

Stability in distribution of stochastic functional differential equations*

Ya Wang^a Fuke Wu^a Xuerong Mao^b

^aSchool of Mathematics and Statistics, Huazhong University of Science and Technology,
Wuhan, Hubei 430074, PR China

^bDepartment of Mathematics and Statistics, University of Strathclyde, Glasgow G1 1XH, UK

Abstract

In this paper we investigate the stability in distribution for a class of stochastic functional differential equations (SFDEs), which include stochastic differential delay equations (SDDEs). Although stability in distribution has been studied by several authors recently, there is so far no stability-in-distribution criterion on SFDEs where the terms involved the delay components are highly nonlinear (not bounded by linear functions). In this paper we will establish the sufficient criteria on the stability in distribution for a class of highly nonlinear SFDEs. Two examples will be given to illustrate our new results. We also explain the reason why the existing stability-in-distribution criteria are not applicable by these two examples.

Key Words: Stochastic functional differential equations; Stability in distribution; Itô formula; Brownian motion.

1 Introduction

In the past few decades, theory of stochastic functional differential equations (SFDEs) has attracted a great deal of attention, see, for example, [1, 2]. In particular, many papers have been devoted to the study of stability of SFDEs since it has wide applications in automatic control, mechanical system, and so on, see, for example, [3–5]. However, most of these papers are concerned with the stability of the trivial solution in the senses of trajectories or the p th moment. Such stabilities are sometimes too strong in the stochastic environment. It is often enough to know whether the probability distribution of the solution will converge to some distribution (but not necessarily to the Dirac delta function). Such stability is known as the stability in distribution and the limit distribution is known as a stationary distribution.

Basak and Bhattacharya [6] studied the stability in distribution for a class of stochastic differential equations (SDEs) with singular diffusions, and then Basak et al. [7] also examined the stationary stability for semi-linear stochastic differential equations. In [8], Basak

*The research was supported in part by the National Natural Science Foundations of China (Grant Nos. 1161101211 and 61873320), and the Royal Society and the Newton Fund (NA160317, Royal Society-Newton Advanced Fellowship).

and this coauthors not only extended their method in [7] to investigate the stationary distribution for nonlinear SDEs with Markovian switching but also derived the functional central limit theorem and the law of iterated logarithm. Later, Yuan et al. [9] and Dang [12] also investigated the stationary distribution for nonlinear SDEs with Markovian switching under weaker conditions. Along this line, considering the continuous function space, Yuan et al. [10] and Du et al. [11] discussed stability in distribution for SDEs with constant time lag by different methods, while Hu et al. [13] discussed the stability of neutral SFDEs by using the Lyapunov function method. By choosing special Lyapunov function, Bao and his coauthors established existence and uniqueness of the invariant measure for several classes of SFDEs under the dissipative condition in [14, 15] and without dissipativity in [16]. For non-degenerate SFDEs, by an asymptotic coupling method, Hairer et al. [17] showed a class of SFDEs has at most one invariant measure under certain assumptions that do not ensure the existence of an invariant measure. Butkovsky [18] and Butkovsky and Scheutzow [19] developed Hairer's approach, and gave the sufficient conditions for existence and uniqueness of an invariant measure for SFDEs. However, to our best knowledge, there is so far no stability-in-distribution criterion on SFDEs when the coefficients involving the delay components are highly nonlinear (not bounded by linear functions). For example, for the following special SFDE (the stochastic integro-differential equation)

$$dx(t) = \left(a - bx(t) - cx^3(t) + \int_{-\tau}^0 x^2(t + \theta) \mu(d\theta) \right) dt + d \int_{-\tau}^0 |x(t + \theta)|^{\frac{3}{2}} \mu(d\theta) dB(t), \quad (1.1)$$

where a, b, c, d are constant, the existing results cannot cover the existence and uniqueness of stationary distribution since the drift and diffusion coefficients involve the highly nonlinear delay terms. This paper aims to establish the stationary distribution criterion for this class of SFDEs.

Let us consider an SFDEs of the form

$$dx(t) = f(x_t)dt + g(x_t)dB(t) \quad (1.2)$$

on $t \geq 0$ with the initial data $x_0 = \xi \in C([- \tau, 0]; \mathbb{R}^n)$, where $x_t = x_t(\theta) \triangleq \{x(t + \theta), -\tau \leq \theta \leq 0\}$, $f : C([- \tau, 0]; \mathbb{R}^n) \rightarrow \mathbb{R}^n$ and $g : C([- \tau, 0]; \mathbb{R}^n) \rightarrow \mathbb{R}^{n \times m}$ are continuous maps, $B(t)$ is an m -dimensional Brownian motion. The main aim of this paper is to consider stability in distribution under the highly nonlinear conditions.

The existing criteria often require the function (namely \mathcal{W}_2 in Assumption **(H3)** below, which is used to restrict the Lyapunov function with respect to the difference between two solutions starting from different initial data) depends on only the difference $\phi - \varphi$, where $\phi, \varphi \in C([- \tau, 0]; \mathbb{R}^n)$. However, in many real applications, such function often rely on ϕ and φ themselves. Motivated by these observations, we hope to obtain the more general criteria for stability in distribution such that they can cover more nonlinear SFDEs (including (1.1)) and improve the existing results.

The rest of this paper is organized as follows. Section 2 gives the necessary notations and definitions. In Section 3, we present five key lemmas which lay good foundation for our main result in the following, and then we give the proof of stability in distribution. In Section 4, two examples will be given to illustrate our new results. We also explain reason why the existing stability-in-distribution criteria are not applicable by these two examples.

2 Preliminaries

Throughout this paper, unless otherwise specified, we use the following notations. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t \geq 0}$ satisfying the usual condition, that is, it is right continuous and increasing while \mathcal{F}_0 contains all \mathbb{P} -null sets. Let $B(t)$ be an m -dimensional Brownian motion defined on this probability space. For $\tau > 0$, let $\mathcal{C} \triangleq C([-\tau, 0]; \mathbb{R}^n)$ denote the space of continuous functions from $[-\tau, 0]$ to \mathbb{R}^n endowed with the uniform norm $\|\xi\| \triangleq \sup_{-\tau \leq \theta \leq 0} |\xi(\theta)|$. If $x(t)$ is an \mathbb{R}^n -value stochastic process, define $x_t = x_t(\theta) \triangleq \{x(t + \theta) : -\tau \leq \theta \leq 0\}$ for $t \geq 0$. Denote by \mathbb{R}^n the n -dimensional Euclidean space and $|\cdot|$ the Euclidean norm. If a and b are vectors, their inner product is denoted by $a \cdot b$. If A is a vector or a matrix, its transpose is denoted by A^T . For a matrix A , denote its trace norm by $|A| = \sqrt{\text{trace}(A^T A)}$. For a set A , A^c represents its complementary set. Let $\mathbf{1}_G$ denote the indicator function of the set G . Let $[b]$ denote the integer part of a real b . If $a, b \in \mathbb{R}$, $a \wedge b$ denotes the minimum of a and b and $a \vee b$ represents their maximum. We denote by \mathcal{M} the family of all probability measures on $[-\tau, 0]$. Let $\mathbb{R}_+ = [0, \infty)$. Let $C^2(\mathbb{R}^n; \mathbb{R}_+)$ denote the family of nonnegative functions V on \mathbb{R}^n which are continuously twice differentiable, and define an operator $LV : \mathcal{C} \rightarrow \mathbb{R}$ by

$$LV(\varphi) = V_x(\varphi(0)) \cdot f(\varphi) + \frac{1}{2} \text{trace}(g^T(\varphi) V_{xx}(\varphi(0)) g(\varphi)),$$

where

$$V_x(x) = \left(\frac{\partial V(x)}{\partial x_1}, \dots, \frac{\partial V(x)}{\partial x_n} \right), V_{xx}(x) = \left(\frac{\partial^2 V(x)}{\partial x_i \partial x_j} \right)_{n \times n}.$$

Denote by $x(t; \xi)$ the solution of (1.2) with the initial data $x_0 = \xi \in \mathcal{C}$ and by $x_t(\xi)$ the solution mapping (or the segment process). Under appropriate conditions, we can prove that the solution mapping $x_t(\xi)$ of (1.2) is a time homogeneous Markov process, see [21]. Let $p(t, \xi, \cdot)$ denote the transition probability function of process $x_t(\xi)$. Now, we introduce the concept of stability in distribution for convenience.

Definition 2.1. *The solution mapping process x_t is said to be stable in distribution if there exists a probability measure $\pi(\cdot)$ on \mathcal{C} such that the transition probability function $p(t, \xi, \cdot)$ of $x_t(\xi)$ converges weakly to $\pi(\cdot)$ as $t \rightarrow \infty$ for any $\xi \in \mathcal{C}$.*

Since the natural projection function is continuous, by the Mapping Theorem (see [22]), if the solution mapping process x_t of (1.2) is stable in distribution, the solution process $x(t, \xi)$ is also stable in distribution. It is also clear that the stability in distribution of x_t implies the existence of a unique invariant probability measure for x_t .

3 Stability in distribution

In this section, we present the sufficient conditions for the existence and uniqueness of the global solution of (1.2), and then give several key lemmas and prove the stability in distribution of (1.2). To examine the existence and uniqueness of the solution, we need the following assumptions.

(H1) For any $k > 0$, there exists a constant $b_k > 0$ such that for those $\phi, \varphi \in \mathcal{C}$ with $\|\phi\| \vee \|\varphi\| \leq k$

$$2\langle f(\phi) - f(\varphi), \phi(0) - \varphi(0) \rangle + |g(\phi) - g(\varphi)|^2 \leq b_k \|\phi - \varphi\|^2,$$

where $\langle \cdot, \cdot \rangle$ denotes the standard inner product on \mathbb{R}^n .

(H2) There are nonnegative numbers λ_1, λ_2 with $\lambda_1 > \lambda_2$, L_1, L_2 , a probability measure $\mu_1 \in \mathcal{M}$ and functions $V \in C^2(\mathbb{R}^n, \mathbb{R}_+)$, $\mathcal{W}_1 \in C(\mathbb{R}^n, \mathbb{R}_+)$ such that

$$\lim_{|x| \rightarrow \infty} V(x) = \infty, \quad V(x) \leq L_1(1 + \mathcal{W}_1(x))$$

for any $x \in \mathbb{R}^n$, and

$$LV(\phi) \leq L_2 - \lambda_1 \mathcal{W}_1(\phi(0)) + \lambda_2 \int_{-\tau}^0 \mathcal{W}_1(\phi(\theta)) \mu_1(d\theta) \quad (3.1)$$

for all $\phi \in \mathcal{C}$.

(H3) Let f and g be bounded on bounded subsets of \mathcal{C} .

Remark 1. Note that **(H2)** is actually a Lyapunov-type condition. When we choose $\mathcal{W}_1(x) = V(x)$, condition (3.1) is the classical Lyapunov-type condition (see [24] Theorem A.1)

Lemma 3.1. *Let Assumptions **(H1)** and **(H2)** hold. Then, for any initial data $\xi \in \mathcal{C}$, there exists a unique global solution $x(t, \xi)$ to (1.2). Moreover, for any compact set $K \subseteq \mathcal{C}$, $T \geq 0$, $\varepsilon > 0$, there exists a positive constant $H = H(K, T, \varepsilon)$ such that for any $t \geq 0$*

$$\mathbb{P}\{\|x_s(\xi)\| \leq H, \forall s \in [t, t + T]\} \geq 1 - \varepsilon, \quad \forall \xi \in K. \quad (3.2)$$

Proof. We will divide the proof into two steps.

Step 1: Existence and uniqueness of the global solution. It follows from Theorem 2.2 in [23] that under assumption **(H1)** (1.2) has a unique maximal local strong solution. To show this solution is global, we only need to prove that $\rho_e = \infty$ a.s., where ρ_e is the explosion time. Define a stopping time $\rho_k = \inf\{t \in [0, \rho_e) : |x(t, \xi)| > k\}$ with the traditional setting $\inf \emptyset = \infty$, where \emptyset denotes the empty set. Clearly, ρ_k is increasing as $k \rightarrow \infty$ and $\rho_k \rightarrow \rho_\infty \leq \rho_e$ a.s.. If we can show $\rho_\infty = \infty$ a.s., then $\rho_e = \infty$ a.s., which implies that $x(t, \xi)$ is actually global. This is equivalent to proving that for any $t > 0$, $\mathbb{P}\{\rho_k \leq t\} \rightarrow 0$ as $k \rightarrow \infty$. Let λ be a positive number to be determined later. Applying the Itô formula to $e^{\lambda t} V(x)$ yields

$$\mathbb{E} e^{\lambda(t \wedge \rho_k)} V(x(t \wedge \rho_k, \xi)) = V(\xi(0)) + \mathbb{E} \int_0^{t \wedge \rho_k} e^{\lambda s} (\lambda V(x(s, \xi)) + LV(x_s(\xi))) ds.$$

By **(H2)**, we obtain

$$\begin{aligned}
\mathbb{E}e^{\lambda(t \wedge \rho_k)}V(x(t \wedge \rho_k, \xi)) &\leq V(\xi(0)) + (\lambda L_1 + L_2)\mathbb{E} \int_0^{t \wedge \rho_k} e^{\lambda s} ds \\
&\quad + (\lambda L_1 - \lambda_1)\mathbb{E} \int_0^{t \wedge \rho_k} e^{\lambda s} \mathcal{W}_1(x(s, \xi)) ds \\
&\quad + \lambda_2 \mathbb{E} \int_0^{t \wedge \rho_k} \int_{-\tau}^0 e^{\lambda s} \mathcal{W}_1(x(s + \theta, \xi)) \mu_1(d\theta) ds. \tag{3.3}
\end{aligned}$$

Note that \mathcal{W}_1 is nonnegative. Applying the Tonelli theorem and a substitution technique gives

$$\begin{aligned}
&\int_0^{t \wedge \rho_k} \int_{-\tau}^0 e^{\lambda s} \mathcal{W}_1(x(s + \theta, \xi)) \mu_1(d\theta) ds \\
&= \int_{-\tau}^0 \int_0^{t \wedge \rho_k} e^{\lambda s} \mathcal{W}_1(x(s + \theta, \xi)) ds \mu_1(d\theta) \\
&= \int_{-\tau}^0 \int_{\theta}^{t \wedge \rho_k + \theta} e^{\lambda(s-\theta)} \mathcal{W}_1(x(s, \xi)) ds \mu_1(d\theta) \\
&\leq \int_{-\tau}^0 \int_{-\tau}^{t \wedge \rho_k} e^{\lambda(s-\theta)} \mathcal{W}_1(x(s, \xi)) ds \mu_1(d\theta) \\
&\leq \int_{-\tau}^0 \int_{-\tau}^0 e^{\lambda(s-\theta)} \mathcal{W}_1(x(s, \xi)) ds \mu_1(d\theta) + \int_{-\tau}^0 \int_0^{t \wedge \rho_k} e^{\lambda(s-\theta)} \mathcal{W}_1(x(s, \xi)) ds \mu_1(d\theta) \\
&\leq \int_{-\tau}^0 e^{-\lambda\theta} \mu_1(d\theta) \int_{-\tau}^0 e^{\lambda s} \mathcal{W}_1(x(s, \xi)) ds + \int_{-\tau}^0 e^{-\lambda\theta} \mu_1(d\theta) \int_0^{t \wedge \rho_k} e^{\lambda s} \mathcal{W}_1(x(s, \xi)) ds \\
&\leq e^{\lambda\tau} \int_{-\tau}^0 \mathcal{W}_1(\xi(s)) ds + e^{\lambda\tau} \int_0^{t \wedge \rho_k} e^{\lambda s} \mathcal{W}_1(x(s, \xi)) ds. \tag{3.4}
\end{aligned}$$

Substituting (3.4) into (3.3) yields

$$\begin{aligned}
\mathbb{E}e^{\lambda(t \wedge \rho_k)}V(x(t \wedge \rho_k, \xi)) &\leq V(\xi(0)) + \frac{\lambda L_1 + L_2}{\lambda} e^{\lambda t} + \lambda_2 e^{\lambda\tau} \int_{-\tau}^0 \mathcal{W}_1(\xi(s)) ds \\
&\quad - (\lambda_1 - \lambda L_1 - \lambda_2 e^{\lambda\tau}) \mathbb{E} \int_0^{t \wedge \rho_k} e^{\lambda s} \mathcal{W}_1(x(s, \xi)) ds. \tag{3.5}
\end{aligned}$$

Let us define $h(\lambda) = \lambda_1 - \lambda L_1 - \lambda_2 e^{\lambda\tau}$, which is continuous on λ with $h(0) = \lambda_1 - \lambda_2 > 0$ since $\lambda_1 > \lambda_2$. We therefore can find some $\lambda^* > 0$ such that $\lambda_1 - \lambda^* L_1 - \lambda_2 e^{\lambda^*\tau} > 0$, which implies that

$$\mathbb{E}e^{\lambda^*(t \wedge \rho_k)}V(x(t \wedge \rho_k, \xi)) \leq V(\xi(0)) + \frac{\lambda^* L_1 + L_2}{\lambda} e^{\lambda^* t} + \lambda_2 e^{\lambda^*\tau} \int_{-\tau}^0 \mathcal{W}_1(\xi(s)) ds.$$

Define $h_k = \inf \{V(x) : |x| \geq k\}$. Clearly, h_k is increasing as $k \rightarrow \infty$ and $h_k \rightarrow \infty$.

$$\begin{aligned}
h_k \mathbb{P}\{\rho_k \leq t\} &\leq \mathbb{E}V(x(\rho_k, \xi)) \mathbf{1}_{\{\rho_k \leq t\}} \\
&\leq \mathbb{E}e^{\lambda^*(t \wedge \rho_k)} V(x(t \wedge \rho_k, \xi)) \mathbf{1}_{\{\rho_k \leq t\}} \\
&\leq \mathbb{E}e^{\lambda^*(t \wedge \rho_k)} V(x(t \wedge \rho_k, \xi)) \\
&\leq V(\xi(0)) + \frac{\lambda^* L_1 + L_2}{\lambda^*} e^{\lambda^* t} + \lambda_2 e^{\lambda^* \tau} \int_{-\tau}^0 \mathcal{W}_1(\xi(s)) ds,
\end{aligned}$$

which implies $\mathbb{P}\{\rho_k \leq t\} \rightarrow 0$ as $k \rightarrow \infty$. Hence, the existence and uniqueness of a global solution is proved.

Step 2: Proof of (3.2). Letting $k \rightarrow \infty$, we obtain from (3.5) that

$$\begin{aligned}
\mathbb{E}e^{\lambda^* t} V(x(t, \xi)) &\leq V(\xi(0)) + \frac{\lambda^* L_1 + L_2}{\lambda^*} e^{\lambda^* t} + \lambda_2 e^{\lambda^* \tau} \int_{-\tau}^0 \mathcal{W}_1(\xi(s)) ds \\
&\quad - (\lambda_1 - \lambda^* L_1 - \lambda_2 e^{\lambda^* \tau}) \mathbb{E} \int_0^t e^{\lambda^* s} \mathcal{W}_1(x(s, \xi)) ds.
\end{aligned} \tag{3.6}$$

Since $\lambda_1 - \lambda^* L_1 - \lambda_2 e^{\lambda^* \tau} > 0$, we have

$$\mathbb{E}V(x(t, \xi)) \leq \frac{\lambda^* L_1 + L_2}{\lambda^*} + e^{-\lambda^* t} \left(V(\xi(0)) + \lambda_2 e^{\lambda^* \tau} \int_{-\tau}^0 \mathcal{W}_1(\xi(s)) ds \right),$$

and

$$\mathbb{E} \int_0^t e^{\lambda^* s} \mathcal{W}_1(x(s, \xi)) ds \leq C \left(V(\xi(0)) + \frac{\lambda^* L_1 + L_2}{\lambda^*} e^{\lambda^* t} + \lambda_2 e^{\lambda^* \tau} \int_{-\tau}^0 \mathcal{W}_1(\xi(s)) ds \right).$$

where $C = 1/(\lambda_1 - \lambda^* L_1 - \lambda_2 e^{\lambda^* \tau})$. Moreover, putting

$$M = \sup_{\xi \in K} \left\{ \frac{\lambda^* L_1 + L_2}{\lambda^*} + e^{-\lambda^* t} \left(V(\xi(0)) + \lambda_2 e^{\lambda^* \tau} \int_{-\tau}^0 \mathcal{W}_1(\xi(s)) ds \right) \right\},$$

we have

$$\sup_{\xi \in K} \mathbb{E}V(x(t, \xi)) \leq M \tag{3.7}$$

and

$$\sup_{\xi \in K} \mathbb{E} \int_0^t e^{\lambda^* s} \mathcal{W}_1(x(s, \xi)) ds \leq C e^{\lambda^* t} M, \tag{3.8}$$

For any $\xi \in K$, $t \geq 0$ and $k > 0$, define a stopping time

$$\sigma_k^t = \inf \{s \geq t : |x(s, \xi)| > k\}.$$

Similar to (3.5), recalling that $\lambda_1 - \lambda^*L_1 - \lambda_2e^{\lambda^*\tau} > 0$, we have

$$\begin{aligned}
& \mathbb{E}e^{\lambda^*((t+T)\wedge\sigma_k^t)}V(x((t+T)\wedge\sigma_k^t), \xi)) \\
& \leq \mathbb{E}e^{\lambda^*t}V(x(t, \xi)) + \frac{\lambda^*L_1 + L_2}{\lambda^*}e^{\lambda^*(t+T)} + \int_t^{(t+T)\wedge\sigma_k^t} (\lambda^*L_1 - \lambda_1)e^{\lambda^*s}\mathcal{W}_1(x(s, \xi))ds \\
& \quad + \lambda_2\mathbb{E}\int_t^{(t+T)\wedge\sigma_k^t} \int_{-\tau}^0 e^{\lambda^*s}\mathcal{W}_1(x(s+\theta, \xi))\mu_1(d\theta)ds \\
& \leq \mathbb{E}e^{\lambda^*t}V(x(t, \xi)) + \frac{\lambda^*L_1 + L_2}{\lambda^*}e^{\lambda^*(t+T)} + e^{\lambda^*\tau} \int_{-\tau}^0 \mathcal{W}_1(x(s, \xi))ds \\
& \quad + e^{\lambda^*\tau} \int_0^t e^{\lambda^*s}\mathcal{W}_1(x(s, \xi))ds - (\lambda_1 - \lambda^*L_1 - \lambda_2e^{\lambda^*\tau}) \int_t^{(t+T)\wedge\sigma_k^t} e^{\lambda^*s}\mathcal{W}_1(x(s, \xi))ds \\
& \leq \mathbb{E}e^{\lambda^*t}V(x(t, \xi)) + \frac{\lambda^*L_1 + L_2}{\lambda^*}e^{\lambda^*(t+T)} + \lambda_2e^{\lambda^*\tau} \int_{-\tau}^0 \mathcal{W}_1(\xi(s))ds \\
& \quad + \lambda_2e^{\lambda^*\tau}\mathbb{E}\int_0^t e^{\lambda^*s}\mathcal{W}_1(x(s, \xi))ds.
\end{aligned}$$

It follows from (3.7) and (3.8) that

$$\begin{aligned}
& \mathbb{E}e^{\lambda^*((t+T)\wedge\sigma_k^t-t)}V(x((t+T)\wedge\sigma_k^t), \xi)) \\
& \leq \mathbb{E}V(x(t, \xi)) + \frac{\lambda^*L_1 + L_2}{\lambda^*}e^{\lambda^*T} + \lambda_2e^{\lambda^*(\tau-t)} \int_{-\tau}^0 \mathcal{W}_1(\xi(s))ds \\
& \quad + \lambda_2e^{\lambda^*(\tau-t)}\mathbb{E}\int_0^t e^{\lambda^*s}\mathcal{W}_1(x(s, \xi))ds \\
& \leq (1 + C\lambda_2e^{\lambda^*\tau})M + \frac{\lambda^*L_1 + L_2}{\lambda^*}e^{\lambda^*T} + \lambda_2e^{\lambda^*\tau} \int_{-\tau}^0 \mathcal{W}_1(\xi(s))ds. \tag{3.9}
\end{aligned}$$

Recall that $\lim_{|x|\rightarrow\infty} V(x) = \infty$. We can therefore define $\hat{H} = \hat{H}(K, T, \varepsilon)$ such that

$$\inf_{|y|\geq\hat{H}} V(y) > \frac{1}{\varepsilon} \left((1 + C\lambda_2e^{\lambda^*\tau})M + \frac{\lambda^*L_1 + L_2}{\lambda^*}e^{\lambda^*T} + \lambda_2e^{\lambda^*\tau} \int_{-\tau}^0 \mathcal{W}_1(\xi(s))ds \right). \tag{3.10}$$

Employing (3.9) and (3.10) yields

$$\begin{aligned}
& \left(\inf_{|y|\geq\hat{H}} V(y) \right) \mathbb{P}\{\sigma_{\hat{H}}^t < t+T\} \leq \mathbb{E}V(x(\sigma_{\hat{H}}^t), \xi)1_{\{\sigma_{\hat{H}}^t < t+T\}} \\
& \leq \mathbb{E}e^{\lambda^*((t+T)\wedge\sigma_{\hat{H}}^t-t)}V(x((t+T)\wedge\sigma_{\hat{H}}^t), \xi)1_{\{\sigma_{\hat{H}}^t < t+T\}} \\
& \leq (1 + C\lambda_2e^{\lambda^*\tau})M + \frac{\lambda^*L_1 + L_2}{\lambda^*}e^{\lambda^*T} + \lambda_2e^{\lambda^*\tau} \int_{-\tau}^0 \mathcal{W}_1(\xi(s))ds.
\end{aligned}$$

This implies that $\mathbb{P}\{\sigma_{\hat{H}}^t < t+T\} < \varepsilon$, which shows

$$\mathbb{P}\{|x(s, \xi)| \leq \hat{H}, \forall s \in [t, t+T]\} \geq 1 - \varepsilon,$$

for any $t \geq 0$. By the Arzelá-Ascoli theorem, there exists a constant $\tilde{H} > 0$ such that $\|\xi\| \leq \tilde{H}$ for all $\xi \in K$. Letting $H = \hat{H} \vee \tilde{H}$, we have

$$\mathbb{P}\{|x(s, \xi)| \leq H, \forall s \in [t, t + T]\} \geq 1 - \varepsilon, \quad (3.11)$$

for any $t \geq -\tau$. Recalling the definition of $\|\cdot\|$, for any $t \geq 0$ we have

$$\begin{aligned} \mathbb{P}\{\|x_s(\xi)\| \leq H, \forall s \in [t, t + T]\} &= \mathbb{P}\left\{\sup_{-\tau \leq \theta \leq 0} |x(s + \theta, \xi)| \leq H, \forall s \in [t, t + t]\right\} \\ &= \mathbb{P}\{|x(s, \xi)| \leq H, \forall s \in [t - \tau, t + T]\}, \end{aligned}$$

which implies (3.2) combining with (3.11). This proof is completed. \square

Remark 2. This theorem shows that for any $\xi \in \mathcal{C}$, $T \geq 0$, $\varepsilon > 0$, there is a constant $H = H(\xi, T, \varepsilon)$ such that for any $t \geq 0$

$$\mathbb{P}\{\|x_s(\xi)\| \leq H, \forall s \in [t, t + T]\} \geq 1 - \varepsilon.$$

Remark 3. In [10], the linear growth condition is needed, but it can be removed in this paper.

Remark 4. In [11], to prove Lemma 3.1 for the case of Markovian switching SDEs with constant time lag, the following assumption is imposed:

$$c_1 \mathcal{W}_1(x) \leq V(x) \leq \mathcal{W}_1(x)$$

for all $x \in \mathbb{R}^n$ and some $c_1 > 0$. This condition is replaced by $V(x) \leq L_1(1 + \mathcal{W}_1(x))$ for all $x \in \mathbb{R}^n$ and some $L_1 > 0$ in this paper, which does not only weakens the condition for \mathcal{W}_1 but also makes (3.1) to be satisfied more easily.

Lemma 3.2. *Let Assumptions (H1), (H2) and (H3) hold. Then for any $\xi \in \mathcal{C}$, the family $\{p(t, \xi, \cdot) : t \geq 0\}$ is tight, where $p(t, \xi, \cdot)$ is the transition probability of $x_t(\xi)$.*

Proof. Fix $\xi \in \mathcal{C}$ arbitrarily. We firstly show that for any $\varepsilon_1, \varepsilon_2 > 0$, there exists $\delta_0 = \delta_0(\varepsilon_1, \varepsilon_2)$ such that

$$\mathbb{P}\left\{\sup_{t_2 - t_1 \leq \delta_0, t - \tau \leq t_1 \leq t_2 \leq t} |x(t_2, \xi) - x(t_1, \xi)| \geq \varepsilon_1\right\} \leq \varepsilon_2, \quad (3.12)$$

for any $t \geq 0$. It follows from Lemma 3.1 that there exists $H_1 = H_1(\xi, \tau, \varepsilon_2)$ such that for any $t \geq 0$

$$\mathbb{P}\left\{\sup_{t \leq s \leq t + \tau} \|x_s(\xi)\| \leq H_1\right\} \geq 1 - \frac{\varepsilon_2}{2}. \quad (3.13)$$

For each $s \geq 0$, define a stopping time $\tau_s = \inf\{u \geq s : \|x_u(\xi)\| > H_1\}$. Let $0 < \delta \leq \tau$. By (H3), for any $k > 0$ there exists a constant $c_k > 0$ such that

$$|f(\phi)| \vee |g(\phi)| \leq c_k$$

for all $\phi \in \mathcal{C}$ with $\|\phi\| \leq k$. Making use of this property, by the Burkholder-Davis-Gundy inequality and the basic inequality $(|x| + |y|)^4 \leq 8(|x|^4 + |y|^4)$, we derive that for any $t_1 \geq 0$

$$\begin{aligned}
& \mathbb{E} \left(\sup_{t_2 \in [t_1, t_1 + \delta]} |x(t_2 \wedge \tau_{t_1}, \xi) - x(t_1, \xi)|^4 \right) \\
& \leq 8 \mathbb{E} \sup_{t_2 \in [t_1, t_1 + \delta]} \left| \int_{t_1}^{t_2 \wedge \tau_{t_1}} f(x_s(\xi)) ds \right|^4 + 8 \mathbb{E} \sup_{t_2 \in [t_1, t_1 + \delta]} \left| \int_{t_1}^{t_2 \wedge \tau_{t_1}} g(x_s(\xi)) dw(s) \right|^4 \\
& \leq 8 \mathbb{E} \left(\int_{t_1}^{t_1 + \delta} 1_{\{\tau_{t_1} \geq s\}} |f(x_s(\xi))| ds \right)^4 + 8 B_p \mathbb{E} \left(\int_{t_1}^{t_1 + \delta} 1_{\{\tau_{t_1} \geq s\}} |g(x_s(\xi))|^2 ds \right)^2 \\
& \leq 8c_k^4 \delta^4 + 8B_p c_k^2 \delta^2 \\
& \leq (8c_k^4 \tau^2 + 8B_p c_k^2) \delta^2 \\
& = \acute{C}_{H_1} \delta^2,
\end{aligned} \tag{3.14}$$

where $\acute{C}_{H_1} = 8c_k^4 \tau^2 + 8B_p c_k^2$ and B_p is the coefficient of the Burkholder-Davis-Gundy inequality. We can further derive that for any $t \geq 0$,

$$\begin{aligned}
& \mathbb{P} \left\{ \tau_t \geq t + \tau, \sup_{t_2 \in [t_1, (t_1 + \delta) \wedge (t + \tau)], t \leq t_1 \leq t + \tau} |x(t_2, \xi) - x(t_1, \xi)| \geq \varepsilon_1 \right\} \\
& \leq \mathbb{P} \left\{ \tau_t \geq t + \tau, 3 \max_{0 \leq k \leq \lfloor \frac{\tau}{\delta} \rfloor} \sup_{s \in [t + k\delta, (t + (k+1)\delta) \wedge (t + \tau)]} |x(s, \xi) - x(t + k\delta, \xi)| \geq \varepsilon_1 \right\} \\
& = \mathbb{P} \left\{ \tau_t \geq t + \tau, \max_{0 \leq k \leq \lfloor \frac{\tau}{\delta} \rfloor} \sup_{s \in [t + k\delta, (t + (k+1)\delta) \wedge (t + \tau)]} |x(s, \xi) - x(t + k\delta, \xi)| \geq \frac{\varepsilon_1}{3} \right\} \\
& \leq \sum_{k=0}^{\lfloor \frac{\tau}{\delta} \rfloor} \mathbb{P} \left\{ \tau_t \geq t + \tau, \sup_{s \in [t + k\delta, (t + (k+1)\delta) \wedge (t + \tau)]} |x(s, \xi) - x(t + k\delta, \xi)| \geq \frac{\varepsilon_1}{3} \right\} \\
& \leq \sum_{k=0}^{\lfloor \frac{\tau}{\delta} \rfloor} \mathbb{P} \left\{ \tau_{t+k\delta} \geq t + \tau, \sup_{s \in [t + k\delta, (t + (k+1)\delta) \wedge (t + \tau)]} |x(s, \xi) - x(t + k\delta, \xi)| \geq \frac{\varepsilon_1}{3} \right\} \\
& \leq \left(\left\lfloor \frac{\tau}{\delta} \right\rfloor + 1 \right) \frac{81 \acute{C}_{H_1}}{\varepsilon_1^4} \delta^2,
\end{aligned} \tag{3.15}$$

where in the last inequality we use the Chebyshev inequality and (3.14).

Letting $\delta = [\varepsilon_1^4 \varepsilon_2 / (162 \acute{C}_{H_1} (\tau + 1))] \wedge 1 \wedge \tau$, it follows from (3.15) and the definition of the stopping time τ_t that for any $t \geq 0$,

$$\begin{aligned}
& \mathbb{P} \left\{ \sup_{t_2 \in [t_1, (t_1 + \delta) \wedge (t + \tau)], t \leq t_1 \leq t + \tau} |x(t_2, \xi) - x(t_1, \xi)| \geq \varepsilon_1 \right\} \\
& \leq \mathbb{P} \{ \tau_t < t + \tau \} + \mathbb{P} \left\{ \tau_t \geq t + \tau, \sup_{t_2 \in [t_1, (t_1 + \delta) \wedge (t + \tau)], t \leq t_1 \leq t + \tau} |x(t_2, \xi) - x(t_1, \xi)| \geq \varepsilon_1 \right\} \\
& \leq \frac{\varepsilon_2}{2} + \frac{\varepsilon_2}{2} \\
& \leq \varepsilon_2.
\end{aligned} \tag{3.16}$$

Since $\xi \in \mathcal{C}$, there exists a constant $\tilde{\delta} > 0$ such that

$$\sup_{t_2 - t_1 \leq \tilde{\delta}, -\tau \leq t_1 \leq t_2 \leq 0} |x(t_2, \xi) - x(t_1, \xi)| = \sup_{t_2 - t_1 \leq \tilde{\delta}, -\tau \leq t_1 \leq t_2 \leq 0} |\xi(t_2) - \xi(t_1)| < \varepsilon_1. \quad (3.17)$$

Letting $\delta_0 = \delta \wedge \tilde{\delta}$, (3.12) follows from (3.16) and (3.17).

Moreover, it follows from Lemma 3.1 that for any positive $\varepsilon > 0$, there exists $H_0 > 0$ such that

$$\mathbb{P}\{|x(t, \xi)| \geq H_0\} \leq \varepsilon, \quad \forall t \geq 0. \quad (3.18)$$

By (3.12) and (3.18), it follows from the Theorem 7.3 in [22] that the family $\{p(t, \xi, \cdot), t \geq 0\}$ is tight. The proof is completed. \square

In what follows we need to consider the difference between two solutions of (1.2) starting from different initial data, namely

$$x(t, \xi) - x(t, \eta) = \xi(0) - \eta(0) + \int_0^t f(x_s(\xi)) - f(x_s(\eta)) ds + \int_0^t g(x_s(\xi)) - g(x_s(\eta)) dB(s).$$

For a given function $U \in C^2(\mathbb{R}^n, \mathbb{R}_+)$, we define an operator

$$\mathcal{L}U(\phi, \psi) = U_x(\phi(0) - \psi(0)) \cdot [f(\phi) - f(\psi)] + \frac{1}{2} \text{trace} \{ [g(\phi) - g(\psi)]^T U_{xx}(\phi(0) - \psi(0)) [g(\phi) - g(\psi)] \}.$$

For the future use, we shall impose another hypothesis:

(H4) There exists a nonnegative constant λ_4 , a probability measure $\mu_2 \in \mathcal{M}$ and functions $U \in C^2(\mathbb{R}^n, \mathbb{R}_+)$, $\mathcal{W}_2, \lambda_3 \in C(\mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}_+)$ such that $U(\cdot)$ vanishes only at 0, $\mathcal{W}_2(x, y)$ vanishes only when $x = y$ ($x, y \in \mathbb{R}^n$), $\lambda_3(x, y) > \lambda_4$ provided $x \neq y$ and

$$\mathcal{L}U(\phi, \psi) \leq -\lambda_3(\phi(0), \psi(0)) \mathcal{W}_2(\phi(0), \psi(0)) + \lambda_4 \int_{-\tau}^0 \mathcal{W}_2(\phi(\theta), \psi(\theta)) \mu_2(d\theta) \quad (3.19)$$

for any $\phi, \psi \in \mathcal{C}$.

Lemma 3.3. *Let Assumptions (H1), (H2), (H3) and (H4) hold. Then for any $\varepsilon > 0$ and compact set $K \subseteq \mathcal{C}$, there exists $T = T(K, \varepsilon)$ such that for all $t > T$,*

$$\mathbb{P}\{\|x_t(\xi) - x_t(\eta)\| < \varepsilon\} \geq 1 - \varepsilon, \quad \forall \xi, \eta \in K.$$

Proof. We divide the proof into two steps.

Step 1. We will firstly show that for any $h > \sigma/2 > 0$,

$$\lim_{t \rightarrow +\infty} \mathbb{P}\{A_t^{\sigma, h}\} = 0 \quad \text{uniformly in } \xi, \eta \in K,$$

where $A_t^{\sigma, h} = \{\omega : \|x_t(\xi)\| \vee \|x_t(\eta)\| \leq h, |x(t, \xi) - x(t, \eta)| \geq \sigma\}$. To simplify the notation, denote

$$C^{\sigma, h} = \min \{ \mathcal{W}_2(x, y) : |x| \vee |y| \leq h, |x - y| \geq \sigma, x, y \in \mathbb{R}^n \}$$

and

$$\lambda_3^{\sigma,h} = \min\{\lambda_3(x, y) : |x| \vee |y| \leq h, |x - y| \geq \sigma, x, y \in \mathbb{R}^n\}.$$

Since $\mathcal{W}_2(x, y) > 0$ and $\lambda_3(x, y) > \lambda_4$ for any $x \neq y$, $C^{\sigma,h} > 0$ and $\lambda_3^{\sigma,h} > \lambda_4$. Then we have the estimate

$$\begin{aligned} & \mathbb{E}U(x(t, \xi) - x(t, \eta)) \\ &= U(\xi(0) - \eta(0)) + \mathbb{E} \int_0^t \mathcal{L}U(x_s(\xi), x_s(\eta)) ds \\ &\leq U(\xi(0) - \eta(0)) - \mathbb{E} \int_0^t \lambda_3(x(s, \xi), x(s, \eta)) \mathcal{W}_2(x(s, \xi), x(s, \eta)) ds \\ &\quad + \lambda_4 \mathbb{E} \int_0^t \int_{-\tau}^0 \mathcal{W}_2(x(s + \theta, \xi), x(s + \theta, \eta)) \mu_2(d\theta) ds \\ &\leq U(\xi(0) - \eta(0)) + \lambda_4 \int_{-\tau}^0 \mathcal{W}_2(\xi(s), \eta(s)) ds \\ &\quad - \mathbb{E} \int_0^t (\lambda_3(x(s, \xi), x(s, \eta)) - \lambda_4) \mathcal{W}_2(x(s, \xi), x(s, \eta)) ds \\ &= U(\xi(0) - \eta(0)) + \lambda_4 \int_{-\tau}^0 \mathcal{W}_2(\xi(s), \eta(s)) ds \\ &\quad - \mathbb{E} \int_0^t 1_{\{x(s, \xi) \neq x(s, \eta)\}} (\lambda_3(x(s, \xi), x(s, \eta)) - \lambda_4) \mathcal{W}_2(x(s, \xi), x(s, \eta)) ds \\ &\leq U(\xi(0) - \eta(0)) + \lambda_4 \int_{-\tau}^0 \mathcal{W}_2(\xi(s), \eta(s)) ds \\ &\quad - (\lambda_3^{\sigma,h} - \lambda_4) \mathbb{E} \int_0^t 1_{\{A_s^{\sigma,h}\}} \mathcal{W}_2(x(s, \xi), x(s, \eta)) ds, \end{aligned} \tag{3.20}$$

where in the last equality we use the fact that $\mathcal{W}_2(x, x) = 0, \forall x \in \mathbb{R}^n$. Noting that $\lambda_3^{\sigma,h} > \lambda_4$, letting $t \rightarrow +\infty$, we obtain

$$\begin{aligned} \int_0^{+\infty} \mathbb{E} 1_{\{A_s^{\sigma,h}\}} \mathcal{W}_2(x(s, \xi), x(s, \eta)) ds &\leq \frac{1}{\lambda_3^{\sigma,h} - \lambda_4} \left(U(\xi(0) - \eta(0)) + \lambda_4 \int_{-\tau}^0 \mathcal{W}_2(\xi(s), \eta(s)) ds \right) \\ &< +\infty. \end{aligned}$$

Consequently we can derive that for any $h > \sigma/2 > 0$

$$\begin{aligned} \int_0^{+\infty} \mathbb{P}\{A_s^{\sigma,h}\} ds &= \int_0^{+\infty} \mathbb{E} 1_{\{A_s^{\sigma,h}\}} ds \\ &\leq \frac{1}{C^{\sigma,h}} \int_0^{+\infty} \mathbb{E} 1_{\{A_s^{\sigma,h}\}} \mathcal{W}_2(x(s, \xi), x(s, \eta)) ds \\ &< +\infty. \end{aligned} \tag{3.21}$$

We claim that $\lim_{t \rightarrow +\infty} \mathbb{P}\{A_t^{\sigma,h}\} = 0$ for any $h > \sigma/2 > 0$. If not, there exist some $h_0 > \sigma_0/2 > 0$ such that $\limsup_{t \rightarrow +\infty} \mathbb{P}\{A_t^{\sigma_0, h_0}\} > 0$. Thus there exist a constant $l > 0$ and an

increasing sequence $t_n (n = 1, 2, \dots), t_n \uparrow \infty$ such that

$$\mathbb{P}\{A_{t_n}^{\sigma_0, h_0}\} = \mathbb{P}\{\|x_{t_n}(\xi)\| \vee \|x_{t_n}(\eta)\| \leq h_0, |x(t_n, \xi) - x(t_n, \eta)| \geq \sigma_0\} \geq l, \quad \forall n \in \mathbb{N}. \quad (3.22)$$

Applying (3.12) for $\varepsilon_1 = \sigma_0/3$, $\varepsilon_2 = l/8$, there is small enough $\delta_0 > 0$ such that for $\zeta = \xi, \eta \in \mathcal{C}$

$$\mathbb{P}\left\{\sup_{t_n \leq s \leq t_n + \delta_0} |x(s, \zeta) - x(t_n, \zeta)| \geq \frac{\sigma_0}{3}\right\} \leq \frac{l}{8}, \quad \forall n \in \mathbb{N}. \quad (3.23)$$

It follows from (3.22) and (3.23) that for $t_n \leq s \leq t_n + \delta_0$

$$\begin{aligned} & \mathbb{P}\left\{|x(s, \xi) - x(s, \eta)| \geq \frac{\sigma_0}{3}\right\} \\ & \geq \mathbb{P}\{|x(t_n, \xi) - x(t_n, \eta)| \geq \sigma_0\} - \mathbb{P}\left\{\sup_{t_n \leq s \leq t_n + \delta_0} |x(s, \xi) - x(t_n, \xi)| \geq \frac{\sigma_0}{3}\right\} \\ & \quad - \mathbb{P}\left\{\sup_{t_n \leq s \leq t_n + \delta_0} |x(s, \eta) - x(t_n, \eta)| \geq \frac{\sigma_0}{3}\right\} \\ & \geq l - \frac{l}{8} - \frac{l}{8} = \frac{3}{4}l. \end{aligned}$$

In view of Lemma 3.1, we can find $H_2 = H_2(K, l) \geq h_0$ satisfying

$$\mathbb{P}\{\|x_t(\xi)\| \leq H_2\} \geq 1 - \frac{l}{4}, \quad \forall \xi \in K, t \geq 0.$$

We deduce that for $t_n \leq s \leq t_n + \delta_0$,

$$\begin{aligned} & \mathbb{P}\left\{\|x_s(\xi)\| \vee \|x_s(\eta)\| \leq H_2 : |x(s, \xi) - x(s, \eta)| \geq \frac{\sigma_0}{3}\right\} \\ & \geq \mathbb{P}\left\{|x(s, \xi) - x(s, \eta)| \geq \frac{\sigma_0}{3}\right\} - \mathbb{P}\{\|x_s(\xi)\| \vee \|x_s(\eta)\| \leq H_2\}^c \\ & \geq \frac{3}{4}l - 2\frac{l}{4} = \frac{l}{4}. \end{aligned}$$

It means that $\int_{t_n}^{t_n + \delta_0} \mathbb{P}\{A_s^{\frac{\sigma_0}{3}, H_2}\} ds \geq l\delta_0/4, \forall n \in \mathbb{N}$. Consequently,

$$\int_0^{+\infty} \mathbb{P}\{A_s^{\frac{\sigma_0}{3}, H_2}\} ds = +\infty,$$

which contradicts (3.21) since the inequality (3.21) also holds for $H_2, \sigma_0/3$. We therefore conclude that

$$\lim_{t \rightarrow +\infty} \mathbb{P}\{A_t^{\sigma, h}\} = 0. \quad (3.24)$$

Next, we will prove the uniformity of the limit above for $\xi, \eta \in K$, that is, for any $\varepsilon > 0, h > \sigma/2 > 0$, there exists $T_\varepsilon^{\sigma, h} = T_\varepsilon^{\sigma, h}(K) > 0$ such that for all $t > T_\varepsilon^{\sigma, h}$ we have

$$\mathbb{P}\{\|x_t(\xi)\| \vee \|x_t(\eta)\| \leq h, |x(t, \xi) - x(t, \eta)| \geq \sigma\} < \varepsilon, \quad \forall \xi, \eta \in K.$$

In view of Lemma 3.1, we can find a constant $H_3 = H_3(K, \varepsilon) \geq h$ such that

$$\mathbb{P}\{\|x_t(\xi)\| > H_3\} < \frac{\varepsilon}{6}, \quad \forall \xi \in K, t > 0.$$

Put $m^{\sigma, H_3} = \min \{U(x-y) : |x| \vee |y| \leq H_3, |x-y| \geq \sigma\}$. Noting that $U(0) = 0$, $\mathcal{W}_2(x, y) = 0$ at $x = y \in \mathbb{R}^n$ and U, \mathcal{W}_2 are continuous, for any $\varepsilon > 0$ we therefore can choose $\delta_1 > 0$ such that for any $\xi, \eta \in K$

$$U(\xi(0) - \eta(0)) + \lambda_4 \int_{-\tau}^0 \mathcal{W}_2(\xi(s), \eta(s)) ds < \frac{\varepsilon}{6} m^{\sigma, H_3} \quad (3.25)$$

provided that $\|\xi - \eta\| \leq \delta_1$. In fact, if $\xi, \eta \in K$, there exists a constant $r > 0$ such that $\xi(s) \in B_0^n(r)$ and $(\xi(s), \eta(s)) \in B_0^{2n}(\sqrt{2}r)$ for any $s \in [-\tau, 0]$, where $B_0^n(r)$ stands for an n -dimensional ball with radius r centered at the origin in \mathbb{R}^n . Additionally, note that for any $s \in [-\tau, 0]$, $|\xi(0) - \eta(0)| \leq \|\xi - \eta\|$ and $|(\xi(s), \eta(s)) - (\xi(s), \xi(s))| \leq \|\xi(s) - \eta(s)\|$ and $U(0) = 0$, $\mathcal{W}_2(\xi(s), \xi(s)) = 0$. Hence, for any $\varepsilon > 0$, we can find a sufficiently small $\delta_1 > 0$ such that (3.25) holds provided $\|\xi - \eta\| \leq \delta_1$ with $\xi, \eta \in K$. If $\|\xi - \eta\| \leq \delta_1$ with $\xi, \eta \in K$, it follows from (3.20) and (3.25) that for any $t \geq 0$,

$$\begin{aligned} & \mathbb{P} \{ \|x_t(\xi)\| \vee \|x_t(\eta)\| \leq H_3, |x(t, \xi) - x(t, \eta)| \geq \sigma \} \\ & \leq \mathbb{P} \{ |x(t, \xi)| \vee |x(t, \eta)| \leq H_3, |x(t, \xi) - x(t, \eta)| \geq \sigma \} \\ & \leq \frac{1}{m^{\sigma, H_3}} \mathbb{E} U(x(t, \xi) - x(t, \eta)) \leq \frac{\varepsilon}{6}. \end{aligned}$$

By the compactness of K , there exist $\xi_1, \xi_2, \dots, \xi_n \in K$ such that for any $\xi \in K$, we can find $\xi_k, k = \{1, 2, \dots, n\}$ such that $\|\xi - \xi_k\| \leq \delta_1$. By (3.24), there exists $T_\varepsilon^{\sigma, H_3} > 0$ such that for all $1 \leq u, v \leq n$,

$$\mathbb{P} \left\{ \|x_t(\xi_u)\| \vee \|x_t(\xi_v)\| \leq H_3, |x(t, \xi_u) - x(t, \xi_v)| \geq \frac{\sigma}{3} \right\} \leq \frac{\varepsilon}{6}, \quad \forall t \geq T_\varepsilon^{\sigma, H_3}.$$

For any $\xi, \eta \in K$, we can find ξ_u, ξ_v such that $\|\xi - \xi_u\| \leq \delta_1$, $\|\eta - \xi_v\| \leq \delta_1$, as a result, for all $t \geq T_\varepsilon^{\sigma, H_3}$,

$$\begin{aligned} & \mathbb{P} \{ \|x_t(\xi)\| \vee \|x_t(\eta)\| \leq h, |x(t, \xi) - x(t, \eta)| \geq \sigma \} \\ & \leq \mathbb{P} \{ \|x_t(\xi)\| \vee \|x_t(\eta)\| \leq H_3, |x(t, \xi) - x(t, \eta)| \geq \sigma \} \\ & \leq \mathbb{P} \left\{ \|x_t(\xi)\| \vee \|x_t(\xi_u)\| \leq H_3, |x(t, \xi) - x(t, \xi_u)| \geq \frac{\sigma}{3} \right\} \\ & \quad + \mathbb{P} \left\{ \|x_t(\xi_u)\| \vee \|x_t(\xi_v)\| \leq H_3, |x(t, \xi_u) - x(t, \xi_v)| \geq \frac{\sigma}{3} \right\} \\ & \quad + \mathbb{P} \left\{ \|x_t(\xi_v)\| \vee \|x_t(\eta)\| \leq H_3, |x(t, \xi_v) - x(t, \eta)| \geq \frac{\sigma}{3} \right\} \\ & \quad + \mathbb{P} \{ \|x_t(\xi_u)\| > H_3 \} + \mathbb{P} \{ \|x_t(\xi_v)\| > H_3 \} < \varepsilon \end{aligned}$$

as desired.

Step 2. Letting $\varepsilon > 0$ arbitrarily, by virtue of Lemma 3.1, there exists $H_4 = H_4(K, \tau, \varepsilon)$ such that

$$\mathbb{P} \{ \|x_s(\xi)\| \leq H_4, s \in [t, t + \tau] \} \geq 1 - \frac{\varepsilon}{16}, \quad \forall \xi \in K, t \geq 0. \quad (3.26)$$

Defining a stopping time $\tau_t = \inf \{s \geq t : \|x_s(\xi)\| \vee \|x_s(\eta)\| > H_4\}$, it follows from (3.26) that

$$\mathbb{P} \{ \tau_t \geq t + \tau \} = \mathbb{P} \{ \|x_s(\xi)\| \vee \|x_s(\eta)\| \leq H_4, s \in [t, t + \tau] \} \geq 1 - \frac{\varepsilon}{8}.$$

Using the similar argument as (3.14) and the Chebyshev inequality, for any $0 < \delta < \tau$ and $t \leq s_1 \leq t + \tau$, we have

$$\begin{aligned}
& \mathbb{P} \left\{ \tau_t \geq t + \tau, \sup_{s_2 \in [s_1, (s_1 + \delta) \wedge (t + \tau)]} |x(s_2, \xi) - x(s_1, \xi)| \geq \frac{\varepsilon}{3} \right\} \\
& \leq \mathbb{P} \left\{ \tau_{s_1} \geq t + \tau, \sup_{s_2 \in [s_1, (s_1 + \delta) \wedge (t + \tau)]} |x(s_2 \wedge \tau_{s_1}, \xi) - x(s_1, \xi)| \geq \frac{\varepsilon}{3} \right\} \\
& \leq \frac{81}{\varepsilon^4} \mathbb{E} \sup_{s_2 \in [s_1, (s_1 + \delta) \wedge (t + \tau)]} |x(s_2 \wedge \tau_{s_1}, \xi) - x(s_1, \xi)|^4 \\
& \leq \frac{81\acute{C}}{\varepsilon^4} \delta^2, \tag{3.27}
\end{aligned}$$

where \acute{C} is a constant depending on K, H_4, ε .

Let $m \in \mathbb{N}$ such that $81\acute{C}\delta/\varepsilon^4 \leq \varepsilon/8\tau$, $\delta = \tau/m$. In view of (3.27), for $k = 0, 1, \dots, m-1$, we have

$$\mathbb{P} \left\{ \{\tau_t \geq t + \tau\} \cap \left\{ \sup_{s \in [t + k\delta, t + (k+1)\delta]} |x(s, \xi) - x(t + k\delta, \xi)| \geq \frac{\varepsilon}{3} \right\} \right\} \leq \frac{\varepsilon}{8\tau} \delta.$$

As a result,

$$\mathbb{P} \left\{ \{\tau_t \geq t + \tau\} \cap C_t^\xi \right\} \leq \frac{\varepsilon}{8\tau} \delta m = \frac{\varepsilon}{8},$$

where

$$C_t^\xi = \left\{ \exists k \in \{0, 1, 2, \dots, m-1\} : \sup_{s \in [t + k\delta, t + (k+1)\delta]} |x(s, \xi) - x(t + k\delta, \xi)| \geq \frac{\varepsilon}{3} \right\}.$$

Hence

$$\mathbb{P} \left\{ \{\tau_t \geq t + \tau\} \setminus C_t^\xi \right\} \geq 1 - \frac{\varepsilon}{8} - \frac{\varepsilon}{8} = 1 - \frac{\varepsilon}{4}. \tag{3.28}$$

Similarly, we have

$$\mathbb{P} \left\{ \{\tau_t \geq t + \tau\} \setminus C_t^\eta \right\} \geq 1 - \frac{\varepsilon}{4}. \tag{3.29}$$

Owing to the uniform convergence shown in step 1, we can find $T_1 = T_1(K, \varepsilon)$ such that for any $\xi, \eta \in K$, $t > T_1$,

$$\sum_{k=0}^{m-1} \mathbb{P} \left\{ \|x_{t+k\delta}(\xi)\| \vee \|x_{t+k\delta}(\eta)\| \leq H_4 : |x(t + k\delta, \xi) - x(t + k\delta, \eta)| \geq \frac{\varepsilon}{3} \right\} \leq \frac{\varepsilon}{4},$$

which implies that $\mathbb{P}\{D_t\} \leq \varepsilon/4$, where

$$D_t = \left\{ \exists k \in \{0, 1, \dots, m-1\} : \|x_{t+k\delta}(\xi)\| \vee \|x_{t+k\delta}(\eta)\| \leq H_4, |x(t + k\delta, \xi) - x(t + k\delta, \eta)| \geq \frac{\varepsilon}{3} \right\}.$$

Thus, for $t > T_1$,

$$\mathbb{P} \left\{ \{\tau_t \geq t + \tau\} \setminus D_t \right\} \geq 1 - \frac{\varepsilon}{8} - \frac{\varepsilon}{4} = 1 - \frac{3}{8}\varepsilon. \tag{3.30}$$

Note that if the three events $\{\tau_t \geq t + \tau\} \setminus D_t$, $\{\tau_t \geq t + \tau\} \setminus C_t^\xi$, $\{\tau_t \geq t + \tau\} \setminus C_t^\eta$ occur simultaneously, the event $\{\tau_t \geq t + \tau, \sup_{t \leq s \leq t+\tau} |x(s, \xi) - x(s, \eta)| < \varepsilon\}$ will occur. This statement, combining with (3.28)-(3.30) implies that for any $\xi, \eta \in K$ and $t > T_1$,

$$\mathbb{P}\{\tau_t \geq t + \tau, \sup_{t \leq s \leq t+\tau} |x(s, \xi) - x(s, \eta)| < \varepsilon\} \geq 1 - \frac{7\varepsilon}{8}. \quad (3.31)$$

According to the definition of the norm $\|\cdot\|$, we have

$$\begin{aligned} & \mathbb{P}\{\|x_{t+\tau}(\xi) - x_{t+\tau}(\eta)\| < \varepsilon\} \\ &= \mathbb{P}\left\{\sup_{-\tau < \theta \leq 0} |x(t + \tau + \theta, \xi) - x(t + \tau + \theta, \eta)| < \varepsilon\right\} \\ &\geq \mathbb{P}\left\{\tau_t \geq t + \tau, \sup_{t \leq s \leq t+\tau} |x(s, \xi) - x(s, \eta)| < \varepsilon\right\}. \end{aligned} \quad (3.32)$$

Therefore, it follows from (3.31) and (3.32) that for all $t > T_1$

$$\mathbb{P}\{\|x_{t+\tau}(\xi) - x_{t+\tau}(\eta)\| < \varepsilon\} \geq 1 - \varepsilon, \quad \forall \xi, \eta \in K.$$

The proof is completed. \square

Remark 5. Similarly, we can prove that for any $\xi, \eta \in \mathcal{C}$, $\varepsilon > 0$, there exists $T = T(\xi, \eta, \varepsilon)$ such that for all $t > T$,

$$\mathbb{P}\{\|x_t(\xi) - x_t(\eta)\| < \varepsilon\} \geq 1 - \varepsilon.$$

Remark 6. In [11], to establish Lemma 3.3 for Markovian switching SDEs with the constant time lag, the following assumption is imposed: there exists a nonnegative number λ_4 and functions $U \in C^2(\mathbb{R}^n \times S; \mathbb{R}_+)$, $\mathcal{W}_2 \in C(\mathbb{R}^n, \mathbb{R}_+)$, $\lambda_3 \in C(\mathbb{R}^{4n}, \mathbb{R}_+)$ such that $U(\cdot, i)$, \mathcal{W}_2 vanish only at 0 for all $i \in S$, $\lambda_3(x, y, z_1, z_2) > \lambda_4$ provided $x \neq y$ and

$$\mathcal{L}U(x, y, z_1, z_2, i) \leq -\lambda_3(x, y, z_1, z_2)\mathcal{W}_2(x - y) + \lambda_4\mathcal{W}_2(z_1 - z_2),$$

where S is a finite state space of a Markov chain. We replace the previous assumption by **(H4)** in this paper. The two terms on the right hand side of **(H4)** depend not only on $\phi - \psi$ but also on ϕ, ψ themselves, where $\phi, \psi \in \mathcal{C}$. And clearly **(H4)** is weaker than the assumption above.

Let $\mathcal{P}(\mathcal{C})$ denote the family of all probability measures on \mathcal{C} . For $P_1, P_2 \in \mathcal{P}(\mathcal{C})$, define the metric $d_{\mathbb{L}}$ by

$$d_{\mathbb{L}}(P_1, P_2) = \sup_{f \in \mathbb{L}} \left| \int_{\mathcal{C}} f(\xi) P_1(d\xi) - \int_{\mathcal{C}} f(\xi) P_2(d\xi) \right|,$$

where

$$\mathbb{L} = \{f : \mathcal{C} \rightarrow \mathbb{R} : |f(\xi) - f(\eta)| \leq \|\xi - \eta\| \text{ and } |f(\cdot)| \leq 1 \text{ for any } \xi, \eta \in \mathcal{C}\}.$$

Lemma 3.4. *Let Assumptions **(H1)**, **(H2)**, **(H3)** and **(H4)** hold. Then for any $\xi, \eta \in \mathcal{C}$*

$$\lim_{t \rightarrow \infty} d_{\mathbb{L}}(p(t, \xi, \cdot), p(t, \eta, \cdot)) = 0. \quad (3.33)$$

Proof. Fix any $\xi, \eta \in \mathcal{C}, \varepsilon > 0$. For any $f \in \mathbb{L}$, compute

$$\begin{aligned} & |\mathbb{E}f(x_t(\xi)) - \mathbb{E}f(x_t(\eta))| \\ & \leq \mathbb{E}|f(x_t(\xi)) - f(x_t(\eta))| \\ & \leq \mathbb{E}(2 \wedge \|x_t(\xi) - x_t(\eta)\|) \\ & \leq 2\mathbb{P}\left\{\|x_t(\xi) - x_t(\eta)\| \geq \frac{\varepsilon}{3}\right\} + \frac{\varepsilon}{3}\mathbb{P}\left\{\|x_t(\xi) - x_t(\eta)\| < \frac{\varepsilon}{3}\right\}. \end{aligned} \quad (3.34)$$

In view of Remark 5, there exists $T = T(\xi, \eta, \varepsilon)$ such that for all $t > T$

$$\mathbb{P}\left\{\|x_t(\xi) - x_t(\eta)\| \geq \frac{\varepsilon}{3}\right\} \leq \frac{\varepsilon}{3}. \quad (3.35)$$

Substituting (3.35) into (3.34), we obtain

$$|\mathbb{E}f(x_t(\xi)) - \mathbb{E}f(x_t(\eta))| \leq \varepsilon, \quad \forall t > T.$$

Since f is arbitrary, we must have that

$$\sup_{f \in \mathbb{L}} |\mathbb{E}f(x_t(\xi)) - \mathbb{E}f(x_t(\eta))| \leq \varepsilon, \quad \forall t > T,$$

namely,

$$d_{\mathbb{L}}(p(t, \xi, \cdot), p(t, \eta, \cdot)) \leq \varepsilon, \quad \forall t > T.$$

This proof is completed. \square

Lemma 3.5. *Let Assumptions **(H1)**, **(H2)**, **(H3)** and **(H4)** hold. Then for any $\xi \in \mathcal{C}$, the family $\{p(t, \xi, \cdot) : t \geq 0\}$ is Cauchy in the space $\mathcal{P}(\mathcal{C})$ with metric $d_{\mathbb{L}}$.*

Proof. Fix $\xi \in \mathcal{C}$ arbitrarily. We need to show that for any $\varepsilon > 0$, there is $T > 0$ such that

$$d_{\mathbb{L}}(p(t+s, \xi, \cdot), p(t, \xi, \cdot)) \leq \varepsilon, \quad \forall t > T, s > 0.$$

This is equivalent to

$$\sup_{f \in \mathbb{L}} |\mathbb{E}f(x_{t+s}(\xi)) - \mathbb{E}f(x_t(\xi))| \leq \varepsilon, \quad \forall t > T, s > 0.$$

For any $f \in \mathbb{L}$ and $s, t > 0$, using the Markov property, we compute

$$\begin{aligned} & |\mathbb{E}f(x_{t+s}(\xi)) - \mathbb{E}f(x_t(\xi))| \\ & = |\mathbb{E}[\mathbb{E}(f(x_{t+s}(\xi)) | \mathcal{F}_s)] - \mathbb{E}f(x_t(\xi))| \\ & = \left| \int_{\mathcal{C}} \mathbb{E}f(x_t(\eta)) p(s, \xi, d\eta) - \mathbb{E}f(x_t(\xi)) \right| \\ & \leq \int_{\mathcal{C}} \mathbb{E}|f(x_t(\eta)) - f(x_t(\xi))| p(s, \xi, d\eta). \end{aligned} \quad (3.36)$$

By Lemma 3.2, there exists a compact subset $\bar{K} \subseteq \mathcal{C}$ such that

$$p(s, \xi, \bar{K}) > 1 - \frac{\varepsilon}{5}. \quad (3.37)$$

By (3.36) and (3.37), compute

$$\begin{aligned} & |\mathbb{E}f(x_{t+s}(\xi)) - \mathbb{E}f(x_t(\xi))| \\ & \leq \int_{\bar{K}} \mathbb{E}|f(x_t(\eta)) - f(x_t(\xi))| p(s, \xi, d\eta) + \frac{2}{5}\varepsilon \\ & \leq \int_{\bar{K} \cup \{\xi\}} \mathbb{E}(2 \wedge \|x_t(\eta) - x_t(\xi)\|) p(s, \xi, d\eta) + \frac{2}{5}\varepsilon. \end{aligned} \quad (3.38)$$

By Lemma 3.3, for any $\xi, \eta \in \hat{K} = K \cup \{\xi\}$, there exists $T = T(\hat{K}, \varepsilon)$ such that

$$\mathbb{P} \left\{ \|x_t(\eta) - x_t(\xi)\| \geq \frac{\varepsilon}{5} \right\} \leq \frac{\varepsilon}{5}, \quad \forall t > T. \quad (3.39)$$

By virtue of (3.38) and (3.39), we have

$$|\mathbb{E}f(x_{t+s}(\xi)) - \mathbb{E}f(x_t(\xi))| \leq \varepsilon, \quad \forall t > T, s > 0.$$

Since f is arbitrary, the desired assertion follows and the proof is completed. \square

After the preparation of the lemmas above, we can now state our main result.

Theorem 3.6. *Let Assumptions **(H1)**, **(H2)**, **(H3)** and **(H4)** hold. Then the solution mapping process x_t of (1.2) is stable in distribution.*

Proof. Fix $\xi \in \mathcal{C}$ arbitrarily. By Lemma 3.5, $\{p(t, \xi, \cdot) : t \geq 0\}$ is Cauchy in the space $\mathcal{P}(\mathcal{C})$ with metric $d_{\mathbb{L}}$. So there is a unique probability $\pi(\cdot) \in \mathcal{P}(\mathcal{C})$ such that

$$\lim_{t \rightarrow \infty} d_{\mathbb{L}}(p(t, \xi, \cdot), \pi(\cdot)) = 0.$$

Now, for any $\eta \in \mathcal{C}$, Lemma 3.4 shows that $\lim_{t \rightarrow \infty} d_{\mathbb{L}}(p(t, \eta, \cdot), p(t, \xi, \cdot)) = 0$. Therefore

$$\begin{aligned} & \lim_{t \rightarrow \infty} d_{\mathbb{L}}(p(t, \eta, \cdot), \pi(\cdot)) \\ & \leq \lim_{t \rightarrow \infty} d_{\mathbb{L}}(p(t, \eta, \cdot), p(t, \xi, \cdot)) + \lim_{t \rightarrow \infty} d_{\mathbb{L}}(p(t, \xi, \cdot), \pi(\cdot)) \\ & = 0 \end{aligned}$$

This completes the proof. \square

Remark 7. When the diffusion is non-degenerate, [18, Theorem 3.3] and [19, Throrem 2.1] showed that the stochastic differential equation has a unique invariant measure. In Theorem 3.3 of [18], the diffusion coefficient is Lipschitz continuous and the drift coefficient needs to be decomposed into two terms $f(\phi) = f_1(\phi) + f_2(\phi(0))$, $\phi \in \mathcal{C}$, where the functional f_1 is bounded. In Theorem 2.1 of [19], the diffusion g is Lipschitz continuous and the drift f is sublinear, that is, there exist constants $\beta \in [0, 1)$, $C > 0$ such that

$$|f(\phi)| \leq C(1 + \|\phi\|^\beta), \quad \phi \in \mathcal{C}.$$

This paper doesn't need these conditions and can involve more stochastic delay and functional models.

4 Examples

In order to illustrate the results established in the previous section, we present two examples in this section.

Example 4.1. Let us recall the scalar Equation (1.1). For the sake of convenience, write Equation (1.1) again here: For example, consider the scalar SFDEs

$$dx(t) = \left(a - bx(t) - cx^3(t) + \int_{-\tau}^0 x^2(t+\theta)\mu(d\theta) \right) dt + d \int_{-\tau}^0 |x(t+\theta)|^{\frac{3}{2}} \mu(d\theta) dB(t),$$

where a is a constant, $2b - 1 \geq 2c > 3d^2 + 2$, and μ is a probability measure on $[-\tau, 0]$. on $t \geq 0$, where a is a constant, $2b - 1 \geq 2c > 3d^2 + 2$, $\mu \in \mathcal{M}$.

It is obvious that the drift and diffusion of equation (1.1) satisfy assumption **(H1)** and **(H3)**. Let us check **(H2)** and **(H4)**. Putting $V(x) = x^2$, for all $\phi \in \mathcal{C}$ we have

$$\begin{aligned} LV(\phi) &= 2\phi(0) \left(a - b\phi(0) - c\phi^3(0) + \int_{-\tau}^0 \phi^2(\theta)\mu(d\theta) \right) + d^2 \left(\int_{-\tau}^0 |\phi(\theta)|^{\frac{3}{2}} \mu(d\theta) \right)^2 \\ &\leq 2a\phi(0) - (2b - 1)\phi^2(0) - 2c\phi^4(0) + \int_{-\tau}^0 \left(d^2|\phi(\theta)|^2 + (d^2 + 1)|\phi(\theta)|^4 \right) \mu(d\theta) \\ &\leq -\varepsilon\phi^2(0) + 2a\phi(0) - 2c \left(\frac{2b - 1 - \varepsilon}{2c} \phi^2(0) + \phi^4(0) \right) \\ &\quad + (d^2 + 1) \int_{-\tau}^0 \left(\frac{d^2}{d^2 + 1} \phi^2(\theta) + \phi^4(\theta) \right) \mu(d\theta) \\ &\leq L_2 - 2c \left(\frac{2b - 1 - \varepsilon}{2c} \phi^2(0) + \phi^4(0) \right) + (d^2 + 1) \int_{-\tau}^0 \left(\frac{d^2}{d^2 + 1} \phi^2(\theta) + \phi^4(\theta) \right) \mu(d\theta), \end{aligned}$$

where ε is a positive constant satisfying $(2b - 1 - \varepsilon)(d^2 + 1) \geq 2cd^2$, $L_2 = \sup_{x \in \mathbb{R}} \{-\varepsilon x^2 + 2ax\} < \infty$. If the probability measure μ is Dirac measure at $-\tau$, then (1.1) reduces to the SDE with constant time lag which has been discussed in [11]. It is easy to see from inequality above that there is no function \mathcal{W}_1 which makes Assumption 1.1 hold in [11]. Therefore we can not apply the theorem in [11]. However, our Assumption **(H2)** holds by choosing $\mathcal{W}_1(x) = (2b - 1 - \varepsilon)x^2/2c + x^4$, $\lambda_1 = 2c$, $\lambda_2 = d^2 + 1$. Next, applying $U(x) = x^2$, for ϕ ,

$\varphi \in \mathcal{C}$, we have

$$\begin{aligned}
\mathcal{L}U(\phi, \varphi) &= -2b[\phi(0) - \varphi(0)]^2 - 2c[\phi(0) - \varphi(0)]^2[\phi^2(0) + \phi(0)\varphi(0) + \varphi^2(0)] \\
&\quad + d^2 \left(\int_{-\tau}^0 (|\phi(\theta)|^{\frac{1}{2}} - |\varphi(\theta)|^{\frac{1}{2}})(|\phi(\theta)| + |\phi(\theta)\varphi(\theta)|^{\frac{1}{2}} + |\varphi(\theta)|)\mu(d\theta) \right)^2 \\
&\quad + 2 \int_{-\tau}^0 [\phi(0) - \varphi(0)][\phi^2(\theta) - \varphi^2(\theta)]\mu(d\theta) \\
&\leq -[\phi(0) - \varphi(0)]^2[2b - 1 + 2c(\phi^2(0) + \phi(0)\varphi(0) + \varphi^2(0))] \\
&\quad + d^2 \int_{-\tau}^0 (|\phi(\theta)| - |\varphi(\theta)|)^2(|\phi(\theta)| + |\phi(\theta)\varphi(\theta)|^{\frac{1}{2}} + |\varphi(\theta)|)\mu(d\theta) \\
&\quad + \int_{-\tau}^0 [\phi(\theta) - \varphi(\theta)]^2[\phi(\theta) + \varphi(\theta)]^2\mu(d\theta) \\
&\leq -c[\phi(0) - \varphi(0)]^2 \left[\phi^2(0) + \varphi^2(0) + \frac{2b-1}{c} \right] \\
&\quad + \left(\frac{3}{2}d^2 + 1 \right) \int_{-\tau}^0 [\phi(\theta) - \varphi(\theta)]^2[\phi^2(\theta) + \varphi^2(\theta) + 2]\mu(d\theta).
\end{aligned}$$

Similarly, it is easy to see from inequality above that there is also no \mathcal{W}_2 which makes Assumption 1.2 hold in [11]. However, our Assumption **(H4)** holds with $\mathcal{W}_2(x, y) = (x - y)^2(x^2 + y^2 + (2b - 1)/c)$, $\lambda_3 = c$, $\lambda_4 = 3d^2/2 + 1$. Therefore, (1.1) is stable in distribution.

Example 4.2. Consider the following two-dimensional SFDE

$$d \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} a_1 - x_1(t) - b_1 x_1^3(t) \\ a_2 - b_2 x_2^3(t) \end{pmatrix} dt + \begin{pmatrix} c_1 \int_{-\tau}^0 x_2^2(t + \theta)\mu(d\theta) \\ c_2 \int_{-\tau}^0 x_1^2(t + \theta)\mu(d\theta) \end{pmatrix} dB(t) \quad (4.1)$$

with the initial data $x_0 = \xi \in \mathcal{C}$, where $a_1, a_2, b_1, b_2, c_1, c_2$ are constants, $\mu \in \mathcal{M}$, $B(t)$ is a scalar Brownian motion. We will show that if $b_1 > c_2^2$, $b_2 c_2^2 > b_1 c_1^2$ or $b_2 > c_1^2$, $b_1 c_1^2 > b_2 c_2^2$, (4.1) is stable in distribution. An analysis similar to Example 4.1 indicates that the theorem in [11] is not applicable even if the diffusion coefficient is replaced by $(c_1 x_2^2(t - \tau), c_2 x_1^2(t - \tau))^T$ (namely, (4.1) reduces to the SDE in [11]). It is easy to verify that the drift f is continuous and bounded on bounded subsets of \mathcal{C} and the diffusion g is local Lipschitz continuous. We therefore only need to check that the coefficients of (4.1) satisfy **(H2)** and **(H4)**.

Let us verify that f and g satisfy **(H2)** and **(H4)** when $b_1 > c_2^2$, $b_2 c_2^2 > b_1 c_1^2$. Letting

$V(x) = x_1^2 + x_2^2$, we have

$$\begin{aligned}
LV(\phi) &= 2a_1\phi_1(0) - 2\phi_1^2(0) - 2b_1\phi_1^4(0) + 2a_2\phi_2(0) - 2b_2\phi_2^4(0) + c_1^2 \left(\int_{-\tau}^0 \phi_2^2(\theta)\mu(d\theta) \right)^2 \\
&\quad + c_2^2 \left(\int_{-\tau}^0 \phi_1^2(\theta)\mu(d\theta) \right)^2 \\
&\leq [2a_1\phi_1(0) - 2\phi_1^2(0)] + [2a_2\phi_2(0) - \varepsilon\phi_2^4(0)] - 2b_1\phi_1^4(0) - (2b_2 - \varepsilon)\phi_2^4(0) \\
&\quad + c_2^2 \int_{-\tau}^0 \phi_1^4(\theta)\mu(d\theta) + c_1^2 \int_{-\tau}^0 \phi_2^4(\theta)\mu(d\theta) \\
&\leq L_2 - 2b_1 \left[\phi_1^4(0) + \frac{2b_2 - \varepsilon}{2b_1} \phi_2^4(0) \right] + c_2^2 \int_{-\tau}^0 \phi_1^4(\theta) + \left(\frac{c_1}{c_2} \right)^2 \phi_2^4(\theta)\mu(d\theta),
\end{aligned}$$

where ε is a positive constant satisfying $(2b_2 - \varepsilon)c_2^2 > 2b_1c_1^2$, $L_2 = \sup_{x \in \mathbb{R}^2} \{2a_1x_1 - 2x_1^2 + 2a_2x_2 - \varepsilon x_2^4\} < +\infty$. The above ε actually exists, since $b_2c_2^2 > b_1c_1^2$. Hence, the assumption **(H2)** holds with $\lambda_1 = 2b_1, \lambda_2 = c_2^2, \mathcal{W}_1(x) = x_1^4 + \frac{2b_2 - \varepsilon}{2b_1}x_2^4$. Now we check the assumption **(H4)**. Applying $U(x) = x_1^2 + x_2^2$, we obtain

$$\begin{aligned}
\mathcal{L}U(\phi, \varphi) &= -2b_1[\phi_1(0) - \varphi_1(0)][\phi_1^3(0) - \varphi_1^3(0)] - 2b_2[\phi_2(0) - \varphi_2(0)][\phi_2^3(0) - \varphi_2^3(0)] \\
&\quad - 2[\phi_1(0) - \varphi_1(0)]^2 + c_1^2 \left(\int_{-\tau}^0 [\phi_2^2(\theta) - \varphi_2^2(\theta)]\mu(d\theta) \right)^2 \\
&\quad + c_2^2 \left(\int_{-\tau}^0 [\phi_1^2(\theta) - \varphi_1^2(\theta)]\mu(d\theta) \right)^2 \\
&\leq -2 \sum_{i=1}^2 \{b_i[\phi_i(0) - \varphi_i(0)]^2[\phi_i^2(0) + \phi_i(0)\varphi_i(0) + \varphi_i^2(0)]\} \\
&\quad + \sum_{i=1}^2 \left\{ c_i^2 \int_{-\tau}^0 [\phi_i(\theta) - \varphi_i(\theta)]^2[\phi_i(\theta) + \varphi_i(\theta)]^2\mu(d\theta) \right\} \\
&\leq -b_1 \sum_{i=1}^2 \left\{ \frac{b_i}{b_1}[\phi_i(0) - \varphi_i(0)]^2[(\phi_i(0) + \varphi_i(0))^2 + \phi_i^2(0) + \varphi_i^2(0)] \right\} \\
&\quad + c_2^2 \sum_{i=1}^2 \left\{ \left(\frac{c_i}{c_2} \right)^2 \int_{-\tau}^0 [\phi_i(\theta) - \varphi_i(\theta)]^2[(\phi_i(\theta) + \varphi_i(\theta))^2 + \phi_i^2(\theta) + \varphi_i^2(\theta)]\mu(d\theta) \right\},
\end{aligned}$$

which implies the assumption **(H4)** holds with $\lambda_3 = b_1, \lambda_4 = c_2^2$ and

$$\mathcal{W}_2(x, y) = \sum_{i=1}^2 \left\{ \frac{b_i}{b_1}(x_i - y_i)^2((x_i + y_i)^2 + x_i^2 + y_i^2) \right\}.$$

Hence, (4.1) is stable in distribution.

Similarly, we can also show that (4.1) is stable in distribution when $b_2 > c_1^2, b_1c_1^2 > b_2c_2^2$.

References

- [1] S.-E. A. Mohammed, *Stochastic functional differential equations*. Pitman Advanced Publishing Program, 1984, vol. 99.
- [2] X. Mao, *Stochastic differential equations and applications*. Elsevier, 2007.
- [3] X. Mao, Razumikhin-type theorems on exponential stability of stochastic functional differential equations, *Stochastic Processes and their Applications*, vol. 65, no. 2, pp. 233–250, 1996.
- [4] V. B. Kolmanovskii and V. R. Nosov, *Stability of functional differential equations*. Elsevier, 1986, vol. 180.
- [5] X. Mao, *Exponential stability of stochastic differential equations*. Marcel Dekker, 1994.
- [6] G. K. Basak and R. N. Bhattacharya, “Stability in distribution for a class of singular diffusions,” *The Annals of Probability*, vol. 20, no. 1, pp. 312–321, 1992.
- [7] G. K. Basak, A. Bisi, and M. K. Ghosh, “Stability of a random diffusion with linear drift,” *Journal of Mathematical Analysis and Applications*, vol. 202, no. 2, pp. 604–622, 1996.
- [8] G. K. Basak, A. Bisi, M. K. Ghosh, “Stability and functional limit theorems for random degenerate diffusions”. *Sankhyā: The Indian Journal of Statistics, Series A*, pp. 12-35, 1999.
- [9] C. Yuan and X. Mao, “Asymptotic stability in distribution of stochastic differential equations with markovian switching,” *Stochastic processes and their applications*, vol. 103, no. 2, pp. 277–291, 2003.
- [10] C. Yuan, J. Zou, and X. Mao, “Stability in distribution of stochastic differential delay equations with markovian switching,” *Systems & control letters*, vol. 50, no. 3, pp. 195–207, 2003.
- [11] N. H. Du, N. H. Dang, and N. T. Dieu, “On stability in distribution of stochastic differential delay equations with markovian switching,” *Systems & Control Letters*, vol. 65, pp. 43–49, 2014.
- [12] N H. Dang, “A note on sufficient conditions for asymptotic stability in distribution of stochastic differential equations with Markovian switching”. *Nonlinear Analysis: Theory, Methods & Applications*, vol. 95, pp. 625-631, 2014.
- [13] G. Hu and K. Wang, “Stability in distribution of neutral stochastic functional differential equations with markovian switching,” *Journal of Mathematical Analysis and Applications*, vol. 385, no. 2, pp. 757–769, 2012.
- [14] J. Bao, G. Yin, and C. Yuan, “Ergodicity for functional stochastic differential equations and applications,” *Nonlinear Anal.*, vol. 98, pp. 66–82, 2014.

- [15] J. Bao, G. Yin, C. Yuan, and L. Y. Wang, “Exponential ergodicity for retarded stochastic differential equations,” *Appl. Anal.*, vol. 93, no. 11, pp. 2330–2349, 2014.
- [16] J. Bao, G. Yin, and C. Yuan, “Stationary distributions for retarded stochastic differential equations without dissipativity,” *Stochastics*, vol. 89, no. 2, pp. 530–549, 2017.
- [17] M. Hairer, J. C. Mattingly, and M. Scheutzow, “Asymptotic coupling and a general form of Harris’ theorem with applications to stochastic delay equations,” *Probab. Theory Related Fields*, vol. 149, no. 1-2, pp. 223–259, 2011.
- [18] O. Butkovsky, “Subgeometric rates of convergence of Markov processes in the Wasserstein metric,” *Ann. Appl. Probab.*, vol. 24, no. 2, pp. 526–552, 2014.
- [19] O. Butkovsky and M. Scheutzow, “Invariant measures for stochastic functional differential equations,” *Electron. J. Probab.*, vol. 22, pp. Paper No. 98, 23, 2017.
- [20] J. Bao, G. Yin, and C. Yuan, *Asymptotic analysis for functional stochastic differential equations*, ser. SpringerBriefs in Mathematics. Springer, Cham, 2016.
- [21] F. Wu, G. Yin, and H. Mei, “Stochastic functional differential equations with infinite delay: existence and uniqueness of solutions, solution maps, Markov properties, and ergodicity,” *J. Differential Equations*, vol. 262, no. 3, pp. 1226–1252, 2017.
- [22] P. Billingsley, *Convergence of probability measures*. John Wiley & Sons, 2013.
- [23] M.-K. von Renesse and M. Scheutzow, “Existence and uniqueness of solutions of stochastic functional differential equations,” *Random Oper. Stoch. Equ.*, vol. 18, no. 3, pp. 267–284, 2010.
- [24] Y. Shen, Q. Luo, X. Mao, “The improved LaSalle-type theorems for stochastic functional differential equations,” *J. Math. Anal. Appl.*, Vol. 318, no. 1, pp. 134C154, 2006.