The truncated $\theta$-Milstein method for highly nonlinear and nonautonomous stochastic differential delay equations

Shuaibin Gao\textsuperscript{a}, Junhao Hu\textsuperscript{b}, Jie He\textsuperscript{a}, Qian Guo\textsuperscript{a,*}

\textsuperscript{a}Department of Mathematics, Shanghai Normal University, Shanghai 200234, China
\textsuperscript{b}School of Mathematics and Statistics, South-Central University For Nationalities, Wuhan 430074, China

\textbf{Abstract}

This paper focuses on the strong convergence of the truncated $\theta$-Milstein method for a class of nonautonomous stochastic differential delay equations whose drift and diffusion coefficients can grow polynomially. The convergence rate, which is close to one, is given under the weaker assumption than the monotone condition. To verify our theoretical findings, we present a numerical example.

\textbf{Keywords:} The truncated $\theta$-Milstein method; Stochastic differential delay equations; Strong convergence rate; Highly nonlinear

1. Introduction

The research on stochastic differential equations has been widely concerned due to their applications in many fields such as finance, biology, chemistry, and ecology [1, 4, 27, 31]. It is well known that time delay is widespread in nature and occurs in dynamics with a finite propagation time, then the corresponding stochastic differential delay equations (SDDEs) are used more widely in stochastic systems [6, 24, 26, 28, 35, 36, 44]. In many instances, the true solutions of the stochastic equations cannot be expressed explicitly. Hence, it is very meaningful to simulate the true solutions with different numerical algorithms; in this way, scholars can grasp some important properties of the true solutions without knowing the explicit form of the true solutions. One of the most famous numerical methods for a stochastic differential equation is Euler-Maruyama(EM) method, which has been investigated and developed in the past few decades [20, 27, 31]. Unfortunately, Hutzenthaler et al. have showed that the $p$th moment of the EM solutions would diverge to infinity for $p \in [1, \infty)$ when the coefficients grow super-linearly in [18]. To approximate the stochastic equations
with highly nonlinear growing coefficients, many implicit methods have been studied [3, 14, 32, 33, 37, 43]. Furthermore, some modified explicit schemes have been established as well, since they have less computational cost [5, 19, 34, 38]. Particularly, the truncated EM method was initially proposed by Mao in [29] and the convergence rate was obtained in [30]. Since then, some scholars have discussed the stochastic equations whose coefficients can grow super-linearly by using the truncated EM method, and we can refer to [3, 4, 11, 12, 22, 25] and references therein. In order to improve the convergence rate, Guo et al. [12] established the truncated Milstein method to approximate stochastic differential equations with commutative noise. Thereafter, the truncated Milstein method for non-autonomous stochastic differential equations was discussed in [23]. The convergence rate of the truncated Milstein method for SDDEs was investigated in [42]. As for other papers about Milstein methods, we refer the readers to [2, 14, 17, 21, 33, 38, 39, 45] for more detailed discussions.

The aim of this paper is to study the strong convergence rate of the truncated $\theta$-Milstein method for highly nonlinear and nonautonomous SDDEs. The main contributions of this paper are as follows.

- The SDDEs in this paper are nonautonomous, unlike many studies that focus on autonomous case. It is worth noting that the autonomy of the equation will affect the convergence rate of the numerical scheme.

- The truncated $\theta$-Milstein method for nonautonomous SDDEs is established in this paper, and it will degenerate into the truncated Milstein method when $\theta = 0$.

- The convergence rate in $L^2$ sense can be obtained with the more relaxed assumptions than those in [42], which is given in Remark 1.

- We relax the requirements for establishing the numerical scheme, which is shown in Remark 2.

- The convergence rate in $L^q(\bar{q} \geq 2)$ sense is presented in this paper, but [42] only gives the convergence rate in $L^2$ sense, which can be found in Theorems 3.7 and 3.11.

This paper is organized in the following way. Some necessary notations and the structure of the truncated $\theta$-Milstein method are shown in Section 2. Main result are presented in Section 3. Section 4 contains an example. The conclusion and discussion on future research are stated in Section 5.

2. Preliminaries

Throughout this paper, we use the following notations. Let $|x|$ denote its Euclidean norm for $x \in \mathbb{R}^n$. For real numbers $a, b$, denote $a \wedge b = \min\{a, b\}$ and $a \vee b = \max\{a, b\}$. Let $|a|$ be the largest integer which does not exceed $a$. Let the delay constant $\tau > 0$. Then denote by $\mathcal{C}([-\tau, 0]; \mathbb{R})$ the family of all continuous functions $\varphi$ from $[-\tau, 0]$ to $\mathbb{R}$ with the norm $\|\varphi\| = \sup_{-\tau \leq \theta \leq 0}|\varphi(\theta)|$. 
If $S$ is a set, denote by $I_S$ its indicator function; that is, $I_S(x) = 1$ if $x \in S$ and $I_S(x) = 0$ if $x \notin S$. Set $\mathbb{R}_+ = [0, +\infty)$. Let $C$ stand for a generic positive real constant whose value may change in different cases.

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ stand for a complete probability space with a filtration $\{\mathcal{F}_t\}_{t \geq 0}$ satisfying the usual conditions (i.e., it is increasing and right continuous while $\mathcal{F}_0$ contains all P-null sets). Denote by $\mathbb{E}$ the probability expectation with respect to $\mathbb{P}$. For $p > 0$, denote by $\mathcal{L}^p = \mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P})$ the space of random variables $X$ with $\mathbb{E}|X|^p < \infty$. Let $\mathcal{C}^0_{\mathcal{F}_0}(]-\tau, 0]; \mathbb{R})$ stand for the family of all bounded, $\mathcal{F}_0$-measurable, $\mathcal{C}(]-\tau, 0]; \mathbb{R})$-valued random variables. Let $B(t)$ be a 1-dimensional Brownian motion which is defined on this probability space.

In this paper, consider the 1-dimensional highly nonlinear and nonautonomous SDDEs of the form

$$dx(t) = f(t, x(t), x(t - \tau)) \, dt + g(t, x(t), x(t - \tau)) \, dB(t),$$

on $t \geq 0$ with the initial data

$$x_0 = \xi = \{\xi(\theta) : -\tau \leq \theta \leq 0\} \in \mathcal{C}^0_{\mathcal{F}_0}(]-\tau, 0]; \mathbb{R}),$$

where $f : \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ and $g : \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$.

In order to get the strong convergence rate of the truncated $\theta$-Milstein method for highly nonlinear and nonautonomous SDDEs (2.1), the following assumptions need to be imposed.

**Assumption 2.1.** There exist constants $\bar{K} > 0$ and $\beta \geq 0$ such that

$$|f(t, x, y) - f(t, \bar{x}, \bar{y})| \vee |g(t, x, y) - g(t, \bar{x}, \bar{y})| \leq \bar{K}(1 + |x|^{\beta} + |y|^{\beta} + |\bar{x}|^{\beta} + |\bar{y}|^{\beta})(|x - \bar{x}| + |y - \bar{y}|)$$

for any $t \in [0, T]$ and $x, y, \bar{x}, \bar{y} \in \mathbb{R}$.

Before Assumption 2.2, we need more notations. Let $\mathcal{U}$ be the family of all continuous functions $U : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}_+$ such that for any $b > 0$, there is a constant $\kappa_b > 0$, which satisfies

$$U(x, \bar{x}) \leq \kappa_b |x - \bar{x}|^2$$

for any $x, \bar{x} \in \mathbb{R}$ with $|x| \vee |\bar{x}| \leq b$.

**Assumption 2.2.** There exist constants $K_1 > 0$ and $q > 2$ such that

$$(x - \bar{x})^T (f(t, x, y) - f(t, \bar{x}, \bar{y})) + (q - 1)|g(t, x, y) - g(t, \bar{x}, \bar{y})|^2 \leq K_1(|x - \bar{x}|^2 + |y - \bar{y}|^2) - U(x, \bar{x}) + U(y, \bar{y})$$

for any $t \in [0, T]$ and $x, y, \bar{x}, \bar{y} \in \mathbb{R}$. 

3
Assumption 2.6. There exist constants $K$ for all $t, s$.

Assumption 2.5. There exist constants $K$ such that

$$
|f(t, x) - f(t, y)| \leq K(x - y)^2,
$$

for any $t \in [0, 1]$, $x, y \in \mathbb{R}$. One can see that there is no $K$ satisfying

$$
(x - x)^T (f(t, x, y) - f(t, x, y)) + (q - 1)|g(t, x, y) - g(t, x, y)|^2
\leq K_1(|x - x|^2 + |y - y|^2),
$$

but Assumption 2.2 is satisfied. The detailed proof can be found in Section 4.

Assumption 2.3. There exist constants $K_2 > 0$ and $p > q$ such that

$$
x^T f(t, x, y) + (p - 1)|g(t, x, y)|^2 \leq K_2(1 + |x|^2 + |y|^2)
$$

for any $t \in [0, T]$ and $x, y \in \mathbb{R}$.

Assumption 2.4. There exist constants $K_3 > 0$, $\beta \geq 0$ and $\sigma \in (0, 1]$ such that

$$
|f(t_1, x, y) - f(t_2, x, y)| + |g(t_1, x, y) - g(t_2, x, y)| \leq K_3(1 + |x|^{\beta + 1} + |y|^{\beta + 1})|t_1 - t_2|\sigma
$$

for any $t_1, t_2 \in [0, T]$ and $x, y \in \mathbb{R}$.

Assumption 2.5. There exist constants $K_4 > 0$ and $\gamma \in (0, 1]$ such that

$$
|\xi(t) - \xi(s)| \leq K_4|t - s|^\gamma
$$

for all $t, s \in [-\tau, 0]$.

Assumption 2.6. There exist constants $K_5 > 0$ and $\beta \geq 0$ such that

$$
|f_i(t, x, y)| + |g_i(t, x, y)| + |f_{ij}(t, x, y)| + |g_{ij}(t, x, y)| \leq K_5(1 + |x|^{\beta + 1} + |y|^{\beta + 1}),
$$

where

$$
\begin{align*}
    f_i(t, x_1, x_2) &= \frac{\partial f(t, x_1, x_2)}{\partial x_i}, & g_i(t, x_1, x_2) &= \frac{\partial g(t, x_1, x_2)}{\partial x_i}, \\
    f_{ij}(t, x_1, x_2) &= \frac{\partial^2 f(t, x_1, x_2)}{\partial x_i \partial x_j}, & g_{ij}(t, x_1, x_2) &= \frac{\partial^2 g(t, x_1, x_2)}{\partial x_i \partial x_j},
\end{align*}
$$

(2.4)

for any $t \in [0, T]$, $x, y \in \mathbb{R}$ and $i, j = 1, 2$.

The notations in (2.4) will be used in the rest of this paper. The boundedness of the moment of the true solution can be obtained with the standard method, which is stated as the following lemma.

Lemma 2.7. Let Assumptions 2.1-2.3 hold. Then SDDE (2.1) has a unique global solution $x(t)$, which satisfies

$$
\sup_{0 \leq t \leq T} E|x(t)|^p \leq C, \quad \forall T \geq 0.
$$
In order to define the truncated $\theta$-Milstein method, first choose a strictly increasing continuous function $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\varphi(r) \to \infty$ as $r \to \infty$ and
\[
\sup_{0 \leq t \leq T} \sup_{|x| \vee |y| \leq r} (|f(t, x, y)| \vee |g(t, x, y)| \vee |g_1(t, x, y)| \vee |g_2(t, x, y)|) \leq \varphi(r), \quad \forall r \geq 1.
\]

(2.5)

Let $\varphi^{-1}$ be the inverse function of $\varphi$. Thus, $\varphi^{-1}$ is a strictly increasing continuous function from $[\varphi(1), \infty)$ to $[1, \infty)$. Next, choose $K_0 \geq 1 \vee \varphi(1)$ and a strictly decreasing function $\alpha : [0, 1] \to (0, \infty)$ such that
\[
\lim_{\Delta \to 0} \alpha(\Delta) = \infty, \quad \Delta^+ \alpha(\Delta) \leq K_0, \quad \forall \Delta \in (0, 1].
\]

(2.6)

**Remark 2.** The condition (2.6) is weaker than that in [42]. Here, let $K_0 = 1$, then $\Delta \in (0, \frac{1}{\alpha(\Delta)}]$. In [43], $\Delta \in (0, \frac{1}{\alpha(\Delta)}]$. Obviously, the step size $\Delta$ in our paper has a wider range than that in [42], which will lead to less computation, so we have more advantages in algorithm than [42]. As for other details, please refer to Remark 2.4 in [43].

The truncated mapping $\pi_\Delta : \mathbb{R} \to \mathbb{R}$ for the given step size $\Delta \in (0, 1]$ is defined as
\[
\pi_\Delta(x) = (|x| \wedge \varphi^{-1}(\alpha(\Delta)))) \frac{x}{|x|},
\]

(2.7)

where set $x/|x| = 0$ when $x = 0$. One can see that $\pi_\Delta$ can map $x$ to itself for $|x| \leq \varphi^{-1}(\alpha(\Delta))$ and to $\varphi^{-1}(\alpha(\Delta))$ for $|x| > \varphi^{-1}(\alpha(\Delta))$. Then define the truncated functions
\[
f_\Delta(t, x, y) = f(t, \pi_\Delta(x), \pi_\Delta(y)), \quad g_\Delta(t, x, y) = g(t, \pi_\Delta(x), \pi_\Delta(y)),
\]
\[
g_1, \Delta(t, x, y) = g_1(t, \pi_\Delta(x), \pi_\Delta(y)), \quad g_2, \Delta(t, x, y) = g_2(t, \pi_\Delta(x), \pi_\Delta(y)).
\]

By the definition, one has that
\[
|f_\Delta(t, x, y)| \vee |g_\Delta(t, x, y)| \vee |g_1, \Delta(t, x, y)| \vee |g_2, \Delta(t, x, y)| \leq \alpha(\Delta).
\]

(2.8)

The following lemma can be proved in a similar way to [41], so we omit the proof process.

**Lemma 2.8.** By Assumption [43], for any $\Delta \in (0, 1]$, we have
\[
x^T f_\Delta(t, x, y) + (p - 1)|g_\Delta(t, x, y)|^2 \leq K_2(1 + |x|^2 + |y|^2)
\]

for any $t \in [0, T]$ and $x, y \in \mathbb{R}$.

Let us now give the definition of the truncated $\theta$-Milstein numerical scheme to approximate the true solution of (2.1). Suppose that there exist two positive integers $M$ and $M'$ such that $\Delta = \frac{1}{M} = \frac{1}{M'}$. Without losing generality, set $M < M'$. Let $t_k = k\Delta$ for $k = -M, -M + 1, \ldots, 0, \ldots, M'$. Then define

\[
5
\]
There exist two kinds of the continuous-time truncated \( \theta \)-Milstein solutions. The first one is that

\[
X_\Delta(t_k) = \xi(t_k), \quad k = -M, -M + 1, \ldots, 0;
\]

\[
X_\Delta(t_{k+1}) = X_\Delta(t_k) + \theta f_\Delta(t_k, X_\Delta(t_k), X_\Delta(t_{k+1}), X_\Delta(t_{k+1-M})) \Delta + (1 - \theta) f_\Delta(t_k, X_\Delta(t_k), X_\Delta(t_{k-M})) \Delta + g_\Delta(t_k, X_\Delta(t_k), X_\Delta(t_{k-M})) \Delta B_k
\]

\[
+ g_1 \Delta(t_k, X_\Delta(t_k), X_\Delta(t_{k-M})), g_\Delta(t_k, X_\Delta(t_k), X_\Delta(t_{k-M})) Q_1
\]

\[
+ g_2 \Delta(t_k, X_\Delta(t_k), X_\Delta(t_{k-M})), g_\Delta(t_k, X_\Delta(t_k), X_\Delta(t_{k-M}), X_\Delta(t_{k-2M})), Q_2,
\]

\[
k = 0, 1, \ldots, M - 1;
\]

\[
X_\Delta(t_{k+1}) = X_\Delta(t_k) + \theta f_\Delta(t_k, X_\Delta(t_k), X_\Delta(t_{k+1}), X_\Delta(t_{k+1-M})) \Delta + (1 - \theta) f_\Delta(t_k, X_\Delta(t_k), X_\Delta(t_{k-M})) \Delta + g_\Delta(t_k, X_\Delta(t_k), X_\Delta(t_{k-M})) \Delta B_k
\]

\[
+ g_1 \Delta(t_k, X_\Delta(t_k), X_\Delta(t_{k-M})), g_\Delta(t_k, X_\Delta(t_k), X_\Delta(t_{k-M})) Q_1
\]

\[
+ g_2 \Delta(t_k, X_\Delta(t_k), X_\Delta(t_{k-M})), g_\Delta(t_k, X_\Delta(t_k), X_\Delta(t_{k-M}), X_\Delta(t_{k-2M})), Q_2,
\]

\[
k = M, M + 1, \ldots, M' - 1.
\]

where

\[
Q_1 = \int_{t_k}^{t_{k+1}} \int_{t_k}^s dB(t) dB(s) = \frac{(\Delta B_k)^2 - \Delta}{2},
\]

\[
Q_2 = \int_{t_k}^{t_{k+1}} \int_{t_k}^s dB(t - \tau) dB(s), \quad \Delta B_k = B(t_{k+1}) - B(t_k).
\]

There is no doubt that we only need to consider the case when \( T > \tau \), which is equivalent to \( M' > M \). Define

\[
\mu(t) = \sum_{k=0}^{M'-1} t_k \|_{[t_k, t_{k+1})}(t).
\]

(2.9)

There exist two kinds of the continuous-time truncated \( \theta \)-Milstein solutions. The first one is that

\[
\bar{x}_\Delta(t) = \sum_{k=-M}^{M'-1} X_\Delta(t_k) \|_{[t_k, t_{k+1})}(t).
\]

(2.10)

The second one can be defined as

\[
x_\Delta(t) - \theta f_\Delta(t, x_\Delta(t), x_\Delta(t - \tau)) \Delta
\]

\[
= \xi(0) - \theta f_\Delta(0, \xi(0), \xi(-\tau)) \Delta + \int_0^t f_\Delta(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) ds
\]

\[
+ \int_0^t g_\Delta(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) dB(s)
\]

\[
+ \int_0^t g_1 \Delta(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) g_\Delta(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) \Delta \hat{B}(s) dB(s)
\]

\[
+ \int_0^t g_2 \Delta(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau))
\]

\[
\cdot g_\Delta(\mu(s - \tau), \bar{x}_\Delta(s - \tau), \bar{x}_\Delta(s - 2\tau)) \Delta \hat{B}(s - \tau) dB(s),
\]

(2.11)
where \( \Delta \dot{B}(t) = \sum_{k=-M}^{M-1} (B(t) - B(t_k)) I_{(t_k,t_{k+1})}(t) \).

By the monotone operator theory in [40], there exists a unique \( X_\Delta(t_k) \) for the given \( X_\Delta(t_k) \) when \( K_1 \theta \Delta < 1 \) holds. Let \( \Delta^* = 1 \wedge \frac{1}{K_1 \theta} \). In the rest of this paper, let \( \Delta \in (0, \Delta^*) \) and \( \theta \in (0, 1] \). To simplify the notations, set \( \kappa(t) = \lfloor t/\Delta \rfloor, \forall t \in [-r, T] \). Additionally, define

\[
Z_\Delta(t) = x_\Delta(t) - \theta f_\Delta(t, x_\Delta(t), x_\Delta(t - \tau)) \Delta, \quad (2.12)
\]

\[
\dot{Z}_\Delta(t) = \dot{x}_\Delta(t) - \theta f_\Delta(\mu(t), \dot{x}_\Delta(t), \ddot{x}_\Delta(t - \tau)) \Delta. \quad (2.13)
\]

The following Taylor expansion would play an important role in our proof. For more details, please refer to [2].

If \( \psi : \mathbb{R}^3 \to \mathbb{R} \) is a third-order continuously differentiable function, one can see that

\[
\psi(\bar{u}) - \psi(\hat{u}) = \psi'(u) \big|_{u=\bar{u}} (\bar{u} - \hat{u}) + R_\psi(\bar{u}, \hat{u}),
\]

where

\[
R_\psi(\bar{u}, \hat{u}) = \int_{0}^{1} (1 - \vartheta) \psi''(u) \big|_{u=\bar{u} + \vartheta (\hat{u} - \bar{u})} (\bar{u} - \hat{u}) \, d\vartheta,
\]

for any \( u, \bar{u}, \hat{u} \in \mathbb{R}^3 \). Here, \( \psi' \) and \( \psi'' \) are defined as

\[
\psi'(u)(j) = \sum_{i=1}^{3} \frac{\partial \psi}{\partial u_i} j_i, \quad \psi''(u)(j, h) = \sum_{i, m=1}^{3} \frac{\partial^2 \psi}{\partial u_i \partial u_m} j_i h_m.
\]

where \( j = (j_1, j_2, j_3) \), \( h = (h_1, h_2, h_3) \).

Next, set \( \bar{u} = (\eta, \bar{x}, \bar{y}) \) and \( \hat{u} = (\bar{\eta}, \bar{x}, \bar{y}) \) for \( \bar{x}, \bar{y}, \hat{x}, \hat{y} \in \mathbb{R}, \eta \in \mathbb{R}^+ \). Thus, \( \bar{u} - \hat{u} = (0, \bar{x} - \hat{x}, \bar{y} - \hat{y}) \). Then, for \( \psi : \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R} \to \mathbb{R} \), one can see that

\[
\psi(\eta, \bar{x}, \bar{y}) - \psi(\eta, \hat{x}, \hat{y}) = \psi'(\eta, x, y) \big|_{x=\bar{x}, y=\bar{y}} (0, \bar{x} - \hat{x}, \bar{y} - \hat{y}) + R_\psi(\eta, \bar{x}, \bar{y}, \hat{x}, \hat{y}),
\]

where

\[
R_\psi(\eta, \bar{x}, \bar{y}, \hat{x}, \hat{y}) = \int_{0}^{1} (1 - \vartheta) \psi''(\eta, x, y) \big|_{x=\bar{x} + \vartheta(\bar{x} - \hat{x}), y=\bar{y} + \vartheta(\bar{y} - \hat{y})} ((0, \bar{x} - \hat{x}, \bar{y} - \hat{y}), (0, \bar{x} - \hat{x}, \bar{y} - \hat{y})) \, d\vartheta,
\]

for any \( \eta \in \mathbb{R}^+ \) and \( \bar{x}, \bar{y}, \hat{x}, \hat{y} \in \mathbb{R} \). Here, \( \psi' \) and \( \psi'' \) are defined as

\[
\psi'(\eta, x, y)(0, \bar{x} - \hat{x}, \bar{y} - \hat{y}) = \frac{\partial \psi}{\partial x}(\bar{x} - \hat{x}) + \frac{\partial \psi}{\partial y}(\bar{y} - \hat{y}),
\]

\[
\psi''(\eta, x, y)((0, \bar{x} - \hat{x}, \bar{y} - \hat{y}), (0, \bar{x} - \hat{x}, \bar{y} - \hat{y})) \frac{\partial^2 \psi}{\partial x \partial y}(\bar{x} - \hat{x})(\bar{y} - \hat{y}) + \frac{\partial^2 \psi}{\partial y \partial y}((\bar{y} - \hat{y}))^2 + \frac{\partial^2 \psi}{\partial y \partial x}(\bar{y} - \hat{y})(\bar{x} - \hat{x}).
\]
for $\eta \in \mathbb{R}_+$ and $\bar{x}, \bar{y}, \hat{x}, \hat{y} \in \mathbb{R}$. By setting $\eta = \mu(t), \bar{x} = x_\Delta(t), \bar{y} = x_\Delta(t-\tau), \hat{x} = \bar{x}_\Delta(t), \hat{y} = \bar{x}_\Delta(t-\tau)$, we get from (2.16) that
\[
\psi_\mu(t, x_\Delta(t), x_\Delta(t-\tau)) - \psi_\mu(t, \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau)) = \psi_1(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau)) (x_\Delta(t) - \bar{x}_\Delta(t)) \\
+ \psi_2(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau)) (x_\Delta(t-\tau) - \bar{x}_\Delta(t-\tau)) \\
+ R_{\psi}(t, x_\Delta(t), \bar{x}_\Delta(t), x_\Delta(t-\tau), \bar{x}_\Delta(t-\tau)) \\
= \psi_1(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau)) \int_{\kappa(t)}^{t} g_{\Delta}(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s-\tau)) dB(s) \\
+ \psi_2(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau)) \int_{\kappa(t)-\tau}^{t-\tau} g_{\Delta}(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s-\tau)) dB(s) \\
+ \hat{R}_{\psi}(t, x_\Delta(t), \bar{x}_\Delta(t), x_\Delta(t-\tau), \bar{x}_\Delta(t-\tau)),
\]
where
\[
\hat{R}_{\psi}(t, x_\Delta(t), \bar{x}_\Delta(t), x_\Delta(t-\tau), \bar{x}_\Delta(t-\tau)) = \psi_1(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau)) \\
\cdot \left[ \theta f_{\Delta}(t, x_\Delta(t), x_\Delta(t-\tau)) \Delta f_{\Delta}(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau)) \Delta \\
+ \int_{\kappa(t)}^{t} f_{\Delta}(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s-\tau)) ds \\
+ \int_{\kappa(t)}^{t} g_{1,\Delta}(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s-\tau)) g_{\Delta}(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s-\tau)) \Delta \hat{B}(s) dB(s) \\
+ \int_{\kappa(t)}^{t} g_{2,\Delta}(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s-\tau)) \Delta \hat{B}(s) dB(s) \right] \\
+ \psi_2(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau)) \\
\cdot \left[ \theta f_{\Delta}(t-\tau, x_\Delta(t-\tau), x_\Delta(t-2\tau)) \Delta f_{\Delta}(\mu(t-\tau), \bar{x}_\Delta(t-\tau), \bar{x}_\Delta(t-2\tau)) \Delta \\
+ \int_{\kappa(t)-\tau}^{t-\tau} f_{\Delta}(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s-\tau)) ds \\
+ \int_{\kappa(t)-\tau}^{t-\tau} g_{1,\Delta}(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s-\tau)) g_{\Delta}(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s-\tau)) \Delta \hat{B}(s) dB(s) \\
+ \int_{\kappa(t)-\tau}^{t-\tau} g_{2,\Delta}(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s-\tau)) \Delta \hat{B}(s) dB(s) \right] \\
+ R_{\psi}(\mu(t), x_\Delta(t), \bar{x}_\Delta(t), x_\Delta(t-\tau), \bar{x}_\Delta(t-\tau)),
\]
(2.19)
Proof. By (2.11), (2.12) and (2.13), we derive that

Thus,

Lemma 3.1. For any $\Delta \in (0, \Delta^*)$ and $t \in [0, T]$, we have

$$E\left|Z_{\Delta}(t) - \bar{Z}_{\Delta}(t)\right|^\bar{p} \leq C\Delta^\bar{p}\alpha(\Delta)^\bar{p}, \quad \forall \bar{p} > 0,$$

and

$$E\left|x_{\Delta}(t) - \bar{x}_{\Delta}(t)\right|^\bar{p} \leq C\Delta^\bar{p}\alpha(\Delta)^\bar{p}, \quad \forall \bar{p} > 0.$$  

Thus,

$$\lim_{\Delta \to 0} E\left|Z_{\Delta}(t) - \bar{Z}_{\Delta}(t)\right|^\bar{p} = \lim_{\Delta \to 0} E\left|x_{\Delta}(t) - \bar{x}_{\Delta}(t)\right|^\bar{p} = 0, \quad \forall \bar{p} > 0.$$

Proof. By (2.11), (2.12) and (2.13), we derive that

$$Z_{\Delta}(t) = \bar{Z}_{\Delta}(t) + \int_{\kappa(t)}^t f_{\Delta}(\mu(s), \bar{x}_{\Delta}(s), \bar{x}_{\Delta}(s - \tau))ds$$

$$+ \int_{\kappa(t)}^t g_{\Delta}(\mu(s), \bar{x}_{\Delta}(s), \bar{x}_{\Delta}(s - \tau))dB(s)$$

$$+ \int_{\kappa(t)}^t g_{1,\Delta}(\mu(s), \bar{x}_{\Delta}(s), \bar{x}_{\Delta}(s - \tau))g_{\Delta}(\mu(s), \bar{x}_{\Delta}(s), \bar{x}_{\Delta}(s - \tau))\Delta\hat{B}(s)dB(s)$$

$$+ \int_{\kappa(t)}^t g_{2,\Delta}(\mu(s), \bar{x}_{\Delta}(s), \bar{x}_{\Delta}(s - \tau))$$

$$\cdot g_{\Delta}(\mu(s - \tau), \bar{x}_{\Delta}(s - \tau), \bar{x}_{\Delta}(s - 2\tau))\Delta\hat{B}(s - \tau)dB(s).$$

(3.3)

For any fixed $\bar{p} \geq 2$ and $t \in [0, T]$, we can get from the Hölder inequality and the Burkholder-Davis-Gundy inequality that

$$E\left|Z_{\Delta}(t) - \bar{Z}_{\Delta}(t)\right|^\bar{p} \leq C\left[\Delta^{\bar{p}-1}E\int_{\kappa(t)}^t |f_{\Delta}(\mu(s), \bar{x}_{\Delta}(s), \bar{x}_{\Delta}(s - \tau))|^\bar{p}ds$$

$$+ \Delta^{\bar{p}-1}E\int_{\kappa(t)}^t |g_{\Delta}(\mu(s), \bar{x}_{\Delta}(s), \bar{x}_{\Delta}(s - \tau))|^\bar{p}ds$$

$$+ \Delta^{\bar{p}-1}E\int_{\kappa(t)}^t |g_{1,\Delta}(\mu(s), \bar{x}_{\Delta}(s), \bar{x}_{\Delta}(s - \tau))g_{\Delta}(\mu(s), \bar{x}_{\Delta}(s), \bar{x}_{\Delta}(s - \tau))\Delta\hat{B}(s)|^\bar{p}ds$$

$$+ \Delta^{\bar{p}-1}E\int_{\kappa(t)}^t |g_{2,\Delta}(\mu(s), \bar{x}_{\Delta}(s), \bar{x}_{\Delta}(s - \tau))\cdot g_{\Delta}(\mu(s - \tau), \bar{x}_{\Delta}(s - \tau), \bar{x}_{\Delta}(s - 2\tau))\Delta\hat{B}(s - \tau)|^\bar{p}ds$$

9
By Lemma 3.1, we derive that
\[ g_2(\mu(s), \bar{x}_\Delta(s)) \Delta \hat{B}(s) \]
\[ \leq C \left( \Delta^{\frac{\alpha(\Delta)}{\nu}} + \Delta^{\frac{\alpha(\Delta)}{\nu}} + \Delta^{\frac{\alpha(\Delta)}{\nu}} \right) \leq C \Delta^{\frac{\alpha(\Delta)}{\nu}}, \]
where \( \Delta^{\frac{\alpha(\Delta)}{\nu}} \leq K_0^p \Delta^p \) is used. For \( 0 < \bar{p} < 2 \), applying H"older’s inequality can give the desired result \( \text{(3.1)} \). Then combining \( \text{(2.12)}, \text{(2.13)} \) and \( \text{(3.1)} \) yields \( \text{(3.2)} \). The proof is complete.

Lemma 3.2. Let Assumptions \( \text{(2.1)} \) and \( \text{(2.3)} \) hold. Then we derive that
\[ \sup_{0 < \Delta < \Delta^*} \sup_{0 \leq t \leq T} \mathbb{E}|x_\Delta(t)|^p \leq C. \]

Proof. By Itô’s formula and \( \text{(3.3)} \), we get that
\[ \mathbb{E}|Z_\Delta(t)|^p - |Z_\Delta(0)|^p \]
\[ \leq \mathbb{E} \int_0^t p|Z_\Delta(s)|^{p-2} (Z_\Delta(s)^T f_\Delta(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) + \frac{p-1}{2} |g_\Delta(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau))\]
\[ + g_1(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) g_\Delta(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) \Delta \hat{B}(s) \]
\[ + g_2(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) g_\Delta(\mu(s - \tau), \bar{x}_\Delta(s - \tau), \bar{x}_\Delta(s - 2\tau)) \Delta \hat{B}(s - \tau) |^2 \] \[ \leq \mathbb{E} \int_0^t p|Z_\Delta(s)|^{p-2} (Z_\Delta(s) - \bar{x}_\Delta(s))^T f_\Delta(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) ds \]
\[ + \mathbb{E} \int_0^t p|Z_\Delta(s)|^{p-2} (\bar{x}_\Delta(s)^T f_\Delta(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) + (p - 1)p|g_\Delta(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau))|^2 ds \]
\[ + \mathbb{E} \int_0^t p(p - 1)|Z_\Delta(s)|^{p-2} g_1(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) g_\Delta(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) \Delta \hat{B}(s) \]
\[ + g_2(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) g_\Delta(\mu(s - \tau), \bar{x}_\Delta(s - \tau), \bar{x}_\Delta(s - 2\tau)) \Delta \hat{B}(s - \tau) |^2 \] \[ =: I_1 + I_2 + I_3. \]

By Lemma 3.1, we derive that
\[ I_1 \leq (p - 2) \mathbb{E} \int_0^t |Z_\Delta(s)|^p ds + 2 \mathbb{E} \int_0^t |(Z_\Delta(s) - \bar{x}_\Delta(s))^T f_\Delta(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) |^\frac{p}{2} ds \]
\[ \leq C \mathbb{E} \int_0^t |Z_\Delta(s)|^p ds + C \mathbb{E} \int_0^t |(Z_\Delta(s) - \bar{Z}_\Delta(s))^T f_\Delta(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) |^\frac{p}{2} ds \]
\[ + \mathbb{E} \int_0^t |g_\Delta(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) \Delta \hat{B}(s) |^2 ds \]
\[ \leq C \mathbb{E} \int_0^t |Z_\Delta(s)|^p ds + C \mathbb{E} \int_0^t \Delta^p + C \mathbb{E} \int_0^t \Delta^p \alpha(\Delta) p ds \]
\[ \leq C + C \mathbb{E} \int_0^t |Z_\Delta(s)|^p ds. \]

(3.4)
By Lemma 2.8, we get that
\[
I_2 \leq C E \int_0^t |Z_\Delta(s)|^{p-2} (1 + |\bar{x}_\Delta(s)|^2 + |\bar{x}_\Delta(s - \tau)|^2) ds
\]
\[
\leq C E \int_0^t |Z_\Delta(s)|^p ds + C E \int_0^t (1 + |\bar{x}_\Delta(s)|^p + |\bar{x}_\Delta(s - \tau)|^p) ds.
\] (3.5)

Using (2.6) and (2.8) yields that
\[
I_3 \leq C E \int_0^t |Z_\Delta(s)|^p ds + C E \int_0^t \Delta \hat{\psi} \alpha(\Delta)^2 \hat{B}(s) ds
\]
\[
\leq C + C E \int_0^t |Z_\Delta(s)|^p ds.
\] (3.6)

Thus, from (3.5) - (3.6), one can see that
\[
E|Z_\Delta(t)|^p \leq C + C \int_0^t E|Z_\Delta(s)|^p ds + C E \int_0^t (|\bar{x}_\Delta(s)|^p + |\bar{x}_\Delta(s - \tau)|^p) ds.
\] (3.7)

Applying Gronwall's inequality gives that
\[
E|Z_\Delta(t)|^p \leq C + C E \int_0^t (|\bar{x}_\Delta(s)|^p + |\bar{x}_\Delta(s - \tau)|^p) ds.
\]

By (2.12) and the elementary inequality $|a - b|^p \geq 2^{1-p}|a|^p - |b|^p, a, b > 0$, we have
\[
|Z_\Delta(t)|^p \geq 2^{1-p}|x_\Delta(t)|^p - \theta^p \Delta^p |f_\Delta(t, x_\Delta(t), x_\Delta(t - \tau))|^p
\]
\[
\geq 2^{1-p}|x_\Delta(t)|^p - \theta^p \Delta^p \alpha(\Delta)^p.
\]

Thus,
\[
\sup_{0 \leq r \leq t} E|x_\Delta(r)|^p \leq C \left( 1 + \int_0^t \sup_{0 \leq r \leq s} E|Z_\Delta(r)|^p ds \right) \leq C \left( 1 + \int_0^t \sup_{0 \leq r \leq s} E|x_\Delta(r)|^p ds \right).
\]

Using Gronwall's inequality again yields that
\[
\sup_{0 < \Delta \leq \Delta^*} \sup_{0 \leq t \leq T} E|x_\Delta(t)|^p \leq C.
\]

We complete the proof. □
The following lemma could be obtained by using Lemma 3.2.

**Lemma 3.3.** Let Assumptions E.1, E.3, and E.6 hold. Assume that $p \geq 2(1+\beta)q$ for $q > 2$, then for any $\tilde{q} \in [2, q)$, we get that

$$\sup_{0 \leq \Delta \leq \Delta^*} \sup_{0 \leq t \leq T} \left( \mathbb{E} [f(t, x_\Delta(t), x_\Delta(t-\tau))]^{2\tilde{q}} \vee \mathbb{E} [g(t, x_\Delta(t), x_\Delta(t-\tau))]^{2\tilde{q}} \right) \leq C.$$

**Lemma 3.4.** Let Assumptions E.1, E.3, and E.6 hold. Assume that $p \geq 2(1+\beta)q$ for $q > 2$. Then for any $\tilde{q} \in [2, q)$, $\Delta \in (0, \Delta^*)$ and $t \in [0, T]$, we have

$$\mathbb{E} |R_f(t, x_\Delta(t), \bar{x}_\Delta(t), x_\Delta(t-\tau), \bar{x}_\Delta(t-\tau))|^{\tilde{q}} \leq C \Delta^{\tilde{q}} \alpha(\Delta)^{4\tilde{q}}.$$

**Proof.** For any $\tilde{q} \in [2, q)$, $\Delta \in (0, \Delta^*)$ and $t \in [0, T]$, we get from Hölder’s inequality, Lemma 3.2 and (2.17) that

$$\mathbb{E} |R_f(t, x_\Delta(t), \bar{x}_\Delta(t), x_\Delta(t-\tau), \bar{x}_\Delta(t-\tau))|^{\tilde{q}} \leq \int_0^1 \left[ \mathbb{E} |f_{11}(\mu(t), \bar{x}_\Delta(t) + \vartheta(x_\Delta(t) - \bar{x}_\Delta(t)), x_\Delta(t-\tau) + \vartheta(x_\Delta(t-\tau) - \bar{x}_\Delta(t-\tau))]^{2\tilde{q}} \right. \left. \cdot \mathbb{E} |x_\Delta(t) - \bar{x}_\Delta(t)|^{4\tilde{q}} \right]^\frac{1}{2} d\vartheta$$

$$+ \int_0^1 \left[ \mathbb{E} |f_{12}(\mu(t), \bar{x}_\Delta(t) + \vartheta(x_\Delta(t) - \bar{x}_\Delta(t)), x_\Delta(t-\tau) + \vartheta(x_\Delta(t-\tau) - \bar{x}_\Delta(t-\tau))]^{2\tilde{q}} \right. \left. \cdot \mathbb{E} |x_\Delta(t) - \bar{x}_\Delta(t)|^{4\tilde{q}} \right]^\frac{1}{2} d\vartheta$$

$$+ \int_0^1 \left[ \mathbb{E} |f_{21}(\mu(t), \bar{x}_\Delta(t) + \vartheta(x_\Delta(t) - \bar{x}_\Delta(t)), x_\Delta(t-\tau) + \vartheta(x_\Delta(t-\tau) - \bar{x}_\Delta(t-\tau))]^{2\tilde{q}} \right. \left. \cdot \mathbb{E} |x_\Delta(t) - \bar{x}_\Delta(t)|^{4\tilde{q}} \right]^\frac{1}{2} d\vartheta$$

$$+ \int_0^1 \left[ \mathbb{E} |f_{22}(\mu(t), \bar{x}_\Delta(t) + \vartheta(x_\Delta(t) - \bar{x}_\Delta(t)), x_\Delta(t-\tau) + \vartheta(x_\Delta(t-\tau) - \bar{x}_\Delta(t-\tau))]^{2\tilde{q}} \right. \left. \cdot \mathbb{E} |x_\Delta(t-\tau) - \bar{x}_\Delta(t-\tau)|^{4\tilde{q}} \right]^\frac{1}{2} d\vartheta$$

$$\leq C \left( 1 + \sup_{0 \leq \Delta \leq \Delta^*} \sup_{0 \leq t \leq T} \mathbb{E} |x_\Delta(t)|^{2\tilde{q}(1+\beta)} \right)^\frac{1}{2} \Delta^{\frac{\tilde{q}}{2}} \alpha(\Delta)^{4\tilde{q}} \right)^\frac{1}{2} \leq C \Delta^{\tilde{q}} \alpha(\Delta)^{2\tilde{q}}.$$
Then, by (2.19), we have

\[ E[\tilde{R}_f(t, x_\Delta(t), \bar{x}_\Delta(t), x_\Delta(t-\tau), \bar{x}_\Delta(t-\tau))]^{\bar{q}} = C\Delta^{q}\mathbb{E}[f_1(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))f_\Delta(t, x_\Delta(t), x_\Delta(t-\tau))]^{\bar{q}} + C\Delta^{q}\mathbb{E}[f_1(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))f_\Delta(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))]^{\bar{q}} + CE[f_1(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))g_1(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))]
\]

\[ \cdot g_\Delta(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))(\Delta\dot{B}(t))^{2q} - \Delta]\]^{\bar{q}} + CE[f_1(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))g_2(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))]
\]

\[ \cdot g_\Delta(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))(\Delta\dot{B}(t))^{2q} - \Delta\]^{\bar{q}} + CE[R_f(t, x_\Delta(t), \bar{x}_\Delta(t), x_\Delta(t-\tau), \bar{x}_\Delta(t-\tau))]^{\bar{q}}.

Using Hölder’s inequality and Lemma [33] gives the following estimates

\[ E[(\Delta\dot{B}(t))^{2q} - \Delta]^{2q} \leq C\Delta^{2q}, \]

\[ \mathbb{E}[f_1(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))f_\Delta(t, x_\Delta(t), x_\Delta(t-\tau))]^{\bar{q}} \]

\[ \vee \mathbb{E}[f_1(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))f_\Delta(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))]^{\bar{q}} \]

\[ \vee \mathbb{E}[f_2(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))f_\Delta(t-\tau, x_\Delta(t-\tau), x_\Delta(t-2\tau))]^{\bar{q}} \]

\[ \vee \mathbb{E}[f_2(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))g_1(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))]^{\bar{q}} \]

\[ \leq C\alpha(\Delta)^{\bar{q}}, \]

\[ \mathbb{E}[f_1(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))g_1(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))g_\Delta(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))]^{2q} \]

\[ \vee \mathbb{E}[f_1(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))g_2(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))g_\Delta(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-2\tau))]^{2q} \]

\[ \vee \mathbb{E}[f_2(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))g_1(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))g_\Delta(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))]^{2q} \]

\[ \vee \mathbb{E}[f_2(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))g_2(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-\tau))g_\Delta(\mu(t), \bar{x}_\Delta(t), \bar{x}_\Delta(t-2\tau))]^{2q} \]

\[ \leq C\alpha(\Delta)^{4\bar{q}}. \]

Combining these inequalities together with Hölder’s inequality gives that

\[ E[\tilde{R}_f(t, x_\Delta(t), \bar{x}_\Delta(t), x_\Delta(t-\tau), \bar{x}_\Delta(t-\tau))]^{\bar{q}} \leq C\Delta^{q}\alpha(\Delta)^{2q}. \]

Similarly, the other results could be obtained. The proof is complete. \qed
The following lemma can be proved by borrowing the proof techniques in Lemma 2.7 and Lemma 3.2.

Lemma 3.5. Let Assumptions 2.1-2.6 hold. For any real number \( L > ||\xi|| \), define the stopping time

\[
\lambda_L = \inf \{ t \geq 0 : |x(t)| \geq L \}, \quad \lambda_{\Delta, L} = \inf \{ t \geq 0 : |x_\Delta(t)| \geq L \}.
\]

Then we derive that

\[
\mathbb{P}(\lambda_L \leq T) \leq \frac{C}{L^p}, \quad \mathbb{P}(\lambda_{\Delta, L} \leq T) \leq \frac{C}{L^p}.
\]

Lemma 3.6. Let Assumptions 2.1-2.6 hold with \( p \geq 2(1+\beta)q \). Let \( L > ||\xi|| \) be a real number, and let \( \Delta \in (0, \Delta^*) \) be sufficiently small such that \( \varphi^{-1}(\alpha(\Delta)) \geq L \). Then we have

\[
\mathbb{E}|x(T \land \rho_{\Delta, L}) - x_\Delta(T \land \rho_{\Delta, L})|^2 \leq C(\Delta^2 \alpha(\Delta)^4 + \Delta^{2\gamma} + \Delta^{2\sigma}),
\]

where \( \rho_{\Delta, L} := \lambda_L \land \lambda_{\Delta, L} \).

Proof. We write \( \rho_{\Delta, L} = \rho \) for simplicity. Denote \( e_\Delta(t) = x(t) - Z_\Delta(t) \). For \( 0 \leq s \leq t \land \rho \), one can see that

\[
|x(s)| \lor |x(s - \tau)| \lor |\bar{x}_\Delta(s)| \lor |\bar{x}_\Delta(s - \tau)| \leq L \leq \varphi^{-1}(\alpha(\Delta)).
\]

In addition, for \( \tau \leq s \leq t \land \rho \), \( |\bar{x}_\Delta(s - 2\tau)| \leq L \leq \varphi^{-1}(\alpha(\Delta)) \). Then from the previous definitions, we get that

\[
\begin{align*}
f_\Delta(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) &= f(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)), \\
g_1(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) &= g(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)), \\
g_2(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) &= g(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)), \\
g_1(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) &= g(\mu(s - \tau), \bar{x}_\Delta(s - \tau), \bar{x}_\Delta(s - 2\tau)).
\end{align*}
\]

Applying Itô’s formula and (2.18) yields that

\[
\begin{align*}
&\mathbb{E}|e_\Delta(t \land \rho)|^2 \\
\leq &\theta^2 |f_\Delta(0, \xi(0), \xi(-\tau))|^2 \Delta^2 \\
&+ \frac{1}{2} \mathbb{E} \int_{0}^{t \land \rho} 2\left[ c_\Delta^T(s) \left( f(s, x(s), x(s - \tau)) - f(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) \right) \\
&+ \frac{1}{2} \left( g(s, x(s), x(s - \tau)) - g(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) \right) \right] ds
\end{align*}
\]
Using Young’s inequality gives that

\[
\begin{align*}
&\leq C\Delta^2 \alpha(\Delta)^2 \\
&\quad + C E \int_0^{t_\rho} \left[ (x(s) - x_\Delta(s))^T (f(s, x(s), x(s - \tau)) - f(\mu(s), x_\Delta(s), x_\Delta(s - \tau))) \\
&\quad + |g(s, x(s), x(s - \tau)) - g(\mu(s), x_\Delta(s), x_\Delta(s - \tau))|^2 \right] \, ds \\
&\quad + C E \int_0^{t_\rho} |\hat{R}_g(s, x_\Delta(s), x_\Delta(s - \tau), x_\Delta(s - \tau))|^2 \, ds \\
&\quad + C E \int_0^{t_\rho} \theta \Delta f_{\Delta}^T (s, x_\Delta(s), x_\Delta(s - \tau) \right) \\
&\quad \cdot (f(s, x(s), x(s - \tau)) - f(\mu(s), x_\Delta(s), x_\Delta(s - \tau))) \, ds \\
=: &C\Delta^2 \alpha(\Delta)^2 + J_1 + J_2 + J_3.
\end{align*}
\]

Using Young’s inequality gives that

\[
J_1 \leq C E \int_0^{t_\rho} \left[ (x(s) - x_\Delta(s))^T (f(s, x(s), x(s - \tau)) - f(s, x_\Delta(s), x_\Delta(s - \tau))) \\
\quad + (q - 1)|g(s, x(s), x(s - \tau)) - g(s, x_\Delta(s), x_\Delta(s - \tau))|^2 \right] \, ds \\
\quad + C E \int_0^{t_\rho} (x(s) - x_\Delta(s))^T (f(s, x_\Delta(s), x_\Delta(s - \tau)) - f(\mu(s), x_\Delta(s), x_\Delta(s - \tau))) \, ds \\
\quad + C E \int_0^{t_\rho} (x(s) - x_\Delta(s))^T (f(\mu(s), x_\Delta(s), x_\Delta(s - \tau)) - f(\mu(s), x_\Delta(s), x_\Delta(s - \tau))) \, ds \\
\quad + C E \int_0^{t_\rho} \frac{q - 1}{q - 2} |g(s, x_\Delta(s), x_\Delta(s - \tau)) - g(\mu(s), x_\Delta(s), x_\Delta(s - \tau))|^2 \, ds \\
=: &J_{11} + J_{12} + J_{13} + J_{14}.
\]

By Assumption 2.2 we have

\[
\begin{align*}
J_{11} &\leq C E \int_0^{t_\rho} \left[ |x(s) - x_\Delta(s)|^2 + |x(s - \tau) - x_\Delta(s - \tau)|^2 \\
&\quad - U(x(s), x_\Delta(s)) + U(x(s - \tau), x_\Delta(s - \tau)) \right] \, ds \\
&\leq C E \int_0^{t_\rho} |x(s) - x_\Delta(s)|^2 + C E \int_{-\tau}^0 |\xi(s) - \xi(\kappa(s))|^2 \, ds \\
&\quad + C E \int_0^{t_\rho} (-U(x(s), x_\Delta(s)) + U(x(s - \tau), x_\Delta(s - \tau))) \, ds \\
&\leq C E \int_0^{t_\rho} |x(s) - x_\Delta(s)|^2 \, ds + C \Delta^2 \gamma + C \int_{-\tau}^0 U(\xi(s), \xi(\kappa(s))) \, ds \\
&\leq C E \int_0^{t_\rho} |x(s) - x_\Delta(s)|^2 \, ds + C \Delta^2 \gamma + C \int_{-\tau}^0 \kappa(s) - \xi(\kappa(s))|^2 \, ds \\
&\leq C \int_0^t E |x(s \wedge \rho) - x_\Delta(s \wedge \rho)|^2 \, ds + C \Delta^2 \gamma.
\end{align*}
\]
By Assumption 2.4, one can see that

\[ J_{12} \leq C \int_0^t \mathbb{E}|x(s \wedge \rho) - x_\Delta(s \wedge \rho)|^2 ds \]
\[ + C\mathbb{E} \int_0^{t \wedge \rho} (1 + |x_\Delta(s)|^{2(\beta+1)} + |x_\Delta(s - \tau)|^{2(\beta+1)})|s - \kappa(s)|^{2\sigma} ds \]
\[ \leq C \int_0^t \mathbb{E}|x(s \wedge \rho) - x_\Delta(s \wedge \rho)|^2 ds + C\Delta^{2\sigma}. \]

Similar to [42], we have

\[ J_{13} \leq C\mathbb{E} \int_0^{t \wedge \rho} (x(s) - x_\Delta(s))^T \]
\[ \cdot [f_1(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) \int_{\mu(s)}^s g_\Delta(\mu(r), \bar{x}_\Delta(r), \bar{x}_\Delta(r - \tau)) dB(r) \]
\[ + f_2(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)) \int_{\mu(s) - \tau}^{s - \tau} g_\Delta(\mu(r), \bar{x}_\Delta(r), \bar{x}_\Delta(r - \tau)) dB(r) \]
\[ + \bar{R}_f(s, x_\Delta(s), \bar{x}_\Delta(s), x_\Delta(s - \tau), \bar{x}_\Delta(s - \tau)) ds \]
\[ \leq C\mathbb{E} \int_0^{t \wedge \rho} (|x(s) - x_\Delta(s)|^2 + |\bar{R}_f(s, x_\Delta(s), \bar{x}_\Delta(s), x_\Delta(s - \tau), \bar{x}_\Delta(s - \tau))|^2) ds \]
\[ + C\Delta^2 \alpha(\Delta)^4 + C\Delta^{2\gamma} \]
\[ \leq C \int_0^t \mathbb{E}|x(s \wedge \rho) - x_\Delta(s \wedge \rho)|^2 ds + C\Delta^2 \alpha(\Delta)^4 + C\Delta^{2\gamma}. \]

Borrowing the technique in the estimation of \( J_{12} \) gives that

\[ J_{14} \leq C\Delta^{2\sigma}. \]

Thus,

\[ J_1 \leq C \int_0^t \mathbb{E}|x(s \wedge \rho) - x_\Delta(s \wedge \rho)|^2 ds + C\Delta^{2\gamma} + C\Delta^{2\sigma} + C\Delta^2 \alpha(\Delta)^4. \]  \hspace{1cm} (3.9)

By Lemma 3.3, we derive that

\[ J_2 \leq C\Delta^2 \alpha(\Delta)^4. \]  \hspace{1cm} (3.10)

Similar to \( J_1 \), using Young’s inequality gives that

\[ J_3 \leq C \int_0^t \mathbb{E}|x(s \wedge \rho) - x_\Delta(s \wedge \rho)|^2 ds + C\Delta^{2\gamma} + C\Delta^{2\sigma} + C\Delta^2 \alpha(\Delta)^4. \]  \hspace{1cm} (3.11)

Combining (3.9) - (3.10) together yields that

\[ \mathbb{E}|\epsilon_\Delta(t \wedge \rho)|^2 \leq C \int_0^t \mathbb{E}|x(s \wedge \rho) - x_\Delta(s \wedge \rho)|^2 ds + C\Delta^{2\gamma} + C\Delta^{2\sigma} + C\Delta^2 \alpha(\Delta)^4. \]
Thus,
\[
\mathbb{E}|x(t \wedge \rho) - x_\Delta(t \wedge \rho)|^2 \\
\leq C (\Delta^2 \alpha(\Delta)^4 + \mathbb{E}|e_\Delta(t \wedge \rho)|^2) \\
\leq C \left( \int_0^t \mathbb{E}|x(s \wedge \rho) - x_\Delta(s \wedge \rho)|^2 ds + \Delta^2 \gamma + \Delta^2 \alpha(\Delta)^4 \right).
\]

Thanks to the Gronwall inequality, the desired result follows. We complete the proof. \(\square\)

**Theorem 3.7.** Let Assumptions 2.1-2.6 hold with \(p \geq 2(1 + \beta)q\). For any sufficiently small \(\Delta \in (0, \Delta^*)\), assume that
\[
\alpha(\Delta) \geq \varphi \left( (\Delta^2 \alpha(\Delta)^4 \lor \Delta^2(\gamma \lor \sigma))^{p-2} \right). \tag{3.12}
\]
Then for any such small \(\Delta\), we derive that
\[
\mathbb{E}|x(T) - x_\Delta(T)|^2 \leq C(\Delta^2 \alpha(\Delta)^4 \lor \Delta^2(\gamma \lor \sigma)), \tag{3.13}
\]
and
\[
\mathbb{E}|x(T) - \bar{x}_\Delta(T)|^2 \leq C(\Delta^2 \alpha(\Delta)^4 \lor \Delta^2(\gamma \lor \sigma)). \tag{3.14}
\]

**Proof.** Let \(\hat{\Delta}(t) = x(t) - x_\Delta(t)\) for \(t \in [0, T]\) and \(\Delta \in (0, \Delta^*)\). Denote \(\rho = \rho_{\Delta, L}\) for simplicity. One can see that
\[
\mathbb{E}|\hat{\Delta}(T)|^2 = \mathbb{E}\left( |\hat{\Delta}(T)|^2 I_{(\rho > T)} \right) + \mathbb{E}\left( |\hat{\Delta}(T)|^2 I_{(\rho \leq T)} \right).
\]

Let \(\delta > 0\) be arbitrary. By Young’s inequality, we derive that
\[
a^2 b = (\delta a^p)^\frac{p}{p-2} (\frac{b^{p/(p-2)}}{p-2}) \leq \frac{2\delta}{p} a^p + \frac{p-2}{p\delta^{2/(p-2)}} b^{p/(p-2)}, \forall a, b > 0.
\]

Therefore,
\[
\mathbb{E}(|\hat{\Delta}(T)|^2 I_{(\rho \leq T)}) \leq \frac{2\delta}{p} \mathbb{E}|\hat{\Delta}(T)|^p + \frac{p-2}{p\delta^{2/(p-2)}} \mathbb{P}(\rho \leq T).
\]

Using Lemma 2.7 and Lemma 3.2 gives that
\[
\mathbb{E}|\hat{\Delta}(T)|^p \leq C \left( \mathbb{E}|x(T)|^p + \mathbb{E}|x_\Delta(T)|^p \right) \leq C.
\]

Applying Lemma 3.5 yields that
\[
\mathbb{P}(\rho \leq T) \leq \mathbb{P}(\lambda_L \leq T) + \mathbb{P}(\lambda_{\Delta, L} \leq T) \leq \frac{C}{L^p}.
\]

Choose \(\delta = \Delta^2 \alpha(\Delta)^4 \lor \Delta^2(\gamma \lor \sigma)\) and \(L = (\Delta^2 \alpha(\Delta)^4 \lor \Delta^2(\gamma \lor \sigma))^{\frac{1}{p-2}}\). Then we have
\[
\mathbb{E}|\hat{\Delta}(T)|^2 \leq \mathbb{E}|\hat{\Delta}(T \wedge \rho)|^2 + C(\Delta^2 \alpha(\Delta)^4 \lor \Delta^2(\gamma \lor \sigma)).
\]

17
Using the condition (3.12) and the fact that $\varphi^{-1}$ is a strictly increasing function gives that

$$\varphi^{-1}(\alpha(\Delta)) \geq (\Delta^2 \alpha(\Delta)^4 \vee \Delta^{2(\gamma \wedge \sigma)})^{\frac{1}{2}} = L.$$ 

So, we get from Lemma 3.6 that

$$\mathbb{E}[\hat{e}_\Delta(T)]^2 \leq C(\Delta^2 \alpha(\Delta)^4 \vee \Delta^{2(\gamma \wedge \sigma)}).$$

Therefore, the desired result (3.13) is obtained. Then combining Lemma 3.1 and (3.13) gives (3.14). The proof is complete. \hfill \Box

**Corollary 3.8.** Under the assumptions in Theorem 3.7. Let $\varphi(r) = c^* r^{\beta+1}$ for $r \geq 1$, $c^* \geq 0$ and $\alpha(\Delta) = \Delta^{-\frac{\varepsilon}{2}}$ for some $\varepsilon \in (0, \frac{1}{2})$, $\Delta \in (0, \Delta^*)$. Then we have

$$\mathbb{E}|x(T) - x_\Delta(T)|^2 \leq C \Delta^{2(1-\varepsilon)\gamma \wedge \sigma},$$

and

$$\mathbb{E}|x(T) - \bar{x}_\Delta(T)|^2 \leq C \Delta^{2(1-\varepsilon)\gamma \wedge \sigma}.$$ 

**Remark 3.** The better result can be given if we impose the stronger conditions on $\gamma$ and $\sigma$. For instance, suppose that $\gamma, \sigma \in [1-\varepsilon, 1]$. Then we have $(1-\varepsilon)$-order convergence rate in $L^2$ sense.

To get the convergence rate in $L^2$ sense for $\bar{q} > 2$, Assumption 2.2 needs to be replaced by the following stronger assumption.

**Assumption 3.9.** The exist constants $\hat{K}_1 > 0$ and $\hat{q} \in (2, q)$ such that

$$(x-\bar{x})^T (f(t, x, y) - f(t, \bar{x}, \bar{y})) + (\hat{q}-1) \left| g(t, x, y) - g(t, \bar{x}, \bar{y}) \right|^2 \leq \hat{K}_1 (|x-\bar{x}|^2 + |y-\bar{y}|^2)$$

for any $t \in [0, T]$ and $x, \bar{x}, y, \bar{y} \in \mathbb{R}$.

Since many techniques used in Lemma 3.6 are applied here, we mainly state different proof processes in the following lemma, but omit the similar processes.

**Lemma 3.10.** Let Assumptions 2.1, 2.3-2.6 and 3.9 hold with $p \geq 2(1 + \beta)\hat{q}$. Let $L > ||\xi||$ be a real number, and let $\Delta \in (0, \Delta^*)$ be sufficiently small such that $\varphi^{-1}(\alpha(\Delta)) \geq L$. Then we have, for any $\bar{q} \in (2, \hat{q})$,

$$\mathbb{E}|x(T \wedge \rho_{\Delta, L}) - x_\Delta(T \wedge \rho_{\Delta, L})|^{\bar{q}} \leq C(\Delta^{\delta \alpha(\Delta)^{2\hat{q}} \vee \Delta^{\delta(\gamma \wedge \sigma)})\),$$

where $\rho_{\Delta, L} := \lambda_L \wedge \lambda_{\Delta, L}$.

**Proof.** For simplicity, let $\rho_{\Delta, L} = \rho$ and $\varepsilon_\Delta(t) = x(t) - Z_\Delta(t)$. By Itô’s formula,
we have

\[ E|e_\Delta(t \wedge \rho)|^q \]

\[ \leq \theta q |f_\Delta(0, \xi(0), \xi(-\tau))|^q \Delta \]

\[ + E \int_0^{t \wedge \rho} \bar{q}e_\Delta(s)|^{q-2}[e_\Delta^T(s)(f(s, x(s), x(s - \tau)) - f_\Delta(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)))
\]

\[ + \frac{q - 1}{2} |g(s, x(s), x(s - \tau)) - g_\Delta(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau))
\]

\[ - g_1(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau))g_\Delta(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau))\Delta \bar{B}(s)
\]

\[ - g_2(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau))g_\Delta(\mu(s - \tau), \bar{x}_\Delta(s - \tau), \bar{x}_\Delta(s - 2\tau))\Delta \bar{B}(s - \tau)|^2 ds \]

\[ \leq C \Delta^{\bar{q}} \alpha(\Delta)^{\bar{q}} + CE \int_0^{t \wedge \rho} |e_\Delta(s)|^{q-2} [(x(s) - x_\Delta(s))^T
\]

\[ \cdot (f(s, x(s), x(s - \tau)) - f(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau)))
\]

\[ + (q - 1)|g(s, x(s), x(s - \tau)) - g(\mu(s), x_\Delta(s), x_\Delta(s - \tau))|^2 ds
\]

\[ + CE \int_0^{t \wedge \rho} |e_\Delta(s)|^{q-2} \bar{R}_g(s, x_\Delta(s), \bar{x}_\Delta(s), x_\Delta(s - \tau), \bar{x}_\Delta(s - \tau))^{2} ds
\]

\[ + CE \int_0^{t \wedge \rho} |e_\Delta(s)|^{q-2} \theta f_\Delta^T(s, x_\Delta(s), x_\Delta(s - \tau))
\]

\[ \cdot (f(s, x(s), x(s - \tau)) - f(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau))) ds
\]

\[ = C \Delta^{\bar{q}} \alpha(\Delta)^{\bar{q}} + B_1 + B_2 + B_3. \]

For \( q \in (2, \bar{q}) \), note that

\[ (q - 1)|g(s, x(s), x(s - \tau)) - g(\mu(s), x_\Delta(s), x_\Delta((s - \tau)))|^2
\]

\[ \leq (q - 1)|g(s, x(s), x(s - \tau)) - g(s, x_\Delta(s), x_\Delta((s - \tau)))|^2
\]

\[ + \frac{(q - 1)(\bar{q} - 1)}{q - \bar{q}} |g(s, x_\Delta(s), x_\Delta(s - \tau)) - g(\mu(s), x_\Delta(s), x_\Delta(s - \tau))|^2. \]  

From (3.15), one can see that

\[ B_1 \leq CE \int_0^{t \wedge \rho} |e_\Delta(s)|^{q-2} [(x(s) - x_\Delta(s))^T
\]

\[ \cdot (f(s, x(s), x(s - \tau)) - f(s, x_\Delta(s), x_\Delta(s - \tau)))
\]

\[ + (q - 1)|g(s, x(s), x(s - \tau)) - g(s, x_\Delta(s), x_\Delta(s - \tau))|^2 ds
\]

\[ + CE \int_0^{t \wedge \rho} |e_\Delta(s)|^{q-2} (x(s) - x_\Delta(s))^T
\]

\[ \cdot (f(s, x_\Delta(s), x_\Delta(s - \tau)) - f(\mu(s), x_\Delta(s), x_\Delta(s - \tau))) ds
\]

\[ + CE \int_0^{t \wedge \rho} |e_\Delta(s)|^{q-2} (x(s) - x_\Delta(s))^T
\]

\[ \cdot (f(\mu(s), x_\Delta(s), x_\Delta(s - \tau)) - f(\mu(s), \bar{x}_\Delta(s), \bar{x}_\Delta(s - \tau))) ds
\]
By Assumption 3.9, we get that

\[ B_{11} \leq \mathbb{E} \int_0^t |e_\Delta(s)|^{\bar{q}} \left( |x(s) - x_\Delta(s)|^2 + |x(s) - \tau - x_\Delta(s) - \tau|^2 \right) \, ds \]

\[ \leq \mathbb{E} \int_0^t |e_\Delta(s)|^\bar{q} \, ds + \mathbb{E} \int_0^t \left| x(s) - x_\Delta(s) \right|^\bar{q} \, ds + \mathbb{E} \int_0^t |\xi(s) - \xi(s_\Delta)|^\bar{q} \, ds \]

\[ \leq C \int_0^t \mathbb{E}|e_\Delta(s \wedge \rho)|^\bar{q} \, ds + C \int_0^t \mathbb{E}|x(s \wedge \rho) - x_\Delta(s \wedge \rho)|^\bar{q} \, ds + C \Delta^{\bar{q} \gamma}. \]

Using the similar technique in Lemma 3.6 gives that

\[ B_{12} \leq C \int_0^t \mathbb{E}|e_\Delta(s \wedge \rho)|^\bar{q} \, ds + C \int_0^t \mathbb{E}|x(s \wedge \rho) - x_\Delta(s \wedge \rho)|^\bar{q} \, ds + C \Delta^{\bar{q} \sigma}, \]

\[ B_{13} \leq C \int_0^t \mathbb{E}|e_\Delta(s \wedge \rho)|^\bar{q} \, ds + C \int_0^t \mathbb{E}|x(s \wedge \rho) - x_\Delta(s \wedge \rho)|^\bar{q} \, ds \]

\[ + C \Delta^{\bar{q} \alpha(\Delta)} \Delta^{2\bar{q}} + C \Delta^{\bar{q} \gamma}, \]

\[ B_{14} \leq C \int_0^t \mathbb{E}|e_\Delta(s \wedge \rho)|^\bar{q} \, ds + C \Delta^{\bar{q} \sigma}. \]

Hence,

\[ B_1 \leq C \int_0^t \mathbb{E}|e_\Delta(s \wedge \rho)|^\bar{q} \, ds + C \int_0^t \mathbb{E}|x(s \wedge \rho) - x_\Delta(s \wedge \rho)|^\bar{q} \, ds \]

\[ + C \Delta^{\bar{q} \alpha(\Delta)} \Delta^{2\bar{q}} + C \Delta^{\bar{q} \gamma} + C \Delta^{\bar{q} \sigma}. \]

We derive from Lemma 3.4 that

\[ B_2 \leq C \int_0^t \mathbb{E}|e_\Delta(s \wedge \rho)|^\bar{q} \, ds + C \Delta^{\bar{q} \alpha(\Delta)} \Delta^{2\bar{q}}. \]

Moreover, applying the technique in the estimation of \( B_1 \) yields that

\[ B_3 \leq C \int_0^t \mathbb{E}|e_\Delta(s \wedge \rho)|^\bar{q} \, ds + C \int_0^t \mathbb{E}|x(s \wedge \rho) - x_\Delta(s \wedge \rho)|^\bar{q} \, ds \]

\[ + C \Delta^{\bar{q} \alpha(\Delta)} \Delta^{2\bar{q}} + C \Delta^{\bar{q} \gamma} + C \Delta^{\bar{q} \sigma}. \]

Combining (3.16) - (3.18) can give that

\[ \mathbb{E}|e_\Delta(t \wedge \rho)|^\bar{q} \leq C \int_0^t \mathbb{E}|e_\Delta(s \wedge \rho)|^\bar{q} \, ds + C \int_0^t \mathbb{E}|x(s \wedge \rho) - x_\Delta(s \wedge \rho)|^\bar{q} \, ds \]

\[ + C \Delta^{\bar{q} \alpha(\Delta)} \Delta^{2\bar{q}} + C \Delta^{\bar{q} \gamma} + C \Delta^{\bar{q} \sigma}. \]
One can get from Gronwall’s inequality that
\[
E|\epsilon_\Delta(t \land \rho)|^{\tilde{q}} \leq C \int_0^t E|x(s \land \rho) - x_\Delta(s \land \rho)|^{\tilde{q}} ds + C \Delta^{\tilde{q}} \alpha(\Delta)^{2\tilde{q}} + C \Delta^{\tilde{q} \gamma} + C \Delta^{\tilde{q} \sigma}.
\]

Then we derive that
\[
E|x(t \land \rho) - x_\Delta(t \land \rho)|^{\tilde{q}} \\
\leq C \Delta^{\tilde{q}} \alpha(\Delta)^{\tilde{q}} + C E|\epsilon_\Delta(t \land \rho)|^{\tilde{q}} \\
\leq C \int_0^t E|x(s \land \rho) - x_\Delta(s \land \rho)|^{\tilde{q}} ds + C \Delta^{\tilde{q}} \alpha(\Delta)^{2\tilde{q}} + C \Delta^{\tilde{q} \gamma} + C \Delta^{\tilde{q} \sigma}.
\]

The desired result can be obtained due to the Gronwall inequality. We complete the proof.

**Theorem 3.11.** Let Assumptions \(2.1, 2.3-2.6\) and \(3.9\) hold with \(p \geq 2(\beta + 1)^{\tilde{q}}\). For any sufficiently small \(\Delta \in (0, \Delta^*)\), assume that
\[
\alpha(\Delta) \geq \varphi \left( (\Delta^{\tilde{q}} \alpha(\Delta)^{2\tilde{q}} \lor \Delta^{\tilde{q} \gamma} \lor \Delta^{\tilde{q} \sigma}) \right)^{\frac{1}{\tilde{q} - 1}}.
\]

Then, for \(\tilde{q} \in (2, \tilde{q})\) and such small \(\Delta\), we have
\[
E|x(T) - x_\Delta(T)|^{\tilde{q}} \leq C(\Delta^{\tilde{q}} \alpha(\Delta)^{2\tilde{q}} \lor \Delta^{\tilde{q} \gamma} \lor \Delta^{\tilde{q} \sigma}).
\]

**Proof.** Let \(\hat{\epsilon}_\Delta(t) = x(t) - x_\Delta(t)\) and \(\rho = \rho_{\Delta,L}\) for simplicity. Obviously,
\[
E|\hat{\epsilon}_\Delta(T)|^{\tilde{q}} = E \left( |\hat{\epsilon}_\Delta(T)|^{\tilde{q}} 1_{(\rho > T)} \right) + E \left( |\hat{\epsilon}_\Delta(T)|^{\tilde{q}} 1_{(\rho \leq T)} \right).
\]

Let \(\tilde{\delta} > 0\) be arbitrary. Applying Young’s inequality gives that
\[
a^{\frac{p}{\tilde{q}}}b = (\delta a^p)^{\frac{\tilde{q}}{\delta}} \left( \frac{b^{p-\tilde{q}}}{\delta^{p-\tilde{q}}} \right)^{\frac{p}{\tilde{q} - 1}} \leq \frac{\tilde{q}}{\delta} a^p + \frac{p - \tilde{q}}{p\delta^{p-\tilde{q}}} b^{p/(p-\tilde{q})}, \quad \forall a, b > 0.
\]

So,
\[
E \left( |\hat{\epsilon}_\Delta(T)|^{\tilde{q}} 1_{(\rho \leq T)} \right) \leq \frac{\tilde{q}}{\delta} E|\hat{\epsilon}_\Delta(T)|^p + \frac{p - \tilde{q}}{p\delta^{p/(p-\tilde{q})}} \mathbb{P}(\rho \leq T).
\]

Note that \(E|\hat{\epsilon}_\Delta(T)|^p \leq C\) and \(\mathbb{P}(\rho \leq T) \leq \frac{C}{\delta^p}\). Then set \(\tilde{\delta} = \Delta^{\tilde{q}} \alpha(\Delta)^{2\tilde{q}} \lor \Delta^{\tilde{q} \gamma} \lor \Delta^{\tilde{q} \sigma}\),
\[L = (\Delta^{\tilde{q}} \alpha(\Delta)^{2\tilde{q}} \lor \Delta^{\tilde{q} \gamma} \lor \Delta^{\tilde{q} \sigma})^{\frac{1}{\tilde{q} - 1}}.\]

Thus, we get that
\[
E|\hat{\epsilon}_\Delta(T)|^{\tilde{q}} \leq E|\hat{\epsilon}_\Delta(T \land \rho)|^{\tilde{q}} + C(\Delta^{\tilde{q}} \alpha(\Delta)^{2\tilde{q}} \lor \Delta^{\tilde{q} \gamma} \lor \Delta^{\tilde{q} \sigma})).
\]

The condition \(3.19\) means that
\[
\varphi^{-1}(\alpha(\Delta)) \geq (\Delta^{\tilde{q}} \alpha(\Delta)^{2\tilde{q}} \lor \Delta^{\tilde{q} \gamma} \lor \Delta^{\tilde{q} \sigma})^{\frac{1}{\tilde{q} - 1}} = L.
\]
Due to Lemma 3.10, one can see that
\[ E|\tilde{e}_\Delta(T)|^q \leq C(\Delta^{\hat{q}}\alpha(\Delta)^{2\hat{q}} \lor \Delta^{\hat{q}(\gamma+\sigma)}). \]

The desired result (3.20) follows. Then combining Lemma 3.1 and (3.20) gives (3.21). The proof is complete.

**Remark 4.** Similar to Remark 3, the sense can be given if some stronger conditions are added.

### 4. Numerical example

A numerical example is presented to test our theory in this section. Consider the following highly nonlinear and nonautonomous SDDE

\[
 dx(t) = \left( \frac{1}{8} |x(t - \tau)|^{3/2} - 5x^3(t) + 2\zeta_\gamma x(t) \right) dt \\
 + \left( \frac{1}{2} |x(t)|^{3/2} + \zeta_\gamma x(t - \tau) \right) dB(t), \quad t \in [0, 1],
\]

with the initial data \( \xi \) which satisfies Assumption 2.4 and Assumption 2.6 are satisfied with \( \beta > p > U \). This means that Assumption 2.2 holds with \( \beta = p > U \) and \( \zeta_\gamma \) is a scalar Brownian motion. It is easy to verify Assumption 2.5 with \( \gamma = 1 \). By Theorem 3.7, we have

\[
 (x - \bar{x})^T(f(t, x, y) - f(t, \bar{x}, \bar{y})) \\
 \leq 5|x - \bar{x}|^2(-\frac{1}{2}(x^2 + \bar{x}^2)) + 3|x - \bar{x}|^2 + \frac{25}{256}|y - \bar{y}|^2(|y|^\frac{3}{2} + |\bar{y}|^\frac{3}{2})^2 \\
 \leq \frac{5}{2}|x - \bar{x}|^2(|x|^2 + |\bar{x}|^2) + 3|x - \bar{x}|^2 + \frac{25}{64}|y - \bar{y}|^2 + \frac{25}{128}|y - \bar{y}|^2(|y|^2 + |\bar{y}|^2).
\]

Let \( q = 2 \). Similarly,

\[
 (q - 1)|g(t, x, y) - g(t, \bar{x}, \bar{y})|^2 \leq \frac{1}{2}|x|^\frac{3}{2} - |\bar{x}|^\frac{3}{2}|^2 + 2|y - \bar{y}|^2 \\
 \leq \frac{9}{2}|x - \bar{x}|^2 + \frac{9}{4}|x - \bar{x}|^2(|x|^2 + |\bar{x}|^2) + 2|y - \bar{y}|^2.
\]

Combining the above two inequalities together yields that

\[
 (x - \bar{x})^T(f(t, x, y) - f(t, \bar{x}, \bar{y})) + (q - 1)|g(t, x, y) - g(t, \bar{x}, \bar{y})|^2 \\
 \leq 8(|x - \bar{x}|^2 + |y - \bar{y}|^2) - \frac{1}{4}|x - \bar{x}|^2(|x|^2 + |\bar{x}|^2) + \frac{1}{4}|y - \bar{y}|^2(|y|^2 + |\bar{y}|^2).
\]

This means that Assumption 2.2 holds with \( U(x, \bar{x}) = \frac{4}{3}|x - \bar{x}|^2(|x|^2 + |\bar{x}|^2) \). Then for \( p > 2 \), Assumption 2.5 can be verified simply. Additionally, it is easy to show that Assumption 2.4 and Assumption 2.6 are satisfied with \( \beta = 2, \sigma = \frac{3}{4} \). Then, choose \( \varphi(r) = 5r^3 \) for \( r \geq 1 \) and \( \alpha(\Delta) = \Delta^{-\frac{1}{4}} \). By Theorem 3.7, we have

\[
 E|x(T) - x_\Delta(T)|^2 \leq C\Delta^{\frac{5}{4}},
\]
which means that the convergence rate of the truncated $\theta$-Milstein method for SDDE (4.1) is 0.75.

Since the true solution of (4.1) cannot be shown explicitly, we view the truncated $\theta$-Milstein method with step size $2^{-11}$ as the true solution in the numerical experiment. Figure 1 gives the $\mathcal{L}^2$-errors defined by

$$\left( \mathbb{E} |x(T) - x_{\Delta}(T)|^2 \right)^{\frac{1}{2}} \approx \left( \frac{1}{2000} \sum_{i=1}^{2000} |x^i(T) - x^i_{\Delta}(T)|^2 \right)^{\frac{1}{2}},$$

with step sizes $2^{-9}, 2^{-8}, 2^{-7}, 2^{-6}$ at $T = 1$. Two thousand sample paths are simulated. From Figure 1, we can observe that the convergence rate of the truncated $\theta$-Milstein method for (4.1) is approximately 0.75, which means that our theoretical results are reliable.

Remark 5. When delay is vanishing (i.e., $\tau = 0$), SDDE (4.1) will degenerate into nonautonomous stochastic differential equations. The research about the explicit truncated Milstein method (i.e., $\theta = 0$) for non-autonomous stochastic differential equations can be found in [23].

5. Conclusion and future research

In this paper, we establish the truncated $\theta$-Milstein method for a class of highly nonlinear and nonautonomous SDDEs which have practical applications in many fields. The convergence rate is investigated in $\mathcal{L}^q (\bar{q} \geq 2)$ sense under the weaker conditions than the existing results. An example and its numerical
simulation are presented to show the effectiveness of the truncated $\theta$-Milstein method.

The future work is to analyze the long-time asymptotic behavior of the truncated $\theta$-Milstein method for SDDE \((2.1)\). In addition, we are working on the stability of the truncated $\theta$-Milstein method for SDDE \((2.1)\) by adjusting the parameter $\theta$.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (Grant Nos. 61876192 and 11871343) and Science and Technology Innovation Plan of Shanghai, China (Grant No. 20JC1414200).

References

[1] E. Allen, Modeling with Itô Stochastic Differential Equations, Springer, Dordrecht, 2007.

[2] A. Ambrosetti, G. Prodi, A Primer of Nonlinear Analysis, Vol. 34, in: Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 1993.

[3] J.A.D. Appleby, M. Guzowska, C. Kelly, A. Rodkina, Preserving positivity in solutions of discretised stochastic differential equations, Appl. Math. Comput. 217 (2010) 763-774.

[4] L. Arnold, Stochastic Differential Equations: Theory and Applications, John Wiley, New York, 1974.

[5] C.T.H. Baker, E. Buckwar, Numerical analysis of explicit one-step methods for stochastic delay differential equations, J. Comput. Math. 3 (2000) 315-335.

[6] J. Bao, B. Böttcher, X. Mao, C. Yuan, Convergence rate of numerical solutions to SFDEs with jumps, J. Comput. Appl. Math. 236 (2) (2011) 119-131.

[7] A. Calzolari, P. Florchinger, G. Nappo, Nonlinear filtering for stochastic systems with fixed delay: Approximation by a modified Milstein scheme, Comput. Math. Appl. 61 (2011) 2498-2509.

[8] S. Deng, W. Fei, W. Liu, X. Mao, The truncated EM method for stochastic differential equations with Poisson jumps, J. Comput. Appl. Math. 355 (2019) 232-257.

[9] W. Fei, L. Hu, X. Mao, D. Xia. Advances in the truncated Euler-Maruyama method for stochastic differential delay equations, Commun. Pure Appl. Anal. 19 (4) (2020) 2081-2100.
S. Gao, J. Hu, L. Tan, C. Yuan, Strong convergence rate of truncated Euler-Maruyama method for stochastic differential delay equations with Poisson jumps, Front. Math. China 16, (2021) 395-423.

Q. Guo, W. Liu, X. Mao, R. Yue, The partially truncated Euler-Maruyama method and its stability and boundedness, Appl. Numer. Math. 115 (2017) 235-251.

Q. Guo, W. Liu, X. Mao, R. Yue, The truncated Milstein method for stochastic differential equations with commutative noise, J. Comput. Appl. Math. 338 (2018) 298-310.

Q. Guo, X. Mao, R. Yue, The truncated Euler-Maruyama method for stochastic differential delay equations. Numer. Algorithms. 78 (2) (2018) 599-624.

D.J. Higham, X. Mao, A.M. Stuart, Strong convergence of Euler-type methods for nonlinear stochastic differential equations, SIAM J. Numer. Anal. 40 (3) (2002) 1041-1063.

L. Hu, X. Li, X. Mao, Convergence rate and stability of the truncated Euler-Maruyama method for stochastic differential equations, J. Comput. Appl. Math. 337 (2018) 274-289.

Y. Hu, S.E.A. Mohammed, F. Yan, Discrete-time approximations of stochastic delay equations: the Milstein Scheme, Ann. Probab. 32 (2004) 265-314.

N. Hofmann, T. Müller-Gronbach, A modified Milstein scheme for approximation of stochastic delay differential equations with constant time lag, J. Comput. Appl. Math. 197 (2006) 89-121.

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. 467 (2011) 1563-1576.

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.

P.E. Kloeden, E. Platen, Numerical Solution of Stochastic Differential Equations, Springer, Berlin, 1992.

P.E. Kloeden, T. Shardlow, The Milstein scheme for stochastic delay differential equations without using anticipative calculus, J. Sto. Anal. Appl. 30 (2) (2012) 181-202.
[22] 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 (2019) 847-892.

[23] J. Liao, W. Liu, X. Wang, Truncated Milstein method for non-autonomous stochastic differential equations and its modification, J. Comput. Appl. Math. 402 (2022) 113817.

[24] M. Liu, W. Cao, Z. Fan, Convergence and stability of the semi-implicit Euler method for a linear stochastic differential delay equation, J. Comput. Appl. Math. 170 (2004) 255-268.

[25] W. Liu, X. Mao, J. Tang, Y. Wu, Truncated Euler-Maruyama method for classical and time-changed non-autonomous stochastic differential equations, Appl. Numer. Math. 153 (2020) 66-81.

[26] W. Mao, Q. Zhu, X. Mao, Existence, uniqueness and almost surely asymptotic estimations of the solutions to neutral stochastic functional differential equations driven by pure jumps, Appl. Math. Comput. 254 (2015) 252-265.

[27] X. Mao, Stochastic Differential Equations and Applications, 2nd ed. Horwood, Chichester, 2007.

[28] X. Mao, Numerical solutions of stochastic differential delay equations under the generalized Khasminskii-type conditions, Appl. Math. Comput. 217 (12) (2011) 5512-5524.

[29] X. Mao, The truncated Euler-Maruyama method for stochastic differential equations, J. Comput. Appl. Math. 290 (2015) 370-384.

[30] X. Mao, Convergence rates of the truncated Euler-Maruyama method for stochastic differential equations, J. Comput. Appl. Math 296 (2016) 362-375.

[31] X. Mao, C. Yuan, Stochastic Differential Equations with Markovian Switching, Imperial College Press, 2006.

[32] G.N. Milstein, E. Platen, H. Schurz, Balanced implicit methods for stiff stochastic system, SIAM J. Numer. Anal. 35 (1998) 1010-1019.

[33] O.F. Rouz, D. Ahmadian, M. Milev, Exponential mean-square stability of two classes of theta Milstein methods for stochastic delay differential equations, AIP Conf. Proc. 1910 (1) (2017) 060015.

[34] S. Sabanis, A note on tamed Euler approximations, Electron. Commun. Probab. 18 (2013) 1-10.

[35] G. Song, J. Hu, S. Gao, X. Li, The strong convergence and stability of explicit approximations for nonlinear stochastic delay differential equations, Numer. Algorithms. (2021).
[36] M. Song, L. Hu, X. Mao, L. Zhang, Khasminskii-type theorems for stochastic functional differential equations, Discrete Contin. Dyn. Syst. Ser. B 6 (6) (2013) 1697-1714.

[37] X. Wang, S. Gan, The improved split-step backward Euler method for stochastic differential delay equations, Int. J. Comput. Math. 88 (11) (2011) 2359-2378.

[38] 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.

[39] Z. Wang, C. Zhang, An analysis of stability of milstein method for stochastic differential equations with delay, Comput. Math. Appl. 51 (2006) 1445-1452.

[40] E. Zeidler, Nonlinear functional analysis and its applications, Springer New York, New York, 1990.

[41] W. Zhang, M. Song, M. Liu, Strong convergence of the partially truncated Euler-Maruyama method for a class of stochastic differential delay equations, J. Comput. Appl. Math. 335 (2018) 114-128.

[42] W. Zhang, X. Yin, M. Song, M. Liu, Convergence rate of the truncated Milstein method of stochastic differential delay equations, Appl. Math. Comput. 357 (2019) 263-281.

[43] Y. Zhang, M. Song, M. Liu, Convergence and stability of stochastic theta method for nonlinear stochastic differential equations with piecewise continuous arguments, J. Comput. Appl. Math. 403 (2022) 113849.

[44] G. Zhao, M. Liu, Numerical methods for nonlinear stochastic delay differential equations with jumps, Appl. Math. Comput. 233 (2014) 222-231.

[45] X. Zong, F. Wu, G. Xu, Convergence and stability of two classes of theta-Milstein schemes for stochastic differential equations, J. Comput. Appl. Math. 336 (2018) 8-29.