Available online at www.sciencedirect.com

ScienceDirect

Journal of Differential Equations 374 (2023) 154–190

Journal of Differential Equations

www.elsevier.com/locate/jde

Hybrid stochastic functional differential equations with infinite delay: Approximations and numerics

Guozhen Li [a], Xiaoyue Li [b,a], Xuerong Mao [c,*], Guoting Song [d]

[a] *School of Mathematics and Stochastics, Northeast Normal University, Changchun, Jilin, 130024, China*
[b] *School of Mathematical Sciences, Tiangong University, Tianjin, 300387, China*
[c] *Department of Mathematics and Statistics, University of Strathclyde, Glasgow G1 1XH, UK*
[d] *Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, 100190, China*

Received 27 March 2023; revised 16 May 2023; accepted 17 July 2023

Abstract

This paper is to investigate if the solution of a hybrid stochastic functional differential equation (SFDE) with infinite delay can be approximated by the solution of the corresponding hybrid SFDE with finite delay. A positive result is established for a large class of highly nonlinear hybrid SFDEs with infinite delay. Our new theory makes it possible to numerically approximate the solution of the hybrid SFDE with infinite delay, via the numerical solution of the corresponding hybrid SFDE with finite delay.

Keywords: Stochastic functional differential equation; Infinite delay; Finite delay; Approximate solution; Numerical solution

1. Introduction

Many dynamic systems in sciences and industry do not only depend on their current state but also past states due to unavoidable time delays, while they are often subject to various system parameter uncertainties and environmental noise. Moreover, random switching takes place frequently in a finite set resulting in the systems being hybrid, in which continuous dynamics and discrete events coexist and interact. Hybrid stochastic functional differential equations (SFDEs)

* Corresponding author.
E-mail address: x.mao@strath.ac.uk (X. Mao).

https://doi.org/10.1016/j.jde.2023.07.028

have been used to model such dynamic systems. There are two categories: (A) hybrid SFDEs with finite delays; (B) hybrid SFDEs with infinite delays. There is a huge literature on type-(A) SFDEs (see, e.g., [5,6,10,20,23,26,34]) but much less on (B). This is certainly not because type-(B) SFDEs have no use in applications but due to the fact that it is much harder to study (B) than (A). As a matter of fact, long memory, also called long-range dependence or persistence, is a phenomenon that occurs in many fields including ecology, biology, econometrics, linguistics, hydrology, climate, DNA sequencing. It is due to the demand from the research in these fields, research on type-(B) SFDEs has been advanced quickly (see e.g., [14,15,28,29,33]). However, there are many open problems to be solved in order to meet the need of applications.

One of the open problems is: how to obtain numerical solutions to type-(B) SFDEs? In order to introduce the main ideas in this paper clearly, it is necessary to distinguish numerical solutions from approximate ones. By numerical solutions we will mean they can be simulated by computers. For example, the Euler-Maruyama (EM) solutions and Milstein solutions can be simulated by computers using, e.g., the MATLAB programs in [8]. However, approximate solutions may not be simulated. For example, the continuous approximate solutions of type-(B) SFDEs to be defined in Section 3 below are theoretical approximate solutions which do not have explicit forms in general; the EM solutions proposed by Asker in [1] cannot be simulated by computer as they need infinitely-many discrete-time initial data which any computer could not cope with. In other words, they are in fact approximate ones. In this paper we will tackle this open problem by bridging the gap between type-(B) SFDEs and type-(A) SFDEs. More precisely, we will establish new approximation theory between type-(B) and type-(A) SFDEs. Based on our new theory, type-(B) SFDEs can be approximated by the corresponding type-(A) SFDEs (Step 1). Applying numerical methods to type-(A) SFDEs, we can obtain the numerical solutions approximating the corresponding type-(A) SFDEs (Step 2). By this bridge, we also yield the numerical solutions to type-(B) SFDEs. Our approach can be illustrated as follows:

$$\text{type-(B) SFDEs} \Leftarrow_{\text{Step 1}} \text{type-(A) SFDEs} \Leftarrow_{\text{Step 2}} \text{numerical solutions.}$$

To see our approach does not only work but is also useful, we need to make four points clear:

- There are existing numerical methods on type-(A) SFDEs (see, e.g., [31,32]), though the numerical theory in this area is still developing. In detail, Li and Hou [12] discussed the EM method for linear hybrid stochastic delay differential equations (SDDEs). Wu and Mao [30] established the strong mean square convergence of the EM method for neutral SFDEs under the linear growth condition. Recently, numerical methods for superlinear type-(A) SFDEs have been developing quickly. For examples, under the generalized Khasminskii-type condition in terms of Lyapunov functions, Li et al. [13] proved that the EM numerical solutions converge to the exact ones in probability in any finite interval; Guo et al. [7] and Song et al. [25] obtained the strong convergence of the truncated EM numerical solutions for type-(A) SDDEs; Zhang et al. [35] extended the truncated EM method to type-(A) SFDEs; Dareiotis et al. [4] extended the tamed EM to type-(A) SDDEs driven by Lévy noise.
- There are a number of papers where type-(B) SFDEs have been used to model population systems [14,15]. To illustrate their results, the Milstein and EM methods were used respectively to perform some computer simulations but there is no explanation on whether the numerical methods are applicable to their superlinear type-(B) SFDEs. In [21], the authors used the simulations on type-(A) SFDEs to illustrate the results on type-(B) SFDEs but once

again there is lack of theoretical support. We hence see there is an urgent need to rigorously establish the new approximation theory between type-(B) and type-(A) SFDEs.
- There are a couple of papers where discrete-time solutions were proposed to re-produce the stability of some very special type-(B) SFDEs. For example, Asker [1] presented the EM solutions of a neutral SFDE with infinite delays under the globally Lipschitz condition and examined the stability in distribution of numerical solutions. It is noted that the stability analysis is theoretical and does not need to compute the EM solutions numerically. As mentioned above, the EM solutions there are only approximate ones.
- Similar situation has happened to the stability analysis of numerical methods for *deterministic* functional differential equations with infinite delays. Song and Baker [27] discretized the *deterministic* Volterra integro-differential equation with infinite delays using the θ method and proved that for a *small bounded initial function* and a small step size the θ method displays the stability property of the underlying equation. The stability analysis of numerical methods can also be found in [2,3].

Our main contributions can therefore be highlighted as follows:

- A novel approximation method is proposed by the truncation technique. More precisely, for a given type-(B) SFDE, we define a corresponding truncated SFDE with finite time delay k. The general approximation theory is established by showing that the solution of the truncated SFDE approximates the solution of the given type-(B) SFDE in the qth moment provided k is sufficiently large.
- Various approximation principles, including the exponential approximation rate, are given for a number of important type-(B) SFDEs.
- Numerical solutions of the truncated SFDE are shown to be close to the solution of the given type-(B) SFDE for sufficiently large k and small numerical step size.
- For the global Lipschitz case the convergence error between the EM numerical solution of type-(A) SFDE and the exact solution of type-(B) SFDE is given.

The rest of the paper is organized as follows: Section 2 gives some necessary notions and assumptions which ensure the well-posedness of the solutions of type-(B) SFDEs. In section 3, the corresponding truncated SFDE with finite time delay k is defined for a given type-(B) SFDE, while both asymptotic approximation Theorem 3.4 and exponential approximation Theorem 3.7 are established. Section 4 discusses a number of important type-(B) SFDEs to which the approximation theory established in Section 3 is applicable. Section 5 shows that the numerical solutions of the truncated SFDE are close to the solution of the given type-(B) SFDE for sufficiently large k and small numerical step size. Two examples with computer simulations are discussed to illustrate the theory.

2. Preliminaries

Throughout this paper, unless otherwise specified, we let $\mathbb{R}^n$ be the n-dimensional Euclidean space and $\mathcal{B}(\mathbb{R}^n)$ denote the family of all Borel measurable sets in $\mathbb{R}^n$. Let $\mathbb{R}_+ = [0, \infty)$ and $\mathbb{R}_- = (-\infty, 0]$. If $x \in \mathbb{R}^n$, then $|x|$ is its Euclidean norm. If A is a vector or matrix, its transpose is denoted by A^T. If A is a matrix, we let $|A| = \sqrt{\text{trace}(A^T A)}$ be its trace norm and $\|A\| = \max\{|Ax| : |x| = 1\}$ be the operator norm. If A is a symmetric matrix ($A = A^T$), denote by $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$ its smallest and largest eigenvalues, respectively. By $A > 0$ and $A \geq 0$, we mean A is positive and non-negative definite, respectively. If both a, b are real numbers, then $a \wedge b = \min\{a, b\}$ and $a \vee b = \max\{a, b\}$. Let $\mathbb{N}_+$ denote the set of nonnegative integers.

We let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions (i.e. it is right continuous and increasing while $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets). For a subset $\bar{\Omega}$ of Ω, $\mathbf{1}_{\bar{\Omega}}$ denotes its indicator function. Let $B(t) = (B_1(t), \cdots, B_m(t))^T$ be an m-dimensional Brownian motion defined on the probability space. Let $\theta(t)$, $t \geq 0$, be a right-continuous irreducible Markov chain on the probability space taking values in a finite state space $\mathbb{S} = \{1, 2, \cdots, N\}$ with generator $\Gamma = (\gamma_{ij})_{N\times N}$ given by

$$\mathbb{P}\{\theta(t+\Delta) = j|\theta(t) = i\} = \begin{cases} \gamma_{ij}\Delta + o(\Delta) & \text{if } i \neq j, \\ 1 + \gamma_{ii}\Delta + o(\Delta) & \text{if } i = j, \end{cases}$$

where $\Delta > 0$. Here $\gamma_{ij} \geq 0$ is the transition rate from i to j if $i \neq j$ while $\gamma_{ii} = -\sum_{j\neq i}\gamma_{ij}$. We assume that the Markov chain $\theta(\cdot)$ is independent of the Brownian motion $B(\cdot)$.

Denote by $C(\mathbb{R}_-; \mathbb{R}^n)$ the family of continuous functions $\varphi : \mathbb{R}_- \to \mathbb{R}^n$. Other families e.g. $C(\mathbb{R}^n \times \mathbb{R}; \mathbb{R}_+)$ can be defined obviously. Fix a positive number r and define the phase space $\mathcal{C}_r$ with the fading memory by

$$\mathcal{C}_r = \left\{\varphi \in C(\mathbb{R}_-; \mathbb{R}^n) : \sup_{-\infty<u\leq 0} e^{ru}|\varphi(u)| < \infty\right\} \tag{2.1}$$

with its norm $\|\varphi\|_r = \sup_{-\infty<u\leq 0} e^{ru}|\varphi(u)|$. It is well known that $\mathcal{C}_r$ under the norm $\|\cdot\|_r$ forms a Banach space (see [9] for more details on this phase space).

Moreover, denote by $\mathcal{P}_0$ the family of probability measures μ on $\mathbb{R}_-$, while for each $b > 0$, define

$$\mathcal{P}_b = \{\mu \in \mathcal{P}_0 : \int_{-\infty}^{0} e^{-bu}\mu(\mathrm{d}u) < \infty\}.$$

Furthermore, set $\mu^{(b)} := \int_{-\infty}^{0} e^{-bu}\mu(du)$ for each $\mu \in \mathcal{P}_b$. Please note that $\mu^{(b)}$ is a positive number but $\mu(\cdot)$ is a measure. Clearly, $\mathcal{P}_{b_1} \subset \mathcal{P}_b \subset \mathcal{P}_0$ if $b_1 > b > 0$. Moreover, if $\mu \in \mathcal{P}_{b_1}$, then $\mu^{(b)}$ is a strictly increasing and continuous function of b in $[0, b_1]$.

Consider a hybrid stochastic functional differential equation (SFDE) with infinite delay of the form

$$dx(t) = f(x_t, \theta(t), t)\mathrm{d}t + g(x_t, \theta(t), t)\mathrm{d}B(t) \tag{2.2}$$

on $t \geq 0$ with the initial data

$$x_0 = \xi \in \mathcal{C}_r \text{ and } \theta(0) = i_0 \in \mathbb{S}. \tag{2.3}$$

Here $f : \mathcal{C}_r \times \mathbb{S} \times \mathbb{R}_+ \to \mathbb{R}^n$ and $g : \mathcal{C}_r \times \mathbb{S} \times \mathbb{R}_+ \to \mathbb{R}^{n\times m}$ are Borel measurable and, moreover, $x(t)$ is an $\mathbb{R}^n$-valued stochastic process on $t \in (-\infty, \infty)$ while $x_t = \{x(t+u) : u \in \mathbb{R}_-\}$ is a $\mathcal{C}_r$-valued stochastic process on $t \geq 0$. We impose the local Lipschitz condition on the coefficients f and g.

Assumption 2.1. For each number $h > 0$, there is a positive constant $\bar{K}_h$ such that

$$|f(\varphi,i,t)-f(\phi,i,t)|\vee|g(\varphi,i,t)-g(\phi,i,t)|\le\bar{K}_h\|\varphi-\phi\|_r \tag{2.4}$$

for those $\varphi,\phi\in\mathcal{C}_r$ with $\|\varphi\|_r\vee\|\phi\|_r\le h$ and all $(i,t)\in\mathbb{S}\times\mathbb{R}_+$. Moreover,

$$\sup_{t\in\mathbb{R}_+}(|f(0,i,t)|\vee|g(0,i,t)|)<\infty,\quad\forall\, i\in\mathbb{S}.$$

We observe that Assumption 2.1 only guarantees the existence of the unique maximal local solution $x(t)$ of the SFDE (2.2) on $t\in(-\infty,\sigma_e)$, where σ_e is known as the explosion time (see, e.g., [17,28,29,33]). To have a global solution (namely, $\sigma_e=\infty$ a.s.), we need some additional conditions. The classical condition is the linear growth condition (see, e.g., [11,16,22]). However, we will impose a much more general Khasminskii-type condition (see, e.g., [19]). For this purpose, we need a couple of new notations. Let $C(\mathbb{R}^n;\mathbb{R}_+)$ denote the family of all continuous functions from $\mathbb{R}^n$ to $\mathbb{R}_+$. Denote by $C^{2,1}(\mathbb{R}^n\times\mathbb{S}\times\mathbb{R}_+;\mathbb{R}_+)$ the family of all continuous non-negative functions $V(x,i,t)$ defined on $\mathbb{R}^n\times\mathbb{S}\times\mathbb{R}_+$ such that for each $i\in\mathbb{S}$, they are continuously twice differentiable in x and once in t. Given a function $V\in C^{2,1}(\mathbb{R}^n\times\mathbb{S}\times\mathbb{R}_+;\mathbb{R}_+)$, we define the functional $LV:\mathcal{C}_r\times\mathbb{S}\times\mathbb{R}_+\to\mathbb{R}$ by

$$\begin{aligned}LV(\varphi,i,t)&=V_t(\varphi(0),i,t)+V_x(\varphi(0),i,t)f(\varphi,i,t)\\&+\frac{1}{2}\mathrm{trace}\big(g(\varphi,i,t)^TV_{xx}(\varphi(0),i,t)g(\varphi,i,t)\big)+\sum_{j=1}^N\gamma_{ij}V(\varphi(0),j,t),\end{aligned}$$

where

$$V_t(x,i,t)=\frac{\partial V(x,i,t)}{\partial t},\quad V_x(x,i,t)=\Big(\frac{\partial V(x,i,t)}{\partial x_1},\cdots,\frac{\partial V(x,i,t)}{\partial x_n}\Big),$$
$$V_{xx}(x,i,t)=\Big(\frac{\partial^2V(x,i,t)}{\partial x_i\partial x_j}\Big)_{n\times n}.$$

Let us emphasize that LV is defined on $\mathcal{C}_r\times\mathbb{S}\times\mathbb{R}_+$ while V on $\mathbb{R}^n\times\mathbb{S}\times\mathbb{R}_+$. The definition of LV is purely based on the following generalized Itô formula (see, e.g., [20, Theorem 1.45 on p.48])

$$\mathrm{d}V(x(t),\theta(t),t)=LV(x_t,\theta(t),t)\mathrm{d}t+\mathrm{d}M(t), \tag{2.5}$$

where $M(t)$ is a local martingale with $M(0)=0$ (whose form is of no use in this paper). Moreover, for each positive number b, define

$$\mathcal{W}_b=\Big\{W\in C(\mathbb{R}^n\times\mathbb{R};\mathbb{R}_+):\sup_{(x,u)\in\mathbb{R}^n\times\mathbb{R}_-}\frac{W(x,u)}{1+|x|^b}<\infty\Big\}.$$

Please note that although W is defined on $\mathbb{R}^n\times\mathbb{R}$, we only need $W(x,u)/(1+|x|^b)$ to be bounded on $(x,u)\in\mathbb{R}^n\times\mathbb{R}_-$. It is also easy to see that if $W\in\mathcal{W}_b$, then

$$\sup_{-\infty<u\leq 0} e^{rbu}W(\varphi(u),u)<\infty, \quad \forall \varphi \in \mathcal{C}_r. \tag{2.6}$$

Moreover, denote by $\mathcal{K}_\infty$ the family of non-decreasing functions $\beta : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\beta(v) \to \infty$ as $v \to \infty$. We can now form the generalized Khasminskii-type condition.

Assumption 2.2. There are positive constants b_1, b_2, K, functions $V \in C^{2,1}(\mathbb{R}^n \times \mathbb{S} \times \mathbb{R}_+; \mathbb{R}_+)$, $W_1 \in \mathcal{W}_{b_1}$, $W_2 \in \mathcal{W}_{b_2}$, $\beta \in \mathcal{K}_\infty$, and probability measures $\mu_1 \in \mathcal{P}_{rb_1}$, $\mu_2 \in \mathcal{P}_{rb_2}$, such that

$$\beta(|x|) \leq \inf_{0\leq t<\infty} W_1(x,t), \quad \forall x \in \mathbb{R}^n, \tag{2.7}$$

$$W_1(x,t) \leq V(x,i,t), \quad \forall (x,i,t) \in \mathbb{R}^n \times \mathbb{S} \times \mathbb{R}_+, \tag{2.8}$$

and

$$\begin{aligned} LV(\varphi,i,t) \leq K + K\int_{-\infty}^{0} W_1(\varphi(u),t+u)\mu_1(\mathrm{d}u) \\ - W_2(\varphi(0),t) + \int_{-\infty}^{0} W_2(\varphi(u),t+u)\mu_2(\mathrm{d}u) \end{aligned} \tag{2.9}$$

for all $(\varphi,i,t) \in \mathcal{C}_r \times \mathbb{S} \times \mathbb{R}_+$.

The following theorem forms the foundation for this paper.

Theorem 2.3. *Under Assumptions 2.1 and 2.2, the SFDE (2.2) with the initial data (2.3) has a unique solution $x(t)$ on $t \in (-\infty,\infty)$, which has the property that*

$$\sup_{0\leq t\leq T} \mathbb{E}W_1(x(t),t) \leq C_T, \quad \forall T>0, \tag{2.10}$$

where C_T is a positive constant dependent on T.

Proof. To show the unique maximal local solution $x(t)$ on $t \in (-\infty, \sigma_e)$ is global, we need to show $\sigma_e = \infty$ a.s. For each sufficiently large number $h \geq \|\xi\|_r$ with $\beta(h) > 0$, define the stopping time

$$\tau_h = \sigma_e \wedge \inf\{t \in [0,\sigma_e) : |x(t)| \geq h\},$$

where throughout this paper we set $\inf \emptyset = \infty$ (as usual $\emptyset$ stands for the empty set). Obviously, τ_h is increasing in h and $\tau_h \leq \sigma_e$ a.s. It is therefore sufficient if we could show $\lim_{h\to\infty} \tau_h = \infty$ a.s. Fix $T > 0$ arbitrarily. By the generalized Itô formula (see, e.g., [20]) and Assumption 2.2, we can easily obtain that

$$\mathbb{E}W_1(x(t\wedge\tau_h),t\wedge\tau_h)\le K_1+K\mathbb{E}\int_0^{t\wedge\tau_h}\int_{-\infty}^0 W_1(x(s+u),s+u)\mu_1(\mathrm{d}u)\mathrm{d}s$$

$$-\mathbb{E}\int_0^{t\wedge\tau_h} W_2(x(s),s))\mathrm{d}s+\mathbb{E}\int_0^{t\wedge\tau_h}\int_{-\infty}^0 W_2(x(s+u),s+u)\mu_2(\mathrm{d}u)\mathrm{d}s \tag{2.11}$$

for $t\in[0,T]$, where $K_1=V(\xi(0),i_0,0)+KT$. Due to $W_1\in\mathcal{W}_{b_1}$, there exists a constant $\hat K>0$ such that

$$W_1(x,u)\le\hat K(1+|x|^{b_1}),\quad (x,u)\in\mathbb{R}^n\times\mathbb{R}_-. \tag{2.12}$$

This together with the fact $\xi\in\mathcal{C}_r$ implies

$$\sup_{-\infty<u\le 0} e^{rb_1u}W_1(\xi(u),u)\le\hat K\sup_{-\infty<u\le 0}e^{rb_1u}(1+|\xi(u)|^{b_1})\le\hat K(1+\|\xi\|_r^{b_1})=:K_2. \tag{2.13}$$

By the Fubini theorem and property (2.12), we derive that

$$\begin{aligned}
&\int_0^{t\wedge\tau_h}\int_{-\infty}^0 W_1(x(s+u),s+u)\mu_1(\mathrm{d}u)\mathrm{d}s\\
=&\int_0^{t\wedge\tau_h}\Big(\int_{-\infty}^{-s} W_1(x(s+u),s+u)\mu_1(\mathrm{d}u)+\int_{-s}^0 W_1(x(s+u),s+u)\mu_1(\mathrm{d}u)\Big)\mathrm{d}s\\
\le&\Big(\sup_{-\infty<u\le -s}e^{rb_1(s+u)}W_1(\xi(s+u),s+u)\Big)\int_0^{t\wedge\tau_h}\int_{-\infty}^{-s}e^{-rb_1(s+u)}\mu_1(\mathrm{d}u)\mathrm{d}s\\
+&\int_{-(t\wedge\tau_h)}^0\int_{-u}^{t\wedge\tau_h}W_1(x(s+u),s+u)\mathrm{d}s\mu_1(\mathrm{d}u)\\
\le&\Big(\sup_{-\infty<u\le 0}e^{rb_1u}W_1(\xi(u),u)\Big)\Big(\int_0^{t\wedge\tau_h}e^{-rb_1s}\mathrm{d}s\Big)\Big(\int_{-\infty}^0 e^{-rb_1u}\mu_1(\mathrm{d}u)\Big)\\
+&\int_{-\infty}^0\int_0^{t\wedge\tau_h}W_1(x(s),s)\mathrm{d}s\mu_1(\mathrm{d}u)\\
\le&(K_2/rb_1)\mu_1^{(rb_1)}+\int_0^{t\wedge\tau_h}W_1(x(s),s)\mathrm{d}s.
\end{aligned} \tag{2.14}$$

Similarly,

$$\int_0^{t\wedge\tau_h}\int_{-\infty}^0 W_2(x(s+u),s+u)\mu_2(\mathrm{d}u)\mathrm{d}s \le (K_3/rb_2)\mu_2^{(rb_2)} + \int_0^{t\wedge\tau_h} W_2(x(s),s)\mathrm{d}s, \quad (2.15)$$

where $K_3 = \sup_{-\infty<u\le 0} e^{rb_2u}W_2(\xi(u),u)$. Substituting (2.14) and (2.15) into (2.11) yields

$$\begin{aligned}\mathbb{E}W_1(x(t\wedge\tau_h),t\wedge\tau_h) &\le K_4 + K\mathbb{E}\int_0^{t\wedge\tau_h} W_1(x(s),s)\mathrm{d}s \\ &\le K_4 + K\mathbb{E}\int_0^{t} \mathbb{E}W_1(x(s\wedge\tau_h),s\wedge\tau_h)\mathrm{d}s,\end{aligned}$$

where $K_4 = K_1 + (K_2/rb_1)\mu_1^{(rb_1)} + (K_3/rb_2)\mu_2^{(rb_2)}$. The well-known Gronwall inequality shows

$$\mathbb{E}W_1(x(t\wedge\tau_h),t\wedge\tau_h) \le C_T,\ \forall t\in[0,T], \quad (2.16)$$

where $C_T = K_4e^{KT}$. This along with (2.7) implies that $C_T \ge \mathbb{E}\beta(|x(T\wedge\tau_h)|) \ge \beta(h)\mathbb{P}(\tau_h \le T)$, namely

$$\mathbb{P}(\tau_h \le T) \le \frac{C_T}{\beta(h)}. \quad (2.17)$$

Consequently, $\lim_{h\to\infty}\mathbb{P}(\tau_h \le T) = 0$, namely $\lim_{h\to\infty}\tau_h > T$ a.s. Since $T > 0$ is arbitrary, we must have $\lim_{h\to\infty}\tau_h = \infty$ a.s. Letting $h\to\infty$ in (2.16) we also get the required assertion (2.10). The proof is therefore complete. □

Let us highlight that property (2.17) shows that $\sup_{0\le t\le T}|x(t)| \le h$ with probability at least $1 - C_T/\beta(h)$ for all sufficiently large h. In other words, the solution remains within the compact ball $\{x\in\mathbb{R}^n : |x|\le h\}$ for the time interval $[0,T]$ with a large probability for all sufficiently large h. This nice property will play its important role when we discuss the approximate solutions in the next section.

3. Approximations by SFDEs with finite delay

3.1. Truncated SFDEs

The key aim of this paper is to approximate the solution of the SFDE (2.2) by the solution of an SFDE with finite delay. In turn, we can numerically approximate the solution of the SFDE with finite delay, and hence obtain the numerical solution of the original SFDE (2.2) with infinite delay.

We first need to design the corresponding SFDE with finite delay. For each positive integer k, define the truncation mapping $\pi_k : \mathcal{C}_r \to \mathcal{C}_r$ by

$$\pi_k(\varphi)(u) = \begin{cases} \varphi(u) & \text{if } -k \le u \le 0, \\ \varphi(-k) & \text{if } u < -k. \end{cases}$$

Define the truncation functions $f_k : \mathcal{C}_r \times \mathbb{S} \times \mathbb{R}_+ \to \mathbb{R}^n$ and $g_k : \mathcal{C}_r \times \mathbb{S} \times \mathbb{R}_+ \to \mathbb{R}^{n\times m}$ by

$$f_k(\varphi, i, t) = f(\pi_k(\varphi), i, t) \text{ and } g_k(\varphi, i, t) = g(\pi_k(\varphi), i, t)$$

respectively. We observe that both f_k and g_k depend on the values of φ on the finite time interval $[-k, 0]$ but not the values of φ on $(-\infty, -k)$. In other words, f_k and g_k can be regarded as functionals defined on $C([-k, 0); \mathbb{R}^n) \times \mathbb{S} \times \mathbb{R}_+$. Consider the corresponding truncated SFDE

$$\mathrm{d}x^k(t) = f_k(x_t^k, \theta(t), t)\mathrm{d}t + g_k(x_t^k, \theta(t), t)\mathrm{d}B(t) \tag{3.1}$$

on $t \ge 0$ with the initial data $x_0^k = \xi$ and $\theta(0) = i_0$. This is clearly an SFDE with finite delay.

We observe that the truncation functions f_k and g_k preserve Assumption 2.1 perfectly. In fact, for $\varphi, \phi \in \mathcal{C}_r$ with $\|\varphi\|_r \vee \|\phi\|_r \le h$ and $(i, t) \in \mathbb{S} \times \mathbb{R}_+$,

$$\begin{aligned} & |f_k(\varphi, i, t) - f_k(\phi, i, t)| \vee |g_k(\varphi, i, t) - g_k(\phi, i, t)| \\ =& |f(\pi_k(\varphi), i, t) - f_k(\pi_k(\phi), i, t)| \vee |g(\pi_k(\varphi), i, t) - g_k(\pi_k(\phi), i, t)| \\ \le& \bar{K}_h \|\pi_k(\varphi) - \pi_k(\phi)\|_r \le \bar{K}_h \|\varphi - \phi\|_r. \end{aligned} \tag{3.2}$$

We now show that they also preserve Assumption 2.2. In fact, for $V \in C^{2,1}(\mathbb{R}^n \times \mathbb{S} \times \mathbb{R}_+; \mathbb{R}_+)$, the generalized Itô formula shows

$$\mathrm{d}V(x^k(t), \theta(t), t) = L_k V(x_t^k, \theta(t), t)\mathrm{d}t + \mathrm{d}M_k(t),$$

where $M_k(t)$ is a local martingale with $M_k(0) = 0$ and $L_k V : \mathcal{C}_r \times \mathbb{S} \times \mathbb{R}_+ \to \mathbb{R}$ is defined by

$$\begin{aligned} & L_k V(\varphi, i, t) = V_t(\varphi(0), i, t) + V_x(\varphi(0), i, t) f_k(\varphi, i, t) \\ & + \frac{1}{2}\text{trace}\big(g_k(\varphi, i, t)^T V_{xx}(\varphi(0), i, t) g_k(\varphi, i, t)\big) + \sum_{j=1}^N \gamma_{ij} V(\varphi(0), j, t). \end{aligned}$$

Under Assumption 2.2, we then derive that

$$\begin{aligned} & L_k V(\varphi, i, t) = V_t(\pi_k(\varphi)(0), i, t) + V_x(\pi_k(\varphi)(0), i, t) f(\pi_k(\varphi), i, t) \\ & + \frac{1}{2}\text{trace}\big(g(\pi_k(\varphi), i, t)^T V_{xx}(\pi_k(\varphi)(0), i, t) g(\pi_k(\varphi), i, t)\big) + \sum_{j=1}^N \gamma_{ij} V(\pi_k(\varphi)(0), j, t) \\ & \le K + K \int_{-\infty}^0 W_1(\pi_k(\varphi)(u), t + u)\mu_1(\mathrm{d}u) \end{aligned}$$

$$- W_2(\varphi(0), t) + \int_{-\infty}^{0} W_2(\pi_k(\varphi)(u), t+u)\mu_2(\mathrm{d}u). \tag{3.3}$$

We can therefore show the following theorem in the same way as Theorem 2.3 was proved.

Theorem 3.1. *Under Assumptions 2.1 and 2.2, for each integer $k \geq 1$ and each $0 < T \leq k$, the truncated SFDE (3.1) has a unique global solution $x^k(t)$ on $t \in (-\infty, T]$. Moreover, there is a positive constant C_T, which is the same as in Theorem 2.3, such that*

$$\sup_{0 \leq t \leq T} \mathbb{E} W_1(x^k(t), t) \leq C_T, \quad \forall\, 0 < T \leq k, \tag{3.4}$$

and

$$\mathbb{P}(\rho_h^k \leq T) \leq \frac{C_T}{\beta(h)} \tag{3.5}$$

for all sufficiently large number h, where $\rho_h^k = \inf\{t \in [0, \rho_e^k) : |x^k(t)| \geq h\}$.

Proof. By virtue of (3.2) we know that the truncated SFDE (3.1) has a unique solution $x^k(t)$ on $t \in (-\infty, \rho_e^k)$, where ρ_e^k is the explosion time. For any sufficient large number h with $h \geq \|\xi\|_r$, ρ_h^k is defined as above. Obviously, ρ_h^k is increasing with respect to h and $\rho_h^k \leq \rho_e^k$ $a.s$. By the generalized Itô formula and (3.3), we obtain that

$$\begin{aligned} \mathbb{E} V(x^k(t \wedge \rho_h^k), \theta(t \wedge \rho_h^k), t \wedge \rho_h^k) &\leq K_1 + K \mathbb{E} \int_0^{t \wedge \rho_h^k} \int_{-\infty}^{0} W_1(\pi_k(x_s^k)(u), s+u)\mu_1(\mathrm{d}u)\mathrm{d}s \\ &- \mathbb{E} \int_0^{t \wedge \rho_h^k} W_2(x^k(s), s)\mathrm{d}s + \mathbb{E} \int_0^{t \wedge \rho_h^k} \int_{-\infty}^{0} W_2(\pi_k(x_s^k)(u), s+u)\mu_2(\mathrm{d}u)\mathrm{d}s, \end{aligned} \tag{3.6}$$

for $t \geq 0$, where $K_1 = \mathbb{E} V(\xi(0), i_0, 0) + KT$. Recalling the definition of π_k one observes

$$\pi_k(x_s^k)(u) = \begin{cases} x^k(s+u), & -k \leq u \leq 0, \\ x^k(s-k), & u < -k. \end{cases}$$

Then for any $t \in [0, T]$, $T \leq k$, we derive that

$$
\begin{aligned}
&\int_0^{t\wedge\rho_h^k}\int_{-\infty}^{0} W_1(\pi_k(x_s^k)(u), s+u)\mu_1(\mathrm{d}u)\mathrm{d}s\\
&=\int_0^{t\wedge\rho_h^k}\int_{-\infty}^{-s} W_1(\pi_k(x_s^k)(u), s+u)\mu_1(\mathrm{d}u)\mathrm{d}s+\int_0^{t\wedge\rho_h^k}\int_{-s}^{0} W_1(x^k(s+u), s+u)\mu_1(\mathrm{d}u)\mathrm{d}s\\
&\le\int_0^{t}\int_{-\infty}^{-s}\Big(\sup_{0\le s\le t}\sup_{-\infty<u\le -s} e^{rb_1(s+u)}W_1(\pi_k(x_s^k)(u), s+u)\Big)e^{-rb_1(s+u)}\mu_1(\mathrm{d}u)\mathrm{d}s\\
&\quad+\int_0^{t\wedge\rho_h^k}\int_{-s}^{0} W_1(x^k(s+u), s+u)\mu_1(\mathrm{d}u)\mathrm{d}s.
\end{aligned}
\tag{3.7}
$$

It follows from (2.12) and $\xi\in\mathcal{C}_r$ that

$$
\begin{aligned}
&\sup_{0\le s\le t}\sup_{-\infty<u\le -s} e^{rb_1(s+u)}W_1(\pi_k(x_s^k)(u), s+u)\\
&\le \hat{K}\sup_{0\le s\le t}\sup_{-\infty<u\le -s} e^{rb_1(s+u)}(1+|\pi_k(x_s^k)(u)|^{b_1})\\
&\le \hat{K}\sup_{0\le s\le t}\Big(1+\sup_{-k\le u\le -s} e^{rb_1(s+u)}\mathbb{E}|x^k(s+u)|^{b_1}\Big)\\
&\le \hat{K}+\hat{K}\sup_{-k\le u\le 0} e^{rb_1u}|\xi(u)|^{b_1}\le K_2,
\end{aligned}
\tag{3.8}
$$

where K_2 is given in (2.13). Inserting the above inequality into (3.7) yields

$$
\begin{aligned}
&\int_0^{t\wedge\rho_h^k}\int_{-\infty}^{0} W_1(\pi_k(x_s^k)(u), s+u)\mu_1(\mathrm{d}u)\mathrm{d}s\\
&\le K_2\int_0^{t}\int_{-\infty}^{-s} e^{-rb_1(s+u)}\mu_1(\mathrm{d}u)\mathrm{d}s+\int_{-t\wedge\rho_h^k}^{0}\int_{-u}^{t\wedge\rho_h^k} W_1(x^k(s+u), s+u)\mathrm{d}s\mu_1(\mathrm{d}u)\\
&\le K_2\int_0^{t} e^{-rb_1s}\mathrm{d}s\int_{-\infty}^{0} e^{-rb_1u}\mu_1(\mathrm{d}u)+\int_{-\infty}^{0}\int_{0}^{t\wedge\rho_h^k} W_1(x^k(s), s)\mathrm{d}s\mu_1(\mathrm{d}u)\\
&\le (K_2/rb_1)\mu_1^{(rb_1)}+\int_0^{t\wedge\rho_h^k} W_1(x^k(s), s)\mathrm{d}s.
\end{aligned}
\tag{3.9}
$$

Similarly,

$$\int_0^{t\wedge\rho_h^k}\int_{-\infty}^0 W_2(\pi_k(x_s^k)(u), s+u)\mu_2(\mathrm{d}u)\mathrm{d}s \le (K_3/rb_2)\mu_2^{rb_2} + \int_0^{t\wedge\rho_h^k} W_2(x^k(s), s)\mathrm{d}s. \tag{3.10}$$

Substituting (3.9) and (3.10) into (3.6) gives

$$\mathbb{E}V(x^k(t\wedge\rho_h^k), \theta(t\wedge\rho_h^k), t\wedge\rho_h^k) \le K_4 + K\int_0^t \mathbb{E}W_1(x^k(s\wedge\rho_h^k), s\wedge\rho_h^k)\mathrm{d}s. \tag{3.11}$$

According to the Gronwall inequality and $W_1(x,t) \le V(x,i,t)$, we have

$$\mathbb{E}W_1(x^k(t\wedge\rho_h^k), t\wedge\rho_h^k) \le C_T, \quad \forall t \in [0,T], \quad T \le k, \tag{3.12}$$

where $C_T = K_4 e^{KT}$. By the same way as Theorem 2.3, we can get the desired assertions. To avoid the duplication we omit the last proof. □

3.2. *Asymptotic approximations*

Recall that our main aim in this paper is to show

$$\lim_{k\to\infty} \mathbb{E}|x(t) - x^k(t)|^q = 0, \quad \forall t > 0, \tag{3.13}$$

for $q \ge 2$. It is even more desired if an error estimate can be obtained. To guarantee the finite qth moment of the solution, we slightly strengthen Assumption 2.2.

Assumption 3.2. Assumption 2.2 holds with $\beta \in \mathcal{K}_\infty$ being defined by $\beta(u) = u^p$ on $u \in \mathbb{R}_+$ for some number $p > 2$.

We also need a slightly stronger local Lipschitz condition.

Assumption 3.3. Let p be the same as in Assumption 3.2. Assume that there is a probability measure $\mu_3 \in \mathcal{P}_b$ with $b > r$ and a positive constant K_h for each $h > 0$ such that

$$|f(\varphi,i,t) - f(\phi,i,t)| \vee |g(\varphi,i,t) - g(\phi,i,t)| \le K_h \int_{-\infty}^0 |\varphi(u) - \phi(u)|\mu_3(\mathrm{d}u) \tag{3.14}$$

for those $\varphi, \phi \in \mathcal{C}_r$ with $\|\varphi\|_r \vee \|\phi\|_r \le h$ and all $(i,t) \in \mathbb{S} \times \mathbb{R}_+$.

Noting that (3.14) implies

$$|f(\varphi,i,t) - f(\phi,i,t)| \vee |g(\varphi,i,t) - g(\phi,i,t)| \le K_h \mu_3^{(r)} \|\varphi - \phi\|_r, \tag{3.15}$$

we see this assumption implies Assumption 2.1. Theorems 2.3 and 3.1 show that under Assumptions 3.2 and 3.3, the solutions of equations (2.2) and (3.1) satisfy

$$\sup_{0\le t\le T}\Big(\mathbb{E}|x(t)|^p \vee \mathbb{E}|x^k(t)|^p\Big) \le C_T, \quad \forall k \ge T > 0, \tag{3.16}$$

while (2.17) and (3.5) become

$$\mathbb{P}(\tau_h \le T) \le \frac{C_T}{h^p} \ \text{ and } \ \mathbb{P}(\rho_h^k \le T) \le \frac{C_T}{h^p} \tag{3.17}$$

respectively. The following theorem further shows that $x^k(\cdot)$ converge to $x(\cdot)$ in L^q.

Theorem 3.4. *Let Assumptions 3.2 and 3.3 hold. Then, for each $q \in [2, p)$, the solutions of equations (2.2) and (3.1) have the property that*

$$\lim_{k\to\infty}\Big(\sup_{0\le t\le T}\mathbb{E}|x(t)-x^k(t)|^q\Big) = 0, \ \ \forall T > 0. \tag{3.18}$$

Proof. Fix $T > 0$ arbitrarily and let $k \ge T$. For each $h > \|\xi\|_r$, let τ_h and ρ_h^k be the same as defined in the proof of Theorem 2.3 (noting that we have already proved $\sigma_e = \infty$ a.s.) and in the statement of Theorem 3.1, respectively. Set

$$\sigma_h^k = \tau_h \wedge \rho_h^k \quad \text{and} \quad e^k(t) = x(t) - x^k(t) \ \text{ for } t \in (-\infty, T].$$

For any $\delta > 0$ and $t \in [0, T]$, we can derive by the Young inequality that

$$\begin{aligned}\mathbb{E}|e^k(t)|^q &= \mathbb{E}\Big(|e^k(t)|^q \mathbf{1}_{\{\sigma_h^k > T\}}\Big) + \mathbb{E}\Big(|e^k(t)|^q \mathbf{1}_{\{\sigma_h^k \le T\}}\Big)\\ &\le \mathbb{E}\Big(|e^k(t)|^q \mathbf{1}_{\{\sigma_h^k > T\}}\Big) + \frac{q\delta}{p}\mathbb{E}|e^k(t)|^p + \frac{p-q}{p\delta^{q/(p-q)}}\mathbb{P}(\sigma_h^k \le T).\end{aligned} \tag{3.19}$$

By (3.16) and (3.17), we have

$$\mathbb{E}|e^k(t)|^p \le 2^p C_T$$

and

$$\mathbb{P}(\sigma_h^k \le T) \le \mathbb{P}(\tau_h \le T) + \mathbb{P}(\rho_h^k \le T) \le \frac{2C_T}{h^p}.$$

We hence have

$$\mathbb{E}|e^k(t)|^q \le \mathbb{E}\Big(|e^k(t)|^q \mathbf{1}_{\{\sigma_h^k > T\}}\Big) + \frac{2^p C_T q\delta}{p} + \frac{2C_T(p-q)}{ph^p\delta^{q/(p-q)}}. \tag{3.20}$$

Now, let $\varepsilon > 0$ be arbitrary. Choosing δ sufficiently small for $2^p C_T q\delta/p \le \varepsilon/3$ and then h sufficiently large for $2C_T(p-q)/(ph^p\delta^{q/(p-q)}) \le \varepsilon/3$, we see from (3.20) that for this particularly chosen h,

$$\mathbb{E}|e^k(t)|^q \le \mathbb{E}\Big(|e^k(t)|^q \mathbf{1}_{\{\sigma_h^k > T\}}\Big) + \frac{2\varepsilon}{3}. \tag{3.21}$$

If we can show that for all sufficiently large k,

$$\sup_{0\le t\le T} \mathbb{E}\Big(|e^k(t)|^q \mathbf{1}_{\{\sigma_h^k > T\}}\Big) \le \frac{\varepsilon}{3}, \tag{3.22}$$

we then have

$$\lim_{k\to\infty}\Big(\sup_{0\le t\le T}\mathbb{E}|e^k(t)|^q\Big)=0,$$

which is the required assertion (3.18). In other words, to complete our proof, all we need to do from now on is to show (3.22) for the particularly chosen h.

For $t\in[0,T]$, it follows from (2.2) and (3.1) as well as the definition of f_k and g_k that

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{0\le v\le t}|e^k(v\wedge\sigma_h^k)|^q\Big)\\
\le&2^{q-1}\mathbb{E}\Big(\sup_{0\le v\le t}\Big|\int_0^{v\wedge\sigma_h^k}[f(x_s,\theta(s),s)-f(\pi_k(x_s^k),\theta(s),s)]\mathrm{d}s\Big|^q\Big)\\
+&2^{q-1}\mathbb{E}\Big(\sup_{0\le v\le t}\Big|\int_0^{v\wedge\sigma_h^k}[g(x_s,\theta(s),s)-g(\pi_k(x_s^k),\theta(s),s)]\mathrm{d}B(s)\Big|^q\Big).
\end{aligned}$$

By the Hölder inequality, the Burkholder-David-Gundy inequality (see, e.g., [16,18]) as well as Assumption 3.3, it is not difficult to show that

$$\mathbb{E}\Big(\sup_{0\le v\le t}|e^k(v\wedge\sigma_h^k)|^q\Big)\le K_5\mathbb{E}\int_0^{t\wedge\sigma_h^k}\Big(\int_{-\infty}^0|x_s(u)-\pi_k(x_s^k)(u)|\mu_3(\mathrm{d}u)\Big)^q\mathrm{d}s, \tag{3.23}$$

where $K_5=2^{q-1}K_h^q[T^{q-1}+(q^{q+1}/2(q-1)^{q-1})^{q/2}T^{(q-2)/2}]$. But, for $0\le s\le t\wedge\sigma_h^k$ (which is less than $T<k$),

$$\int_{-\infty}^0|x_s(u)-\pi_k(x_s^k)(u)|\mu_3(\mathrm{d}u)=\int_{-s}^0|e^k(s+u)|\mu_3(\mathrm{d}u)+J, \tag{3.24}$$

where

$$J=\int_{-\infty}^{-s}|x(s+u)-\pi_k(x_s^k)(u)|\mu_3(\mathrm{d}u).$$

Noting that $x(s+u)=\xi(s+u)$ for $u\le -s$ while

$$\pi_k(x_s^k)(u) = \begin{cases} \xi(s+u) & \text{if } -k \le u \le -s, \\ \xi(s-k) & \text{if } u < -k. \end{cases}$$

We derive that

$$\begin{aligned} J &= \int_{-\infty}^{-k} |\xi(s+u) - \xi(s-k)| \mu_3(\mathrm{d}u) \\ &\le \int_{-\infty}^{-k} \big(|\xi(s+u)| + |\xi(s-k)|\big) \mu_3(\mathrm{d}u) \\ &\le \int_{-\infty}^{-k} \big(\|\xi\|_r e^{-r(s+u)} + \|\xi\|_r e^{-r(s-k)}\big) \mu_3(\mathrm{d}u) \\ &\le 2\|\xi\|_r \int_{-\infty}^{-k} e^{-r(s+u)} \mu_3(\mathrm{d}u) \\ &\le 2\|\xi\|_r \int_{-\infty}^{-k} e^{-bu+(b-r)u} \mu_3(\mathrm{d}u) \\ &\le 2\|\xi\|_r e^{-(b-r)k} \int_{-\infty}^{-k} e^{-bu} \mu_3(\mathrm{d}u) \le K_6 e^{-(b-r)k}, \end{aligned}$$

where $K_6 = 2\|\xi\|_r \mu_3^{(b)}$. Putting this into (3.24) we obtain

$$\int_{-\infty}^{0} |x_s(u) - \pi_k(x_s^k)(u)| \mu_3(\mathrm{d}u) \le \int_{-s}^{0} |e^k(s+u)| \mu_3(\mathrm{d}u) + K_6 e^{-(b-r)k}.$$

Then

$$\Big(\int_{-\infty}^{0} |x_s(u) - \pi_k(x_s^k)(u)| \mu_3(\mathrm{d}u) \Big)^q \le 2^{q-1} \int_{-s}^{0} |e^k(s+u)|^q \mu_3(\mathrm{d}u) + 2^{q-1} K_6^q e^{-q(b-r)k}. \quad (3.25)$$

Substituting this into (3.23) yields

$$\begin{aligned} &\mathbb{E}\Big(\sup_{0\le v\le t} |e^k(v \wedge \sigma_h^k)|^q \Big) \\ \le & 2^{q-1} K_5 K_6^q T e^{-q(b-r)k} + K_5 \mathbb{E} \int_0^{t\wedge\sigma_h^k} \Big(\int_{-s}^{0} |e^k(s+u)|^q \mu_3(\mathrm{d}u) \Big) \mathrm{d}s. \end{aligned} \quad (3.26)$$

But

$$\int_0^{t\wedge\sigma_h^k}\Big(\int_{-s}^{0}|e^k(s+u)|^q\mu_3(\mathrm{d}u)\Big)\mathrm{d}s=\int_{-(t\wedge\sigma_h^k)}^{0}\Big(\int_{-u}^{t\wedge\sigma_h^k}|e^k(s+u)|^q\mathrm{d}s\Big)\mu_3(\mathrm{d}u)$$

$$\le\int_{-(t\wedge\sigma_h^k)}^{0}\Big(\int_0^{t\wedge\sigma_h^k}|e^k(s)|^q\mathrm{d}s\Big)\mu_3(\mathrm{d}u)\le\int_0^{t\wedge\sigma_h^k}|e^k(s)|^q\mathrm{d}s. \tag{3.27}$$

It therefore follows from (3.26) that

$$\mathbb{E}\Big(\sup_{0\le v\le t}|e^k(v\wedge\sigma_h^k)|^q\Big)\le 2^{q-1}K_5K_6^qTe^{-q(b-r)k}+K_5\mathbb{E}\int_0^{t\wedge\sigma_h^k}|e^k(s)|^q\mathrm{d}s$$

$$\le 2^{q-1}K_5K_6^qTe^{-q(b-r)k}+K_5\int_0^t\mathbb{E}\Big(\sup_{0\le v\le s}|e^k(v\wedge\sigma_h^k)|^q\Big)\mathrm{d}s. \tag{3.28}$$

An application of the Gronwall inequality yields

$$\mathbb{E}\Big(\sup_{0\le v\le T}|e^k(v\wedge\sigma_h^k)|^q\Big)\le 2^{q-1}K_5K_6^qTe^{K_5T-q(b-r)k}.$$

Hence

$$\mathbb{E}\Big(\sup_{0\le v\le T}|e^k(v)|^q\mathbf{1}_{\{\sigma_h^k>T\}}\Big)\le 2^{q-1}K_5K_6^qTe^{K_5T-q(b-r)k}. \tag{3.29}$$

This implies (3.22) of course. The proof is therefore complete. □

3.3. Approximations with exponential convergence order

The convergence of $x^k(\cdot)$ to $x(\cdot)$ in Theorem 3.4 is described in the asymptotic way. The proof itself provides us with a way to estimate the error. Namely, for a given $\varepsilon>0$, we can determine a positive integer $k_0=k_0(\varepsilon)$ (by choosing δ, h first) so that

$$\sup_{0\le t\le T}\mathbb{E}|x(t)-x^k(t)|^q<\varepsilon,\quad\forall k\ge k_0.$$

On the other hand, we observe that (3.29) gives an exponential estimate on the error up to the stopping time $\sigma_h^k\wedge T$. If we can remove the stopping time there, we will have a stronger convergence order of $x^k(\cdot)$ to $x(\cdot)$. In this sub-section, we aim to establish the exponential convergence order described by

$$\limsup_{k\to\infty}\frac{1}{k}\log\Big(\mathbb{E}|x(T)-x^k(T)|^q\Big)<0,\quad\forall T>0.$$

We need a couple of new notations. Given a function $\bar{V} \in C^{2,1}(\mathbb{R}^n \times \mathbb{S} \times \mathbb{R}_+; \mathbb{R}_+)$, we define the functional $\mathcal{L}\bar{V} : \mathcal{C}_r \times \mathcal{C}_r \times \mathbb{S} \times \mathbb{R}_+ \to \mathbb{R}$ by

$$\begin{aligned}\mathcal{L}\bar{V}(\varphi, \phi, i, t) &= \bar{V}_t(\varphi(0) - \phi(0), i, t) + \bar{V}_x(\varphi(0) - \phi(0), i, t)[f(\varphi, i, t) - f(\phi, i, t)] \\ &+ \frac{1}{2}\text{trace}\big([g(\varphi, i, t) - g(\phi, i, t)]^T \bar{V}_{xx}(\varphi(0) - \phi(0), i, t)[g(\varphi, i, t) - g(\phi, i, t)]\big) \\ &+ \sum_{j=1}^{N} \gamma_{ij} \bar{V}(\varphi(0) - \phi(0), j, t).\end{aligned}$$

Let us emphasize that $\mathcal{L}\bar{V}$ is defined on $\mathcal{C}_r \times \mathcal{C}_r \times \mathbb{S} \times \mathbb{R}_+$ while $\bar{V}$ on $\mathbb{R}^n \times \mathbb{S} \times \mathbb{R}_+$. For $\beta > 0$, let $\mathcal{U}_{0,\beta}$ denote the family of continuous functions $U : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ such that $U(x, y) = 0$ whenever $x = y \in \mathbb{R}^n$ while

$$\sup_{x,y\in\mathbb{R}^n} \frac{U(x, y)}{1 + |x|^\beta + |y|^\beta} < \infty.$$

It is easy to see

$$\sup_{-\infty<u\le 0} e^{r\beta u} U(\varphi(u), \phi(u)) < \infty, \quad \forall \varphi, \phi \in \mathcal{C}_r. \tag{3.30}$$

For example, if $U(x, y) = |x - y|(1 + |x| + |y|)$ for $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, then $U \in \mathcal{U}_{0,2}$.

Assumption 3.5. There are positive numbers $\beta, \bar{q}, \bar{K}, b_4, b_5$, functions $\bar{V} \in C^{2,1}(\mathbb{R}^n \times \mathbb{S} \times \mathbb{R}_+; \mathbb{R}_+)$, $U \in \mathcal{U}_{0,\beta}$, as well as two probability measures $\mu_4 \in \mathcal{P}_{b_4}$, $\mu_5 \in \mathcal{P}_{b_5}$, such that $b_4 > r\bar{q}$, $b_5 > r\beta$,

$$\bar{V}(0, i, t) = 0, \quad \forall (i, t) \in \mathbb{S} \times \mathbb{R}_+, \tag{3.31}$$

$$|x|^{\bar{q}} \le \bar{V}(x, i, t), \quad \forall (x, i, t) \in \mathbb{R}^n \times \mathbb{S} \times \mathbb{R}_+, \tag{3.32}$$

and

$$\begin{aligned}\mathcal{L}\bar{V}(\varphi, \phi, i, t) \le{}& \bar{K} \int_{-\infty}^{0} |\varphi(u) - \phi(u)|^{\bar{q}} \mu_4(\mathrm{d}u) \\ &- U(\varphi(0), \phi(0)) + \int_{-\infty}^{0} U(\varphi(u), \phi(u))\mu_5(\mathrm{d}u)\end{aligned} \tag{3.33}$$

for all $(\varphi, \phi, i, t) \in \mathcal{C}_r \times \mathcal{C}_r \times \mathbb{S} \times \mathbb{R}_+$.

Remark 3.6. Assumption 3.5 is the generalized Khasminskii-type condition for the continuous dependence of the solutions on the initial data—any two solutions are close to each other in the $\bar{q}$th moment as long as their corresponding initial data are close to each other. There is a large class of SFDEs satisfying this assumption, for examples, the SFDEs with the global Lipschitz coefficients in Section 4.1 and the stochastic functional volatility equation (5.20) in Example 5.11.

Theorem 3.7. *Let Assumptions 2.1, 2.2 and 3.5 hold. Set $\lambda = (b_4 - r\bar{q}) \wedge (b_5 - r\beta)$. Then the solutions of equations (2.2) and (3.1) have the properties that, for all $T > 0$,*

$$\limsup_{k\to\infty} \frac{1}{k} \log\left(\mathbb{E}|x(T) - x^k(T)|^{\bar{q}}\right) \le -\lambda, \tag{3.34}$$

and

$$\limsup_{k\to\infty} \frac{1}{k} \log(|x(T) - x^k(T)|) \le -\frac{\lambda}{\bar{q}} \quad a.s. \tag{3.35}$$

Proof. Fix any $T > 0$ and integer $k > T$. Let $h > \|\xi\|_r$. Let $e^k(t)$ and σ_h^k be the same as defined in the proof of Theorem 3.4. Obviously, $\sigma_h^k \to \infty$ almost surely as $h \to \infty$. By the generalized Itô formula as well as the definitions of f_k, g_k, π_k, it is straightforward to verify that, for $t \in [0, T]$,

$$\mathbb{E}\bar{V}(e^k(t\wedge\sigma_h^k), \theta(t\wedge\sigma_h^k), t\wedge\sigma_h^k) = \mathbb{E}\int_0^{t\wedge\sigma_h^k} \mathcal{L}\bar{V}(x_s, \pi_k(x_s^k), \theta(s), s)\mathrm{d}s.$$

By Assumption 3.5, we then have

$$\begin{aligned}\mathbb{E}|e^k(t\wedge\sigma_h^k)|^{\bar{q}} \le{}& \bar{K}\mathbb{E}\int_0^{t\wedge\sigma_h^k}\Big(\int_{-\infty}^0 |x_s(u) - \pi_k(x_s^k)(u)|^{\bar{q}}\mu_4(\mathrm{d}u)\Big)\mathrm{d}s\\ &- \mathbb{E}\int_0^{t\wedge\sigma_h^k} U(x(s), x^k(s))\mathrm{d}s\\ &+ \mathbb{E}\int_0^{t\wedge\sigma_h^k}\Big(\int_{-\infty}^0 U(x_s(u), \pi_k(x_s^k)(u))\mu_5(\mathrm{d}u)\Big)\mathrm{d}s.\end{aligned} \tag{3.36}$$

In the same way as (3.25) was proved, we can show that

$$\int_{-\infty}^0 |x_s(u) - \pi_k(x_s^k)(u)|^{\bar{q}}\mu_4(\mathrm{d}u) \le \int_{-s}^0 |e^k(s+u)|^{\bar{q}}\mu_4(\mathrm{d}u) + K_7 e^{-(b_4 - r\bar{q})k} \tag{3.37}$$

and

$$\int_{-\infty}^{0} U(x_s(u), \pi_k(x_s^k)(u))\mu_5(\mathrm{d}u)$$
$$\le \int_{-s}^{0} U(x(s+u), x^k(s+u))\mu_5(\mathrm{d}u) + K_8 e^{-(b_5 - r\beta)k}, \tag{3.38}$$

where K_7 and K_8 are both positive constants independent of k. Substituting these into (3.36) and making use of

$$\int_0^{t\wedge\sigma_h^k} \Big(\int_{-s}^{0} |e^k(s+u)|^{\bar q}\mu_4(\mathrm{d}u)\Big)\mathrm{d}s \le \int_0^{t\wedge\sigma_h^k} |e^k(s)|^{\bar q}\mathrm{d}s$$

and

$$\int_0^{t\wedge\sigma_h^k} \Big(\int_{-s}^{0} U(x(s+u), x^k(s+u))\mu_5(\mathrm{d}u)\Big)\mathrm{d}s \le \int_0^{t\wedge\sigma_h^k} U(x(s), x^k(s))\mathrm{d}s$$

(please see (3.27)), we obtain

$$\mathbb{E}|e^k(t\wedge\sigma_h^k)|^{\bar q} \le K_9 e^{-\lambda k} + \bar K \int_0^t \mathbb{E}|e^k(s\wedge\sigma_h^k)|^{\bar q}\mathrm{d}s, \tag{3.39}$$

where λ has been defined in the statement of the theorem and $K_9 = T(\bar K K_7 + K_8)$. An application of the Gronwall inequality implies

$$\mathbb{E}|e^k(T\wedge\sigma_h^k)|^{\bar q} \le K_{10} e^{-\lambda k},$$

where $K_{10} = K_9 e^{\bar K T}$. Letting $h \to \infty$ yields

$$\mathbb{E}|e^k(T)|^{\bar q} \le K_{10} e^{-\lambda k}, \tag{3.40}$$

which implies immediately the first required assertion (3.34).

To show the second assertion (3.35), we let $\varepsilon \in (0, \lambda)$ be arbitrary. It follows from (3.40) that

$$\mathbb{P}\{|e^k(T)|^{\bar q} > e^{-(\lambda-\varepsilon)k}\} \le \frac{\mathbb{E}|e^k(T)|^{\bar q}}{e^{-(\lambda-\varepsilon)k}} \le K_{10} e^{-\varepsilon k}, \quad \forall k > T.$$

By the well-known Borel-Cantelli lemma (see, e.g., [18, Lemma 2.4 on page 7]), we can find a subset $\Omega_0 \subset \Omega$ with $\mathbb{P}(\Omega_0) = 1$ so that for each $\omega \in \Omega_0$, there is an integer $k_0(\omega)$ such that

$$|e^k(T,\omega)|^{\bar q} \le e^{-(\lambda-\varepsilon)k}, \quad \forall k \ge k_0(\omega).$$

This yields

$$\lim_{k\to\infty} \frac{1}{k} \log(|e^k(T,\omega)|) \le -\frac{(\lambda-\varepsilon)}{\bar{q}}, \quad \forall \omega \in \Omega_0.$$

This further implies the another assertion (3.35) as ε is arbitrary while $\mathbb{P}(\Omega_0) = 1$. The proof is hence complete. □

Remark 3.8. We observe that both Assumptions 2.1 and 2.2 were not used explicitly in the proof of Theorem 3.7. But they were in fact used to guarantee both SFDEs (2.2) and (3.1) have their own unique solution. In other words, Theorem 3.7 holds if both Assumptions 2.1 and 2.2 are replaced by the condition that both SFDEs (2.2) and (3.1) have their own unique solution.

4. Important classes of SFDEs

To show the power of the general approximation theory established in the previous section, we will study a couple of important classes of SFDEs and their approximations in this section.

4.1. Global Lipschitz

We start with the class of SFDEs under the global Lipschitz condition.

Assumption 4.1. There are two constants $c_1 > 0$, $p > 2$ and a probability measure $\mu_6 \in \mathcal{P}_{rp}$ such that

$$|f(\varphi,i,t) - f(\phi,i,t)| \vee |g(\varphi,i,t) - g(\phi,i,t)| \le c_1 \int_{-\infty}^{0} |\varphi(u) - \phi(u)|\mu_6(\mathrm{d}u) \tag{4.1}$$

for all $\varphi, \phi \in \mathcal{C}_r$ and $(i,t) \in \mathbb{S} \times \mathbb{R}_+$. Moreover,

$$\sup_{t\in\mathbb{R}_+} \big(|f(0,i,t)| \vee |g(0,i,t)|\big) < \infty, \quad \forall\, i \in \mathbb{S}. \tag{4.2}$$

This assumption implies Assumption 3.3 obviously. It is also easy to verify that Assumption 3.2 is satisfied with $V(x,i,t) = |x|^p$, $W_1(x,t) = |x|^p$, $W_2(x,t) = 0$ etc. To verify Assumption 3.5, we let $\bar{V}(x,i,t) = |x|^{\bar{q}}$ for $\bar{q} \in [2,p)$. From now on, we also let $\delta_0(\cdot)$ be the Dirac measure at 0, which can be regarded as a probability measure on $\mathbb{R}_-$ and it belongs to $\bigcap_{b\ge 1} \mathcal{P}_b$. For $\varphi, \phi \in \mathcal{C}_r$ and $(i,t) \in \mathbb{S} \times \mathbb{R}_+$, we can then derive that

$$\begin{aligned}
&\mathcal{L}\bar{V}(\varphi,\phi,i,t)\\
\le&\bar{q}|\varphi(0)-\phi(0)|^{\bar{q}-2}\Big((\varphi(0)-\phi(0))^T(f(\varphi,i,t)-f(\phi,i,t)) + \frac{\bar{q}-1}{2}|g(\varphi,i,t)-g(\phi,i,t)|^2\Big)\\
\le&0.5\bar{q}(\bar{q}-1)|\varphi(0)-\phi(0)|^{\bar{q}} + \bar{q}\big(|f(\varphi,i,t)-f(\phi,i,t)|^{\bar{q}} \vee |g(\varphi,i,t)-g(\phi,i,t)|^{\bar{q}}\big)\\
\le&0.5\bar{q}(\bar{q}-1)\int_{-\infty}^{0} |\varphi(u)-\phi(u)|^{\bar{q}}\delta_0(\mathrm{d}u) + \bar{q}c_1^{\bar{q}}\int_{-\infty}^{0} |\varphi(u)-\phi(u)|^{\bar{q}}\mu_6(\mathrm{d}u)
\end{aligned}$$

$$\leq [0.5\bar{q}(\bar{q}-1)+\bar{q}c_1^{\bar{q}}]\int_{-\infty}^{0} |\varphi(u)-\phi(u)|^{\bar{q}}(\delta_0(\mathrm{d}u)+\mu_6(\mathrm{d}u))$$

$$=\bar{q}[\bar{q}-1+2c_1^{\bar{q}}]\int_{-\infty}^{0} |\varphi(u)-\phi(u)|^{\bar{q}}\bar{\mu}_6(\mathrm{d}u), \tag{4.3}$$

where $\bar{\mu}_6(\cdot)=0.5(\mu_6(\cdot)+\delta_0(\cdot))\in\mathcal{P}_{rp}$. This shows that Assumption 3.5 is satisfied with $U(x)=0$ etc. By Theorem 3.7, we can therefore conclude that under Assumption 4.1, the solutions of equations (2.2) and (3.1) have properties (3.34) and (3.35) with $\lambda=(p-\bar{q})r$ for any $\bar{q}\in[2,p)$.

4.2. Khasminiskii case

A much wider class of SFDEs is covered by the Khasminskii condition than the global Lipschitz condition. We here propose a special Khasminskii-type condition.

Assumption 4.2. There are constants $c_2>0$, $\bar{p}>2$ and a probability measure $\mu_7\in\mathcal{P}_{r\bar{p}}$ such that

$$(\varphi(0)-\phi(0))^T[f(\varphi,i,t)-f(\phi,i,t)]+\frac{\bar{p}-1}{2}|g(\varphi,i,t)-g(\phi,i,t)|^2$$
$$\leq c_2\int_{-\infty}^{0} |\varphi(u)-\phi(u)|^2\mu_7(\mathrm{d}u) \tag{4.4}$$

for all $\varphi,\phi\in\mathcal{C}_r$ and $(i,t)\in\mathbb{S}\times\mathbb{R}_+$. Moreover, condition (4.2) holds.

This assumption does not in general imply the local Lipschitz continuity of both f and g, we hence need to retain Assumption 2.1. But this assumption does imply Assumptions 2.2 and 3.5. Let us verify Assumption 2.2 first. Let $V(x,i,t)=|x|^p$ for any $p\in(2,\bar{p})$. For $(\varphi,i,t)\in\mathcal{C}_r\times\mathbb{S}\times\mathbb{R}_+$, we have

$$LV(\varphi,i,t)\leq p|\varphi(0)|^{p-2}\Big(\varphi(0)^T f(\varphi,i,t)+\frac{p-1}{2}|g(\varphi,i,t)|^2\Big).$$

Making use of the elementary inequality

$$|a+b|^2\leq\frac{\bar{p}-1}{p-1}a^2+\frac{\bar{p}-1}{\bar{p}-p}b^2,\quad \forall a,b\geq 0$$

as well as Assumption 4.2, we can get

$$LV(\varphi,i,t)\leq p|\varphi(0)|^{p-2}\Big(c_3+c_3|\varphi(0)|+c_2\int_{-\infty}^{0} |\varphi(u)|^2\mu_7(\mathrm{d}u)\Big),$$

where c_3 and the following c_4, c_5 are all positive constants. By the well-known Young inequality, we can further get

$$
\begin{aligned}
LV(\varphi, i, t) &\le c_4\Big(1 + \int_{-\infty}^{0} |\varphi(u)|^p \delta_0(\mathrm{d}u) + \int_{-\infty}^{0} |\varphi(u)|^p \mu_7(\mathrm{d}u)\Big) \\
&\le 2c_4\Big(1 + \int_{-\infty}^{0} |\varphi(u)|^p \bar{\mu}_7(\mathrm{d}u)\Big),
\end{aligned}
\tag{4.5}
$$

where $\bar{\mu}_7(\cdot) = 0.5(\delta_0(\cdot) + \mu_7(\cdot)) \in \mathcal{P}_{r\bar{p}}$. This shows that Assumption 2.2 is satisfied with $W_1(x, t) = |x|^p$, $W_2(x, t) = 0$, etc. To verify Assumption 3.5, we let $\bar{V}(x, i, t) = |x|^{\bar{q}}$ for $\bar{q} \in [2, \bar{p})$. For $\varphi, \phi \in \mathcal{C}_r$ and $(i, t) \in \mathbb{S} \times \mathbb{R}_+$, it then follows from the first inequality in (4.3) and Assumption 4.2 that

$$
\begin{aligned}
&\mathcal{L}\bar{V}(\varphi, \phi, i, t) \\
\le& \bar{q}c_2|\varphi(0) - \phi(0)|^{\bar{q}-2} \int_{-\infty}^{0} |\varphi(u) - \phi(u)|^2 \mu_7(\mathrm{d}u) \\
\le& c_5\Big(\int_{-\infty}^{0} |\varphi(u) - \phi(u)|^{\bar{q}} \delta_0(\mathrm{d}u) + \int_{-\infty}^{0} |\varphi(u) - \phi(u)|^{\bar{q}} \mu_7(\mathrm{d}u)\Big) \\
=& 2c_5 \int_{-\infty}^{0} |\varphi(u) - \phi(u)|^{\bar{q}} \bar{\mu}_7(\mathrm{d}u).
\end{aligned}
\tag{4.6}
$$

This shows that Assumption 3.5 is satisfied with $U(x) = 0$ etc. By Theorem 3.7, we can conclude under Assumptions 2.1 and 4.2, the assertions of Theorem 3.7 hold with $\lambda = (\bar{p} - \bar{q})r$ for any $\bar{q} \in [2, \bar{p})$.

4.3. Highly nonlinear SFDEs

In this subsection we will consider a class of highly nonlinear SFDEs in the form

$$
\mathrm{d}x(t) = F(x(t), \psi_1(x_t), \theta(t), t)\mathrm{d}t + G(x(t), \psi_2(x_t), \theta(t), t)\mathrm{d}B(t) \tag{4.7}
$$

on $t \ge 0$ with the initial data (2.3). Here $F : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{S} \times \mathbb{R}_+ \to \mathbb{R}^n$ and $G : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{S} \times \mathbb{R}_+ \to \mathbb{R}^{n\times m}$ are Borel measurable while $\psi_1, \psi_2 : \mathcal{C}_r \to \mathbb{R}^n$ are defined by

$$
\psi_1(\varphi) = \int_{-\infty}^{0} \varphi(u)\mu_8(\mathrm{d}u) \quad \text{and} \quad \psi_2(\varphi) = \int_{-\infty}^{0} \varphi(u)\mu_9(\mathrm{d}u) \tag{4.8}
$$

with $\mu_8, \mu_9 \in \mathcal{P}_0$ (which will be strengthened later). Equation (4.7) becomes our underlying SFDE (2.2) if we define $f : \mathcal{C}_r \times \mathbb{S} \times \mathbb{R}_+ \to \mathbb{R}^n$ and $g : \mathcal{C}_r \times \mathbb{S} \times \mathbb{R}_+ \to \mathbb{R}^{n\times m}$ by

$$f(\varphi, i, t) = F(\varphi(0), \psi_1(\varphi), i, t) \ \text{ and } \ g(\varphi, i, t) = G(\varphi(0), \psi_2(\varphi), i, t) \tag{4.9}$$

respectively. Moreover, the corresponding truncated SFDE (3.1) becomes

$$\begin{aligned} \mathrm{d}x^k(t) &= f(\pi_k(x_t^k), \theta(t), t)\mathrm{d}t + g(\pi_k(x_t^k), \theta(t), t)\mathrm{d}B(t) \\ &= F(x^k(t), \psi_1(\pi_k(x_t^k)), \theta(t), t)\mathrm{d}t + G(x^k(t), \psi_2(\pi_k(x_t^k)), \theta(t), t)\mathrm{d}B(t) \end{aligned} \tag{4.10}$$

on $t \ge 0$ with initial data $x_0^k = \pi_k(\xi)$ and $\theta(0) = i_0$.

To apply our theory established in the previous sections, we impose the local Lipschitz condition on F and G.

Assumption 4.3. For each $h > 0$, there is a positive constant $\tilde{K}_h$ such that

$$|F(x, y, i, t) - F(\bar{x}, \bar{y}, i, t)| \vee |G(x, y, i, t) - G(\bar{x}, \bar{y}, i, t)| \le \tilde{K}_h(|x - \bar{x}| + |y - \bar{y}|) \tag{4.11}$$

for all $x, y, \bar{x}, \bar{y} \in \mathbb{R}^n$ with $|x| \vee |y| \vee |\bar{x}| \vee |\bar{y}| \le h$ and $(i, t) \in \mathbb{S} \times \mathbb{R}_+$.

We also impose a generalized Khasminskii-type condition (see, e.g., [19]).

Assumption 4.4. There are nonnegative constants c_6 - c_9, p, $\bar{p}$ such that $c_7 \ge c_8 + c_9$, $p > 2$, $\bar{p} > 0$ and

$$\begin{aligned} &x^T F(x, y, i, t) + \frac{p-1}{2}|G(x, z, i, t)|^2 \\ \le& c_6(1 + |x|^2 + |y|^2 + |z|^2) - c_7|x|^{2+\bar{p}} + c_8|y|^{2+\bar{p}} + c_9|z|^{2+\bar{p}} \end{aligned} \tag{4.12}$$

for all $(x, y, z, i, t) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{S} \times \mathbb{R}_+$. Moreover, $\mu_8, \mu_9 \in \mathcal{P}_{r(p+\bar{p})}$.

We first verify Assumption 3.2 under Assumption 4.4. Let $V(x, i, t) = |x|^p$. Then

$$LV(\varphi, i, t) \le p|\varphi(0)|^{p-2}\Big(\varphi(0)^T F(\varphi(0), \psi_1(\varphi), i, t) + \frac{p-1}{2}|G(\varphi(0), \psi_2(\varphi), i, t)|^2\Big),$$

for $(\varphi, i, t) \in \mathcal{C}_r \times \mathbb{S} \times \mathbb{R}_+$. By Assumption 4.4,

$$\begin{aligned} LV(\varphi, i, t) \le& p|\varphi(0)|^{p-2}\Big(c_6(1 + |\varphi(0)|^2 + |\psi_1(\varphi)|^2 + |\psi_2(\varphi)|^2) \\ &- c_7|\varphi(0)|^{2+\bar{p}} + c_8|\psi_1(\varphi)|^{2+\bar{p}} + c_9|\psi_2(\varphi)|^{2+\bar{p}}\Big). \end{aligned}$$

By the Young inequality, we can then easily show that

$$\begin{aligned} LV(\varphi, i, t) \le& 2pc_6(1 + |\varphi(0)|^p + |\psi_1(\varphi)|^p + |\psi_2(\varphi)|^p) \\ &- \bar{c}_7|\varphi(0)|^{p+\bar{p}} + \bar{c}_8|\psi_1(\varphi)|^{p+\bar{p}} + \bar{c}_9|\psi_2(\varphi)|^{p+\bar{p}}, \end{aligned}$$

where $\bar{c}_7 = pc_7 - p(c_8 + c_9)(p-2)/(p+\bar{p})$, $\bar{c}_8 = pc_8(2+\bar{p})/(p+\bar{p})$ and $\bar{c}_9 = pc_9(2+\bar{p})/(p+\bar{p})$ so $\bar{c}_7 \geq \bar{c}_8 + \bar{c}_9$. Using $|\psi_1(\varphi)|^p \leq \int_{-\infty}^0 |\varphi(u)|^p \mu_8(\mathrm{d}u)$, we further get

$$
\begin{aligned}
LV(\varphi, i, t) \leq & 4pc_6\Big(1 + \int_{-\infty}^0 |\varphi(u)|^p \delta_0(\mathrm{d}u) + \int_{-\infty}^0 |\varphi(u)|^p \mu_8(\mathrm{d}u) + \int_{-\infty}^0 |\varphi(u)|^p \mu_9(\mathrm{d}u)\Big) \\
& - \bar{c}_7|\varphi(0)|^{p+\bar{p}} + \bar{c}_8 \int_{-\infty}^0 |\varphi(u)|^{p+\bar{p}} \mu_8(\mathrm{d}u) + \bar{c}_9 \int_{-\infty}^0 |\varphi(u)|^{p+\bar{p}} \mu_9(\mathrm{d}u) \\
\leq & 4pc_6\Big(1 + 3\int_{-\infty}^0 |\varphi(u)|^p \mu_{10}(\mathrm{d}u)\Big) \\
& - \bar{c}_7|\varphi(0)|^{p+\bar{p}} + \bar{c}_7 \int_{-\infty}^0 |\varphi(u)|^{p+\bar{p}} \mu_{11}(\mathrm{d}u),
\end{aligned}
$$

where $\mu_{10}(\cdot) = [\delta_0(\cdot) + \mu_8(\cdot) + \mu_9(\cdot)]/3 \in \mathcal{P}_{r(p+\bar{p})}$ and $\mu_{11}(\cdot) = [\bar{c}_8\mu_8(\cdot) + \bar{c}_9\mu_9(\cdot)]/(\bar{c}_8 + \bar{c}_9) \in \mathcal{P}_{r(p+\bar{p})}$. We therefore see that Assumption 3.2 is satisfied with $W_1(x,t) = |x|^p$, $W_2(x,t) = \bar{c}_7|x|^{p+\bar{p}}$ etc.

Let us now verify Assumption 3.3. Let $h > 0$ and set $\bar{h} = h(\mu_8^{(r)} \vee \mu_9^{(r)})$. For $\varphi, \phi \in \mathcal{C}_r$ with $\|\varphi\|_r \vee \|\phi\|_r \leq h$ and $(i,t) \in \mathbb{S} \times \mathbb{R}_+$, we derive from Assumption 4.3 that

$$
\begin{aligned}
& |F(\varphi(0), \psi_1(\varphi), i, t) - F(\phi(0), \psi_1(\phi), i, t)| \vee |G(\varphi(0), \psi_2(\varphi), i, t) - G(\phi(0), \psi_2(\phi), i, t)| \\
\leq & \tilde{K}_h\Big(\int_{-\infty}^0 |\varphi(u) - \phi(u)|\delta_0(\mathrm{d}u) + \int_{-\infty}^0 |\varphi(u) - \phi(u)|\mu_8(\mathrm{d}u) + \int_{-\infty}^0 |\varphi(u) - \phi(u)|\mu_9(\mathrm{d}u)\Big) \\
= & 3\tilde{K}_h \int_{-\infty}^0 |\varphi(u) - \phi(u)|\mu_{10}(\mathrm{d}u),
\end{aligned}
$$

where μ_{10} has been defined above. We can therefore conclude from Theorem 3.4 that under Assumptions 4.3 and 4.4, the solutions of equations (4.7) and (4.10) have property (3.18) for each $q \in [2, p)$.

5. Numerical methods

The theory established in the previous sections enables us to obtain numerical approximate solutions to SFDEs with infinite delay. More precisely, in order to numerically approximate the solution of the SFDE (2.2), we can now obtain the numerical solution of the corresponding truncated SFDE (3.1) for a sufficiently large k.

As pointed out before, the truncated SFDE (3.1) is an SFDE with finite delay. Although numerical methods for the SFDEs with finite delay have been studied by many authors (see, e.g., [7,30–32,35]), the existing results can not be applied directly to the SFDE (3.1) due to its special truncated feature. Fortunately, numerical methods can be modified to obtain the numerical solutions of the truncated SFDE (3.1). To demonstrate the idea, we will concentrate on obtaining the numerical solutions of the truncated SFDE (4.10) and hence the numerical solutions of the SFDE (4.7).

5.1. Lipschitz case

How the existing numerical analysis can be modified to obtain the numerical solutions of the truncated SFDE (4.10) is best illustrated in the globally Lipschitz case.

Assumption 5.1. There exists a constant $L_1 > 0$ such that

$$|F(x, y, i, t) - F(\bar{x}, \bar{y}, i, t)| \vee |G(x, y, i, t) - G(\bar{x}, \bar{y}, i, t)| \le L_1(|x - \bar{x}| + |y - \bar{y}|)$$

for all $x, y, \bar{x}, \bar{y} \in \mathbb{R}^n$, $t \in \mathbb{R}_+$ and $i \in \mathbb{S}$. Moreover,

$$\sup_{t\in\mathbb{R}_+} (|F(0, 0, i, t)| \vee |G(0, 0, i, t)|) < \infty, \quad \forall\, i \in \mathbb{S}.$$

This assumption of course guarantees that the truncated SFDE (4.10) has a unique global solution $x^k(t)$. Let us now fix a sufficiently large k and apply the EM method (see, e.g., [30]) to the truncated SFDE (4.10) to obtain its numerical solutions. Let k_1 be a positive integer and set the step size $\Delta = 1/k_1$. Let $t_j = j\Delta$ for $j = -kk_1, -(kk_1 - 1), \cdots, -1, 0, 1, \cdots$. We first need to form discrete-time numerical approximations $X^k_\Delta(t_j) \approx x^k(t_j)$ for $j \ge 0$ given the initial data $x^k_0 = \pi_k(\xi)$ and $\theta(0) = i_0$. Recall that $\pi_k(\xi)$ depends only on the values $\xi(u)$ for $u \in [-k, 0]$. Accordingly, we set $X^k_\Delta(t_j) = \xi(t_j)$ for $j = -kk_1, \cdots, -1, 0$ and form $X^k_\Delta(t_{j+1})$ for $j \ge 0$ by

$$X^k_\Delta(t_{j+1}) = X^k_\Delta(t_j) + F(X^k_\Delta(t_j), \psi_{1j}, \theta(t_j), t_j)\Delta + G(X^k_\Delta(t_j), \psi_{2j}, \theta(t_j), t_j)\Delta B_j, \tag{5.1}$$

where $\Delta B_j = B(t_{j+1}) - B(t_j)$, and

$$\psi_{1j} = \sum_{h=-kk_1}^{-1} X^k_\Delta(t_{j+h})\mu_8([t_h, t_{h+1})) + X^k_\Delta(t_{j-kk_1})\mu_8((-\infty, -k)).$$

Here ψ_{2j} is defined as ψ_{1j} by replacing μ_8 with μ_9. Note that $\{\theta(t_j)\}_{j\ge 0}$ is a discrete-time Markov chain starting from $\theta(0) = r_0$ with the one-step transition probability matrix $e^{\Delta\Gamma}$. The numerical simulation of $\{\theta(t_j)\}_{j\ge 0}$ can be performed in the way as described in [20, p.112]. We next form the continuous-time numerical solution

$$X^k_\Delta(t) = \sum_{j=0}^{\infty} X^k_\Delta(t_j)\mathbf{1}_{[t_j, t_{j+1})}(t), \quad t \ge 0. \tag{5.2}$$

Although this is the numerical solution we usually compute in practice, the numerical analysis is carried out via the continuous auxiliary process defined by

$$\bar{X}^k_\Delta(t) := \xi(0) + \int_0^t F(X^k_\Delta(s), \psi_1(s), \bar{\theta}(s), \tau(s))\mathrm{d}s + \int_0^t G(X^k_\Delta(s), \psi_2(s), \bar{\theta}(s), \tau(s))\mathrm{d}B(s), \tag{5.3}$$

for $t \geq 0$ while set $\bar{X}^k_\Delta(t) = \xi(t)$ for $t \in [-k, 0]$, where

$$\tau(t) := t_j, \quad \bar{\theta}(t) := \theta(t_j), \quad \psi_1(t) := \psi_{1j}, \quad \psi_2(t) := \psi_{2j}, \qquad t \in [t_j, t_{j+1}).$$

Note that $\bar{X}^k_\Delta(t_j) = X^k_\Delta(t_j)$ for all j. That is, $\bar{X}^k_\Delta(t)$ coincides with the numerical solution $X^k_\Delta(t)$ at the grid-points. We need two more assumptions.

Assumption 5.2. There exist constants $\alpha \in [1/2, 1]$ and $L_2 > 0$ such that

$$|F(x, y, i, t_1) - F(x, y, i, t_2)| \vee |G(x, y, i, t_1) - G(x, y, i, t_2)| \leq L_2(1 + |x| + |y|)|t_1 - t_2|^\alpha$$

for all $x, y \in \mathbb{R}^n$, $t_1, t_2 \in \mathbb{R}_+$ and $i \in \mathbb{S}$.

Assumption 5.3. There exist constants $\beta \geq 1/2$ and $L_3 > 0$ such that the initial function ξ satisfies

$$|\xi(s_1) - \xi(s_2)| \leq L_3|s_1 - s_2|^\beta, \quad \forall s_1, s_2 \in (-\infty, 0].$$

In the remaining of this section we fix $p \geq 2$, $T > 0$ and $k > T$ arbitrarily and let C stand for a universal positive constant dependent on p, T, ξ etc. but independent of Δ and k. Let us present a number of useful lemmas.

Lemma 5.4. *Suppose that Assumption 5.1 holds and $\mu_8, \mu_9 \in \mathcal{P}_r$ (please recall (4.8) regarding μ_8 and μ_9). Then*

$$\sup_{0<\Delta\leq 1} \mathbb{E}\Big(\sup_{t\in[0,T]} |\bar{X}^k_\Delta(t)|^p\Big) \leq C. \tag{5.4}$$

Proof. Fix $\Delta \in (0, 1]$ arbitrarily. It is standard (see, e.g., [18,20]) to show from (5.3) along with Assumption 5.1 that for $t \in [0, T]$,

$$\mathbb{E}\Big(\sup_{0\leq u\leq t} |\bar{X}^k_\Delta(u)|^p\Big) \leq C + C\mathbb{E}\int_0^t (|X^k_\Delta(s)|^p + |\psi_1(s)|^p + |\psi_2(s)|^p)\mathrm{d}s. \tag{5.5}$$

For each $s \in [0, T]$, there is a unique j such that $s \in [t_j, t_{j+1})$ and $\psi_1(s) = \psi_{1j}$. By the definition of ψ_{1j} and $\mu_8 \in \mathcal{P}_r$, we further derive

$$\begin{aligned}
|\psi_{1j}| &= |\sum_{h=-kk_1}^{-1} X^k_\Delta(t_{j+h})\mu_8([t_h, t_{h+1})) + X^k_\Delta(t_{j-kk_1})\mu_8((-\infty, -k))| \\
&\leq \sum_{h=-kk_1}^{-1} |X^k_\Delta(t_{j+h})|\mu_8([t_h, t_{h+1})) + |X^k_\Delta(t_{j-kk_1})|\mu_8((-\infty, -k))
\end{aligned}$$

$$
\begin{aligned}
&\leq \sum_{h=-kk_1}^{-1} e^{rt_h}|X_\Delta^k(t_{j+h})|e^{-rt_h}\mu_8([t_h,t_{h+1})) + e^{rt_{-kk_1}}|X_\Delta^k(t_{j-kk_1})|e^{-rt_{-kk_1}}\mu_8((-\infty,-k)) \\
&\leq \Big(\sup_{h\leq 0} e^{rt_h}|X_\Delta^k(t_{j+h})|\Big)\Big(\sum_{h=-kk_1}^{-1} e^{-rt_h}\mu_8([t_h,t_{h+1})) + e^{-rt_{-kk_1}}\mu_8((-\infty,-k))\Big) \\
&\leq \Big(\sup_{h\leq 0} e^{rt_h}|X_\Delta^k(t_{j+h})|\Big)\int_{-\infty}^{0} e^{r\Delta - ru}\mu_8(\mathrm{d}u) \\
&\leq \Big(\sup_{h\leq 0} e^{rt_h}|X_\Delta^k(t_{j+h})|\Big)e^r\mu_8^{(r)}.
\end{aligned} \tag{5.6}
$$

Note that

$$
\begin{aligned}
&\sup_{h\leq 0} e^{rt_h}|X_\Delta^k(t_{j+h})| = e^{-rt_j}\sup_{h\leq j} e^{rt_h}|X_\Delta^k(t_h)| \\
\leq\ & e^{-rt_j}\Big(\sup_{h\leq 0} e^{rt_h}|X_\Delta^k(t_h)| + \sup_{0\leq h\leq j} e^{rt_h}|X_\Delta^k(t_h)|\Big) \\
\leq\ & e^{-rt_j}\Big(\sup_{\theta\leq 0} e^{r\theta}|\xi(\theta)| + \sup_{0\leq h\leq j} e^{rt_h}|X_\Delta^k(t_h)|\Big) \\
\leq\ & C + \sup_{0\leq h\leq j} e^{r(t_h-t_j)}|X_\Delta^k(t_h)| \leq C + \sup_{0\leq h\leq j}|X_\Delta^k(t_h)|
\end{aligned}
$$

Inserting the above inequality into (5.6) gives

$$
|\psi_{1j}| \leq C + \sup_{0\leq h\leq j}|X_\Delta^k(t_h)|.
$$

Consequently

$$
|\psi_1(s)|^p \leq C\Big(1 + \sup_{0\leq u\leq s}|\bar{X}_\Delta^k(u)|^p\Big). \tag{5.7}
$$

Similarly

$$
|\psi_2(s)|^p \leq C\Big(1 + \sup_{0\leq u\leq s}|\bar{X}_\Delta^k(u)|^p\Big). \tag{5.8}
$$

Substituting these into (5.5), we obtain

$$
\mathbb{E}\big(\sup_{0\leq u\leq t}|\bar{X}_\Delta^k(u)|^p\big) \leq C + C\int_0^t \mathbb{E}\big(\sup_{0\leq u\leq s}|\bar{X}_\Delta^k(u)|^p\big)\mathrm{d}s.
$$

An application of the Gronwall inequality implies

$$\mathbb{E}\Big(\sup_{0\le u\le T}|\bar{X}^k_\Delta(u)|^p\Big)\le C.$$

As Δ is arbitrary, we must have the desired assertion (5.4). □

Remark 5.5. By virtue of Lemma 5.4, it follows from (5.7) and (5.8) that

$$\sup_{t\in[0,T]}\mathbb{E}|\psi_i(t)|^p\le C,\quad i=1,2.$$

Lemma 5.6. *Suppose that all conditions of Lemma 5.4 hold. Then for any $\Delta\in(0,1]$,*

$$\mathbb{E}|\bar{X}^k_\Delta(t)-X^k_\Delta(t)|^p\le C\Delta^{p/2},\ \forall\, t\in[0,T]$$

Proof. Fix any $\Delta\in(0,1]$. For each $t\in[0,T]$, there exists a unique integer $j\ge 0$ such that $t_j\le t<t_{j+1}$. Recalling the definitions of $\bar{X}^k_\Delta(\cdot)$ and $X^k_\Delta(\cdot)$ we derive from (5.3) along with Assumption 5.1 easily that

$$\begin{aligned}&\mathbb{E}|\bar{X}^k_\Delta(t)-X^k_\Delta(t)|^p\\ &\le 2^{p-1}\Bigg(\mathbb{E}\Big|\int_{t_j}^t F(X^k_\Delta(s),\psi_1(s),\bar\theta(s),\tau(s))\mathrm{d}s\Big|^p+\mathbb{E}\Big|\int_{t_j}^t G(X^k_\Delta(s),\psi_2(s),\bar\theta(s),\tau(s))\mathrm{d}B(s)\Big|^p\Bigg)\\ &\le C\Delta^{(p-1)/2}\int_{t_j}^t\big(1+\mathbb{E}|X^k_\Delta(s)|^p+\mathbb{E}|\psi_1(s)|^p+\mathbb{E}|\psi_2(s)|^p\big)\mathrm{d}s.\end{aligned}$$

By Lemma 5.4 and Remark 5.5 we obtain the assertion. □

Lemma 5.7. *Let Assumptions 5.1, 5.2, 5.3 hold and $\mu_8,\mu_9\in\mathcal{P}_r$. Then for any $\Delta\in(0,1]$*

$$\mathbb{E}|x^k(t)-\bar{X}^k_\Delta(t)|^2\le C\Delta,\quad \forall\, t\in[0,T]. \tag{5.9}$$

Proof. Fix $\Delta\in(0,1]$ arbitrarily. Let $e_\Delta(t)=x^k(t)-\bar{X}^k_\Delta(t)$ for $t\in[0,T]$. It is straightforward to see that

$$\mathbb{E}|e_\Delta(t)|^2\le C(I(t)+J(t)), \tag{5.10}$$

where

$$I(t):=\mathbb{E}\int_0^t|F(x^k(s),\psi_1(\pi_k(x^k_s)),\theta(s),s)-F(X^k_\Delta(s),\psi_1(s),\bar\theta(s),\tau(s))|^2\mathrm{d}s,$$

$$J(t):=\mathbb{E}\int_0^t|G(x^k(s),\psi_2(\pi_k(x^k_s)),\theta(s),s)-G(X^k_\Delta(s),\psi_2(s),\bar\theta(s),\tau(s))|^2\mathrm{d}s.$$

It is also easy to see that

$$I(t) \le C(I_1(t) + I_2(t) + I_3(t)), \tag{5.11}$$

where

$$I_1(t) := \mathbb{E}\int_0^t |F(x^k(s), \psi_1(\pi_k(x_s^k)), \theta(s), s) - F(X_\Delta^k(s), \psi_1(s), \theta(s), s)|^2 \mathrm{d}s,$$

$$I_2(t) := \mathbb{E}\int_0^t |F(X_\Delta^k(s), \psi_1(s), \theta(s), s) - F(X_\Delta^k(s), \psi_1(s), \bar{\theta}(s), s)|^2 \mathrm{d}s,$$

$$I_3(t) := \mathbb{E}\int_0^t |F(X_\Delta^k(s), \psi_1(s), \bar{\theta}(s), s) - F(X_\Delta^k(s), \psi_1(s), \bar{\theta}(s), \tau(s))|^2 \mathrm{d}s.$$

By Assumption 5.1 and Lemma 5.6, we have

$$\begin{aligned} I_1(t) &\le 2L_1^2\mathbb{E}\int_0^t \Big(|x^k(s) - X_\Delta^k(s)|^2 + |\psi_1(\pi_k(x_s^k)) - \psi_1(s)|^2\Big)\mathrm{d}s \\ &\le C\Delta + C\mathbb{E}\int_0^t |e_\Delta(s)|^2 \mathrm{d}s + C\mathbb{E}\int_0^t |\psi_1(\pi_k(x_s^k)) - \psi_1(s)|^2 \mathrm{d}s. \end{aligned} \tag{5.12}$$

For each $t > 0$, let $N = \lfloor t/\Delta \rfloor$ and $t_{N+1} = t$ for a meanwhile. Recalling the definition of π_k, using the Hölder inequality and Assumption 5.3, we derive that

$$\begin{aligned} &\mathbb{E}\int_0^t |\psi_1(\pi_k(x_s^k)) - \psi_1(s)|^2 \mathrm{d}s = \sum_{j=0}^N \mathbb{E}\int_{t_j}^{t_{j+1}} |\psi_1(\pi_k(x_s^k)) - \psi_{1j}|^2 \mathrm{d}s \\ &= \sum_{j=0}^N \mathbb{E}\int_{t_j}^{t_{j+1}} \Big| \int_{-k}^0 x^k(s+u)\mu_8(\mathrm{d}u) + x^k(s-k)\mu_8((-\infty, -k)) \\ &\quad - \sum_{h=-kk_1}^{-1} X_\Delta^k(t_{j+h})\mu_8(t_h, t_{h+1}) - X_\Delta^k(t_j - k)\mu_8((-\infty, -k))\Big|^2 \mathrm{d}s \\ &\le \sum_{j=0}^N \mathbb{E}\int_{t_j}^{t_{j+1}} \Big(\sum_{h=-kk_1}^{-1} \int_{t_h}^{t_{h+1}} |x^k(s+u) - X_\Delta^k(t_j + t_h)|\mu_8(\mathrm{d}u) \\ &\quad + |x^k(s-k) - X_\Delta^k(t_j - k)|\mu_8((-\infty, -k))\Big)^2 \mathrm{d}s \end{aligned}$$

$$\begin{aligned}
&\leq 2\sum_{j=0}^{N}\mathbb{E}\int_{t_j}^{t_{j+1}}\Big(\sum_{h=-kk_1}^{-1}\int_{t_h}^{t_{h+1}}|x^k(s+u)-X_\Delta^k(t_j+u)|\mu_8(\mathrm{d}u)\Big)^2\mathrm{d}s\\
&\quad+2\sum_{j=0}^{N}\mathbb{E}\int_{t_j}^{t_{j+1}}\Big(|\xi(s-k)-\xi(t_j-k)|\mu_8((-\infty,-k))\Big)^2\mathrm{d}s\\
&\leq 2\sum_{j=0}^{N}\mathbb{E}\int_{t_j}^{t_{j+1}}\int_{-k}^{0}|x^k(s+u)-X_\Delta^k(t_j+u)|^2\mu_8(\mathrm{d}u)\mathrm{d}s+2TL_3^2\Delta^{2\beta}\\
&\leq 4\sum_{j=0}^{N}\mathbb{E}\int_{t_j}^{t_{j+1}}\int_{-k}^{-s}|\xi(s+u)-\xi(t_j+u)|^2\mu_8(\mathrm{d}u)\mathrm{d}s+2TL_3^2\Delta^{2\beta}\\
&\quad+4\sum_{j=0}^{N}\mathbb{E}\int_{t_j}^{t_{j+1}}\int_{-s}^{0}|x^k(s+u)-X_\Delta^k(t_j+u)|^2\mu_8(\mathrm{d}u)\mathrm{d}s\\
&\leq C\Delta^{2\beta}+4\mathbb{E}\int_0^t\int_{-s}^{0}|x^k(s+u)-X_\Delta^k(\tau(s)+u)|^2\mu_8(\mathrm{d}u)\mathrm{d}s.
\end{aligned}\tag{5.13}$$

But we obviously have

$$\begin{aligned}
&\mathbb{E}\int_0^t\int_{-s}^{0}|x^k(s+u)-X_\Delta^k(\tau(s)+u)|^2\mu_8(\mathrm{d}u)\mathrm{d}s\\
&\leq 3\mathbb{E}\int_0^t\int_{-s}^{0}|e_\Delta(s+u)|^2\mu_8(\mathrm{d}u)\mathrm{d}s+3\mathbb{E}\int_0^t\int_{-s}^{0}|\bar{X}_\Delta^k(s+u)-\bar{X}_\Delta^k(\tau(s)+u)|^2\mu_8(\mathrm{d}u)\mathrm{d}s\\
&\quad+3\mathbb{E}\int_0^t\int_{-s}^{0}|\bar{X}_\Delta^k(\tau(s)+u)-X_\Delta^k(\tau(s)+u)|^2\mu_8(\mathrm{d}u)\mathrm{d}s.
\end{aligned}$$

In the same way Lemma 5.6 was proved, we can show that $\mathbb{E}|\bar{X}_\Delta^k(s+u)-\bar{X}_\Delta^k(\tau(s)+u)|^2\leq C\Delta$ for any $s\in[0,T]$. Applying this and Lemma 5.6 to the inequality above yields

$$\begin{aligned}
&\mathbb{E}\int_0^t\int_{-s}^{0}|x^k(s+u)-X_\Delta^k(\tau(s)+u)|^2\mu_8(\mathrm{d}u)\mathrm{d}s\\
&\leq C\Delta+\mathbb{E}\int_0^t\int_{-s}^{0}|e_\Delta(s+u)|^2\mu_8(\mathrm{d}u)\mathrm{d}s
\end{aligned}$$

$$\leq C\Delta + \mathbb{E}\int_{-\infty}^{0}\int_{0}^{t}|e_\Delta(s)|^2\mathrm{d}s\mu_8(\mathrm{d}u)$$
$$= C\Delta + \mathbb{E}\int_{0}^{t}|e_\Delta(s)|^2\mathrm{d}s.$$

Substituting this into (5.13) yields

$$\mathbb{E}\int_{0}^{t}|\psi_1(\pi_k(x_s^k)) - \psi_1(s)|^2\mathrm{d}s \leq C\Delta + C\mathbb{E}\int_{0}^{t}|e_\Delta(s)|^2\mathrm{d}s.$$

Consequently, inserting this into (5.12) we arrive at

$$I_1(t) \leq C\Delta + C\mathbb{E}\int_{0}^{t}|e_\Delta(s)|^2\mathrm{d}s. \tag{5.14}$$

By the Markov property of $\theta(\cdot)$, Assumption 5.1, Lemma 5.4 and Remark 5.5, we derive that

$$\begin{aligned}
I_2(t) &= \sum_{j=0}^{N}\mathbb{E}\int_{t_j}^{t_{j+1}}|F(X_\Delta^k(t_j),\psi_{1j},\theta(s),t_j) - F(X_\Delta^k(t_j),\psi_{1j},\theta(t_j),t_j)|^2\mathrm{d}s\\
&= \sum_{j=0}^{N}\mathbb{E}\int_{t_j}^{t_{j+1}}|F(X_\Delta^k(t_j),\psi_{1j},\theta(s),t_j) - F(X_\Delta^k(t_j),\psi_{1j},\theta(t_j),t_j)|^2\mathbf{1}_{\{\theta(s)\neq\theta(t_j)\}}\mathrm{d}s\\
&\leq C\sum_{j=0}^{N}\int_{t_j}^{t_{j+1}}\mathbb{E}\Big[\big(1+|X_\Delta^k(t_j)|^2+|\psi_{1j}|^2\big)\mathbb{E}\big(\mathbf{1}_{\{\theta(s)\neq\theta(t_j)\}}|\mathcal{F}_{t_j}\big)\Big]\mathrm{d}s\\
&\leq C\Delta\sum_{j=0}^{N}\int_{t_j}^{t_{j+1}}\mathbb{E}\Big[\big(1+|X_\Delta^k(t_j)|^2+|\psi_{1j}|^2\big)\Big]\mathrm{d}s\\
&\leq C\Delta.
\end{aligned} \tag{5.15}$$

It follows from Assumption 5.2, Lemma 5.4 and Remark 5.5 that

$$I_3(t) \leq C\mathbb{E}\int_{0}^{t}\Big(1+|\bar{X}_\Delta^k(s)|^2+|\psi_1(s)|^2\Big)\Delta^{2\alpha}\mathrm{d}s \leq C\Delta^{2\alpha}. \tag{5.16}$$

Inserting (5.14)-(5.16) into (5.11), we obtain that $I(t) \leq C\Delta + C\int_0^t \mathbb{E}|e_\Delta(s)|^2\mathrm{d}s$. Similarly, we can show $J(t) \leq C\Delta + C\int_0^t \mathbb{E}|e_\Delta(s)|^2\mathrm{d}s$. Putting these into (5.10) gives

$$\mathbb{E}|e_\Delta(t)|^2 \le C\Delta + C\int_0^t \mathbb{E}|e_\Delta(s)|^2 \mathrm{d}s.$$

An application of the Gronwall inequality yields that

$$\mathbb{E}|e_\Delta(t)|^2 \le C\Delta$$

as required. The proof is hence complete. □

Combining Lemmas 5.6 and 5.7, we obtain the strong convergence of the EM numerical solutions to the true solution of the truncated SFDE (4.10).

Theorem 5.8. *Let Assumptions 5.1, 5.2, 5.3 hold and $\mu_8, \mu_9 \in \mathcal{P}_r$. Then for any $\Delta \in (0, 1]$*

$$\mathbb{E}|x^k(t) - X^k_\Delta(t)|^2 \le C\Delta, \quad \forall\, t \in [0, T]. \tag{5.17}$$

On the other hand, in a similar way as Theorem 3.7 was proved, we can show the following corollary.

Corollary 5.9. *Suppose that Assumption 5.1 holds and $\mu_8, \mu_9 \in \mathcal{P}_b$ with $b > r$. Then the solution $x^k(t)$ of the truncated SFDE (4.10) approximates the solution $x(t)$ of the given SFDE (4.7) in the sense that*

$$\mathbb{E}|x(t) - x^k(t)|^p \le Ce^{-(b-r)pk}, \ \forall\, t \in [0, T]. \tag{5.18}$$

Consequently, we obtain the following strong convergence result of the EM numerical solutions to the true solution of the given SFDE (4.7).

Theorem 5.10. *Suppose that Assumptions 5.1, 5.2, 5.3 hold and $\mu_8, \mu_9 \in \mathcal{P}_b$ with $b > r$. Then for any $\Delta \in (0, 1]$ and any integer $k > T$,*

$$\mathbb{E}|x(t) - X^k_\Delta(t)|^2 \le C(e^{-2(b-r)k} + \Delta), \ \forall\, t \in [0, T]. \tag{5.19}$$

5.2. Highly nonlinear case

In Section 4.3, we already showed that the solution $x^k(t)$ of the truncated SFDE (4.10) approximates the solution of the given SFDE (4.7) under Assumptions 4.3 and 4.4. To obtain the numerical solution of the truncated SFDE (4.10) under these assumptions, we can apply the truncated EM method (see, e.g., [7,35]). Due to the page limit, we leave the details to the reader but discuss a couple of examples and carry out some simulations using MATLAB to illustrate the idea.

Example 5.11. In 1973, the well-known Black-Scholes model with constant volatility was presented. But tests of this model on real market data have questioned the assumption of constant volatility in the stock dynamics. For this reason, several variants of the Black-Scholes model with non-constant volatility such as stochastic functional volatility equations have been proposed [24].

In this example, we consider a scalar stochastic functional volatility equation with infinite delay of the form

$$\mathrm{d}x(t) = f(x_t, \theta(t), t)\mathrm{d}t + g(x_t, \theta(t), t)\mathrm{d}B(t), \quad t \geq 0, \tag{5.20}$$

where the coefficients are defined by

$$f(\varphi, i, t) = \begin{cases} 1 + 4\varphi(0) - 4\varphi^3(0), & i = 1, \\ 2 + 3\varphi(0) - 5\varphi^3(0), & i = 2, \end{cases}$$

$$g(\varphi, i, t) = \begin{cases} \int_{-\infty}^{0} \varphi^2(u)\mu(\mathrm{d}u), & i = 1, \\ \frac{1}{2}\int_{-\infty}^{0} \varphi^2(u)\mu(\mathrm{d}u), & i = 2, \end{cases}$$

for $\varphi \in \mathcal{C}_{1/5}$, the probability measure $\mu(\cdot)$ has its probability density function e^u on $(-\infty, 0]$ (i.e., $\mu(\mathrm{d}u) = e^u\mathrm{d}u$), $B(t)$ is a scalar Brownian motion and $\theta(t)$ is a Markov chains on the state space $\mathbb{S} = \{1, 2\}$ with its generator

$$\Gamma = \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix}.$$

Let the initial data $\xi(u) = e^u \in \mathcal{C}_{1/5}$ and $\theta(0) = 1$. Recalling the definition of truncation mapping π_k we get the corresponding approximation SFDEs

$$\mathrm{d}x^k(t) = f_k(x_t^k, \theta(t), t)\mathrm{d}t + g_k(x_t^k, \theta(t), t)\mathrm{d}B(t). \tag{5.21}$$

Here f_k and g_k are defined by

$$f_k(\varphi, i, t) = \begin{cases} 1 + 4\varphi(0) - 4\varphi^3(0), & i = 1, \\ 2 + 3\varphi(0) - 5\varphi^3(0), & i = 2, \end{cases}$$

and

$$g_k(\varphi, i, t) = \begin{cases} e^{-k}\varphi^2(-k) + \int_{-k}^{0} \varphi^2(u)\mu(\mathrm{d}u), & i = 1, \\ \frac{1}{2}e^{-k}\varphi^2(-k) + \frac{1}{2}\int_{-k}^{0} \varphi^2(u)\mu(\mathrm{d}u), & i = 2, \end{cases}$$

for $\varphi \in \mathcal{C}_{1/5}$. One observes that f and g satisfy Assumption 2.1 owing to $\mu \in \mathcal{P}_\gamma$ for any $\gamma \in [0, 1)$ and Assumption 2.2 with $W_1(x, t) = V(x, i, t) = x^4$ and $W_2(x, t) = x^6$. Thus, (5.20) has a unique global solution. Furthermore, Assumption 3.5 holds with $\bar{q} = 2$, $\bar{V}(x, t) = x^2$, $\mu_4 = \delta_0 \in \mathcal{P}_\gamma$ for any $\gamma > 0$, $\mu_5 = \mu$ and $U(x, y) = |x - y|^2|x + y|^2 \in \mathcal{U}_{0,4}$. By virtue of Theorem 3.7, $x^k(T)$ converges to $x(T)$ exponentially for any $T > 0$.

In order to test the efficiency of the result in Theorem 3.7 we carry out some numerical experiments by MATLAB. We use the truncated EM numerical solution of (5.21) with $k = 200$ and $\Delta = 2^{-6}$ as the exact solution of (5.20) and plot the mean square error $\mathbb{E}|x(10) - x^k(10)|^2$ for 1000 sample points between the solution of (5.20) and that of (5.21) as function of k when $k \in \{10, 11, 12, 13, 14\}$. Furthermore, in order to characterize the error between the exact solu-

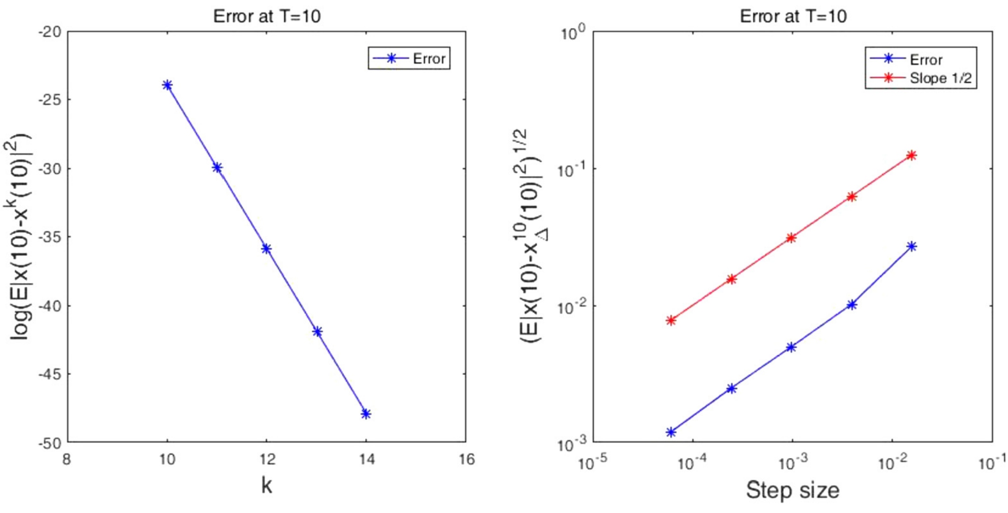

Fig. 1. (a) The mean square error for 1000 sample points between $x(10)$ and $x^k(10)$ as the function of $k \in \{10, 11, 12, 13, 14\}$. (b) The root mean square errors for 1000 sample points between $x(10)$ and $X^{50}_{\Delta}(10)$ as the function of $\Delta \in \{2^{-6}, 2^{-8}, 2^{-10}, 2^{-12}, 2^{-14}\}$.

tion and numerical solution with respect to Δ, we take the truncated EM numerical solution of (5.21) with $k = 200$ and $\Delta = 2^{-18}$ as the exact solution of (5.20). Fig. 1 depicts the root mean square error $(\mathbb{E}|x(10) - X^k_{\Delta}(10)|^2)^{1/2}$ between the exact solution and the numerical solution of (5.20) with $k = 50$, as a function of $\Delta \in \{2^{-6}, 2^{-8}, 2^{-10}, 2^{-12}, 2^{-14}\}$ for 1000 sample points.

Example 5.12. The delay Lotka-Volterra systems have received great attention owing to their extensive applications. The author of [14] analyzed the global asymptotic stability of the generic stochastic Lokta-Volterra systems with infinite delay. Let us consider such a system with special coefficients described by

$$\begin{cases} dx_1(t) = x_1(t)[(0.5 + 0.1\sin t) - 0.8x_1(t) - 0.2x_2(t)]dt + 0.5x_1(t)dB_1(t), \\ dx_2(t) = x_2(t)\Big[(0.3 + 0.2\sin 2t) - 0.6x_2(t) \\ \qquad - 0.12\displaystyle\int_{-\infty}^{0} x_1(t+u)\mu(du)\Big]dt + 0.5x_2(t)dB_2(t), \end{cases} \tag{5.22}$$

and the initial data are given by $\xi_1(u) = 0.8e^u$ for $u \le 0$ and $x_2(0) = 0.6$. Here $(B_1(t), B_2(t))$ is a 2-dimensional Brownian motion and μ is the same probability measure as in Example 5.11. According to [14], (5.22) has a unique global positive solution $x(t) = (x_1(t), x_2(t))$ for $t \ge 0$. One observes that Assumption 3.2 holds with $\beta(x) = W_1(x,t) = V(x,i,t) = |x|^4$, $W_2(x,t) = 0$ and $\mu_1 = \delta_0$. Assumption 3.3 holds with $\mu_3(\cdot) = 0.5(\delta_0(\cdot) + \mu(\cdot))$. For any $k > 0$, the corresponding approximation equation is

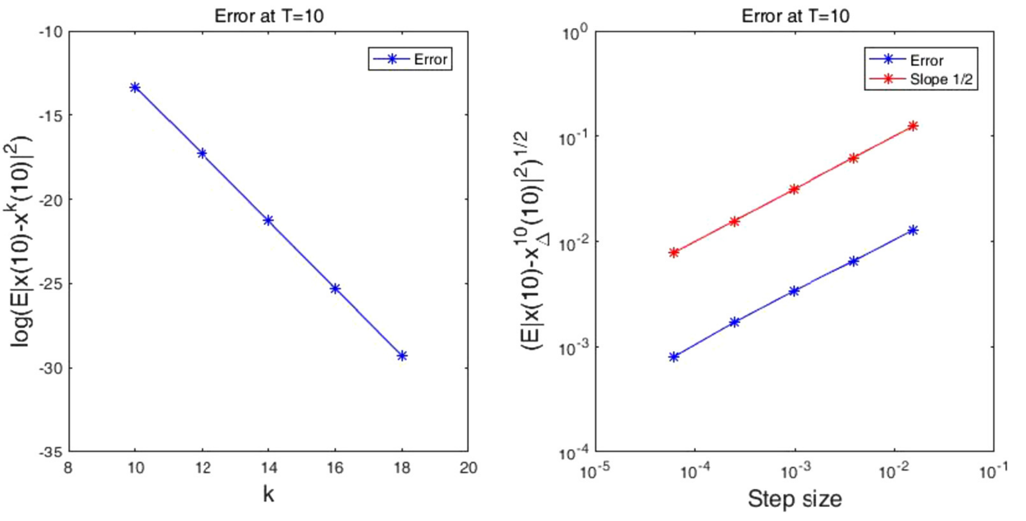

Fig. 2. (a) The mean square error for 1000 sample points between $x(10)$ and $x^k(10)$ as the function of $k \in \{10, 12, 14, 16, 18\}$. (b) The root mean square error for 1000 sample points between $x(10)$ and $X_\Delta^{30}(10)$ as the function of $\Delta \in \{2^{-6}, 2^{-8}, 2^{-10}, 2^{-12}, 2^{-14}\}$.

$$\begin{cases} dx_1^k(t) = x_1^k(t)[(0.5 + 0.1\sin t) - 0.8x_1^k(t) - 0.2x_2^k(t)]dt + 0.5x_1^k(t)dB_1(t), \\ dx_2^k(t) = x_2^k(t)\big[(0.3 + 0.2\sin 2t) - 0.6x_2^k(t) - 0.12e^{-k}x_1(t-k) \\ \qquad - 0.12\int_{-k}^{0} x_1^k(t+u)\mu(du)\big]dt + 0.5x_2^k(t)dB_2(t), \end{cases} \tag{5.23}$$

which has a unique global solution $x^k(t) = (x_1^k(t), x_2^k(t))$ for $t \geq 0$. Thus, by virtue of Theorem 3.4, $x^k(t)$ converges to the solution $x(t)$ in L^2. For illustration, we carry out some numerical experiments using MATLAB. Due to the unsolvability of (5.23) we regard the numerical solution of the truncated EM scheme with $\Delta = 2^{-6}$ and $k = 200$ as the exact $x(t)$ of (5.22), while for $k \in \{10, 12, 14, 16, 18\}$, we regard the numerical solution of the truncated EM scheme with $\Delta = 2^{-6}$ as the exact $x^k(t)$ of (5.23). Furthermore, we view the truncated EM numerical solution of (5.23) with $k = 200$ and $\Delta = 2^{-18}$ as the exact solution of (5.22). Let $k = 30$ and $\Delta \in \{2^{-6}, 2^{-8}, 2^{-10}, 2^{-12}, 2^{-14}\}$. Fig. 2 depicts the root mean square error $(\mathbb{E}|x(10) - X_\Delta^k(10)|^2)^{1/2}$ between the exact solution and the numerical solution of (5.22) for 1000 sample points, as functions of Δ for 1000 sample points.

Data availability

No data was used for the research described in the article.

Acknowledgments

The authors would like to thank the editor and reviewer for their very helpful comments and suggestions. The authors would also like to thank the National Natural Science Foundation of China (No. 11971096), the Natural Science Foundation of Jilin Province (No. YDZJ202101ZYTS154), the National Key R&D Program of China (2020YFA0714102), the Fundamental Research Funds for the Central Universities, the Royal Society of Edinburgh (RSE1832) and the Royal Society (WM160014, Royal Society Wolfson Research Merit Award) for their financial support.

References

[1] H.K. Asker, Stability in distribution of numerical solution of neutral stochastic functional differential equations with infinite delay, J. Comput. Appl. Math. 396 (2021) 113625.
[2] C.T.H. Baker, A. Tang, Stability analysis of continuous implicit Runge-Kutta methods for Volterra integro-differential systems with unbounded delays, Appl. Numer. Math. 24 (1997) 153–173.
[3] H. Brunner, The numerical treatment of Volterra integro-differential equations with unbounded delay, J. Comput. Appl. Math. 28 (1989) 5–23.
[4] K. Dareiotis, C. Kumar, S. Sabanis, On tamed Euler approximations of SDEs driven by Lévy noise with applications to delay equations, SIAM J. Numer. Anal. 54 (2016) 1840–1872.
[5] N.H. Du, D.H. Nguyen, N.N. Nguyen, G. Yin, Stability of stochastic functional differential equations with random switching and applications, Automatica J. IFAC 125 (2021) 109410.
[6] L. Feng, S. Li, X. Mao, Asymptotic stability and boundedness of stochastic functional differential equations with Markovian switching, J. Franklin Inst. 353 (2016) 4924–4949.
[7] Q. Guo, X. Mao, R. Yue, The truncated Euler-Maruyama method for stochastic differential delay equations, Numer. Algorithms 78 (2018) 599–624.
[8] D.J. Higham, An algorithmic introduction to numerical simulation of stochastic differential equations, SIAM Rev. 43 (2001) 525–546.
[9] Y. Hino, S. Murakami, T. Naito, Functional-Differential Equations with Infinite Delay, Springer-Verlag, Berlin, 1991.
[10] L. Huang, X. Mao, SMC design for robust H_∞ control of uncertain stochastic delay systems, Automatica J. IFAC 46 (2010) 405–412.
[11] V.B. Kolmanovskii, V.R. Nosov, Stability of Functional-Differential Equations, Academic Press, London, 1986.
[12] R. Li, Y. Hou, Convergence and stability of numerical solutions to SDDEs with Markovian switching, Appl. Math. Comput. 175 (2006) 1080–1091.
[13] X. Li, X. Mao, Y. Shen, Approximate solutions of stochastic differential delay equations with Markovian switching, J. Differ. Equ. Appl. 16 (2010) 195–207.
[14] M. Liu, Global asymptotic stability of stochastic Lotka-Volterra systems with infinite delays, IMA J. Appl. Math. 80 (2015) 1431–1453.
[15] M. Liu, K. Wang, Stochastic logistic equation with infinite delay, Math. Methods Appl. Sci. 35 (2012) 812–827.
[16] X. Mao, Stability of Stochastic Differential Equations with Respect to Semimartingales, Longman Scientific and Technical, New York, 1991.
[17] X. Mao, Exponential Stability of Stochastic Differential Equations, Marcel Dekker, New York, 1994.
[18] X. Mao, Stochastic Differential Equations and Their Applications, second edition, Horwood, Chichester, 2007.
[19] X. Mao, M.J. Rassias, Khasminskii-type theorems for stochastic differential delay equations, Stoch. Anal. Appl. 23 (2005) 1045–1069.
[20] X. Mao, C. Yuan, Stochastic Differential Equations with Markovian Switching, Imperial College Press, London, 2006.
[21] C. Mei, C. Fei, W. Fei, X. Mao, Exponential stabilization by delay feedback control for highly nonlinear hybrid stochastic functional differential equations with infinite delay, Nonlinear Anal. Hybrid Syst. 40 (2021) 101026.
[22] S.E.A. Mohammed, Stochastic Functional Differential Equations, Pitman, Boston, 1984.
[23] D.H. Nguyen, G. Yin, Stability of stochastic functional differential equations with Regime-Switching: analysis using Dupire's functional Itô formula, Potential Anal. 53 (2020) 247–265.
[24] F. Sancier, S. Mohammed, An option pricing model with memory, Commun. Stoch. Anal. 11 (2017) 505–531.
[25] G. Song, J. Hu, S. Gao, X. Li, The strong convergence and stability of explicit approximations for nonlinear stochastic delay differential equations, Numer. Algorithms 89 (2022) 855–883.
[26] M. Song, X. Mao, Almost sure exponential stability of hybrid stochastic functional differential equations, J. Math. Anal. Appl. 458 (2018) 1390–1408.

[27] Y. Song, C.T.H. Baker, Qualitative behaviour of numerical approximations to Volterra integro-differential equations, J. Comput. Appl. Math. 172 (2004) 101–115.
[28] Y. Wang, F. Wu, X. Mao, E. Zhu, Advances in the Lasalle-type theorems for stochastic functional differential equations with infinite delay, Discrete Contin. Dyn. Syst., Ser. B 25 (2020) 287–300.
[29] F. Wei, K. Wang, The existence and uniqueness of the solution for stochastic functional differential equations with infinite delay, J. Math. Anal. Appl. 331 (2007) 516–531.
[30] F. Wu, X. Mao, Numerical solutions of neutral stochastic functional differential equations, SIAM J. Numer. Anal. 46 (2008) 1821–1841.
[31] F. Wu, X. Mao, P.E. Kloeden, Almost sure exponential stability of the Euler–Maruyama approximations for stochastic functional differential equations, Random Oper. Stoch. Equ. 19 (2011) 165–186.
[32] F. Wu, X. Mao, L. Szpruch, Almost sure exponential stability of numerical solutions for stochastic delay differential equations, Numer. Math. 115 (2010) 681–697.
[33] F. Wu, G. Yin, H. Mei, Stochastic functional differential equations with infinite delay: existence and uniqueness of solutions, solution maps, Markov properties and ergodicity, J. Differ. Equ. 262 (2017) 1226–1252.
[34] S. Xu, J. Lam, X. Mao, Y. Zou, A new LMI condition for delay-dependent robust stability of stochastic time-delay systems, Asian J. Control 7 (2005) 419–423.
[35] W. Zhang, M. Song, M. Liu, Strong convergence of the truncated Euler-Maruyama method for stochastic functional differential equations, Int. J. Comput. Math. 95 (2018) 2363–2387.