019).
13. Ikeda, N. and Watanabe, S.: *Stochastic Differential Equations and Diffusion Processes*. Vol. 24. North-Holland and Kodansha Ltd, 1981.
14. Kloeden, P. and Platen, E.: *Numerical Solution of Stochastic Differential Equations*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2013.
15. Li, X., Ma, Q., Yang, H., and Yuan, C.: The Numerical Invariant Measure of Stochastic Differential Equations With Markovian Switching. *SIAM Journal on Numerical Analysis* **56** (2018), 1435–1455.
16. Li, X., Mao, X., and Shen, Y.: Approximate Solutions of Stochastic Differential Delay Equations with Markovian Switching. *Journal of Difference Equations and Applications* **16** (2010), 195–207.
17. Longstaff, F. A.: A Nonlinear General Equilibrium Model of The Term Structure of Interest Rates. *Journal of Financial Economics* **23** (1989), 195–224.
18. Mao, X. and Yuan, C.: *Stochastic Differential Equations with Markovian Switching*. World Scientific, 2006.
19. Mao, X., Yuan, C., and Yin, G.: Approximations of Euler-Maruyama Type for Stochastic Differential Equations with Markovian Switching, Under Non-Lipschitz Conditions. *Journal of Computational and Applied Mathematics* **205** (2007), 936–948.
20. Milstein, G. N. and Tretyakov, M. V.: *Stochastic Numerics for Mathematical Physics*. 1st ed. Springer Berlin Heidelberg, 2004.
21. Ngo, H. L. and Taguchi, D.: Strong rate of convergence for the EulerMaruyama approximation of stochastic differential equations with irregular coefficients. *Math. Comput.* **85** (2016), 1793–1819.
22. Nguyen, D. T., Nguyen, S. L., Hoang, T. A., and Yin, G.: Tamed-Euler Method for Hybrid Stochastic Differential Equations with Markovian Switching. *Nonlinear Analysis: Hybrid Systems* **30** (2018), 14–30.
23. Nguyen, S. L., Hoang, T. A., Nguyen, D. T., and Yin, G.: Milstein-Type Procedures for Numerical Solutions of Stochastic Differential Equations with Markovian Switching. *SIAM Journal on Numerical Analysis* **55** (2017), 953–979.
24. Nguyen, S. L. and Yin, G.: Pathwise Convergence Rates for Numerical Solutions of Markovian Switching Stochastic Differential Equations. *Nonlinear Analysis: Real World Applications* **13** (2012), 1170–1185.
25. Shao, J.: Invariant Measures and Euler–Maruyama's Approximations of State-Dependent Regime-Switching Diffusions. *SIAM Journal on Control and Optimization* **56** (2018), 3215–3238.

26. Yao, D. D., Zhang, Q., and Zhou, X. Y.: A Regime-Switching Model for European Options. *Stochastic Processes, Optimization, and Control Theory: Applications in Financial Engineering, Queueing Networks, and Manufacturing Systems: A Volume in Honor of Suresh Sethi*. Springer US, 2006, 281–300.
27. Yin, G., Mao, X., Yuan, C., and Cao, D.: Approximation Methods for Hybrid Diffusion Systems with State-Dependent Switching Processes: Numerical Algorithms and Existence and Uniqueness of Solutions. *SIAM Journal on Mathematical Analysis* **41** (2010), 2335–2352.
28. Yin, G., Song, Q. S., and Zhang, Z.: Numerical Solutions for Jump-Diffusions with Regime Switching. *Stochastics An International Journal of Probability and Stochastic Processes* **77** (2005), 61–79.
29. Yin, G. and Zhang, Q.: *Continuous-Time Markov Chains and Applications: a Two-Time-Scale Approach*. Vol. 37. Springer Science & Business Media, 2013.
30. Yin, G. and Zhu, C.: *Hybrid Switching Diffusions: Properties and Applications*. Vol. 63. Springer, New York, 2009.
31. Yuan, C. and Mao, X.: Convergence of The Euler-Maruyama Method for Stochastic Differential Equations with Markovian Switching. *Mathematics and Computers in Simulation* **64** (2004), 223–235.
32. Zong, X., Wu, F., and Huang, C.: The Moment Exponential Stability Criterion of Nonlinear Hybrid Stochastic Differential Equations and Its Discrete Approximations. *Proceedings of The Royal Society of Edinburgh: Section A Mathematics* **146** (2016), 1303–1328.
33. Zong, X., Wu, F., and Tian, T.: Stability and Stochastic Stabilization of Numerical Solutions of Regime-Switching Jump Diffusion Systems. *Journal of Difference Equations and Applications* **19** (2013), 1733–1757.

Dung T. Nguyen: Department of Applied Mathematics, Faculty of Applied Science, HCMC University of Technology, Vietnam
E-mail address: `dungnt@hcmut.edu.vn`

Son L. Nguyen: Department of Mathematics, University of Puerto Rico-Rio Piedras, San Juan, PR 00936
E-mail address: `sonluu.nguyen@upr.edu`