# Strong Convergence of Neutral Stochastic Functional Differential Equations with Two Time-Scales

Junhao Hu *        Chenggui Yuan[†]

December 8, 2018

## Abstract

The purpose of this paper is to discuss the strong convergence of neutral stochastic functional differential equations (NSFDEs) with two time-scales. The existence and uniqueness of invariant measure of the fast component is proved by using Wasserstein distance and the stability-in-distribution argument. The strong convergence between the slow component and the averaged component is also obtained by the the averaging principle in the spirit of Khasminskii's approach.

**Keywords:** Two time-scales, neutral functional differential equations, exponential ergodicity, invariant measure, averaging principle

**AMS Subject Classification:** 60H15, 60J25, 60H30, 39B82

---

*College of Mathematics and Statistics, South-Central University for Nationalities, Wuhan, 430074, PR China, junhaohu74@163.com

[†]Department of Mathematics, Swansea University, SA2 8PP, UK, C.Yuan@swansea.ac.uk

# 1  Introduction

It is well-known that time-delay cannot be avoided in practise and it often results in instability and poor performance. Delays become more and more prevalent in physical, cyber-physical and biological systems, and hence delay dynamical systems are used in a lot of models of science and engineering. In many applications, the properties being modelled will also need to satisfy other constraints, for example in order to describe vibrating masses attached to an elastic bar, Drive [7] consider a system of neutral type as follows

$$(1.1) \qquad dX(t) + h(X(t), X(t-\tau))dX(t-\tau) = f(X(t), X(t-\tau)),$$

where $\tau$ is a positive constant. Eq. (1.1) involves derivative with delay, and is called neutral delay equation. Taking stochastic perturbations into account, a neutral stochastic functional differential equation has the following form

$$(1.2) \qquad d[X(t) + D(X_t)] = f(X_t, t)dt + g(X_t)dW(t),$$

Here $X(t)$ denotes the value of the stochastic process $X$ at the time $t$, while $X_t = (X(t+\theta) : -\tau \le \theta \le 0)$ which is called the segment process of $X$ during the delay interval $[t-\tau, t]$, $D(X_t)$ is called the neutral term. This system includes derivative with delay and is driven by a standard Brownian motion $W(t)$. However such a model does not take into account the rates of changes of the systems or different rates of interactions of subsystems and components. To describe such scenarios a singularly perturbed system is often used. Generally speaking, a singularly perturbed system exhibits multi-scale behavior, which is reflected by a slow subsystem and a fast subsystem. Due to the multi-scale property, it is frequently difficult to deal with such systems using a direct approach, and the averaging principle method pioneered by Khasminskii [18] is used in many papers, for example: [13, 22, 23, 27, 4, 5, 12, 19, 10, 14, 20, 21], and reference therein. In particular, the averaging principle method has been studied for stochastic functional differential equations in [1], our aim is to extend results in [1] to neutral stochastic functional differential equations with two-time-scales. Because of the neutral term, we can see that the techniques in the present paper are much more complicated and different from those of [1].

In this paper, we shall bring delay, neutral, multi-scale and noise together, and investigate the strong convergence of neutral stochastic functional differential equations (NSFDEs) with two time-scales. The rest of the paper is organized as follows. Section 2 presents the setup of the problems and the main results we wish to study. The proof of the ergodicity of a frozen equation with memory is obtained in Section 3. In Section 4 after we constructs some auxiliary neutral functional stochastic systems with two time-scales and provides a number of lemmas, the main result is then proved.

# 2  Problem Formulation and Main Results

Throughout the paper, let $\mathbb{R}^n$ be an $n$-dimensional Euclidean space, and $\mathbb{R}^{n \times m}$ denote the collection of all $n \times m$ matrices with real entries. For an $A \in \mathbb{R}^{n \times m}$, $\|A\|$ stands for its

1 Hilbert-Schimidt norm. For a fixed $\tau > 0$, let $\mathscr{C} = C([-\tau, 0]; \mathbb{R}^n)$ denote the family of
2 all continuous functions from $[-\tau, 0] \mapsto \mathbb{R}^n$, endowed with the uniform norm $\| \cdot \|_\infty$. For
3 $h(\cdot) \in C([-\tau, \infty); \mathbb{R}^n)$ and $t \geq 0$, define the segment $h_t \in \mathscr{C}$ by $h_t(\theta) = h(t + \theta)$, $\theta \in [-\tau, 0]$.
4 Generic constants will be denoted by $c$, we use the shorthand notation $a \lesssim b$ to mean $a \leq cb$,
5 we use $a \lesssim_T b$ to emphasize the constant $c$ depends on $T$.

Let $\varepsilon \in (0, 1)$, we consider a class of NSFDEs with two time-scales

$$(2.1) \qquad \mathrm{d}[X^\varepsilon(t) - D_1(X_t^\varepsilon)] = b_1(X_t^\varepsilon, Y_t^\varepsilon)\mathrm{d}t + \sigma_1(X_t^\varepsilon)\mathrm{d}W_1(t), \quad t > 0, \quad X_0^\varepsilon = \xi \in \mathscr{C},$$

and

$$(2.2) \qquad \begin{aligned} \mathrm{d}[Y^\varepsilon(t) - D_2(Y^\varepsilon(t - \tau))] &= \frac{1}{\varepsilon}b_2(X_t^\varepsilon, Y^\varepsilon(t), Y^\varepsilon(t - \tau))\mathrm{d}t \\ &\quad + \frac{1}{\sqrt{\varepsilon}}\sigma_2(X_t^\varepsilon, Y^\varepsilon(t), Y^\varepsilon(t - \tau))\mathrm{d}W_2(t), \ t > 0 \end{aligned}$$

6 with the initial value $Y_0^\varepsilon = \eta \in \mathscr{C}$, where $b_1 : \mathscr{C} \times \mathscr{C} \to \mathbb{R}^n$, $b_2 : \mathscr{C} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$,
7 $\sigma_1 : \mathscr{C} \to \mathbb{R}^{n \times m}$, $\sigma_2 : \mathscr{C} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{n \times m}$ are Gâteaux differentiable, $D_1 : \mathscr{C} \to \mathbb{R}^n$,
8 $D_2 : \mathbb{R}^n \to \mathbb{R}^n$ are measurable, locally bounded and continuous, $(W_1(t))_{t \geq 0}$ and $(W_2(t))_{t \geq 0}$
9 are two mutually independent $m$-dimensional Brownian motions defined on a probability
10 space $(\Omega, \mathscr{F}, \mathbb{P})$, equipped with a reference family $(\mathscr{F}_t)_{t \geq 0}$ satisfying the usual conditions
11 (i.e., for each $t \geq 0$, $\mathscr{F}_t = \mathscr{F}_{t+} := \bigcap_{s > t} \mathscr{F}_s$, and $\mathscr{F}_0$ contains all $\mathbb{P}$-null sets). $X^\varepsilon(t)$ is called
12 the slow component, while $Y^\varepsilon(t)$ is called the fast component.

13 Throughout the paper, for any $\chi, \phi, \overline{\chi}, \overline{\phi} \in \mathscr{C}$ and $x, x', y, y' \in \mathbb{R}^n$, we assume that

(**A1**) $|\nabla_{(\overline{\chi}, \overline{\phi})} b_1(\chi, \phi)| \leq L_1(1 + \|\overline{\chi}\|_\infty + \|\overline{\phi}\|_\infty)$ for some $L_1 > 0$, and there exists an $L > 0$ such that

$$|b_1(\chi, \phi)| \leq L(1 + \|\chi\|_\infty) \quad \text{and} \quad \|\sigma_1(\phi) - \sigma_1(\chi)\| \leq L\|\phi - \chi\|_\infty.$$

14 (**A2**) There exists $L_2 > 0$ such that $|\nabla_{(\phi, x', y')} b_2(\chi, x, y)| \leq L_2(1 + \|\phi\|_\infty + |x'| + |y'|)$ and
15 $\|\nabla_{(\phi, x', y')} \sigma_2(\chi, x, y)\| \leq L_2(1 + \|\phi\|_\infty + |x'| + |y'|)$.

(**A3**) There exist $\lambda_1 > \lambda_2 > 0$, independent of $\chi$, such that

$$\begin{aligned} 2\langle x - x' - (D_2(y) - D_2(y')), b_2(\chi, x, y) - b_2(\chi, x', y')\rangle + \|\sigma_2(\chi, x, y) - \sigma_2(\chi, x', y')\|^2 \\ \leq -\lambda_1|x - x'|^2 + \lambda_2|y - y'|^2. \end{aligned}$$

(**A4**) $D_1(0) = 0$, $D_2(0) = 0$, and there exist $\kappa_1, \kappa_2 \in (0, 1)$ such that

$$|D_1(\phi) - D_1(\chi)|^2 \leq \kappa_1\|\phi - \chi\|_\infty^2 \quad \text{and} \quad |D_2(y) - D_2(y')|^2 \leq \kappa_2|y - y'|^2.$$

(**A5**) For the initial value $X_0^\varepsilon = \xi \in \mathscr{C}$ of (2.1), there exists $\lambda_3 > 0$ such that

$$|\xi(t) - \xi(s)| \leq \lambda_3|t - s|, \quad s, t \in [-\tau, 0].$$

3

*Remark* 2.1. From (**A1**) and (**A2**), the gradient operators $\nabla b_1$, $\nabla b_2$, and $\nabla \sigma_2$ are bounded, respectively, these imply that $b_1$, $b_2$, and $\sigma_2$ are Lipschitz, so under (**A1**), (**A2**) and (**A4**), both (2.1) and (2.2) are well posed (see, e.g., [24, Theorem 2.2, P.204]). Here (**A3**) is imposed to analyze the ergodic property of the frozen equation (see Theorem 2.1 below), guarantee the Lipschitz property of $\bar{b}_1$ (see Corollary 2.2 below), defined in (2.5), and provide a uniform bound of the segment process $(Y_t^\varepsilon)_{t\in[0,T]}$ (see Lemma 4.3 below). Next, (**A5**) ensures that the displacement of the segment process $(X_t^\varepsilon)_{t\in[0,T]}$ is continuous in the mean $L^p$-norm sense (see Lemma 4.1 below).

Consider a neutral stochastic differential delay equation (NSDDE) associated with the fast motion while with the frozen slow component in the form

$$(2.3) \quad \begin{aligned} \mathrm{d}[Y(t) - D_2(Y(t-\tau))] &= b_2(\zeta, Y(t), Y(t-\tau))\mathrm{d}t \\ &\quad + \sigma_2(\zeta, Y(t), Y(t-\tau))\mathrm{d}W_2(t), \quad t > 0, \quad Y_0 = \eta \in \mathscr{C}. \end{aligned}$$

Under (**A2**) and (**A4**), Eq. (2.3) has a unique strong solution $(Y(t))_{t\geq-\tau}$ (see, e.g., [24, Theorem 2.2, P.204]). To highlight the initial value $\eta \in \mathscr{C}$ and the frozen segment $\zeta \in \mathscr{C}$, we write the corresponding solution process $(Y^\zeta(t,\eta))_{t\geq-\tau}$ and the segment process $(Y_t^\zeta(\eta))_{t\geq0}$ instead of $(Y(t))_{t\geq-\tau}$ and $(Y_t)_{t\geq0}$, respectively.

The first main result in this paper is stated as below which is concerned with ergodicity of the frozen Eq. (2.3) .

**Theorem 2.1.** Under (**A1**)-(**A4**), $Y_t^\zeta(\eta)$ has a unique invariant measure $\mu^\zeta$, and there exists $\lambda > 0$ such that

$$(2.4) \quad |\mathbb{E}b_1(\zeta, Y_t^\zeta(\eta)) - \bar{b}_1(\zeta)| \lesssim \mathrm{e}^{-\lambda t}(1 + \|\eta\|_\infty + \|\zeta\|_\infty), \quad t \geq 0, \quad \eta \in \mathscr{C},$$

where

$$(2.5) \quad \bar{b}_1(\zeta) := \int_{\mathscr{C}} b_1(\zeta, \varphi)\mu^\zeta(\mathrm{d}\varphi), \quad \zeta \in \mathscr{C}.$$

The next proposition, which plays a crucial role in discussing strong convergence for the averaging principle, states that $\bar{b}_1$ enjoys a Lipschitz property.

**Proposition 2.2.** Under (**A1**)-(**A4**), $\bar{b}_1 : \mathscr{C} \mapsto \mathbb{R}^n$, defined in (2.5), is Lipschitz.

Our main aim is to discuss the strong deviation between the slow component $X^\varepsilon(t)$ and the averaged component $\overline{X}(t)$, which satisfies the following NSFDE

$$(2.6) \quad \mathrm{d}[\overline{X}(t) - D_1(\overline{X}_t)] = \bar{b}_1(\overline{X}_t)\mathrm{d}t + \sigma_1(\overline{X}_t)\mathrm{d}W_1(t), \quad \overline{X}_0 = \xi \in \mathscr{C},$$

**Theorem 2.3.** Under (**A1**)-(**A5**), one has

$$\lim_{\varepsilon\to0}\mathbb{E}\left(\sup_{0\leq t\leq T}|X^\varepsilon(t) - \overline{X}(t)|^p\right) = 0, \quad p > 0.$$

4

# 3 Proofs of Theorem 2.1 and Proposition 2.2

Firstly, we give a proof for Theorem 2.1.

*Proof of Theorem 2.1.* The proof is rather technical so we divide it into six steps.

*Step* 1. We claim that there exists a positive number $\lambda'$ such that

$$(3.1) \qquad \mathbb{E}|Y^\zeta(t,\eta)|^2 \lesssim \mathrm{e}^{-\lambda' t}\|\eta\|_\infty^2 + 1 + \|\zeta\|_\infty^2, \quad t > 0.$$

By (**A2**), there exists $\alpha > 0$ such that

$$(3.2) \qquad \|\sigma_2(\chi, x, y) - \sigma_2(\chi, x', y')\| \leq \alpha(|x - x'| + |y - y'|),$$

and

$$(3.3) \qquad |b_2(\chi, 0, 0)| + \|\sigma_2(\chi, 0, 0)\| \leq \alpha(1 + \|\chi\|_\infty),$$

for any $\chi \in \mathscr{C}$ and $x, x', y, y' \in \mathbb{R}^n$. Accordingly, (3.2) and (3.3), together with (**A3**), yield that there exist $\lambda'_1 > \lambda'_2 > 0$, independent of $\chi$, such that

$$(3.4) \qquad 2\langle x - D_2(y), b_2(\chi, x, y)\rangle + \|\sigma_2(\chi, x, y)\|^2 \leq -\lambda'_1|x|^2 + \lambda'_2|y|^2 + c(1 + \|\chi\|_\infty^2)$$

for any $\chi \in \mathscr{C}$ and $x, y \in \mathbb{R}^n$. For a sufficiently small $\lambda' > 0$, applying Itô's formula, we infer from (3.4) that

$$
\begin{aligned}
(3.5) \quad & \mathrm{e}^{\lambda' t}\mathbb{E}|Y^\zeta(t,\eta) - D_2(Y^\zeta(t - \tau, \eta))|^2 \\
& = \mathbb{E}|\eta(0) - D_2(\eta(-\tau))|^2 + \lambda' \int_0^t \mathrm{e}^{\lambda' s}\mathbb{E}|Y^\zeta(s,\eta) - D_2(Y^\zeta(s - \tau, \eta))|^2 \mathrm{d}s \\
& \quad + \int_0^t \mathrm{e}^{\lambda' s}\mathbb{E}\Big[2\langle Y^\zeta(s,\eta) - D_2(Y^\zeta(s - \tau, \eta)), b_2(\zeta, Y^\zeta(s,\eta), Y^\zeta(s - \tau, \eta))\rangle \\
& \quad + \|\sigma_2(\zeta, Y^\zeta(s,\eta), Y^\zeta(s - \tau, \eta))\|^2\Big]\mathrm{d}s \\
& \leq \mathbb{E}|\eta(0) - D_2(\eta)|^2 + \lambda' \int_0^t \mathrm{e}^{\lambda' s}\mathbb{E}|Y^\zeta(s,\eta) - D_2(Y^\zeta(s - \tau, \eta))|^2 \mathrm{d}s \\
& \quad + \int_0^t \mathrm{e}^{\lambda' s}\mathbb{E}\Big[-\lambda'_1|Y^\zeta(s,\eta)|^2 + \lambda'_2|Y^\zeta(s - \tau, \eta)|^2 + c(1 + \|\zeta\|_\infty^2)\Big]\mathrm{d}s.
\end{aligned}
$$

By the elementary inequality:

$$(3.6) \qquad (a + b)^2 \leq (1 + \epsilon)a^2 + \frac{1 + \epsilon}{\epsilon}b^2, \quad a, b \in \mathbb{R}, \quad \epsilon > 0,$$

it follows from (**A4**) that

$$
\begin{aligned}
& \lambda' \int_0^t \mathrm{e}^{\lambda' s}\mathbb{E}|Y^\zeta(s,\eta) - D_2(Y^\zeta(s - \tau, \eta))|^2 \mathrm{d}s \\
& \leq \lambda'(1 + \kappa_2)\int_0^t \mathrm{e}^{\lambda' s}\mathbb{E}|Y^\zeta(s,\eta)|^2 \mathrm{d}s + \lambda'(1 + \kappa_2)\mathrm{e}^{\lambda' \tau}\int_{-\tau}^t \mathrm{e}^{\lambda' s}\mathbb{E}|Y^\zeta(s,\eta)|^2 \mathrm{d}s \\
& \leq \lambda'(1 + \kappa_2)\frac{\mathrm{e}^{\lambda' \tau}}{\lambda'}\|\eta\|_\infty^2 + \lambda'(1 + \kappa_2)(1 + \mathrm{e}^{\lambda' \tau})\int_0^t \mathrm{e}^{\lambda' s}\mathbb{E}|Y^\zeta(s,\eta)|^2 \mathrm{d}s.
\end{aligned}
$$

5

Substituting this into (3.5), along with (**A4**), it gives that

(3.7)
$$
\begin{aligned}
&\mathrm{e}^{\lambda' t}\mathbb{E}|Y^\zeta(t,\eta) - D_2(Y^\zeta(t-\tau,\eta))|^2 \\
&\qquad \lesssim \|\eta\|_\infty^2 + \mathrm{e}^{\lambda' t}(1 + \|\zeta\|_\infty^2) \\
&\qquad - \left(\lambda_1' - \lambda'(1+\kappa_2) - (\lambda_2' + \lambda'(1+\kappa_2)\mathrm{e}^{\lambda'\tau})\right)\int_0^t \mathrm{e}^{\lambda' s}\mathbb{E}|Y^\zeta(s,\eta)|^2 \mathrm{d}s.
\end{aligned}
$$

On the other hand, using the inequality (3.6) again and taking (**A4**) into consideration, for any $\epsilon > 0$, we have

$$
\mathrm{e}^{\lambda' t}\mathbb{E}|Y^\zeta(t,\eta)|^2 \leq (1+\epsilon)\mathrm{e}^{\lambda' t}\mathbb{E}|Y^\zeta(t,\eta) - D_2(Y^\zeta(t-\tau,\eta))|^2 + \frac{(1+\epsilon)\kappa_2}{\epsilon}\mathrm{e}^{\lambda' t}\mathbb{E}|Y^\zeta(t-\tau,\eta)|^2.
$$

Therefore,

$$
\begin{aligned}
\sup_{t \geq -\tau}(\mathrm{e}^{\lambda' t}\mathbb{E}|Y^\zeta(t,\eta)|^2) &\leq \sup_{0 \geq t \geq -\tau}(\mathrm{e}^{\lambda' t}\mathbb{E}|Y^\zeta(t,\eta)|^2) + \sup_{t \geq 0}(\mathrm{e}^{\lambda' t}\mathbb{E}|Y^\zeta(t,\eta)|^2) \\
&\leq \|\eta\|_\infty^2 + (1+\epsilon)\sup_{t\geq 0}(\mathrm{e}^{\lambda' t}\mathbb{E}|Y^\zeta(t,\eta) - D_2(Y^\zeta(t-\tau,\eta))|^2) \\
&\quad + \frac{(1+\epsilon)\kappa_2}{\epsilon}\mathrm{e}^{\lambda'\tau}\sup_{t\geq -\tau}(\mathrm{e}^{\lambda' t}\mathbb{E}|Y^\zeta(t,\eta)|^2).
\end{aligned}
$$

Due to $\lambda_1' > \lambda_2' > 0$, we can find some $\lambda' \in (0,1)$ sufficiently small and $\epsilon > 0$ sufficiently large such that $\frac{(1+\epsilon)\kappa_2}{\epsilon}\mathrm{e}^{\lambda'\tau} < 1$ and $\lambda_1' - \lambda'(1+\kappa_2) - (\lambda_2' + \lambda'(1+\kappa_2))\mathrm{e}^{\lambda'\tau} > 0$. Thus, if follows from (3.7) that for $t \geq -\tau$,

$$
\sup_{t\geq 0}(\mathrm{e}^{\lambda' t}\mathbb{E}|Y^\zeta(t,\eta) - D_2(Y^\zeta(t-\tau,\eta))|^2) \lesssim \|\eta\|_\infty^2 + \mathrm{e}^{\lambda' t}(1 + \|\zeta\|_\infty^2).
$$

These imply for $t \geq -\tau$

$$
\left(1 - \frac{(1+\epsilon)\kappa_2}{\epsilon}\mathrm{e}^{\lambda'\tau}\right)\sup_{t\geq -\tau}(\mathrm{e}^{\lambda' t}\mathbb{E}|Y^\zeta(t,\eta)|^2) \lesssim \|\eta\|_\infty^2 + \mathrm{e}^{\lambda' t}(1 + \|\zeta\|_\infty^2).
$$

In consequence, one arrives at

(3.8)
$$
\mathbb{E}|Y^\zeta(t,\eta)|^2 \lesssim \mathrm{e}^{-\lambda' t}\|\eta\|_\infty^2 + 1 + \|\zeta\|_\infty^2, \quad t > 0.
$$

*Step* 2. We now give an estimate for the segment process, i.e. we shall show there exists a positive number $\lambda'$ such that

(3.9)
$$
\mathbb{E}\|Y_t^\zeta(\eta)\|_\infty^2 \lesssim \mathrm{e}^{-\lambda' t}\|\eta\|_\infty^2 + 1 + \|\zeta\|_\infty^2.
$$

According to the Itô formula and (3.4), for $t \geq \tau, -\tau \leq \theta \leq 0$, we have

$$
\begin{aligned}
&|Y^\zeta(t+\theta,\eta) - D_2(Y^\zeta(t-\tau+\theta,\eta))|^2 \\
&= |Y^\zeta(t-\tau,\eta) - D_2(Y^\zeta(t-2\tau,\eta))|^2 \\
&\quad + \int_{t-\tau}^{t+\theta} \Big[ 2\langle Y^\zeta(s,\eta) - D_2(Y^\zeta(s-\tau,\eta)), b_2(\zeta, Y^\zeta(s,\eta), Y^\zeta(s-\tau,\eta)) \rangle \\
&\quad + \|\sigma_2(\zeta, Y^\zeta(s,\eta), Y^\zeta(s-\tau,\eta))\|^2 \Big] \mathrm{d}s + M(t,\theta) \\
&\leq |Y^\zeta(t+\theta,\eta) - D_2(Y^\zeta(t-\tau+\theta,\eta))|^2 + M(t,\theta) \\
&\quad + \int_{t-\tau}^{t+\theta} \Big[ -\lambda_1'|Y^\zeta(s,\eta)|^2 + \lambda_2'|Y^\zeta(s-\tau,\eta)|^2 + c(1+\|\zeta\|_\infty^2) \Big] \mathrm{d}s,
\end{aligned}
$$

where

$$
M(t,\theta) = \int_{t-\tau}^{t+\theta} 2\langle Y^\zeta(s,\eta) - D_2(Y^\zeta(s-\tau,\eta)), \sigma_2(\zeta, Y^\zeta(s,\eta), Y^\zeta(s-\tau,\eta)) \mathrm{d}W_2(s) \rangle.
$$

By the Burkhold-Davis-Gundy (B-D-G for abbreviation) inequality, we derive from (**A2**) that there exists some positive constant $c$ such that

$$
\begin{aligned}
&\mathbb{E}\Big( \sup_{-\tau \leq \theta \leq 0} M(t,\theta) \Big) \\
&\leq c\mathbb{E}\Big( \sup_{-\tau \leq \theta \leq 0} |Y^\zeta(s+\theta,\eta) - D_2(Y^\zeta(s-\tau+\theta,\eta))|^2 \int_{t-\tau}^t \|\sigma_2(\zeta, Y^\zeta(s,\eta), Y^\zeta(s-\tau,\eta))\|^2 \mathrm{d}s \Big)^{\frac{1}{2}} \\
&\leq \frac{1}{2}\mathbb{E}\Big( \sup_{-\tau \leq \theta \leq 0} |Y^\zeta(s+\theta,\eta) - D_2(Y^\zeta(s-\tau+\theta,\eta))|^2 \Big) \\
&\quad + c\int_{t-\tau}^t \mathbb{E}\|\sigma_2(\zeta, Y^\zeta(s,\eta), Y^\zeta(s-\tau,\eta))\|^2 \mathrm{d}s \\
&\leq \frac{1}{2}\mathbb{E}\Big( \sup_{-\tau \leq \theta \leq 0} |Y^\zeta(s+\theta,\eta) - D_2(Y^\zeta(s-\tau+\theta,\eta))|^2 \Big) + c\|\zeta\|_\infty^2 + c\int_{t-2\tau}^t \mathbb{E}|Y^\zeta(s,\eta)|^2 \mathrm{d}s.
\end{aligned}
$$

Substituting this into (3.10), combining with (3.1) and (**A4**), it gives that for $t \geq \tau$

$$
\begin{aligned}
&\mathbb{E}\Big( \sup_{-\tau \leq \theta \leq 0} |Y^\zeta(t+\theta,\eta) - D_2(Y^\zeta(t-\tau+\theta,\eta))|^2 \Big) \\
&\leq 2\mathbb{E}|Y^\zeta(t-\tau,\eta) - D_2(Y^\zeta(t-2\tau,\eta))|^2 + c(1+\|\zeta\|_\infty^2) + c\int_{t-2\tau}^t |Y^\zeta(s,\eta)|^2 \mathrm{d}s \\
&\leq 4\mathbb{E}|Y^\zeta(t-\tau,\eta)|^2 + 4\kappa_2 \mathbb{E}|Y^\zeta(t-2\tau,\eta)|^2 + c(1+\|\zeta\|_\infty^2) + c\int_{t-2\tau}^t \mathbb{E}|Y^\zeta(s,\eta)|^2 \mathrm{d}s \\
&\lesssim \mathrm{e}^{-\lambda' t}\|\eta\|_\infty^2 + 1 + \|\zeta\|_\infty^2.
\end{aligned}
$$

7

On the other hand, following an argument to achieve (3.11), one has for $t \in [0, \tau]$

$$
\begin{aligned}
\mathbb{E}\Big( \sup_{-\tau \leq \theta \leq 0} |Y^\zeta(t + \theta, \eta) - D_2(Y^\zeta(t - \tau + \theta, \eta))|^2 \Big) & \\
\end{aligned}
$$

(3.12)
$$
\leq c(1 + \|\zeta\|_\infty^2 + \|\eta\|_\infty^2) + c \int_0^t \mathbb{E}|Y^\zeta(s, \eta)|^2 \mathrm{d}s
$$

$$
\lesssim 1 + \|\zeta\|_\infty^2 + \mathrm{e}^{-\lambda' t} \|\eta\|_\infty^2.
$$

By using the inequality (3.6) and (**A4**), for any $\epsilon > 0$, we have

$$
\mathbb{E}\|Y_t^\zeta(\eta)\|_\infty^2 = \mathbb{E}\Big( \sup_{t - \tau \leq s \leq t} |Y^\zeta(s, \eta)|^2 \Big)
$$

(3.13)
$$
\leq \frac{1}{1 - \sqrt{\kappa_2}} \mathbb{E}\Big( \sup_{t - \tau \leq s \leq t} |Y^\zeta(s, \eta) - D_2(Y^\zeta(s - \tau, \eta))|^2 \Big)
$$

$$
+ \sqrt{\kappa_2} \, \mathbb{E}\Big( \sup_{t - \tau \leq s \leq t} |Y^\zeta(s - \tau, \eta)|^2 \Big).
$$

Hence, combining (3.11) with (3.12) and (3.13), we derive the desired assertion.

*Step* 3. For any $t_2 > t_1 > \tau$ and the frozen segment $\zeta \in \mathscr{C}$, consider the following

(3.14)
$$
\begin{aligned}
\mathrm{d}[\overline{Y}(t) - D_2(\overline{Y}(t - \tau))] &= b_2(\zeta, \overline{Y}(t), \overline{Y}(t - \tau))\mathrm{d}t \\
&\quad + \sigma_2(\zeta, \overline{Y}(t), \overline{Y}(t - \tau))\mathrm{d}W_2(t), \ \ t \in [t_2 - t_1, t_2]
\end{aligned}
$$

with the initial value $\overline{Y}_{t_2 - t_1} = \eta$. The solution process and the segment process associated with (3.14) are denoted by $(\overline{Y}^\zeta(t, \eta))$ and $(\overline{Y}_t^\zeta(\eta))$, respectively.

Let $\Gamma^\zeta(t, \eta) := Y^\zeta(t, \eta) - \overline{Y}^\zeta(t, \eta), t \in [t_2 - t_1, t_2]$, we claim that

(3.15)
$$
\|\Gamma_{t_2}^\zeta(\eta)\| \lesssim \mathrm{e}^{-\lambda t_1}(1 + \|\eta\|_\infty^2 + \|\zeta\|_\infty^2), \quad t \geq 0.
$$

Observe that the laws of $Y_{t_2}^\zeta(\eta)$ and $\overline{Y}_{t_2}^\zeta(\eta)$ are $P_{t_2}^{\zeta, \eta}$ and $P_{t_1}^{\zeta, \eta}$, respectively. Again, owing to $\lambda_1 > \lambda_2 > 0$, we find some $\widetilde{\lambda} \in (0, 1)$ sufficiently small such that $\lambda_1 - \widetilde{\lambda}(1 + \kappa_2) - (\lambda_2 + \widetilde{\lambda}(1 + \kappa_2))\mathrm{e}^{\widetilde{\lambda}\tau} > 0$. Taking $\lambda = \lambda' \wedge \widetilde{\lambda}$, by the Itô formula and the inequality (3.6), it follows

8

from (**A3**) and (**A4**) that

$$
\begin{aligned}
\mathrm{e}^{\lambda t}\mathbb{E}|\Gamma^{\zeta}(t,\eta) &- (D_2(Y^{\zeta}(t-\tau,\eta)) - D_2(\overline{Y}^{\zeta}(t-\tau,\eta)))|^2 \\
&\leq \mathrm{e}^{\lambda(t_2-t_1)}\mathbb{E}|\Gamma^{\zeta}(t_2-t_1,\eta) - (D_2(Y^{\zeta}(t_2-t_1-\tau,\eta)) - D_2(\overline{Y}^{\zeta}(t_2-t_1-\tau,\eta)))|^2 \\
&\quad + \int_{t_2-t_1}^{t} \lambda\mathrm{e}^{\lambda s}\mathbb{E}|\Gamma^{\zeta}(s,\eta) - (D_2(Y^{\zeta}(s-\tau,\eta)) - D_2(\overline{Y}^{\zeta}(s-\tau,\eta)))|^2\mathrm{d}s \\
&\quad + \int_{t_2-t_1}^{t} \mathrm{e}^{\lambda s}\mathbb{E}\{-\lambda_1|\Gamma^{\zeta}(s,\eta)|^2 + \lambda_2|\Gamma^{\zeta}(s-\tau,\eta)|^2\}\mathrm{d}s \\
&\leq \mathrm{e}^{\lambda(t_2-t_1)}(1+\kappa_2)(\mathbb{E}|\Gamma^{\zeta}(t_2-t_1,\eta)|^2 + \mathbb{E}|\Gamma^{\zeta}(t_2-t_1-\tau,\eta)|^2) \\
&\quad - \left(\lambda_1 - \lambda(1+\kappa_2) - (\lambda_2 + \lambda(1+\kappa_2))\mathrm{e}^{\lambda\tau}\right)\int_{t_2-t_1}^{t}\mathrm{e}^{\lambda s}\mathbb{E}|\Gamma^{\zeta}(s,\eta)|^2\mathrm{d}s \\
&\quad + \mathrm{e}^{\lambda\tau}[\lambda(1+\kappa_2)+\lambda_2]\int_{t_2-t_1-\tau}^{t_2-t_1}\mathrm{e}^{\lambda s}\mathbb{E}|\Gamma^{\zeta}(s,\eta)|^2\mathrm{d}s \\
&\lesssim \mathrm{e}^{\lambda(t_2-t_1)}\|\eta\|_{\infty}^2 + \mathrm{e}^{\lambda(t_2-t_1)}\mathbb{E}\|Y_{t_2-t_1}^{\zeta}(\eta)\|_{\infty}^2.
\end{aligned}
$$

By carrying out a similar argument to obtain (3.11), we have

$$
\begin{aligned}
(3.16) \qquad \mathbb{E}\left(\sup_{-\tau\theta\leq 0} \|\Gamma^{\zeta}(t+\theta,\eta) - (D_2(Y^{\zeta}(t-\tau+\theta,\eta)) - D_2(\overline{Y}^{\zeta}(t-\tau+\theta,\eta)))\|^2\right) \\
\lesssim \mathrm{e}^{-\lambda(t+t_1-t_2)}(1+\|\eta\|_{\infty}^2 + \|\zeta\|_{\infty}^2), \quad t\geq 0.
\end{aligned}
$$

In the same way as (3.12) and (3.13), we arrive at

$$
\mathbb{E}\|\Gamma_{t_2}^{\zeta}(\eta)\|_{\infty}^2 \lesssim \mathrm{e}^{-\lambda t_1}(1+\|\eta\|_{\infty}^2 + \|\zeta\|_{\infty}^2), \quad t\geq 0,
$$

1   as required.

   *Step* 4. Let $\eta, \eta' \in \mathscr{C}$, we prove that

$$
(3.17) \qquad \mathbb{E}\|Y_t^{\zeta}(\eta) - Y_t^{\zeta}(\eta')\|_{\infty}^2 \lesssim \mathrm{e}^{-\lambda t}\|\eta - \eta'\|_{\infty}^2.
$$

   Consider the difference of the solution process of (2.3) starting from differential initial value. It follows that

$$
\begin{aligned}
Y^{\zeta}(t,\eta) - Y^{\zeta}(t,\eta') &- [D_2(Y^{\zeta}(t-\tau,\eta)) - D_2(Y^{\zeta}(t-\tau,\eta'))] \\
&= \eta(0) - \eta'(0) - [D_2(\eta(-\tau)) - D_2(\eta'(-\tau))] \\
&\quad + \int_0^t \left(b_2(\zeta, Y^{\zeta}(s,\eta), Y^{\zeta}(s-\tau,\eta)) - b_2(\zeta, Y^{\zeta}(s,\eta'), Y^{\zeta}(s-\tau,\eta'))\right)\mathrm{d}s \\
&\quad + \int_0^t \left(\sigma_2(\zeta, Y^{\zeta}(s,\eta), Y^{\zeta}(s-\tau,\eta)) - \sigma_2(\zeta, Y^{\zeta}(s,\eta'), Y^{\zeta}(s-\tau,\eta'))\right)\mathrm{d}W_2(t).
\end{aligned}
$$

9

By the Itô formula and the fundamental inequality (3.6), it follows from (**A3**) and (**A4**) that

$$
\begin{aligned}
\mathrm{e}^{\lambda t}\mathbb{E}\big|Y^\zeta(t,\eta) &- Y^\zeta(t,\eta') - [D_2(Y^\zeta(t-\tau,\eta)) - D_2(Y^\zeta(t-\tau,\eta'))]\big|^2 \\
&\le \mathbb{E}\big|\eta(0) - \eta'(0) - [D_2(\eta(-\tau)) - D_2(\eta'(-\tau))]\big|^2 \\
&\quad + \lambda\int_0^t \mathrm{e}^{\lambda s}\mathbb{E}\big|Y^\zeta(s,\eta) - Y^\zeta(s,\eta') - [D_2(Y^\zeta(s-\tau,\eta)) - D_2(Y^\zeta(s-\tau,\eta'))]\big|^2 \mathrm{d}s \\
&\quad + \int_0^t \mathrm{e}^{\lambda s}\mathbb{E}\Big[ -\lambda_1|Y^\zeta(s,\eta) - Y^\zeta(s,\eta')|^2 + \lambda_2[Y^\zeta(s-\tau,\eta) - Y^\zeta(s-\tau,\eta')|^2\Big]\mathrm{d}s \\
&\le \mathbb{E}\big|\eta(0) - \eta'(0) - [D_2(\eta(-\tau)) - D_2(\eta'(-\tau))]\big|^2 \\
&\quad - \Big(\lambda_1 - \lambda(1+\kappa_2) - (\lambda_2 + \lambda(1+\kappa_2))\mathrm{e}^{\lambda\tau}\Big)\int_0^t \mathrm{e}^{\lambda s}\mathbb{E}|Y^\zeta(s,\eta) - Y^\zeta(s,\eta')|^2 \mathrm{d}s \\
&\quad + \mathrm{e}^{\lambda\tau}[\lambda(1+\kappa_2) + \lambda_2]\int_{-\tau}^0 \mathrm{e}^{\lambda s}\mathbb{E}|Y^\zeta(s,\eta) - Y^\zeta(s,\eta')|^2 \mathrm{d}s \\
&\lesssim \|\eta - \eta'\|_\infty^2.
\end{aligned}
$$

Following the steps of (3.1), we obtain

$$
\mathbb{E}|Y^\zeta(t,\eta) - Y^\zeta(t,\eta')|^2 \lesssim \mathrm{e}^{-\lambda t}\|\eta - \eta'\|_\infty^2, \quad t \ge 0.
$$

Also, by the Itô formula and the B-D-G inequality, one gives

(3.18)
$$
\begin{aligned}
\mathbb{E}\big(\sup_{-\tau\le\theta\le 0} &|Y^\zeta(t+\theta,\eta) - Y^\zeta(t+\theta,\eta') - (D_2(Y^\zeta(t-\tau+\theta,\eta)) - D_2(Y^\zeta(t-\tau+\theta,\eta')))|^2\big) \\
&\lesssim \mathrm{e}^{-\lambda t}\|\eta - \eta'\|_\infty^2, \quad t \ge 0.
\end{aligned}
$$

In the same way as (3.12) and (3.13), we derive for any $t \ge 0$,

$$
\mathbb{E}\|Y_t^\zeta(\eta) - Y_t^\zeta(\eta')\|_\infty^2 \lesssim \mathrm{e}^{-\lambda t}\|\eta - \eta'\|_\infty^2.
$$

*Sept* 5. We shall show the existence and uniqueness of invariant measure of $Y_t^\zeta$.

Let $\mathscr{P}(\mathscr{C})$ be the set of all probability measures on $\mathscr{C}$. $d_2$ denotes the $L^2$-Wasserstein distance on $\mathscr{P}(\mathscr{C})$ induced by the bounded distance $\rho(\xi,\eta) := 1 \wedge \|\xi - \eta\|_\infty$, i.e.,

$$
d_2(\mu_1,\mu_2) = \inf_{\pi\in\mathscr{C}(\mu_1,\mu_2)} \big(\pi(\rho^2)\big)^{\frac{1}{2}}, \quad \mu_1,\mu_2 \in \mathscr{P}(\mathscr{C}),
$$

where $\mathscr{C}(\mu_1,\mu_2)$ is the set of all coupling probability measures with marginals $\mu_1$ and $\mu_2$. It is well known that $\mathscr{P}(\mathscr{C})$ is a complete metric space w.r.t. the distance $d_2$ (see, e.g., [6, Lemma 5.3, P.174] and [6, Theorem 5.4, P.175]), and the convergence in $d_2$ is equivalent to the weak convergence (see, e.g., [6, Theorem 5.6, P.179]). Let $P_t^{\zeta,\eta}$ be the law of the segment process $Y_t^\zeta(\eta)$. According to the Krylov-Bogoliubov existence theorem (see, e.g., [9, Theorem 3.1.1, P.21]), if $P_t^{\zeta,\eta}$ converges weakly to a probability measure $\mu_\eta^\zeta$, then $\mu_\eta^\zeta$ is an invariant measure.

By (3.15), we have

$$
\begin{aligned}
d_2(P_{t_1}^{\zeta,\eta}, P_{t_2}^{\zeta,\eta}) &\leq \mathbb{E}\{1 \wedge \|Y_{t_1}^{\zeta}(\eta) - \overline{Y}_{t_2}^{\zeta}(\eta)\|_\infty^2\}^{1/2} \\
&\lesssim \mathbb{E}\{1 + \|Y_{t_1}^{\zeta}(\eta) - \overline{Y}_{t_2}^{\zeta}(\eta)\|_\infty^2\} \\
&\lesssim e^{-(p_1 \wedge 1)\lambda t_1}(1 + \|\eta\|_\infty^2 + \|\zeta\|_\infty^2),
\end{aligned}
$$

which goes to zero as $t_1$ (hence $t_2$) tends to $\infty$. Therefore, $\{P_t^{\zeta,\eta}\}_{t \geq 0}$ is a Cauchy sequence w.r.t. the distance $d_2$. By the completeness of $\mathscr{P}(\mathscr{C})$ w.r.t. the distance $d_2$, there is $\mu_\eta^\zeta \in \mathscr{P}(\mathscr{C})$ such that

$$
(3.19) \qquad \lim_{t \to \infty} d_2(P_t^{\zeta,\eta}, \mu_\eta^\zeta) = 0.
$$

Moreover, for fixed $\zeta \in \mathscr{C}$ and arbitrary $\eta, \eta \in \mathscr{C}$, observing that

$$
(3.20) \qquad d_2(\mu_\eta^\zeta, \mu_{\eta'}^\zeta) \leq d_2(P_t^{\zeta,\eta}, \mu_\eta^\zeta) + d_2(P_t^{\zeta,\eta'}, \mu_{\eta'}^\zeta) + d_2(P_t^{\zeta,\eta}, P_t^{\zeta,\eta'}),
$$

and using (3.17), we obtain for any $\eta, \eta' \in \mathscr{C}$ and frozen $\zeta \in \mathscr{C}$

$$
(3.21) \qquad d_2(\mu_\eta^\zeta, \mu_{\eta'}^\zeta) = 0.
$$

The existence and uniqueness of invariant measure of $Y_t^\zeta$ follows by (3.19) and (3.21).

*Step* 6. We are now going to prove (2.4).

By virtue of (3.9) and the invariance of $\mu^\zeta$, it then follows that

$$
\int_{\mathscr{C}} \|\psi\|_\infty^2 \mu^\zeta(\mathrm{d}\psi) \leq c\Big\{1 + \|\zeta\|_\infty^2 + e^{-\lambda t} \int_{\mathscr{C}} \|\psi\|_\infty^2 \mu^\zeta(\mathrm{d}\psi)\Big\}.
$$

Thus, choosing $t > 0$ sufficiently large such that $\overline{\delta} := ce^{-\lambda t} < 1$, one finds that

$$
(3.22) \qquad \int_{\mathscr{C}} \|\psi\|_\infty^2 \mu^\zeta(\mathrm{d}\psi) \lesssim 1 + \|\zeta\|_\infty^2.
$$

Next, with the aid of the invariance of $\mu^\zeta$, (3.17), and (3.22), we deduce from (**A1**) that

$$
\begin{aligned}
|\mathbb{E}b_1(\zeta, Y_t^\zeta(\eta)) - \overline{b}_1(\zeta)| &\lesssim \int_{\mathscr{C}} \mathbb{E}\|Y_t^\zeta(\eta) - Y_t^\zeta(\psi)\|_\infty \mu^\zeta(\mathrm{d}\psi) \lesssim e^{-\frac{(p_1 \wedge 1)\lambda t}{2}} \int_{\mathscr{C}} \|\eta - \psi\|_\infty \mu^\zeta(\mathrm{d}\psi) \\
&\lesssim e^{-\frac{(p_1 \wedge 1)\lambda t}{2}}(1 + \|\eta\|_\infty + \|\zeta\|_\infty).
\end{aligned}
$$

As a result, (2.4) follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

We now complete

*Proof of Proposition 2.2.* For arbitrary $\phi, \zeta \in \mathscr{C}$, let

$$
\nabla_\phi \overline{b}_1(\zeta) = \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \overline{b}_1(\zeta + \varepsilon\phi)\Big|_{\varepsilon=0}
$$

11

be the direction derivative of $\bar{b}_1$ at $\zeta$ along the direction $\phi$. By Theorem 2.1, we have

$$
\begin{aligned}
\nabla_\phi \bar{b}_1(\zeta) &= \lim_{\varepsilon \to 0} \frac{\bar{b}_1(\zeta + \varepsilon\phi) - \bar{b}_1(\zeta)}{\varepsilon} \\
&= \lim_{t \to \infty} \lim_{\varepsilon \to 0} \frac{\mathbb{E}\{b_1(\zeta + \varepsilon\phi, Y_t^\zeta) - b_1(\zeta, Y_t^\zeta)\}}{\varepsilon} \\
&= \lim_{t \to \infty} \mathbb{E}\nabla_\phi b_1(\zeta, Y_t^\zeta(\eta)) \\
&= \lim_{t \to \infty} \mathbb{E}\Big\{ (\nabla_\phi^{(1)} b_1)(\zeta, Y_t^\zeta(\eta)) + \Big(\nabla_{\nabla_\phi Y_t^\zeta(\eta)}^{(2)} b_1\Big)(\zeta, Y_t^\zeta(\eta)) \Big\}, \quad \phi, \zeta, \eta \in \mathscr{C}.
\end{aligned}
$$

According to (**A1**), to verify that $\bar{b}_1 : \mathscr{C} \mapsto \mathbb{R}^n$ is Lipschitz, it remains to verify

(3.23)
$$
\sup_{t \geq 0} \mathbb{E}\|\nabla_\phi Y_t^\zeta(\eta)\|_\infty^2 < \infty.
$$

Observe that $\nabla_\phi Y^\zeta(t, \eta)$ satisfies the following NSDDE

$$
\begin{aligned}
\mathrm{d}(\nabla_\phi(Y^\zeta(t, \eta) &- D_2(Y^\zeta(t - \tau, \eta)))) \\
&= \Big\{ (\nabla_\phi^{(1)} b_2)(\zeta, Y^\zeta(t, \eta), Y^\zeta(t - \tau, \eta)) \\
&\quad + \Big(\nabla_{\nabla_\phi Y^\zeta(t, \eta)}^{(2)} b_2\Big)(\zeta, Y^\zeta(t, \eta), Y^\zeta(t - \tau, \eta)) \\
&\quad + \Big(\nabla_{\nabla_\phi Y^\zeta(t - \tau, \eta)}^{(3)} b_2\Big)(\zeta, Y^\zeta(t, \eta), Y^\zeta(t - \tau, \eta)) \Big\}\mathrm{d}t \\
&\quad + \Big\{ (\nabla_\phi^{(1)} \sigma_2)(\zeta, Y^\zeta(t, \eta), Y^\zeta(t - \tau, \eta)) \\
&\quad + \Big(\nabla_{\nabla_\phi Y^\zeta(t, \eta)}^{(2)} \sigma_2\Big)(\zeta, Y^\zeta(t, \eta), Y^\zeta(t - \tau, \eta)) \\
&\quad + \Big(\nabla_{\nabla_\phi Y^\zeta(t - \tau, \eta)}^{(3)} \sigma_2\Big)(\zeta, Y^\zeta(t, \eta), Y^\zeta(t - \tau, \eta)) \Big\}\mathrm{d}W_2(t), \quad t > 0
\end{aligned}
$$

with the initial datum $\nabla_\phi Y_0^\zeta(\eta) = 0$. In the sequel, let $\chi \in \mathscr{C}$ and $x, x', y, y' \in \mathbb{R}^n$. For any $\varepsilon > 0$, it is trivial to see from (**A3**) that

$$
2\varepsilon\langle x - D_2(y), b_2(\chi, x' + \varepsilon x, y' + \varepsilon y) - b_2(\chi, x', y')\rangle + \|\sigma_2(\chi, x' + \varepsilon x, y' + \varepsilon y) - \sigma_2(\chi, x', y')\|^2
$$
$$
\leq -\lambda_1 \varepsilon^2 |x|^2 + \lambda_2 \varepsilon^2 |y|^2.
$$

Multiplying $\varepsilon^{-2}$ on both sides, followed by taking $\varepsilon \downarrow 0$, gives that

(3.24)
$$
\begin{aligned}
2\langle x - D_2(y), (\nabla_x^{(2)} b_2)(\chi, x', y') &+ (\nabla_y^{(3)} b_2)(\chi, x', y')\rangle \\
+ \|(\nabla_x^{(2)} \sigma_2)(\chi, x', y') &+ (\nabla_y^{(3)} \sigma_2)(\chi, x', y')\|^2 \\
&\leq -\lambda_1 |x|^2 + \lambda_2 |y|^2.
\end{aligned}
$$

On the other hand, by virtue of (3.2), for any $\varepsilon > 0$, one has

$$
\|\sigma_2(\chi, x' + \varepsilon x, y' + \varepsilon y) - \sigma_2(\chi, x', y')\|^2 \leq \alpha\varepsilon^2(|x|^2 + |y|^2),
$$

12

which further yields that

$$(3.25) \qquad \|(\nabla_x^{(2)}\sigma_2)(\chi, x', y') + (\nabla_y^{(3)}\sigma_2)(\chi, x', y')\|^2 \leq \alpha(|x|^2 + |y|^2).$$

Also, for any $\varepsilon > 0$, we have from **(A4)** that

$$|D_2(y' + \varepsilon y) - D_2(y')|^2 \leq \kappa_2 \varepsilon^2 |y|^2,$$

which further yields that

$$(3.26) \qquad |(\nabla_y D_2)(y')|^2 \leq \kappa_2 |y|^2.$$

Thus, with (3.24), (3.25) and (3.26) in hand, (3.23) holds by repeating an argument to derive (3.9). $\qquad\square$

# 4 Proof of Theorem 2.3

In order to prove our main result, we need to construct some auxiliary two time-scales stochastic systems with memory and provide a number of lemmas.

Let $T > 0$ be fixed and set $\delta := \frac{\tau}{N} \in (0,1)$ for a sufficiently large positive integer $N$. For any $t \in [0, T]$, consider the following auxiliary two-time-scale systems of NSFDEs

$$(4.1) \qquad d[\widetilde{X}^\varepsilon(t) - D_1(X_{t_\delta}^\varepsilon)] = b_1(X_{t_\delta}^\varepsilon, \widetilde{Y}_t^\varepsilon)dt + \sigma_1(X_{t_\delta}^\varepsilon)dW_1(t), \quad X_0^\varepsilon = \xi \in \mathscr{C},$$

and

$$(4.2)$$
$$\begin{cases} d[\widetilde{Y}^\varepsilon(t) - D_2(\widetilde{Y}^\varepsilon(t - \tau)))] = \frac{1}{\varepsilon}b_2(X_{t_\delta}^\varepsilon, \widetilde{Y}^\varepsilon(t), \widetilde{Y}^\varepsilon(t - \tau))dt + \frac{1}{\sqrt{\varepsilon}}\sigma_2(X_{t_\delta}^\varepsilon, \widetilde{Y}^\varepsilon(t), \widetilde{Y}^\varepsilon(t - \tau))dW_2(t), \\ \widetilde{Y}^\varepsilon(t_\delta) = Y^\varepsilon(t_\delta) \end{cases}$$

with the initial value $\widetilde{Y}_0^\varepsilon = Y_0^\varepsilon = \eta \in \mathscr{C}$, where $t_\delta := \lfloor t/\delta \rfloor \delta$, the nearest breakpoint preceding $t$, with $\lfloor t/\delta \rfloor$ being the integer part of $t/\delta$.

To proceed, we present several preliminary lemmas. The first lemma concerns the continuity in the mean $L^p$-norm sense for the displacement of the segment process $(X_t^\varepsilon)_{t\in[0,T]}$.

**Lemma 4.1.** Under **(A1)** and **(A4)**,

$$\sup_{t\in[0,T]} \mathbb{E}\|X_t^\varepsilon - X_{t_\delta}^\varepsilon\|_\infty^p \lesssim_T \delta^{\frac{p-2}{2}}, \quad p > 2.$$

*Proof.* Using **(A1)** and noting [24, Theorem 4.5, P.213], we have

$$(4.3) \qquad \mathbb{E}\Big(\sup_{0\leq t\leq T} \|X_t^\varepsilon\|_\infty^p\Big) \lesssim_T 1 + \|\xi\|_\infty^p.$$

13

Observe that

$$\mathbb{E}\Big( \sup_{-\tau \le \theta \le 0} |X^\varepsilon(t+\theta) - D_1(X^\varepsilon_{t+\theta}) - X^\varepsilon(t_\delta + \theta) + D_1(X^\varepsilon_{t_\delta+\theta})|^p \Big)$$

$$\le \sum_{m=0}^{N-1} \mathbb{E}\Big( \sup_{-(m+1)\delta \le \theta \le -m\delta} |X^\varepsilon(t+\theta) - D_1(X^\varepsilon_{t+\theta}) - X^\varepsilon(t_\delta + \theta) + D_1(X^\varepsilon_{t_\delta+\theta})|^p \Big)$$

$$=: \sum_{m=0}^{N-1} J_p(t, m, \delta),$$

where $N = \tau/\delta$ by the definition of $\delta$. For any $t \in [0, T]$, take $k \ge 0$ such that $t \in [k\delta, (k+1)\delta)$. Thus, for any $\theta \in [-(m+1)\delta, -m\delta]$, one has

$$t + \theta \in [(k-m-1)\delta, (k+1-m)\delta] \quad \text{and} \quad t_\delta + \theta \in [(k-m-1)\delta, (k-m)\delta].$$

In what follows, we separate the following three cases to show

(4.4) $$J_p(t, m, \delta) \lesssim_T \delta^{\frac{p}{2}}.$$

**Case 1:** $m \le k - 1$. Invoking Hölder's inequality and B-D-G's inequality, we obtain from (**A1**) and (4.3) that

(4.5)

$$J_p(t, m, \delta) \lesssim \delta^{p-1} \int_{(k-m-1)\delta}^{t-m\delta} \mathbb{E}|b_1(X^\varepsilon_s, Y^\varepsilon_s)|^p \mathrm{d}s + \mathbb{E}\Big( \sup_{-(m+1)\delta \le \theta \le -m\delta} \Big| \int_{k\delta+\theta}^{t+\theta} \sigma_1(X^\varepsilon_s)\mathrm{d}W_1(s) \Big|^p \Big)$$

$$\lesssim \delta^{p-1} \int_{(k-m-1)\delta}^{t-m\delta} \mathbb{E}|b_1(X^\varepsilon_s, Y^\varepsilon_s)|^p \mathrm{d}s + \mathbb{E}\Big( \Big| \int_{(k-m-1)\delta}^{t-(m+1)\delta} \sigma_1(X^\varepsilon_s)\mathrm{d}W_1(s) \Big|^p \Big)$$

$$+ \mathbb{E}\Big( \sup_{-(m+1)\delta \le \theta \le -m\delta} \Big| \int_{t-(m+1)\delta}^{t+\theta} \sigma_1(X^\varepsilon_s)\mathrm{d}W_1(s) \Big|^p \Big)$$

$$+ \mathbb{E}\Big( \sup_{-(m+1)\delta \le \theta \le -m\delta} \Big| \int_{(k-m-1)\delta}^{k\delta+\theta} \sigma_1(X^\varepsilon_s)\mathrm{d}W_1(s) \Big|^p \Big)$$

$$\lesssim \delta^{p-1} \int_{(k-m-1)\delta}^{t-m\delta} \mathbb{E}|b_1(X^\varepsilon_s, Y^\varepsilon_s)|^p \mathrm{d}s + \delta^{\frac{p-2}{2}} \mathbb{E}\Big( \int_{(k-m-1)\delta}^{t-(m+1)\delta} \|\sigma_1(X^\varepsilon_s)\|^p \mathrm{d}s \Big)$$

$$+ \mathbb{E}\Big( \int_{t-(m+1)\delta}^{t-m\delta} \|\sigma_1(X^\varepsilon_s)\|^2 \mathrm{d}s \Big)^{p/2} + \mathbb{E}\Big( \int_{(k-m-1)\delta}^{(k-m)\delta} \|\sigma_1(X^\varepsilon_s)\|^2 \mathrm{d}s \Big)^{p/2}$$

$$\lesssim_T \delta^{\frac{p}{2}}.$$

**Case 2:** $m \ge k + 1$. In view of (**A4**) and (**A5**), it follows that

$$J_p(t, m, \delta) \lesssim |\xi(t+\theta) - \xi(t_\delta + \theta)|^p + |D_1(\xi_{t+\theta}) - D_1(\xi_{t_\delta+\theta})|^p$$

$$\lesssim |\xi(t+\theta) - \xi(t_\delta + \theta)|^p + \sup_{-\tau \le \theta' \le 0} |\xi(t+\theta+\theta') - \xi(t_\delta + \theta + \theta')|^p$$

$$\lesssim \delta^p.$$

**Case 3:** $m = k$. Also, by Hölder's inequality and B-D-G's inequality, we deduce from $(\mathbf{A1})$, and (4.3) that

$$
\begin{aligned}
J_p(t, m, \delta) = &\mathbb{E}\Big( \sup_{-(k+1)\delta \leq \theta \leq -k\delta} |X^\varepsilon(t+\theta) - D_1(X^\varepsilon_{t+\theta}) - X^\varepsilon(k\delta+\theta) + D_1(X^\varepsilon_{k\delta+\theta})|^p \Big) \\
&\lesssim \delta^p + \mathbb{E}\Big( \sup_{-(k+1)\delta \leq \theta \leq -k\delta} (|X^\varepsilon(t+\theta) - X^\varepsilon(0)|^p \mathbf{1}_{\{t+\theta>0\}}) \Big) \\
&\lesssim \delta^p + \mathbb{E}\Big( \sup_{-t \leq \theta \leq -k\delta} \Big| \int_0^{t+\theta} b_1(X^\varepsilon_s, Y^\varepsilon_s) \mathrm{d}s \Big|^p \Big) \\
&\quad + \mathbb{E}\Big( \sup_{-t \leq \theta \leq -k\delta} \Big| \int_0^{t+\theta} \sigma_1(X^\varepsilon_s) \mathrm{d}W_1(s) \Big|^p \Big) \\
&\lesssim \delta^p + \delta^{p-1} \int_0^{t-k\delta} \mathbb{E}|b_1(X^\varepsilon_s, Y^\varepsilon_s)|^p \mathrm{d}s + \delta^{\frac{p-2}{2}} \int_0^{t-k\delta} \mathbb{E}\|\sigma_1(X^\varepsilon_s)\|^p \mathrm{d}s \\
&\lesssim_T \delta^{\frac{p}{2}},
\end{aligned}
$$

On the other hand, using the inequality:

$$
(4.7) \qquad (a+b)^p \leq \Big[ 1 + \widehat{\epsilon}^{\frac{1}{p-1}} \Big]^{p-1} \Big( a^p + \frac{b^p}{\widehat{\epsilon}} \Big), \quad a, b > 0, \;\; \widehat{\epsilon} > 0, \quad p > 1,
$$

we have

$$
\begin{aligned}
\mathbb{E}\|X^\varepsilon_t - X^\varepsilon_{t_\delta}\|^p_\infty = &\mathbb{E}\Big( \sup_{-\tau \leq \theta \leq 0} |X^\varepsilon(t+\theta) - X^\varepsilon(t_\delta+\theta)|^p_\infty \Big) \\
\leq &\Big[ 1 + \widehat{\epsilon}^{\frac{1}{p-1}} \Big]^{p-1} \Big( \frac{1}{\widehat{\epsilon}} \mathbb{E}\Big( \sup_{-\tau \leq \theta \leq 0} |D_1(X^\varepsilon_{t+\theta}) - D_1(X^\varepsilon_{t_\delta+\theta})|^p \Big) \\
&+ \mathbb{E}\Big( \sup_{-\tau \leq \theta \leq 0} |X^\varepsilon(t+\theta) - D_1(X^\varepsilon_{t+\theta}) - X^\varepsilon(t_\delta+\theta) + D_1(X^\varepsilon_{t_\delta+\theta})|^p \Big) \Big).
\end{aligned}
$$

Letting $\widehat{\epsilon} = \Big[ \frac{\sqrt{\kappa_1}}{1-\sqrt{\kappa_1}} \Big]^{p-1}$ and using (4.4) we see that

$$
\begin{aligned}
\mathbb{E}\|X^\varepsilon_t - X^\varepsilon_{t_\delta}\|^p_\infty \leq &\sqrt{\kappa_1} \mathbb{E}\Big( \sup_{-\tau \leq \theta \leq 0} \|X^\varepsilon_{t+\theta} - X^\varepsilon_{t_\delta+\theta}\|^p_\infty \Big) + \frac{c_T \delta^{\frac{p-2}{2}}}{(1-\sqrt{\kappa_1})^{p-1}} \\
\leq &\sqrt{\kappa_1} \mathbb{E}\Big( \|X^\varepsilon_t - X^\varepsilon_{t_\delta}\|^p_\infty \Big) + \sqrt{\kappa_1} \mathbb{E}\Big( \|X^\varepsilon_{t-\tau} - X^\varepsilon_{t_\delta-\tau}\|^p_\infty \Big) + \frac{c_T \delta^{\frac{p-2}{2}}}{(1-\sqrt{\kappa_1})^{p-1}}
\end{aligned}
$$

holds for all $0 \leq t \leq T$. Consequently, for any $t \geq 0$, there exists an integer $n \geq 1$ such that $t \in [(n-1)\tau, n\tau)$

$$
\begin{aligned}
\mathbb{E}\|X^\varepsilon_t - X^\varepsilon_{t_\delta}\|^p_\infty \leq &\frac{\sqrt{\kappa_1}}{1-\sqrt{\kappa_1}} \mathbb{E}\Big( \|X^\varepsilon_{t-\tau} - X^\varepsilon_{t_\delta-\tau}\|^p_\infty \Big) + \frac{c_T \delta^{\frac{p-2}{2}}}{(1-\sqrt{\kappa_1})^p} \\
\leq &\Big( \frac{\sqrt{\kappa_1}}{1-\sqrt{\kappa_1}} \Big)^{n-1} \delta^p + \frac{c_T \delta^{\frac{p-2}{2}}}{(1-\sqrt{\kappa_1})^p} \Big\{ 1 + \frac{\sqrt{\kappa_1}}{1-\sqrt{\kappa_1}} + \cdots + \Big( \frac{\sqrt{\kappa_1}}{1-\sqrt{\kappa_1}} \Big)^{n-2} \Big\} \\
\lesssim_T &\delta^{\frac{p-2}{2}}.
\end{aligned}
$$

$\square$

The lemma below provides an error bound of the difference in the strong sense between the slow component $(X^\varepsilon(t))$ and its approximation $(\widetilde{X}^\varepsilon(t))$.

**Lemma 4.2.** Assume that **(A1)** and **(A2)** hold and suppose further $\varepsilon/\delta \in (0,1)$. Then, there exists $\beta > 0$ such that

$$\mathbb{E}\left(\sup_{0\le s\le T}|X^\varepsilon(t) - \widetilde{X}^\varepsilon(t)|^p\right) \lesssim_T \delta^{\frac{p-2}{2}}(1 + \varepsilon^{-1}\mathrm{e}^{\frac{\beta\delta}{\varepsilon}}), \quad p > 2.$$

*Proof.* Note that

$$\mathbb{E}\left(\sup_{0\le s\le T}|X^\varepsilon(s) - \widetilde{X}^\varepsilon(s)|^p\right)$$
$$\le 2^{p-1}\mathbb{E}\left(\sup_{0\le s\le T}|D_1(X_s^\varepsilon) - D_1(X_{s_\delta}^\varepsilon)|^p\right)$$
$$+ 2^{p-1}\mathbb{E}\left(\sup_{0\le s\le T}|X^\varepsilon(s) - D_1(X_s^\varepsilon) - \widetilde{X}^\varepsilon(s) + D_1(X_{s_\delta}^\varepsilon)|^p\right).$$

In view of Hölder's inequality, B-D-G's inequality, it follows from (2.1) and **(A4)** that

$$\mathbb{E}\left(\sup_{0\le s\le T}|D_1(X_s^\varepsilon) - D_1(X_{s_\delta}^\varepsilon)|^p\right)$$
$$\le \kappa_1^{\frac{p}{2}}\mathbb{E}\left(\sup_{0\le s\le T}|X^\varepsilon(s) - X^\varepsilon(s_\delta)|^p\right)$$
$$\lesssim_T \delta^{\frac{p}{2}}.$$

By using Lemma 4.1 and **(A1)**, one gives

$$\mathbb{E}\left(\sup_{0\le s\le t}|X^\varepsilon(s) - D_1(X_s^\varepsilon) - \widetilde{X}^\varepsilon(s) + D_1(X_{s_\delta}^\varepsilon)|^p\right)$$
$$\lesssim_T \int_0^t \mathbb{E}\{\|X_s^\varepsilon - X_{s_\delta}^\varepsilon\|_\infty^p + \|Y_s^\varepsilon - \widetilde{Y}_s^\varepsilon\|_\infty^p\}\mathrm{d}s$$
$$\lesssim_T \delta^{\frac{p-2}{2}} + \int_0^t \mathbb{E}\|Y_s^\varepsilon - \widetilde{Y}_s^\varepsilon\|_\infty^p\mathrm{d}s, \quad t \in (0,T].$$

Therefore, to finish the argument of Lemma 4.2, it suffices to show that there exists $\beta > 0$ such that

(4.8)
$$\sup_{t\in[0,T]}\mathbb{E}\|Y_t^\varepsilon - \widetilde{Y}_t^\varepsilon\|_\infty^p \lesssim_T \varepsilon^{-1}\delta^{\frac{p-2}{2}}\mathrm{e}^{\frac{\beta\delta}{\varepsilon}}.$$

In the sequel, we shall claim (4.8) by an induction argument. For any $t \in [0,\tau)$, due to $Y_0^\varepsilon = \widetilde{Y}_0^\varepsilon = \eta$, it is readily to check that

$$\mathbb{E}\|Y_t^\varepsilon - \widetilde{Y}_t^\varepsilon\|_\infty^p \le \sum_{j=0}^{\lfloor t/\delta\rfloor}\mathbb{E}\left(\sup_{j\delta\le s\le((j+1)\delta)\wedge t}|Y^\varepsilon(s) - \widetilde{Y}^\varepsilon(s)|^p\right) =: I(t,\delta).$$

16

By means of Itô's formula and B-D-G's inequality, together with $\widetilde{Y}^\varepsilon(t_\delta) = Y^\varepsilon(t_\delta)$, we obtain from (**A2**) that

$$\mathbb{E}\Big(\sup_{j\delta \leq s \leq ((j+1)\delta)\wedge t} |Y^\varepsilon(s) - \widetilde{Y}^\varepsilon(s)|^p\Big)$$

$$= \mathbb{E}\Big(\sup_{j\delta \leq s \leq ((j+1)\delta)\wedge t} |Y^\varepsilon(s) - D_2(Y^\varepsilon(s-\tau)) - \widetilde{Y}^\varepsilon(s) + D_2(\widetilde{Y}^\varepsilon(s-\tau))|^p\Big)$$

$$\leq \frac{c}{\varepsilon} \int_{j\delta}^{((j+1)\delta)\wedge t} \{\mathbb{E}\|X_s^\varepsilon - X_{s_\delta}^\varepsilon\|_\infty^p + \mathbb{E}|Y^\varepsilon(s) - \widetilde{Y}^\varepsilon(s)|^p\}\mathrm{d}s$$

$$+ \frac{1}{2}\mathbb{E}\Big(\sup_{j\delta \leq s \leq ((j+1)\delta)\wedge t} |Y^\varepsilon(s) - \widetilde{Y}^\varepsilon(s)|^p\Big), \quad t \in [0, \tau].$$

Consequently, we conclude that

(4.9)
$$I(t,\delta) \lesssim \frac{1}{\varepsilon} \int_0^t \mathbb{E}\|X_s^\varepsilon - X_{s_\delta}^\varepsilon\|_\infty^p \mathrm{d}s + \frac{1}{\varepsilon} \int_0^\delta \sum_{j=0}^{\lfloor t/\delta \rfloor} \mathbb{E}\Big(\sup_{j\delta \leq r \leq ((j\delta+s))\wedge t} |Y^\varepsilon(r) - \widetilde{Y}^\varepsilon(r)|^p\Big)\mathrm{d}s$$

$$\lesssim \frac{1}{\varepsilon} \int_0^t \mathbb{E}\|X_s^\varepsilon - X_{s_\delta}^\varepsilon\|_\infty^p \mathrm{d}s + \frac{1}{\varepsilon} \int_0^\delta I(t,s)\mathrm{d}s.$$

This, combining Lemma 4.1 with Gronwall's inequality, gives that

(4.10)
$$\mathbb{E}\|Y_t^\varepsilon - \widetilde{Y}_t^\varepsilon\|_\infty^p \lesssim \varepsilon^{-1}\delta^{\frac{p-2}{2}}\mathrm{e}^{\frac{c\delta}{\varepsilon}}, \qquad t \in [0, \tau)$$

for some $c > 0$. Next, for any $t \in [\tau, 2\tau)$, thanks to (4.10), it is immediate to note that

$$\mathbb{E}\|Y_t^\varepsilon - \widetilde{Y}_t^\varepsilon\|_\infty^p \leq \mathbb{E}\Big(\|Y_\tau^\varepsilon - \widetilde{Y}_\tau^\varepsilon\|_\infty^p\Big) + \mathbb{E}\Big(\sup_{\tau \leq s \leq t} |Y^\varepsilon(s) - \widetilde{Y}^\varepsilon(s)|^p\Big)$$

$$\leq c\Big\{\varepsilon^{-1}\delta^{\frac{p-2}{2}}\mathrm{e}^{\frac{c\delta}{\varepsilon}} + \sum_{j=0}^{\lfloor \frac{t-\tau}{\delta} \rfloor} \mathbb{E}\Big(\sup_{(N+j)\delta \leq s \leq ((N+j+1)\delta)\wedge t} |Y^\varepsilon(s) - \widetilde{Y}^\varepsilon(s)|^p\Big)\Big\}$$

$$=: c\{\varepsilon^{-1}\delta^{\frac{p-2}{2}}\mathrm{e}^{\frac{c\delta}{\varepsilon}} + M(t,\tau,\delta)\}.$$

By using Itô's formula and B-D-G's inequality again, for any $t \in [\tau, 2\tau)$, we deduce from (4.10) that

$$\mathbb{E}\Big(\sup_{j\delta \leq s \leq ((j+1)\delta)\wedge t} |Y^\varepsilon(s) - D_2(Y^\varepsilon(s-\tau)) - \widetilde{Y}^\varepsilon(s) + D_2(\widetilde{Y}^\varepsilon(s-\tau))|^p\Big)$$

$$\leq \frac{c}{\varepsilon} \int_{j\delta}^{((j+1)\delta)\wedge t} \{\mathbb{E}\|X_s^\varepsilon - X_{s_\delta}^\varepsilon\|_\infty^p + \mathbb{E}|Y^\varepsilon(s) - \widetilde{Y}^\varepsilon(s)|^p + \mathbb{E}|Y^\varepsilon(s-\tau) - \widetilde{Y}^\varepsilon(s-\tau)|^p\}\mathrm{d}s$$

$$+ \frac{1}{2}\mathbb{E}\Big(\sup_{j\delta \leq s \leq ((j+1)\delta)\wedge t} |Y^\varepsilon(s) - D_2(Y^\varepsilon(s-\tau)) - \widetilde{Y}^\varepsilon(s) + D_2(\widetilde{Y}^\varepsilon(s-\tau))|^p\Big),$$

17

and,

$$M(t,\tau,\delta) \le 2^{p-1} \sum_{j=0}^{\lfloor \frac{t-\tau}{\delta} \rfloor} \mathbb{E}\Big( \sup_{j\delta \le s \le ((j+1)\delta)\wedge t} |D_2(Y^\varepsilon(s-\tau)) - D_2(\widetilde{Y}^\varepsilon(s-\tau))|^p \Big)$$

$$+ 2^{p-1} \sum_{j=0}^{\lfloor \frac{t-\tau}{\delta} \rfloor} \mathbb{E}\Big( \sup_{j\delta \le s \le ((j+1)\delta)\wedge t} |Y^\varepsilon(s) - D_2(Y^\varepsilon(s-\tau)) - \widetilde{Y}^\varepsilon(s) + D_2(\widetilde{Y}^\varepsilon(s-\tau))|^p \Big)$$

$$\lesssim \varepsilon^{-1}\delta^{\frac{p-2}{2}} e^{\frac{c\delta}{\varepsilon}} + \frac{1}{\varepsilon}\int_\tau^t \mathbb{E}\|X_s^\varepsilon - X_{s_\delta}^\varepsilon\|_\infty^p \mathrm{d}s$$

$$+ \frac{1}{\varepsilon}\int_0^\delta \sum_{j=0}^{\lfloor \frac{t-\tau}{\delta} \rfloor} \mathbb{E}\Big( \sup_{(N+j)\delta \le r \le ((N+j)\delta + s)\wedge t} |Y^\varepsilon(r) - \widetilde{Y}^\varepsilon(r)|^p \Big)\mathrm{d}s$$

$$+ \frac{1}{\varepsilon}\int_0^\delta \sum_{j=0}^{\lfloor \frac{t-\tau}{\delta} \rfloor} \mathbb{E}\Big( \sup_{j\delta \le s \le ((j+1)\delta)\wedge(t-\tau)} |Y^\varepsilon(s) - \widetilde{Y}^\varepsilon(s)|^p \Big)\mathrm{d}s$$

$$\lesssim \frac{\delta^{\frac{p-2}{2}}}{\varepsilon} + \Big(\frac{\delta}{\varepsilon} + 1\Big) \cdot \frac{\delta^{\frac{p-2}{2}}}{\varepsilon} e^{\frac{c\delta}{\varepsilon}} + \frac{1}{\varepsilon}\int_0^\delta M(t,\tau,s)\mathrm{d}s.$$

Thus, the Gronwall inequality reads

$$M(t,\tau,\delta) \lesssim \Big\{ \frac{\delta^{\frac{p-2}{2}}}{\varepsilon} + \Big(\frac{\delta}{\varepsilon} + 1\Big) \cdot \frac{\delta^{\frac{p-2}{2}}}{\varepsilon} e^{\frac{c\delta}{\varepsilon}} \Big\} e^{\frac{c\delta}{\varepsilon}} \lesssim \frac{\delta}{\varepsilon} \cdot \frac{\delta^{\frac{p-2}{2}}}{\varepsilon} e^{\frac{c\delta}{\varepsilon}} \lesssim \frac{\delta^{\frac{p-2}{2}}}{\varepsilon} e^{\frac{c\delta}{\varepsilon}},$$

where we have used $\varepsilon/\delta \in (0,1)$ in the second step . Finally, (4.8) follows by repeating the previous procedure. □

The following consequence explores a uniform estimate w.r.t. the parameter $\varepsilon$ for the segment process associated with the auxiliary fast motion.

**Lemma 4.3.** Assume that (**A1**) and (**A3**) hold. Then, there exists $C_T > 0$, independent of $\varepsilon$, such that

(4.11) $$\sup_{t\in[0,T]} \mathbb{E}\|\widetilde{Y}_t^\varepsilon\|_\infty^2 \le C_T.$$

*Proof.* From (2.2), it follows that

(4.12)
$$Y^\varepsilon(t) = \eta(0) - D_2(\eta) + D_2(Y^\varepsilon(t-\tau)) + \int_0^{t/\varepsilon} b_2(X_{\varepsilon s}^\varepsilon, Y^\varepsilon(\varepsilon s), Y^\varepsilon(\varepsilon s - \tau))\mathrm{d}t$$

$$+ \int_0^{t/\varepsilon} \sigma_2(X_{\varepsilon s}^\varepsilon, Y^\varepsilon(\varepsilon s), Y^\varepsilon(\varepsilon s - \tau))\mathrm{d}\overline{W}_2(s), \quad t > 0,$$

where we used the fact that $\overline{W}(t) := \frac{1}{\sqrt{\varepsilon}}W_2(\varepsilon t)$ is a Brownian motion. For fixed $\varepsilon > 0$ and $t \ge 0$, let $\overline{Y}^\varepsilon(t+\theta) = Y^\varepsilon(\varepsilon t + \theta), \theta \in [-\tau, 0]$. So, one has $\overline{Y}_t^\varepsilon = Y_{\varepsilon t}^\varepsilon$. Observe that (4.12) can

18

be rewritten as

$$\overline{Y}^{\varepsilon}(t/\varepsilon) = \eta(0) - D_2(\eta) + D_2(Y^{\varepsilon}(t/\varepsilon - \tau)) + \int_0^{t/\varepsilon} b_2(X_{\varepsilon s}^{\varepsilon}, \overline{Y}^{\varepsilon}(s), \overline{Y}^{\varepsilon}(s - \tau))\mathrm{d}s$$

$$+ \int_0^{t/\varepsilon} \sigma_2(X_{\varepsilon s}^{\varepsilon}, \overline{Y}^{\varepsilon}(s), \overline{Y}^{\varepsilon}(s - \tau))\mathrm{d}\overline{W}_2(s).$$

Then, following an argument to deduce (3.9), for any $s > 0$ we can deduce that

$$\mathbb{E}\|\overline{Y}_s^{\varepsilon}\|_{\infty}^2 \lesssim 1 + \|\eta\|_{\infty}^2 \mathrm{e}^{-\lambda s} + \mathbb{E}\Big( \sup_{0 \le r \le \varepsilon s} \|X_r^{\varepsilon}\|_{\infty}^2 \Big).$$

This, together with $\overline{Y}_t^{\varepsilon} = Y_{\varepsilon t}^{\varepsilon}$, gives that

$$\mathbb{E}\|Y_{\varepsilon s}^{\varepsilon}\|_{\infty}^2 \lesssim 1 + \|\eta\|_{\infty}^2 \mathrm{e}^{-\lambda s} + \mathbb{E}\Big( \sup_{0 \le r \le \varepsilon s} \|X_r^{\varepsilon}\|_{\infty}^2 \Big).$$

In particular, taking $s = t/\varepsilon$ we arrive at

$$\mathbb{E}\|Y_t^{\varepsilon}\|_{\infty}^2 \lesssim 1 + \|\eta\|_{\infty}^2 + \mathbb{E}\Big( \sup_{0 \le r \le t} \|X_r^{\varepsilon}\|_{\infty}^2 \Big).$$

This, together with (4.3), yields that

$$\sup_{t \in [0,T]} \mathbb{E}\|Y_t^{\varepsilon}\|_{\infty}^2 \le C_T$$

for some $C_T > 0$. Observe from (4.8) and Höder's inequality that

$$\mathbb{E}\|\widetilde{Y}_t^{\varepsilon}\|_{\infty}^2 \le 2\mathbb{E}\|Y_t^{\varepsilon} - \widetilde{Y}_t^{\varepsilon}\|_{\infty}^2 + 2\mathbb{E}\|Y_t^{\varepsilon}\|_{\infty}^2$$

$$\lesssim_T 1 + \Big( \varepsilon^{-1} \delta^{\frac{p-2}{2}} \mathrm{e}^{\frac{\beta\delta}{\varepsilon}} \Big)^{2/p}, \quad p > 4.$$

Next, taking $\delta = \varepsilon(-\ln\varepsilon)^{\frac{1}{2}}$ in the estimate above and letting $y = (-\ln\varepsilon)^{\frac{1}{2}}$, we have

$$\mathbb{E}\|\widetilde{Y}_t^{\varepsilon}\|_{\infty}^2 \lesssim_T 1 + \Big( \mathrm{e}^{y^2}(\mathrm{e}^{-y^2}y)^{\frac{p-2}{2}} \mathrm{e}^{\beta y} \Big)^{2/p}, \quad p > 4.$$

Then, the desired assertion follows since the leading term $\mathrm{e}^{y^2}(\mathrm{e}^{-y^2}y)^{\frac{p-2}{2}}\mathrm{e}^{\beta y} \to 0$ as $y \uparrow \infty$ whenever $p > 4$.

$\square$

Equipped with several lemmas above, we are in position to show our main result as below.

*Proof of Theorem 2.3.* For any $t \in [0,T]$ and $p > 0$, set

$$\Lambda(t) := \mathbb{E}\Big( \sup_{0 \le s \le t} |X^{\varepsilon}(s) - \overline{X}(s)|^p \Big) \quad \text{and} \quad \Gamma(t) := \mathbb{E}\Big( \sup_{0 \le s \le t} |\widetilde{X}^{\varepsilon}(s) - D_1(X_{s_\delta}^{\varepsilon}) - \overline{X}(s) + D_1(\overline{X}_s)|^p \Big).$$

By Hölder's inequality, it is sufficient to verify that

$$\lim_{\varepsilon \to 0} \Lambda(T) = 0, \quad p > 4. \tag{4.13}$$

In what follows, let $t \in [0, T]$ be arbitrary and assume $p > 4$. By using the inequality (4.7), we have that for any $\widehat{\epsilon} > 0$

$$
\begin{aligned}
\Lambda(t) &= \mathbb{E}\Big( \sup_{0 \le s \le t} |X^\varepsilon(s) - \widetilde{X}^\varepsilon(s) + \widetilde{X}^\varepsilon(s) - D_1(X^\varepsilon_{s\delta}) - \overline{X}(s) + D_1(\overline{X}_s) + D_1(X^\varepsilon_{s\delta}) - D_1(\overline{X}_s)|^p \Big) \\
&\le [1 + \widehat{\epsilon}^{\frac{1}{p-1}}]^{p-1}\Big( \mathbb{E}\Big( \sup_{0 \le s \le t} |X^\varepsilon(s) - \widetilde{X}^\varepsilon(s) + \widetilde{X}^\varepsilon(s) - D_1(X^\varepsilon_{s\delta}) - \overline{X}(s) + D_1(\overline{X}_s)|^p \Big) \\
&\quad + \frac{1}{\widehat{\epsilon}}\mathbb{E}\Big( \sup_{0 \le s \le t} |D_1(X^\varepsilon_{s\delta}) - D_1(\overline{X}_s)|^p \Big)\Big).
\end{aligned}
$$

Letting $\widehat{\epsilon} = [\frac{\sqrt{\kappa_1}}{1 - \sqrt{\kappa_1}}]^{p-1}$, one gives

$$
\begin{aligned}
\Lambda(t) &\le \frac{1}{(1 - \sqrt{\kappa_1})^{p-1}}\mathbb{E}\Big( \sup_{0 \le s \le t} |X^\varepsilon(s) - \widetilde{X}^\varepsilon(s) + \widetilde{X}^\varepsilon(s) - D_1(X^\varepsilon_{s\delta}) - \overline{X}(s) + D_1(\overline{X}_s)|^p \Big) \\
&\quad + \sqrt{\kappa_1}\,\mathbb{E}\Big( \sup_{0 \le s \le t} \|X^\varepsilon_{s\delta} - \overline{X}_s\|_\infty^p \Big) \\
&\le \frac{2^{p-1}}{(1 - \sqrt{\kappa_1})^{p-1}}\Big( \mathbb{E}\Big( \sup_{0 \le s \le t} |X^\varepsilon(s) - \widetilde{X}^\varepsilon(s)|^p \Big) + \Gamma(t) \Big) + \sqrt{\kappa_1}\,\mathbb{E}\Big( \sup_{0 \le s \le t} \|X^\varepsilon_{s\delta} - \overline{X}_s\|_\infty^p \Big) \\
&\le \frac{2^{p-1}}{(1 - \sqrt{\kappa_1})^{p-1}}\Big( \mathbb{E}\Big( \sup_{0 \le s \le t} |X^\varepsilon(s) - \widetilde{X}^\varepsilon(s)|^p \Big) + \Gamma(t) \Big) + \sqrt{\kappa_1}\,\Lambda(t)
\end{aligned}
$$

Therefore, it follows that from Lemma 4.2 that

$$\Lambda(t) \lesssim \mathbb{E}\Big( \sup_{0 \le s \le t} |X^\varepsilon(s) - \widetilde{X}^\varepsilon(s)|^p \Big) + \Gamma(t) \lesssim \delta^{\frac{p-2}{2}}\Big( 1 + \frac{1}{\varepsilon}e^{\frac{\beta\delta}{\varepsilon}} \Big) + \Gamma(t). \tag{4.14}$$

Next, if we can show that

$$\Gamma(t) \lesssim \delta^{\frac{p-2}{2}}\Big( 1 + \frac{1}{\varepsilon}e^{\frac{\beta\delta}{\varepsilon}} \Big) + \Big( \frac{\varepsilon}{\delta} \Big)^\nu + \int_0^t \Lambda(s)\mathrm{d}s \tag{4.15}$$

for some $\nu \in (0, 1)$, inserting (4.15) back into (4.14) and utilizing Gronwall's inequality, we deduce that

$$\Lambda(t) \lesssim \delta^{\frac{p-2}{2}}\Big( 1 + \frac{1}{\varepsilon}e^{\frac{\beta\delta}{\varepsilon}} \Big) + \Big( \frac{\varepsilon}{\delta} \Big)^\nu.$$

Thus, the desired assertion (4.13) follows by, in particular, choosing $\delta = \varepsilon(-\ln \varepsilon)^{\frac{1}{2}}$. Ineed, it is easy to see that $\varepsilon/\delta \in (0, 1)$, which is prerequisite in Lemma 4.2, for $\varepsilon \in (0, 1)$ small enough, and that $\delta \to 0$ as $\varepsilon \downarrow 0$. Furthermore, let $y = (-\ln \varepsilon)^{\frac{1}{2}}$ (hence $\varepsilon = e^{-y^2}$), which goes into infinity as $\varepsilon$ tends to zero. Then, we have

$$\Lambda(t) \lesssim (e^{-y^2}y)^{\frac{p-2}{2}}\Big( 1 + e^{y^2 + \beta y} \Big) + y^{-\nu},$$

20

which goes to zero by taking $p > 4$ and letting $y \uparrow \infty$.

Next, we intend to claim (4.15). Set

$$\Gamma_p(t, \delta, \varepsilon) := \mathbb{E}\Big( \sup_{0 \leq s \leq t} \Big| \int_0^s \{b_1(X_{r_\delta}^\varepsilon, \widetilde{Y}_r^\varepsilon) - \bar{b}_1(X_{r_\delta}^\varepsilon)\} \mathrm{d}r \Big|^p \Big), \quad t \in [0, T].$$

Applying Hölder's inequality, B-D-G's inequality, Lipschitz property of $\bar{b}_1$ due to Corollary 2.2, and Lemma 4.1, we derive that

$$\Gamma(t) \lesssim \mathbb{E}\Big( \sup_{0 \leq s \leq t} \Big| \int_0^t \{b_1(X_{s_\delta}^\varepsilon, \widetilde{Y}_s^\varepsilon) - \bar{b}_1(\overline{X}_s)\} \mathrm{d}s \Big|^p \Big) + \int_0^t \mathbb{E}\|\sigma_1(X_{s_\delta}^\varepsilon) - \sigma_1(\overline{X}_s)\|^p \mathrm{d}s$$

$$\lesssim \Gamma_p(t, \delta, \varepsilon) + \int_0^t \mathbb{E}|\bar{b}_1(X_{s_\delta}^\varepsilon) - \bar{b}_1(X_s^\varepsilon)|^p \mathrm{d}s + \int_0^t \mathbb{E}|\bar{b}_1(X_s^\varepsilon) - \bar{b}_1(\widetilde{X}_s^\varepsilon)|^p \mathrm{d}s$$

$$+ \int_0^t \mathbb{E}|\bar{b}_1(\widetilde{X}_s^\varepsilon) - \bar{b}_1(\overline{X}_s)|^p \mathrm{d}s + \int_0^t \mathbb{E}\|\sigma_1(X_{s_\delta}^\varepsilon) - \sigma_1(\overline{X}_s)\|^p \mathrm{d}s$$

$$\lesssim \Gamma_p(t, \delta, \varepsilon) + \int_0^t \mathbb{E}\|X_s^\varepsilon - \widetilde{X}_s^\varepsilon\|_\infty^p \mathrm{d}s + \int_0^t \mathbb{E}\|X_{s_\delta}^\varepsilon - X_s^\varepsilon\|_\infty^p \mathrm{d}s$$

$$+ \int_0^t \mathbb{E}\Big( \sup_{0 \leq r \leq s} |D_1(X_{r_\delta}^\varepsilon) - D_1(\overline{X}_r)|^p \Big) \mathrm{d}s + \int_0^t \Gamma(s) \mathrm{d}s + \int_0^t \Lambda(s) \mathrm{d}s$$

$$\lesssim \delta^{\frac{p-2}{2}} + \frac{1}{\varepsilon}\delta^{\frac{p-2}{2}} \mathrm{e}^{\frac{c\delta}{\varepsilon}} + \Gamma_p(t, \delta, \varepsilon) + \int_0^t \Gamma(s) \mathrm{d}s + \int_0^t \Lambda(s) \mathrm{d}s,$$

which, together with Gronwall's inequality, leads to

$$(4.16) \qquad \Gamma(t) \lesssim \delta^{\frac{p-2}{2}} \Big( 1 + \frac{1}{\varepsilon} \mathrm{e}^{\frac{\beta\delta}{\varepsilon}} \Big) + \Gamma_p(t, \delta, \varepsilon) + \int_0^t \Lambda(s) \mathrm{d}s,$$

where we have utilized the fact that $\Gamma_p(t, \delta, \varepsilon)$ is nondecreasing with respect to $t$. By a comparison (4.15) with (4.16), we need only to prove

$$(4.17) \qquad \Gamma_p(t, \delta, \varepsilon) \lesssim \Big( \frac{\varepsilon}{\delta} \Big)^\nu$$

for some $\nu \in (0, 1)$.

Let

$$\Upsilon_p(k, \delta, \varepsilon) = \mathbb{E}\Big( \Big| \int_{k\delta}^{((k+1)\delta) \wedge t} \{b_1(X_{k\delta}^\varepsilon, \widetilde{Y}_s^\varepsilon) - \bar{b}_1(X_{k\delta}^\varepsilon)\} \mathrm{d}s \Big|^p \Big) \quad \text{for any } p > 0.$$

21

In the sequel, we show that (4.17) holds. By Hölder's inequality, we obtain that

$$\Gamma_p(t, \delta, \varepsilon) = \mathbb{E}\Big( \sup_{0 \leq s \leq t} \Big| \sum_{k=0}^{\lfloor s/\delta \rfloor} \int_{k\delta}^{((k+1)\delta) \wedge t} \{b_1(X_{k\delta}^\varepsilon, \widetilde{Y}_r^\varepsilon) - \bar{b}_1(X_{k\delta}^\varepsilon)\} \mathrm{d}r \Big|^p \Big)$$

(4.18)
$$\leq \mathbb{E}\Big( \sup_{0 \leq s \leq t} \Big( (\lfloor s/\delta \rfloor + 1)^{p-1} \sum_{k=0}^{\lfloor s/\delta \rfloor} \Upsilon_p(k, \delta, \varepsilon) \Big) \Big)$$

$$\leq (\lfloor t/\delta \rfloor + 1)^{p-1} \sum_{k=0}^{\lfloor t/\delta \rfloor} \Upsilon_p(k, \delta, \varepsilon)$$

$$\leq (\lfloor t/\delta \rfloor + 1)^p \max_{0 \leq k \leq \lfloor t/\delta \rfloor} \Upsilon_p(k, \delta, \varepsilon).$$

For any $p' \in (1, 2)$, by Hölder's inequality, (**A1**), and (4.3), observe that

$$\Upsilon_p(k, \delta, \varepsilon) \leq \Upsilon_2(k, \delta, \varepsilon)^{\frac{p'}{2}} \Big( \mathbb{E}\Big( \Big| \int_{k\delta}^{((k+1)\delta) \wedge t} \{b_1(X_{k\delta}^\varepsilon, \widetilde{Y}_s^\varepsilon) - \bar{b}_1(X_{k\delta}^\varepsilon)\} \mathrm{d}s \Big|^{\frac{2(p-p')}{2-p'}} \Big) \Big)^{\frac{2-p'}{2}}$$

$$\leq \Upsilon_2(k, \delta, \varepsilon)^{\frac{p'}{2}} \Big( \delta^{\frac{2(p-p')}{2-p'}-1} \mathbb{E}\Big( \Big| \int_{k\delta}^{((k+1)\delta) \wedge t} |b_1(X_{k\delta}^\varepsilon, \widetilde{Y}_s^\varepsilon) - \bar{b}_1(X_{k\delta}^\varepsilon)|^{\frac{2(p-p')}{2-p'}} \mathrm{d}s \Big| \Big) \Big)^{\frac{2-p'}{2}}$$

$$\lesssim \Upsilon_2(k, \delta, \varepsilon)^{\frac{p'}{2}} \delta^{\frac{2(p-p')}{2-p'} \times \frac{2-p'}{2}}$$

$$\lesssim \Upsilon_2(k, \delta, \varepsilon)^{\frac{p'}{2}} \delta^{p-p'}, \quad p > 4.$$

Substituting this into (4.18), we arrive at

$$\Gamma_p(t, \delta, \varepsilon) \lesssim \Upsilon_2(k, \delta, \varepsilon)^{\frac{p'}{2}} \delta^{-p'}.$$

Thus, to complete the argument, it remains to show that

$$\Upsilon_2(k, \delta, \varepsilon) \lesssim \varepsilon\delta.$$

Also, by virtue of Hölder's inequality, (**A1**), and (4.3), we derive that

(4.19)
$$\Upsilon_2(k, \delta, \varepsilon)$$
$$= 2 \int_{k\delta}^{((k+1)\delta) \wedge t} \int_s^{((k+1)\delta) \wedge t} \mathbb{E}\langle b_1(X_{k\delta}^\varepsilon, \widetilde{Y}_s^\varepsilon) - \bar{b}_1(X_{k\delta}^\varepsilon), b_1(X_{k\delta}^\varepsilon, \widetilde{Y}_r^\varepsilon) - \bar{b}_1(X_{k\delta}^\varepsilon) \rangle \mathrm{d}r \mathrm{d}s$$

$$\lesssim \int_{k\delta}^{(k+1)\delta} \int_s^{(k+1)\delta} (\mathbb{E}|\mathbb{E}((b_1(X_{k\delta}^\varepsilon, \widetilde{Y}_r^\varepsilon) - \bar{b}_1(X_{k\delta}^\varepsilon))|\mathscr{F}_s)|^2)^{1/2} \mathrm{d}r \mathrm{d}s.$$

For any $r \in [k\delta, (k+1)\delta)$, by the definition of $\widetilde{Y}^\varepsilon$, defined as in (4.2), it follows that

(4.20)
$$
\begin{aligned}
\widetilde{Y}^\varepsilon(r) &- D_2(\widetilde{Y}^\varepsilon(r-\tau)) \\
&= \widetilde{Y}^\varepsilon(k\delta) - D_2(\widetilde{Y}^\varepsilon(k\delta - \tau)) + \frac{1}{\varepsilon}\int_{k\delta}^r b_2(X_{k\delta}^\varepsilon, \widetilde{Y}^\varepsilon(u), \widetilde{Y}^\varepsilon(u-\tau))\mathrm{d}u \\
&\quad + \frac{1}{\sqrt{\varepsilon}}\int_{k\delta}^r \sigma_2(X_{k\delta}^\varepsilon, \widetilde{Y}^\varepsilon(u), \widetilde{Y}^\varepsilon(u-\tau))\mathrm{d}W_2(u) \\
&= \widetilde{Y}^\varepsilon(k\delta) - D_2(\widetilde{Y}^\varepsilon(k\delta - \tau)) + \int_0^{\frac{r-k\delta}{\varepsilon}} b_2(X_{k\delta}^\varepsilon, \widetilde{Y}^\varepsilon(k\delta + \varepsilon u), \widetilde{Y}^\varepsilon(k\delta + \varepsilon u - \tau))\mathrm{d}u \\
&\quad + \int_0^{\frac{r-k\delta}{\varepsilon}} \sigma_2(X_{k\delta}^\varepsilon, \widetilde{Y}^\varepsilon(k\delta + \varepsilon u - \tau))\mathrm{d}\widetilde{W}_2(u),
\end{aligned}
$$

where $\widetilde{W}_2(u) := (W_2(\varepsilon u + k\delta) - W(k\delta))/\sqrt{\varepsilon}$, which is also a Wiener process. For fixed $\varepsilon > 0$ and $u \geq 0$, let

$$
\overline{Y}^{X_{k\delta}^\varepsilon}(u + \theta) = \widetilde{Y}^\varepsilon(k\delta + \varepsilon u + \theta), \quad \theta \in [-\tau, 0].
$$

Then (4.20) can be rewritten as

$$
\begin{aligned}
\overline{Y}^{X_{k\delta}^\varepsilon}&\Big(\frac{r-k\delta}{\varepsilon}\Big) - D_2\Big(\overline{Y}^{X_{k\delta}^\varepsilon}\Big(\frac{r-k\delta}{\varepsilon} - \tau\Big)\Big) \\
&= \widetilde{Y}^\varepsilon(k\delta) - D_2(\widetilde{Y}^\varepsilon(k\delta - \tau)) + \int_0^{\frac{r-k\delta}{\varepsilon}} b_2\Big(X_{k\delta}^\varepsilon, \overline{Y}^{X_{k\delta}^\varepsilon}(u), \overline{Y}^{X_{k\delta}^\varepsilon}(u-\tau)\Big)\,\mathrm{d}u \\
&\quad + \int_0^{\frac{r-k\delta}{\varepsilon}} \sigma_2\Big(X_{k\delta}^\varepsilon, \overline{Y}^{X_{k\delta}^\varepsilon}(u), \overline{Y}^{X_{k\delta}^\varepsilon}(u-\tau)\Big)\,\mathrm{d}\widetilde{W}_2(u).
\end{aligned}
$$

Consequently, by the weak uniqueness of solution, we arrive at

(4.21)
$$
\mathscr{L}(\widetilde{Y}_r^\varepsilon) = \mathscr{L}\Big(Y_{(r-k\delta)/\varepsilon}^{X_{k\delta}^\varepsilon}(\widetilde{Y}_{k\delta}^\varepsilon)\Big),
$$

where $\mathscr{L}(\zeta)$ denotes the law of random variable $\zeta$. Finally, we obtain from (2.4), (4.19), (4.21), and Lemma 4.3 that

$$
\begin{aligned}
\Upsilon_2(k, \delta, \varepsilon) &\lesssim (1 + \mathbb{E}\|X_{k\delta}^\varepsilon\|_\infty^2 + \mathbb{E}\|\widetilde{Y}_{k\delta}^\varepsilon\|_\infty^2) \int_{k\delta}^{(k+1)\delta} \int_s^{(k+1)\delta} \exp\Big(-\frac{c(r-k\delta)}{\varepsilon}\Big)\mathrm{d}r\mathrm{d}s \\
&\lesssim \varepsilon\delta.
\end{aligned}
$$

The whole proof is therefore complete. $\square$

## Acknowledgments

23

# References

[1] Bao, J., Song, Q., Yin G., Yuan, C., Ergodicity and strong limit results for two-time-scale functional stochastic differential equations, arXiv:1508.07288v1, 2015.

[2] Bao, J., Yin, G., Yuan, C., Ergodicity for functional stochastic differential equations and applications, *Nonlinear Anal.*, **98** (2014), 66–82.

[3] Bao, J., Yin, G., Yuan, C., Stationary Distributions for Retarded Stochastic Differential Equations without Dissipativity, arXiv:1308.2018.

[4] Blömker, D., Hairer, M., Pavliotis, G. A., Multiscale analysis for stochastic partial differential equations with quadratic nonlineaities, *Nonlinearity*, **20** (2007), 1721–1744.

[5] Bréhier, C.-E., Strong and weak orders in averaging for SPDEs, *Stochastic Process. Appl.*, **122** (2012), 2553–2593.

[6] M.-F. Chen, *From Markov Chains to Non-Equilibrium Particle Systems,* Second Edition, World Scientific, Singapore, 2004.

[7] Drive, R.D., A functional differential system of neutral type arising in a two-body problem of classical electrodynamics, in "Nonlinear Differential Equations and Nonlinear Mechanics," 474–484, Academic Press, New York, 1963.

[8] Es-Sarhir, A., Scheutzow, M., van Gaans, O., Invariant measures for stochastic functional differential equations with superlinear drift term, *Differential Integral Equations*, **23** (2010), 189–200.

[9] Da Prato, G., Zabczyk, J., *Ergodicity for infinite-dimensional systems*, In: London Mathematical Society. Lecture Note Series, vol. 229, Cambridge University Press, Cambridge, 1996.

[10] E, W., Liu, D., Vanden-Eijnden, E., Analysis of multiscale methods for stochastic differential equations, *Comm. Pure Appl. Math.*, **58** (2005), 1544–1585.

[11] Freidlin, M. I., Wentzell, A. D., *Random Perturbations of Dynamical Systems*, second ed., New York, Springer-Verlag, 1998.

[12] Fu, H., Wan, L., Liu, J., Strong convergence in averaging principle for stochastic hyperbolic-parabolic equations with two time-scales, *Stochastic Process. Appl.*, **125** (2015), 3255–3279.

[13] Givon, D., Strong convergence rate for two-time-scale jump-diffusion stochastic differential systems, *SIAM Multiscale Model. Simul.*, **6** (2007), 577–594.

[14] Givon, D., Kevrekidis, I. G., Kupferman, R., Strong convergence of projective integration schemes for singularly perturbed stochastic differential systems, *Commun. Math. Sci.*, **4** (2006), 707–729.

[15] Hairer, M., Mattingly, J. C., Scheutzow, M., Asymptotic coupling and a general form of Harris' theorem with applications to stochastic delay equations, *Probab. Theory Related Fields*, **149** (2011), 223–259.

[16] Kabanov, Y., Pergamenshchikov, S., *Two-Scale Stochastic Systems*, Springer, Berlin, 2003.

[17] Kinnally, M. S., Williams, R. J., On existence and uniqueness of stationary distributions for stochastic delay differential equations with positivity constraints, *Electron. J. Probab.*, **15** (2010), 409–451.

[18] Khasminskii, R. Z., On an averaging principle for Itô stochastic differential equations, *Kibernetica*, **4** (1968), 260-279.

[19] Kuksin, S. B., Piatnitski, A. L., Khasminski-Whitman averaging for randonly perturbed KdV equations, *J Math Pures Appl.*, **89** (2008), 400–428.

[20] Kushner, H. J. Large deviations for two-time-scale diffusions, with delays, *Appl. Math. Optim.*, **62** (2010), 295–322.

[21] Kushner, H. J., *Weak Convergence Methods and Singularly Perturbed Stochastic Control and Filtering Problems*, Birkhäuser, Boston, MA, 1990.

[22] Liu, D., Strong convergence of principle of averaging for multiscale stochastic dynamical systems, *Commun. Math. Sci.*, **8** (2010), 999–1020.

[23] Liu, Y., Yin, G., Asymptotic expansions of transition densities for hybrid jump-diffusions, *Acta Math. Appl. Sin. Engl. Ser.*, **20** (2004), 1–18.

[24] Mao, X., *Stochastic differential equations and applications*, 2nd Ed., Horwood Publishing Limited, Chichester, 2008.

[25] Mohammed, S-E. A., *Stochastic Functional Differential Equations*, Pitman, Boston, 1984.

[26] Reiβ, M., Riedle, M., van Gaans, O., Delay differential equations driven by Lévy processes: stationarity and Feller properties, *Stochastic Process. Appl.*, **116** (2006), 1409–1432.

[27] Yin, G., Zhang, Q., *Continuous-time Markov Chians and Applications: A Two-Time-Scale Approach*, 2nd Ed., Springer, New York, NY, 2013.