SIAM J. CONTROL OPTIM.
Vol. 48, No. 5, pp. 3589–3622

CONTINUOUS-TIME STOCHASTIC AVERAGING ON THE INFINITE INTERVAL FOR LOCALLY LIPSCHITZ SYSTEMS*

SHU-JUN LIU† AND MIROSLAV KRSTIC‡

Abstract. We investigate stochastic averaging on the infinite time interval for a class of continuous-time nonlinear systems with stochastic perturbation and remove or weaken several restrictions present in existing results: global Lipschitzness of the nonlinear vector field, equilibrium preservation under the stochastic perturbation, global exponential stability of the average system, and compactness of the state space of the perturbation process. If an equilibrium of the average system is exponentially stable, we show that the original system is exponentially practically stable in probability. If, in addition, the original system has the same equilibrium as the average system, then the equilibrium of the original system is locally asymptotically stable in probability. These results extend the deterministic general averaging for aperiodic functions to the stochastic case.

Key words. stochastic averaging, stability in probability, stochastic differential equations

AMS subject classifications. 60H10, 93E15

DOI. 10.1137/090758970

1. Introduction. The basic idea of averaging theory—either deterministic or stochastic—is to approximate the original system (time-varying and periodic or almost periodic, or randomly perturbed) by a simpler (average) system (time-invariant, deterministic) or some approximating diffusion system (a stochastic system simpler than the original one). Starting with considerations driven by applications, the averaging principle has been developed in mechanics/dynamics [4, 28, 29, 32, 38] as well as in rigorous mathematical framework [3, 7, 8, 10, 11, 12, 31] for deterministic dynamics [4, 10, 29, 30] as well as stochastic dynamics [7, 12, 19, 37]. Stochastic averaging has been the cornerstone of many control and optimization methods, such as in stochastic approximation and adaptive algorithms [2, 20, 23, 33, 34]. Stochastic averaging is also a key tool in the newly emerging algorithms for stochastic extremum seeking and source localization [24, 35], which extend deterministic extremum seeking [1, 36].

Compared with mature theoretical results for the deterministic averaging principle, stochastic averaging offers a much broader spectrum of possibilities for developing averaging theorems (due to multiple notions of convergence and stability, as well as multiple possibilities for noise processes), which are far from being exhausted. On a finite time interval, in which case one does not study stability but only approximation accuracy, there have been many averaging theorems about weak convergence [7, 13, 21, 31], convergence in probability [7, 22], and almost sure convergence [8, 21]. However, the study of the stochastic averaging principle on the infinite time interval is not complete compared to complete results for the deterministic case [10, 30].

In general, the averaging principle on the infinite time interval is considered under the stability condition of average systems or diffusion approximation. The stability of stochastic systems with wide-band noise disturbances under diffusion approximation

*Received by the editors May 13, 2009; accepted for publication (in revised form) December 17, 2009; published electronically March 3, 2010.
http://www.siam.org/journals/sicon/48-5/75897.html
†Department of Mathematics, Southeast University, Nanjing 210096, China (sjliu@seu.edu.cn).
‡Department of Mechanical and Aerospace Engineering, University of California, San Diego, La Jolla, CA 92093-0411 (krstic@ucsd.edu).

conditions was stated by [3]. The stability of dynamic systems with Markov perturbations under the stability condition of the average system was studied in [14]. Under a condition on a diffusion approximation of a dynamical system with Markov perturbations, the problem of stability was solved in [15]. Under conditions of averaging and diffusion approximation, the stability of dynamic systems in a semi-Markov medium was studied in [16]. However, all these results are established under all or almost all of the following conditions: (a) the average system or approximating diffusion system is globally exponentially stable; (b) the nonlinear vector field of the original system has bounded derivative or is dominated by some form of Lyapunov function of the average system; (c) the nonlinear vector field of the original system vanishes at the origin for any value of the perturbation process (equilibrium condition); (d) the state space of the perturbation process is a compact space. These conditions largely limit the application of existing stochastic averaging theorems.

In this paper, we remove or weaken the above restrictions and develop stochastic averaging theorems for studying the stability of a general class of nonlinear systems with a stochastic perturbation. If the perturbation process satisfies a uniform strong ergodic condition and the equilibrium of the average system is exponentially stable, we show that the original system is exponentially practically stable in probability. Under the condition that the equilibrium of the average system is exponentially stable, if the perturbation process is ϕ-mixing with an exponential mixing rate and exponentially ergodic, and the original system satisfies an equilibrium condition, we show that the equilibrium of the original system is asymptotically stable in probability. For the case where the average system is globally exponentially stable and all the other assumptions are valid globally, a global result is obtained for the original system.

A reader familiar with the deterministic averaging theory should view our result as an extension to the stochastic case of the so-called general averaging for aperiodic functions (rather than of the standard averaging for periodic functions).

The rest of the paper is organized as follows. Section 2 describes the problem investigated. Section 3 presents results for two cases: a uniform strong ergodic perturbation process, and an exponentially ϕ-mixing and exponentially ergodic perturbation process, respectively. In section 4, we give the detailed proofs for the results in section 3. In section 5 we give three examples. Section 6 contains concluding remarks.

2. Problem formulation. Consider the system

$$\frac{dX_t^\epsilon}{dt} = a(X_t^\epsilon, Y_{t/\epsilon}), \quad X_0^\epsilon = x, \tag{2.1}$$

where $X_t^\epsilon \in \mathbb{R}^n$, and the stochastic perturbation $Y_t \in \mathbb{R}^m$ is a time homogeneous continuous Markov process defined on a complete probability space $(\Omega, \mathcal{F}, P)$, where Ω is the sample space, $\mathcal{F}$ is the σ-field, P is the probability measure, and ϵ is a small positive parameter, where $\epsilon \in (0, \epsilon_0)$ for some fixed $\epsilon_0 > 0$.

The average system corresponding to system (2.1) can be defined in various ways, depending on assumptions on the perturbation process $(Y_t,\, t \geq 0)$, for example, as

$$\frac{d\bar{X}_t}{dt} = \bar{a}(\bar{X}_t), \quad \bar{X}_0 = x, \tag{2.2}$$

where $\bar{a}(x)$ is a function such that for any $\delta > 0$ and $x \in \mathbb{R}^n$,

$$\lim_{T\to\infty} P\left\{\left|\frac{1}{T}\int_t^{t+T} a(x, Y_s)ds - \bar{a}(x)\right| > \delta\right\} = 0$$

uniformly in $t \geq 0$.

The assertion that the trajectory X_t^ϵ is close to $\bar{X}_t$ for sufficiently small ϵ is called the averaging principle [7]. For the system with random perturbation, Theorem 7.9.1 of [7] gives a clear result about the averaging principle when t is in a finite time interval $[0, T]$: for any $T > 0$ and $\delta > 0$

$$\lim_{\epsilon\to 0} P\left\{\sup_{0\le t\le T} |X_t^\epsilon - \bar{X}_t| > \delta\right\} = 0. \tag{2.3}$$

In this paper, we will explore the averaging principle when t belongs to the infinite time interval $[0, \infty)$. First, in the case where the original stochastic system may not have an equilibrium, but the average system has an exponentially stable equilibrium at the origin, a stability-like property of the original system is established for ϵ sufficiently small. Second, when $a(0, y) \equiv 0$, namely, when the original system (2.1) maintains an equilibrium at the origin, despite the presence of noise, we establish the stability of this equilibrium for sufficiently small ϵ.

3. Main results.

3.1. Uniform strong ergodic perturbation process. In the time scale $s = t/\epsilon$, define $Z_s^\epsilon = X_{\epsilon s}^\epsilon = X_t^\epsilon$, $Y_s = Y_{t/\epsilon}$. Then we transform system (2.1) into

$$\frac{dZ_s^\epsilon}{ds} = \epsilon\, a(Z_s^\epsilon, Y_s), \tag{3.1}$$

with the initial value $Z_0^\epsilon = x$. Let S_Y be the living space of the perturbation process $(Y_t,\, t \ge 0)$. Notice that S_Y may be a proper (e.g., compact) subset of $\mathbb{R}^m$.

Assumption 1. The vector field $a(x, y)$ is separable; i.e., it can be written as $a(x, y) = \sum_{i=1}^{l} a_i(x) b_i(y)$, where the functions $b_i : O_Y \to \mathbb{R}$, $i = 1, \dots, l$, are continuous (the set O_Y, which contains S_Y, is an open subset of $\mathbb{R}^n$) and bounded on S_Y; the functions $a_i : D \to \mathbb{R}^n$, $i = 1, \dots, l$, and their partial derivatives up to the second order are continuous on some domain (open connected set) $D \subset \mathbb{R}^n$.

Assumption 2. For $i = 1, \dots, l$, there exists a constant $\bar{b}_i$ such that

$$\lim_{T\to\infty} \frac{1}{T}\int_t^{t+T} b_i(Y_s)ds = \bar{b}_i \quad \text{a.s. uniformly in } t \in [0, \infty). \tag{3.2}$$

By Assumption 2 we obtain the average system of (3.1) as $d\bar{Z}_s^\epsilon/ds = \epsilon\bar{a}(\bar{Z}_s^\epsilon)$, with the initial value $\bar{Z}_0^\epsilon = x$, where $\bar{a}(x) = \sum_{i=1}^{l} a_i(x)\bar{b}_i$.

THEOREM 3.1. *Suppose that Assumptions* 1 *and* 2 *hold. If the origin* $\bar{Z}_s^\epsilon \equiv 0$ *is an exponentially stable equilibrium point of the average system,* $K \subset D$ *is a compact subset of its region of attraction, and* $\bar{Z}_0^\epsilon = x \in K$*, then for any* $\varsigma \in (0, 1)$*, there exist a measurable set* $\Omega_\varsigma \subset \Omega$ *with* $P(\Omega_\varsigma) > 1 - \varsigma$*, a class* $\mathcal{K}$ *function* α_ς*, and a constant* $\epsilon^*(\varsigma) > 0$ *such that if* $Z_0^\epsilon - \bar{Z}_0^\epsilon = O(\alpha_\varsigma)$*, then for all* $0 < \epsilon < \epsilon^*(\varsigma)$*,*

$$Z_s^\epsilon(\omega) - \bar{Z}_s^\epsilon = O(\alpha_\varsigma(\epsilon)) \quad \forall s \in [0, \infty)$$

uniformly in $\omega \in \Omega_\varsigma$*, which implies*

$$P\left\{\sup_{s\in[0,\infty)} \left|Z_s^\epsilon(\omega) - \bar{Z}_s^\epsilon\right| = O(\alpha_\varsigma(\epsilon))\right\} > 1 - \varsigma.$$

Next we extend the finite-time result (2.3) of [7, Theorem 7.9.1] to infinite time.

THEOREM 3.2. *Suppose that Assumptions* 1 *and* 2 *hold. If the origin* $\bar{Z}_s^\epsilon \equiv 0$ *is an exponentially stable equilibrium of the average system,* $K \subset D$ *is a compact subset of its region of attraction, and* $\bar{Z}_0^\epsilon = Z_0^\epsilon = x \in K$*, then for any* $\delta > 0$,

$$\lim_{\epsilon \to 0} P\left\{\sup_{s\in[0,\infty)} |Z_s^\epsilon(\omega) - \bar{Z}_s^\epsilon| > \delta\right\} = 0; \tag{3.3}$$

i.e., $\sup_{s\in[0,\infty)} |Z_s^\epsilon(\omega) - \bar{Z}_s^\epsilon|$ *converges to* 0 *in probability as* $\epsilon \to 0$.

The above two theorems are about systems in the time scale $s = t/\epsilon$. Now we turn to the X-system (2.1) and its average system (2.2), where $\bar{X}_t = \bar{Z}_{t/\epsilon}^\epsilon$, and $X_t^\epsilon = Z_{t/\epsilon}^\epsilon$. Theorems 3.1 and 3.2 yield the following corollaries.

COROLLARY 3.3. *If the origin* $\bar{X}_t = 0$ *is an exponentially stable equilibrium point of the average system* (2.2)*,* $K \subset D$ *is a compact subset of its region of attraction, and* $\bar{X}_0 = x \in K$*, then for any* $\varsigma \in (0,1)$*, there exist a class* $\mathcal{K}$ *function* α_ς *and a constant* $\epsilon^*(\varsigma) > 0$ *such that if* $X_0^\epsilon - \bar{X}_0 = O(\alpha_\varsigma)$*, then for all* $0 < \epsilon < \epsilon^*(\varsigma)$,

$$P\left\{\sup_{t\in[0,\infty)} \left|X_t^\epsilon(\omega) - \bar{X}_t\right| = O(\alpha_\varsigma(\epsilon))\right\} > 1 - \varsigma.$$

COROLLARY 3.4. *If the origin* $\bar{X}_t = 0$ *is an exponentially stable equilibrium point of the average system* (2.2)*,* $K \subset D$ *is a compact subset of its region of attraction, and* $X_0^\epsilon = \bar{X}_0 = x \in K$*, then for any* $\delta > 0$,

$$\lim_{\epsilon \to 0} P\left\{\sup_{t\in[0,\infty)} |X_t^\epsilon(\omega) - \bar{X}_t| > \delta\right\} = 0. \tag{3.4}$$

From Theorem 3.1 and the definition of exponential stability of deterministic systems, we obtain the following stability result.

THEOREM 3.5. *Suppose that Assumptions* 1 *and* 2 *hold. If the origin* $\bar{X}_t \equiv 0$ *is an exponentially stable equilibrium point of the average system* (2.2)*,* $K \subset D$ *is a compact subset of its region of attraction, and* $\bar{X}_0 = x \in K$*, then for any* $\varsigma \in (0,1)$*, there exist a measurable set* $\Omega_\varsigma \subset \Omega$ *with* $P(\Omega_\varsigma) > 1 - \varsigma$*, a class* $\mathcal{K}$ *function* α_ς*, and a constant* $\epsilon^*(\varsigma) > 0$ *such that if* $X_0^\epsilon - \bar{X}_0 = O(\alpha_\varsigma(\epsilon))$*, then for all* $0 < \epsilon < \epsilon^*(\varsigma)$,

$$|X_t^\epsilon(\omega)| \le c|x|e^{-\gamma t} + O(\alpha_\varsigma(\epsilon)) \quad \forall t \in [0,\infty) \tag{3.5}$$

uniformly in $\omega \in \Omega_\varsigma$ *for some constants* $\gamma, c > 0$.

Remark 3.6. Notice that for any given $\varsigma \in (0,1)$, $\alpha_\varsigma(\epsilon) \in \mathcal{K}$. Then by (3.5), we obtain that for any $\delta > 0$ and any $\varsigma > 0$, there exists a constant $\epsilon^*(\varsigma, \delta) > 0$ such that for all $0 < \epsilon < \epsilon^*(\varsigma, \delta)$,

$$P\left\{|X_t^\epsilon(\omega)| \le c|x|e^{-\gamma t} + \delta \ \forall t \in [0,\infty)\right\} > 1 - \varsigma \tag{3.6}$$

for $X_0^\epsilon = \bar{X}_0 = x \in K$ and some positive constants γ, c. This can be viewed as a form of exponential practical stability in probability.

Remark 3.7. Since Y_t is a time homogeneous continuous Markov process, if $a(x,y)$ is globally Lipschitz in (x,y), then the solution of (2.1) exists with probability 1 for any $x \in \mathbb{R}^n$ and it is defined uniquely for all $t \ge 0$ (see section 2 of Chapter 7 of [7]). Here, by Assumption 1, $a(x,y)$ is, in general, locally Lipschitz instead of globally Lipschitz. Notice that the solution of (2.1) can be defined for every trajectory of the stochastic process $(Y_s, s \geq 0)$. Then by Corollary 3.3, for any sufficiently small positive number ς, there exist a measurable set $\Omega_\varsigma \subset \Omega$ and a positive number $\epsilon^*(\varsigma)$ such that $P(\Omega_\varsigma) > 1-\varsigma$ (which can be sufficiently close to 1) and for any $0 < \epsilon < \epsilon^*(\varsigma)$ and any $\omega \in \Omega_\varsigma$, the solution $\{X_t^\epsilon(\omega), t \in [0,\infty)\}$ exists. The uniqueness of $\{X_t^\epsilon(\omega), t \in [0,\infty)\}$ is ensured by the local Lipschitzness of $a(x,y)$ with respect to x.

Remark 3.8. Assumptions 1 and 2 guarantee that there exists a deterministic vector function $\bar{a}(x)$ such that

$$\lim_{T\to\infty} \frac{1}{T} \int_t^{t+T} a(x, Y_s(\omega))ds = \bar{a}(x) \quad \text{a.s.} \tag{3.7}$$

uniformly in $(t,x) \in [0,\infty) \times D_0$ for any compact subset $D_0 \subset D$. This uniform convergence condition is critical in the proof, and a similar condition is required in the deterministic general averaging on the infinite time interval for aperiodic functions [10].

In a weak convergence method of stochastic averaging on a finite time interval, some uniform convergence with respect to (t,x) of some integral of $a(x,Y_s)$ is required [11, equation (3.2)], [7, equation (9.3), p. 263], and there the boundedness of $a(x,y)$ is assumed. Here we do not need the boundedness of $a(x,y)$ but need a stronger convergence (3.7) to obtain a better result—"exponential practical stability" on the infinite time interval.

The separable form in Assumption 1 is to guarantee that the limit (3.7) is uniform with respect to x, while the uniform convergence (3.2) in Assumption 2 is to guarantee that the limit (3.7) is uniform with respect to t. For the following stochastic processes $(Y_s,\ s \geq 0)$, we can verify that the uniform convergence (3.2) holds:

1. $dY_s = pY_s ds + qY_s dw_s,\ p < \frac{q^2}{2}$;
2. $dY_s = -pY_s ds + qe^{-s} dw_s,\ p, q > 0$;
3. $Y_s = e^{\xi_s} + c$, where c is a constant and ξ_s satisfies $d\xi_s = -ds + dw_s$.

In these three examples, w_s is a 1-dimensional standard Brownian motion defined on some complete probability space and Y_0 is independent of $(w_s,\ s \geq 0)$. In fact, for these three kinds of stochastic processes, it holds that $\lim_{s\to\infty} Y_s = c$ a.s. for some constant c, which, together with the fact that $\lim_{T\to\infty} \frac{1}{T}\int_t^{t+T} b_i(Y_s)ds = \lim_{s\to\infty} b_i(Y_s)$ a.s. when the latter limit exists, gives that for any continuous function b_i, $\lim_{T\to\infty} \frac{1}{T}\int_t^{t+T} b_i(Y_s)ds = \lim_{s\to\infty} b_i(Y_s) = b_i(c)$ a.s. uniformly in $t \in [0,\infty)$. If b_i has the form $b_i(y_1+y_2) = b_{i1}(y_1) + b_{i2}(y_2) + b_{i3}(y_1)b_{i4}(y_2)$ for any $y_1, y_2 \in S_Y$ and $b_{ij},\ j = 1,\dots,4$, are continuous functions, and $Y_s = \sin(s) + g(s)\sin(\xi_s)$, where $(\xi_s,\ s \geq 0)$ is any continuous stochastic process and $g(s)$ is a function decaying to zero, e.g., e^{-s}, $\frac{1}{1+s}$, then

$$\begin{aligned}
&\lim_{T\to\infty} \frac{1}{T} \int_t^{t+T} b_i(Y_s)ds \\
&= \lim_{T\to\infty} \frac{1}{T} \left\{ \int_t^{t+T} [b_{i1}(\sin(s)) + b_{i2}(g(s)\sin(\xi_s)) + b_{i3}(\sin(s))b_{i4}(g(s)\sin(\xi_s))]ds \right\} \\
&= \frac{1}{2\pi} \int_0^{2\pi} b_{i1}(\sin(s))ds + b_{i2}(0) + b_{i4}(0) \cdot \frac{1}{2\pi} \int_0^{2\pi} b_{i3}(\sin(s))ds \quad \text{a.s.}
\end{aligned}$$

uniformly in $t \in [0,\infty)$.

If the process $(Y_s,\ s \geq 0)$ is ergodic with invariant measure μ, then (cf., e.g., Theorem 3 on page 9 of [31])

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T b_i(Y_s)ds = \bar{b}_i \quad \text{a.s.,} \tag{3.8}$$

where $\bar{b}_i = \int_{S_Y} b_i(y)\mu(dy)$. While one might expect the averaging under condition (3.8) to be applicable on the infinite interval, this is not true. A stronger condition (3.2) on the perturbation process is needed (note the difference between the integration limits; that is the reason why we refer to this kind of perturbation process as "uniform strong ergodic"). Uniform convergence, as opposed to ergodicity, is essential for the averaging principle on the infinite time interval. The same requirement of uniformity in time is needed for general averaging on the infinite time in the deterministic case.

In sections 5.1 and 5.2 we give examples illustrating the theorems of this section.

3.2. ϕ-mixing perturbation process. Let $\mathcal{F}_t^s$ denote the smallest σ-algebra that measures $\{Y_u,\, t\le u\le s\}$. If there is a function $\phi(s)\to 0$ as $s\to\infty$ such that $\sup_{A\in\mathcal{F}_{t+s}^\infty,\, B\in\mathcal{F}_0^t}|P\{A|B\}-P\{B\}|\le\phi(s)$, then $(Y_u,\, u\ge 0)$ is said to be ϕ-mixing with mixing rate $\phi(\cdot)$ (see [18]).

In this subsection, we assume that the perturbation $(Y_t,\, t\ge 0)$ is ϕ-mixing and also ergodic with invariant measure μ. The average system of (2.1) is (2.2), where

$$\bar{a}(x) = \int_{S_Y} a(x,y)\mu(dy), \tag{3.9}$$

and S_Y is the living space of the perturbation process $(Y_t,\, t\ge 0)$.

Assumption 3. The process $(Y_t,\, t\ge 0)$ is continuous, ϕ-mixing with exponential mixing rate $\phi(t)$, and also exponentially ergodic with invariant measure μ.

Remark 3.9. (i) In the weak convergence methods (see, e.g., [18]), the perturbation process is usually assumed to be ϕ-mixing with mixing rate $\phi(t)$ ($\int_0^\infty \phi^{\frac{1}{2}}(s)ds<\infty$). Here we consider the infinite time horizon, so exponential ergodicity is needed.

(ii) According to [26], ergodic Markov processes on compact state space are examples of ϕ-mixing processes with an exponential mixing rate, e.g., Brownian motion on the unit circle [6] $(Y_t,\, t\ge 0)$: $dY_t = -\frac{1}{2}Y_t dt + BY_t dW_t$, $Y_0 = [\cos(\vartheta), \sin(\vartheta)]^T$, for all $\vartheta\in\mathbb{R}$, where $B = \left[\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right]$ and W_t is a 1-dimensional standard Brownian motion.

Assumption 4. For the average system (2.2), there exist a function $V(x)\in\mathbb{C}^2$ and positive constants c_i $(i=1,\ldots,4)$, δ, γ such that for $|x|\le\delta$,

$$c_1|x|^2 \le V(x) \le c_2|x|^2, \tag{3.10}$$

$$\left|\frac{\partial V(x)}{\partial x}\right| \le c_3|x|, \tag{3.11}$$

$$\left|\frac{\partial^2 V(x)}{\partial x^2}\right| \le c_4, \tag{3.12}$$

$$\frac{dV(x)}{dt} = \left(\frac{\partial V(x)}{\partial x}\right)^T \bar{a}(x) \le -\gamma V(x); \tag{3.13}$$

i.e., the average system (2.2) is exponentially stable.

Assumption 5. The vector field $a(x,y)$ satisfies the following:

1. $a(x,y)$ and its first-order partial derivatives with respect to x are continuous and $a(0,y) \equiv 0$;
2. for any compact set $D \subset \mathbb{R}^n$, there is a constant $k_D > 0$ such that for all $x \in D$ and $y \in S_Y$, $\left|\frac{\partial a(x,y)}{\partial x}\right| \le k_D$.

THEOREM 3.10. *Consider the system* (2.1) *satisfying Assumptions* 3, 4, *and* 5. *Then there exists* $\epsilon^* > 0$ *such that for* $0 < \epsilon \le \epsilon^*$, *the solution* $X_t^\epsilon \equiv 0$ *of the original system is asymptotically stable in probability; i.e., for any* $r > 0$ *and* $\varsigma > 0$, *there is a constant* $\delta_0 > 0$ *such that if* $|X_0^\epsilon| = |x| < \delta_0$, *then*

$$P\left\{\sup_{t\ge 0}|X_t^\epsilon| \le r\right\} \ge 1-\varsigma, \tag{3.14}$$

$$\lim_{x\to 0} P\left\{\lim_{t\to\infty}|X_t^\epsilon| = 0\right\} = 1. \tag{3.15}$$

Remark 3.11. This is the first local stability result based on the stochastic averaging approach for locally Lipschitz nonlinear systems, which is an extension from the deterministic general averaging for aperiodic functions [30].

If the local conditions in Theorem 3.10 hold globally, we get global results under the following set of assumptions.

Assumption 6. The average system (2.2) is globally exponentially stable; i.e., Assumption 4 holds with "for $|x| \le \delta$" replaced by "for any $x \in \mathbb{R}^n$."

Assumption 7. The vector field $a(x,y)$ satisfies the following:

1. $a(x,y)$ and its first-order partial derivatives with respect to x are continuous and $a(0,y) \equiv 0$;
2. there is a constant $k > 0$ such that for all $x \in \mathbb{R}^n$ and $y \in S_Y$, $\left|\frac{\partial a(x,y)}{\partial x}\right| \le k$.

Assumption 8. The vector field $a(x,y)$ satisfies the following:

1. $a(x,y)$ and its first-order partial derivatives with respect to x are continuous and $\sup_{y\in S_Y}|a(0,y)| < \infty$;
2. there is a constant $k > 0$ such that for all $x \in \mathbb{R}^n$ and $y \in S_Y$, $\left|\frac{\partial a(x,y)}{\partial x}\right| \le k$.

THEOREM 3.12. *Consider the system* (2.1) *satisfying Assumptions* 3, 6, *and* 7. *Then there exists* $\epsilon^* > 0$ *such that for* $0 < \epsilon \le \epsilon^*$, *the solution* $X_t^\epsilon \equiv 0$ *of the original system is globally asymptotically stable in probability; i.e., for any* $\eta_1 > 0$ *and* $\eta_2 > 0$, *there is a constant* $\delta_0 > 0$ *such that if* $|X_0^\epsilon| = |x| < \delta_0$, *then* $P\left\{|X_t^\epsilon| \le \eta_2 e^{-\tilde{\gamma}t},\ t \ge 0\right\} \ge 1-\eta_1$ *with a constant* $\tilde{\gamma} > 0$, *and, moreover, for any* $x \in \mathbb{R}^n$, $P\{\lim_{t\to\infty}|X_t^\epsilon| = 0\} = 1$.

If, on the other hand, (2.1) has no equilibrium, we obtain the following result.

THEOREM 3.13. *Consider the system* (2.1) *satisfying Assumptions* 3, 6, *and* 8. *Then there exists* $\epsilon^* > 0$ *such that for* $0 < \epsilon \le \epsilon^*$, *the solution process* X_t^ϵ *of the original system is bounded in probability, i.e.,* $\lim_{r\to\infty}\sup_{t\ge 0} P\{|X_t^\epsilon| > r\} = 0$.

Remark 3.14. Theorems 3.12 and 3.13 are aimed at globally Lipschitz systems and can be viewed as an extension from the deterministic averaging principle [30] to the stochastic case. We present the results for the global case not only for the sake of completeness but also because of the novelty relative to [3]: (i) an ergodic Markov process on some compact space is replaced by an exponential ϕ-mixing and exponentially ergodic process; (ii) for the case without equilibrium condition the weak convergence is considered in [3], while here we obtain the result on boundedness in probability.

In section 5.3 we present an example that illustrates the theorems of this section.

4. Proofs of the results.

4.1. Proofs for the case of uniform strong ergodic perturbation process.

4.1.1. Technical lemma. To prove Theorems 3.1 and 3.2, we first prove one technical lemma. Towards that end, denote $F_i(T,\lambda,\omega) = \frac{1}{T}\int_\lambda^{\lambda+T} b_i(Y_u(\omega))du$ for $T > 0$, $\lambda \geq 0$, $\omega \in \Omega$, $i = 1,\dots,l$. We can verify that $F_i(T,\lambda,\omega)$ is continuous with respect to (T,λ) for any $i = 1,\dots,l$.

LEMMA 4.1. *Suppose that Assumptions* 1 *and* 2 *hold. Then, for any* $\varsigma > 0$, *there exists a measurable set* $\Omega_\varsigma \subset \Omega$ *such that* $P(\Omega_\varsigma) > 1-\varsigma$, *and for any* $i = 1,\dots,l$,

$$\lim_{T\to\infty} \frac{1}{T}\int_\lambda^{\lambda+T} b_i(Y_u(\omega))du = \bar{b}_i \textit{ uniformly} \quad \textit{in } (\omega,\lambda) \in \Omega_\varsigma \times [0,\infty). \tag{4.1}$$

Moreover, there exists a strictly decreasing, continuous, bounded function $\sigma^\varsigma(T)$ *such that* $\sigma^\varsigma(T) \to 0$ *as* $T \to \infty$, *and for any compact subset* $D_0 \subset D$,

$$\left|\frac{1}{T}\int_\lambda^{\lambda+T} a(x,Y_u(\omega))du - \bar{a}(x)\right| \leq k_{D_0}\sigma^\varsigma(T) \quad \forall(\omega,\lambda,x) \in \Omega_\varsigma \times [0,\infty) \times D_0, \tag{4.2}$$

where k_{D_0} *is a positive constant.*

Proof. Step 1 *(proof of* (4.1)*).* From (3.2) we know that for any $i = 1,\dots,l$,

$$\text{for a.e. } \omega \in \Omega, \quad \lim_{T\to\infty} F_i(T,\lambda,\omega) = \bar{b}_i \text{ uniformly} \quad \text{in } \lambda \geq 0. \tag{4.3}$$

Noticing that $\{\omega \mid \lim_{T\to\infty} F_i(T,\lambda,\omega) = \bar{b}_i \text{ uniformly in } \lambda \geq 0\} = \bigcap_{k=1}^\infty \bigcup_{t>0} \bigcap_{T\geq t}\bigcap_{\lambda\geq 0}\left\{|F_i(T,\lambda,\omega) - \bar{b}_i| < \frac{1}{k}\right\}$, by (4.3), we get that

$$P\left(\bigcup_{k=1}^\infty \bigcap_{t>0}\bigcup_{T\geq t}\bigcup_{\lambda\geq 0}\left\{|F_i(T,\lambda,\omega) - \bar{b}_i| \geq \frac{1}{k}\right\}\right) = 0. \tag{4.4}$$

Since $F_i(T,\lambda,\omega)$ is continuous with respect to (T,λ), we can easily prove that for all $k \geq 1$, for all $t > 0$, the sets $\bigcup_{\lambda\geq 0}\left\{|F_i(T,\lambda,\omega) - \bar{b}_i| \geq \frac{1}{k}\right\}$, $\bigcup_{T\geq t}\bigcup_{\lambda\geq 0}\left\{|F_i(T,\lambda,\omega) - \bar{b}_i| \geq \frac{1}{k}\right\}$, and $\bigcap_{t>0}\bigcup_{T\geq t}\bigcup_{\lambda\geq 0}\left\{|F_i(T,\lambda,\omega) - \bar{b}_i| \geq \frac{1}{k}\right\}$ are measurable. Then by (4.4) we obtain that for any $k \geq 1$,

$$P\left(\bigcap_{t>0}\bigcup_{T\geq t}\bigcup_{\lambda\geq 0}\left\{|F_i(T,\lambda,\omega) - \bar{b}_i| \geq \frac{1}{k}\right\}\right) = 0. \tag{4.5}$$

Since the set $\bigcup_{T\geq t}\bigcup_{\lambda\geq 0}\left\{|F_i(T,\lambda,\omega) - \bar{b}_i| \geq \frac{1}{k}\right\}$ is decreasing as t increases, it follows from (4.5) that $\lim_{t\to\infty} P\left(\bigcup_{T\geq t}\bigcup_{\lambda\geq 0}\left\{|F_i(T,\lambda,\omega) - \bar{b}_i| \geq \frac{1}{k}\right\}\right) = 0$. Thus, for any $\varsigma > 0$ and any $k \geq 1$, there exists $t_k^{(i)} > 0$ such that

$$P\left(\bigcup_{T\geq t_k^{(i)}}\bigcup_{\lambda\geq 0}\left\{|F_i(T,\lambda,\omega) - \bar{b}_i| \geq \frac{1}{k}\right\}\right) < \frac{\varsigma}{2^k l}. \tag{4.6}$$

Define $\Omega_\varsigma = \bigcap_{i=1}^l \bigcap_{k=1}^\infty \bigcap_{T\geq t_k^{(i)}}\bigcap_{\lambda\geq 0}\left\{|F_i(T,\lambda,\omega) - \bar{b}_i| < \frac{1}{k}\right\}$. Then by (4.6), $P(\Omega_\varsigma) \geq 1-\varsigma$. Further, by the construction of Ω_ς, we know that for any $i = 1,\dots,l$,

$$\lim_{T\to\infty} \frac{1}{T}\int_\lambda^{\lambda+T} b_i(Y_u(\omega))du = \bar{b}_i \text{ uniformly} \quad \text{in } (\omega,\lambda) \in \Omega_\varsigma \times [0,\infty); \tag{4.7}$$

i.e., (4.1) holds.

Step 2 *(proof of* (4.2)*).* By (4.7), for any $k \geq 1$, there exists $t_k(\varsigma) > 0$ (without loss of generality, we can assume that $t_k(\varsigma)$ is increasing with respect to k) such that for any $T \geq t_k(\varsigma)$, any $(\omega, \lambda) \in \Omega_\varsigma \times [0, \infty)$, and any $i = 1, \ldots, l$, we have that

$$\left| \frac{1}{T} \int_\lambda^{\lambda+T} b_i(Y_u(\omega))du - \bar{b}_i \right| < \frac{1}{k}. \tag{4.8}$$

By Assumption 1 and (3.2), there exists a constant $M > 1$ such that for any $i = 1, \ldots, l$, $\sup_{y \in S_Y} |b_i(y)| \leq M$ and $|\bar{b}_i| \leq M$. Now we define a function $H^\varsigma(T)$ as

$$H^\varsigma(T) = \begin{cases} 2M & \text{if} \quad T \in [0, t_1(\varsigma)); \\ \frac{1}{k} & \text{if} \quad T \in [t_k(\varsigma), t_{k+1}(\varsigma)), \quad k = 1, 2, \ldots. \end{cases}$$

Then by (4.8), for any $(\omega, \lambda) \in \Omega_\varsigma \times [0, \infty)$, and any $i = 1, \ldots, l$, we have

$$\left| \frac{1}{T} \int_\lambda^{\lambda+T} b_i(Y_u(\omega))du - \bar{b}_i \right| \leq H^\varsigma(T), \tag{4.9}$$

and $H^\varsigma(T) \downarrow 0$ as $T \to \infty$. Noticing that the function $H^\varsigma(T)$ is a piecewise constant (and thus piecewise continuous) function, we construct a strictly decreasing, continuous, bounded function $\sigma^\varsigma(T)$:

$$\sigma^\varsigma(T) = \begin{cases} -\dfrac{1}{t_1(\varsigma)} T + (2M+1) & \text{if} \quad T \in [0, t_1(\varsigma)); \\ -\dfrac{2M-1}{t_2(\varsigma) - t_1(\varsigma)} (T - t_1(\varsigma)) + 2M & \text{if} \quad T \in [t_1(\varsigma), t_2(\varsigma)); \\ -\dfrac{\frac{1}{k-1} - \frac{1}{k}}{t_{k+1}(\varsigma) - t_k(\varsigma)} (T - t_k(\varsigma)) + \dfrac{1}{k-1} & \text{if} \quad \begin{array}{l} T \in [t_k(\varsigma), t_{k+1}(\varsigma)), \\ k = 2, 3, \ldots, \end{array} \end{cases}$$

which satisfies $\sigma^\varsigma(T) \geq H^\varsigma(T)$, for all $T \geq 0$, and $\sigma^\varsigma(T) \downarrow 0$ as $T \to \infty$.

For any compact set $D_0 \subset D$, by Assumption 1, there exists a positive constant $M_{D_0} > 0$ such that for any $i = 1, \ldots, l$,

$$|a_i(x)| \leq M_{D_0} \quad \forall x \in D_0. \tag{4.10}$$

Define $k_{D_0} = l M_{D_0}$. Then, by Assumption 1, (4.9), (4.10), and the fact that $\bar{a}(x) = \sum_{i=1}^l a_i(x) \bar{b}_i$ and $\sigma^\varsigma(T) \geq H^\varsigma(T)$, for all $T \geq 0$, we get that for all $(\omega, \lambda, x) \in \Omega_\varsigma \times [0, \infty) \times D_0$,

$$\begin{aligned} \left| \frac{1}{T} \int_\lambda^{\lambda+T} a(x, Y_u(\omega))du - \bar{a}(x) \right| &= \left| \sum_{i=1}^l a_i(x) \left(\frac{1}{T} \int_\lambda^{\lambda+T} b_i(Y_u(\omega))du - \bar{b}_i \right) \right| \\ &\leq \sum_{i=1}^l |a_i(x)| \left| \frac{1}{T} \int_\lambda^{\lambda+T} b_i(Y_u(\omega))du - \bar{b}_i \right| \leq k_{D_0} \sigma^\varsigma(T); \end{aligned} \tag{4.11}$$

i.e., (4.2) holds. ◻

4.1.2. Proof of Theorem 3.1. The basic idea of the proof comes from [10, section 10.6]. Fix ς and Ω_ς as in Lemma 4.1. For any $\omega \in \Omega_\varsigma$, define $\hat{a}(s,x,\omega) = a(x, Y_s(\omega))$. Then we simply rewrite the system (3.1) as

$$\frac{dz}{ds} = \epsilon\, \hat{a}(s,z,\omega). \tag{4.12}$$

Let

$$h(s,z,\omega) = \hat{a}(s,z,\omega) - \bar{a}(z), \tag{4.13}$$

$$w(s,z,\omega,\eta) = \int_0^s h(\tau,z,\omega)\exp[-\eta(s-\tau)]d\tau \tag{4.14}$$

for some $\eta > 0$. For any compact set $D_0 \subset D$, by (4.11), we get that for $z \in D_0$,

$$\begin{aligned} |w(s+\delta,z,\omega,0) - w(s,z,\omega,0)| &= \left| \int_0^{s+\delta} h(\tau,z,\omega)d\tau - \int_0^s h(\tau,z,\omega)d\tau \right| \\ &= \left| \int_s^{s+\delta} h(\tau,z,\omega)d\tau \right| \le k_{D_0}\delta\, \sigma^\varsigma(\delta). \end{aligned} \tag{4.15}$$

This implies, in particular, that

$$|w(s,z,\omega,0)| \le k_{D_0}\, s\, \sigma^\varsigma(s) \quad \forall (s,z) \in (0,\infty) \times D_0, \tag{4.16}$$

since $w(0,z,\omega,0) = 0$. Integrating the right-hand side of (4.14) by parts, we obtain

$$\begin{aligned} &w(s,z,\omega,\eta) \\ &\quad = w(s,z,\omega,0) - \eta \int_0^s \exp[-\eta(s-\tau)]w(\tau,z,\omega,0)d\tau \\ &\quad = \exp(-\eta s)w(s,z,\omega,0) - \eta \int_0^s \exp[-\eta(s-\tau)][w(\tau,z,\omega,0) - w(s,z,\omega,0)]d\tau, \end{aligned}$$

where the second equality is obtained by adding and subtracting $\eta \int_0^s \exp[-\eta(s-\tau)]d\tau\, w(s,z,\omega,0)$ to and from the right-hand side. Using (4.15) and (4.16), we obtain that

$$|w(s,z,\omega,\eta)| \le k_{D_0}\, s\exp(-\eta s)\, \sigma^\varsigma(s) + k_{D_0}\, \eta \int_0^s \exp[-\eta(s-\tau)](s-\tau)\, \sigma^\varsigma(s-\tau)\, d\tau. \tag{4.17}$$

For (4.17), we now show that there is a class $\mathcal{K}$ function α_ς such that

$$\eta|w(s,z,\omega,\eta)| \le k_{D_0}\, \alpha_\varsigma(\eta) \quad \forall (s,z,\omega) \in [0,\infty) \times D_0 \times \Omega_\varsigma. \tag{4.18}$$

Let $z \in D_0$. First, for $s \le \frac{1}{\sqrt{\eta}}$, by (4.17) and the property of the function σ^ς,

$$\begin{aligned} &\eta|w(s,z,\omega,\eta)| \\ &\le k_{D_0}\left(\eta s e^{-\eta s}\, \sigma^\varsigma(s) + \eta^2 \int_0^s \exp[-\eta(s-\tau)](s-\tau)\, \sigma^\varsigma(s-\tau)d\tau\right) \\ &= k_{D_0}\left(\eta s e^{-\eta s}\, \sigma^\varsigma(s) + \eta^2 \int_0^s \exp(-\eta u)u\, \sigma^\varsigma(u)du\right) \\ &\le k_{D_0}\left(\sqrt{\eta}\sigma^\varsigma(0) + \sqrt{\eta}\left(1 - e^{-\sqrt{\eta}}\right)\sigma^\varsigma(0)\right) \le k_{D_0}\,(2\sqrt{\eta}\sigma^\varsigma(0)). \end{aligned} \tag{4.19}$$

Then, for $s \geq \frac{1}{\sqrt{\eta}}$, by (4.17), (4.19), and the property of the function σ^ς, we obtain

$$\begin{aligned}
\eta|w(s,z,\omega,\eta)| \\
&\leq k_{D_0}\left\{\eta s e^{-\eta s}\,\sigma^\varsigma(s) + \eta^2\int_0^s \exp[-\eta(s-\tau)](s-\tau)\,\sigma^\varsigma(s-\tau)d\tau\right\} \\
&= k_{D_0}\left\{\eta s e^{-\eta s}\,\sigma^\varsigma(s) + \eta^2\int_0^s \exp(-\eta u)u\,\sigma^\varsigma(u)du\right\} \\
&= k_{D_0}\left\{\eta s e^{-\eta s}\,\sigma^\varsigma(s) + \eta^2\left[\int_0^{\frac{1}{\sqrt{\eta}}} \exp(-\eta u)u\,\sigma^\varsigma(u)du \right.\right. \\
&\qquad\qquad \left.\left. + \int_{\frac{1}{\sqrt{\eta}}}^s \exp(-\eta u)u\,\sigma^\varsigma(u)du\right]\right\} \\
&\leq k_{D_0}\left(\sqrt{\eta}\,\sigma^\varsigma(0) + \sigma^\varsigma\left(\frac{1}{\sqrt{\eta}}\right)\right).
\end{aligned} \tag{4.20}$$

Thus we define

$$\alpha_\varsigma(\eta) = \begin{cases} 2\sqrt{\eta}\sigma^\varsigma(0) + \sigma^\varsigma\left(\frac{1}{\sqrt{\eta}}\right) & \text{if } \eta > 0; \\ 0 & \text{if } \eta = 0. \end{cases}$$

Then $\alpha_\varsigma(\eta)$ is a class $\mathcal{K}$ function of η, and for $\eta \in [0,1]$, $\alpha_\varsigma(\eta) \geq 2\sigma^\varsigma(0)\eta$. By (4.19) and (4.20), we obtain that for any $\eta \geq 0$, (4.18) holds.

The partial derivatives $\frac{\partial w}{\partial s}$ and $\frac{\partial w}{\partial z}$ are given by

$$\begin{aligned}
\frac{\partial w(s,z,\omega,\eta)}{\partial s} &= h(s,z,\omega) - \eta w(s,z,\omega,\eta), \\
\frac{\partial w(s,z,\omega,\eta)}{\partial z} &= \int_0^s \frac{\partial h}{\partial z}(\tau,z,\omega)\exp[-\eta(s-\tau)]d\tau.
\end{aligned} \tag{4.21}$$

Noticing that $\frac{\partial \bar{a}(x)}{\partial x} = \sum_{i=1}^l \frac{\partial a_i(x)}{\partial x}\bar{b}_i = \sum_{i=1}^l \frac{\partial a_i(x)}{\partial x}\lim_{T\to\infty}\int_t^{t+T} b_i(Y_s)ds = \lim_{T\to\infty}\int_t^{t+T}\frac{\partial a(x,Y_s)}{\partial x}ds$ a.s., we can build results similar to (4.1) and (4.2) in Lemma 4.1 for $\left(\frac{\partial a(x,y)}{\partial x}, \frac{\partial \bar{a}(x)}{\partial x}\right)$ instead of $(a(x,y),\bar{a}(x))$. Furthermore, for $\varsigma > 0$, we can take the same measurable set $\Omega_\varsigma \subset \Omega$. Hence, for $\frac{\partial \hat{a}(s,z,\omega)}{\partial z} = \frac{\partial a(z,Y_s(\omega))}{\partial z}$, we can obtain the same property (4.11) as $\hat{a}(s,z,\omega) = a(z,Y_s(\omega))$. Consequently, $\frac{\partial h}{\partial z}(s,z,\omega) = \frac{\partial \hat{a}}{\partial z}(s,z,\omega) - \frac{\partial \bar{a}}{\partial z}(z)$ possesses the same properties as $h(s,z,\omega)$. Thus we can repeat the above derivations to obtain that (4.18) also holds for $\frac{\partial w}{\partial z}$, i.e.,

$$\eta\left|\frac{\partial w}{\partial z}(s,z,\omega,\eta)\right| \leq k_{D_0}\,\alpha_\varsigma(\eta) \quad \forall (s,z,\omega) \in [0,\infty)\times D_0\times\Omega_\varsigma. \tag{4.22}$$

There is no loss of generality in using the same positive constant k_{D_0} in both (4.18) and (4.22). Since $k_{D_0} = l\,M_{D_0}$ will differ only in the bound M_{D_0} in (4.10), we can define M_{D_0} by using the larger of the two constants.

Define the change of variable

$$z = \zeta + \epsilon\, w(s,\zeta,\omega,\epsilon), \tag{4.23}$$

where $\epsilon w(s,\zeta,\omega,\epsilon)$ is of order$O(\alpha_\varsigma(\epsilon))$ by (4.18). By (4.22), for sufficiently small ϵ, the matrix $\left[I+\epsilon\frac{\partial w}{\partial\zeta}\right]$ is nonsingular. Differentiating both sides with respect to s, we obtain $\frac{dz}{ds}=\frac{d\zeta}{ds}+\epsilon\frac{\partial w(s,\zeta,\omega,\epsilon)}{\partial s}+\epsilon\frac{\partial w(s,\zeta,\omega,\epsilon)}{\partial\zeta}\frac{d\zeta}{ds}$. Substituting for $\frac{dz}{ds}$ from (4.12), by (4.23), (4.21), and (4.13), we find that the new state variable ζ satisfies the equation

$$\begin{aligned}(4.24)\qquad \left[I+\epsilon\frac{\partial w}{\partial\zeta}\right]\frac{d\zeta}{ds}&=\epsilon\,\hat{a}(s,\zeta+\epsilon w,\omega)-\epsilon\,\frac{\partial w(s,\zeta,\omega,\epsilon)}{\partial s}\\&=\epsilon\hat{a}(s,\zeta+\epsilon w,\omega)-\epsilon[\hat{a}(s,\zeta,\omega)-\bar{a}(\zeta)]+\epsilon^2 w(s,\zeta,\omega,\epsilon)\\&=\epsilon\bar{a}(\zeta)+p(s,\zeta,\omega,\epsilon),\end{aligned}$$

where $p(s,\zeta,\omega,\epsilon)=\epsilon\,[\hat{a}(s,\zeta+\epsilon w,\omega)-\hat{a}(s,\zeta,\omega)]+\epsilon^2 w(s,\zeta,\omega,\epsilon)$. Using the mean value theorem, there exists a function f such that $p(s,\zeta,\omega,\epsilon)$ is expressed as

$$\begin{aligned}(4.25)\qquad p(s,\zeta,\omega,\epsilon)&=\epsilon^2 f(s,\zeta,\epsilon w,\omega)w(s,\zeta,\omega,\epsilon)+\epsilon^2 w(s,\zeta,\omega,\epsilon)\\&=\epsilon^2[f(s,\zeta,\epsilon w,\omega)+1]\,w(s,\zeta,\omega,\epsilon).\end{aligned}$$

Notice that $\left[I+\epsilon\frac{\partial w}{\partial\zeta}\right]^{-1}=I+O(\alpha_\varsigma(\epsilon))$, and $\alpha_\varsigma(\epsilon)\geq 2\sigma^\varsigma(0)\,\epsilon$ for $\epsilon\in[0,1]$. Then by (4.24) and (4.25), the state equation for ζ is given by

$$\begin{aligned}(4.26)\qquad \frac{d\zeta}{ds}&=[I+O(\alpha_\varsigma(\epsilon))]\times\left[\epsilon\bar{a}(\zeta)+\epsilon^2(f(s,\zeta,\epsilon w,\omega)+1)\,w(s,\zeta,\omega,\epsilon)\right]\\&\triangleq\epsilon\bar{a}(\zeta)+\epsilon\alpha_\varsigma(\epsilon)q(s,\zeta,\omega,\epsilon),\end{aligned}$$

where $q(s,\zeta,\omega,\epsilon)$ is uniformly bounded on $[0,\infty)\times D_0\times\Omega_\varsigma$ for sufficiently small ϵ. The system (4.26) is a perturbation of the average system $d\zeta/ds=\epsilon\,\bar{a}(\zeta)$. Notice that for any compact set $D_0\subset D$, $q(s,\zeta,\omega,\epsilon)$ is uniformly bounded on $[0,\infty)\times D_0\times\Omega_\varsigma$ for sufficiently small ϵ. Then by the definition of Ω_ς and the averaging principle of deterministic systems (see Theorems 10.5 and 9.1 of [10]), we obtain the result of Theorem 3.1. The proof is completed.

4.1.3. Proof of Theorem 3.2. For any $\varsigma>0$, by Theorem 3.1, there exist a measurable set $\Omega_\varsigma\subset\Omega$ with $P(\Omega_\varsigma)>1-\varsigma$, a class $\mathcal{K}$ function α_ς, and a constant $\epsilon^*(\varsigma)>0$ such that for all $0<\epsilon<\epsilon^*(\varsigma)$, $\sup_{s\in[0,\infty)}\left|Z_s^\epsilon(\omega)-\bar{Z}_s^\epsilon\right|=O(\alpha_\varsigma(\epsilon))$ uniformly in $\omega\in\Omega_\varsigma$. So there exists a positive constant $C_\varsigma>0$ such that for any $\omega\in\Omega_\varsigma$ and any $0<\epsilon<\epsilon^*(\varsigma)$,

$$\sup_{s\in[0,\infty)}\left|Z_s^\epsilon(\omega)-\bar{Z}_s^\epsilon\right|\leq C_\varsigma\cdot\alpha_\varsigma(\epsilon).$$

Since $\alpha_\varsigma(\epsilon)$ is continuous and $\alpha_\varsigma(0)=0$, for any $\delta>0$, there exists an $\epsilon'(\varsigma)>0$ such that for any $0<\epsilon<\epsilon'(\varsigma)$, $C_\varsigma\cdot\alpha_\varsigma(\epsilon)<\delta$. Denote $\bar{\epsilon}(\varsigma)=\min\{\epsilon^*(\varsigma),\epsilon'(\varsigma)\}$. Then for any $\omega\in\Omega_\varsigma$ and any $0<\epsilon<\bar{\epsilon}(\varsigma)$, it holds that

$$\sup_{s\in[0,\infty)}\left|Z_s^\epsilon(\omega)-\bar{Z}_s^\epsilon\right|<\delta,$$

which means that $\left\{\sup_{s\in[0,\infty)}\left|Z_s^\epsilon(\omega)-\bar{Z}_s^\epsilon\right|>\delta\right\}\subset(\Omega\backslash\Omega_\varsigma)$. Thus, we obtain that for any $0<\epsilon<\bar{\epsilon}(\varsigma)$, $P\left\{\sup_{s\in[0,\infty)}|Z_s^\epsilon-\bar{Z}_s^\epsilon|>\delta\right\}\leq P(\Omega\backslash\Omega_\varsigma)<\varsigma$. Hence the limit (3.3) holds. The proof is completed.

4.2. Proofs for the case of ϕ-mixing perturbation process.

4.2.1. Proof of Theorem 3.10. Throughout this part, we suppose that the initial value $X_0^\epsilon = x$ satisfies $|x| < \delta$ (δ is stated in Assumption 4). Define $D_\delta = \{x' \in \mathbb{R}^n : |x'| \le \delta\}$. For any $\epsilon > 0$ and $t \ge 0$, define two stopping times τ_δ^ϵ and $\tau_\delta^\epsilon(t)$ by

$$\tau_\delta^\epsilon = \inf\{s \ge 0 : X_s^\epsilon \notin D_\delta\} = \inf\{s \ge 0 : |X_s^\epsilon| > \delta\} \quad \text{and} \quad \tau_\delta^\epsilon(t) = \tau_\delta^\epsilon \wedge t. \tag{4.27}$$

Hereafter, we make the convention that $\inf \emptyset = \infty$.

Define the truncated processes $X_t^{\epsilon,\delta}$ by

$$X_t^{\epsilon,\delta} = X_{t\wedge\tau_\delta^\epsilon}^\epsilon = X_{\tau_\delta^\epsilon(t)}^\epsilon, \quad t \ge 0. \tag{4.28}$$

Then for any $t \ge 0$, we have that $X_t^{\epsilon,\delta} = x + \int_0^{\tau_\delta^\epsilon(t)} a(X_s^\epsilon, Y_{s/\epsilon})ds$. For any $t \ge 0$, define a σ-field $\mathcal{F}_t^{\epsilon,\delta}$ as follows:

$$\mathcal{F}_t^{\epsilon,\delta} = \sigma\left\{X_s^{\epsilon,\delta}, Y_{s/\epsilon} : 0 \le s \le t\right\} = \sigma\left\{Y_{s/\epsilon} : 0 \le s \le t\right\} \triangleq \mathcal{F}_{t/\epsilon}^Y. \tag{4.29}$$

Since $\mathcal{F}_t^{\epsilon,\delta} = \mathcal{F}_{t/\epsilon}^Y$ is independent of δ, for simplicity, throughout the rest part of this paper we use $\mathcal{F}_t^\epsilon$ instead of $\mathcal{F}_t^{\epsilon,\delta}$.

Step 1 *(Lyapunov estimates for Theorem* 3.10*).* For any $x \in \mathbb{R}^n$ with $|x| \le \delta$, and $t \ge 0$, define $V^\epsilon(x,t)$ by

$$V^\epsilon(x,t) = V(x) + V_1^\epsilon(x,t), \tag{4.30}$$

where

$$\begin{aligned} V_1^\epsilon(x,t) &= \int_{\tau_\delta^\epsilon(t)}^{\tau_\delta^\epsilon} \left(\frac{\partial V(x)}{\partial x}\right)^T E\left[a(x,Y_{s/\epsilon}) - \bar{a}(x)|\mathcal{F}_t^\epsilon\right] ds \\ &= \epsilon \int_{\frac{\tau_\delta^\epsilon(t)}{\epsilon}}^{\frac{\tau_\delta^\epsilon}{\epsilon}} \left(\frac{\partial V(x)}{\partial x}\right)^T E\left[a(x,Y_u) - \bar{a}(x)|\mathcal{F}_t^\epsilon\right] du \\ &= \epsilon \int_{\frac{\tau_\delta^\epsilon(t)}{\epsilon}}^{\frac{\tau_\delta^\epsilon}{\epsilon}} \left(\frac{\partial V(x)}{\partial x}\right)^T \Bigg[E[a(x,Y_u)|\mathcal{F}_t^\epsilon] \\ &\qquad - \int_{S_Y} a(x,y)[P_u(dy) - P_u(dy) + \mu(dy)]\Bigg] du \\ &= \epsilon \int_{\frac{\tau_\delta^\epsilon(t)}{\epsilon}}^{\frac{\tau_\delta^\epsilon}{\epsilon}} \left(\frac{\partial V(x)}{\partial x}\right)^T (E[a(x,Y_u)|\mathcal{F}_t^\epsilon] - E[a(x,Y_u)])\, du \\ &\qquad + \epsilon \int_{\frac{\tau_\delta^\epsilon(t)}{\epsilon}}^{\frac{\tau_\delta^\epsilon}{\epsilon}} \left(\frac{\partial V(x)}{\partial x}\right)^T \left(\int_{S_Y} a(x,y)(P_u(dy) - \mu(dy))\right) du \\ &\triangleq \epsilon V_{1,1}^\epsilon(x,t) + \epsilon V_{1,2}^\epsilon(x,t), \end{aligned} \tag{4.31}$$

and where P_u is the distribution of the random variable Y_u. Next we give some estimates of $\epsilon V_{1,1}^\epsilon(x,t)$ and $\epsilon V_{1,2}^\epsilon(x,t)$, which imply that $V_1^\epsilon(x,t)$ is well defined.

By Assumption 5, there exists a positive constant k_δ such that for any $x \in \mathbb{R}^n$ with $|x| \le \delta$, and $y \in S_Y$,

$$a(0,y) \equiv 0, \quad \left|\frac{\partial a(x,y)}{\partial x}\right| \le k_\delta. \tag{4.32}$$

Then by Taylor's expansion and (3.9), for any $x \in \mathbb{R}^n$ with $|x| \leq \delta$ and $y \in S_Y$,

$$|a(x,y)| \leq k_\delta |x|, \quad |\bar{a}(x)| \leq k_\delta |x|. \tag{4.33}$$

Without loss of generality, we assume that the initial condition $Y_0 = y$ is deterministic. By Assumption 3, we have

$$\mathrm{var}(P_t - \mu) \leq c_5\, e^{-\alpha t} \tag{4.34}$$

for two positive constants c_5 and α, where "var" denotes the total variation norm of a signed measure over the Borel σ-field, and the mixing rate function $\phi(\cdot)$ of the process Y_t satisfies $\phi(s) = c_6\, e^{-\beta s}$ for two positive constants c_6 and β.

Thus, by (4.29), (3.11), (4.33), Lemma B.1, and the mixing rate function $\phi(s) = c_6 e^{-\beta s}$ of the process Y_t, we obtain that for $t < \tau_\delta^\epsilon$,

$$\begin{aligned}\epsilon \left|V_{1,1}^\epsilon(x,t)\right| &\leq \epsilon \int_{\frac{t}{\epsilon}}^{\frac{\tau_\delta^\epsilon}{\epsilon}} \left|\frac{\partial V(x)}{\partial x}\right| \cdot \left|E\left[a(x,Y_u)|\mathcal{F}_{t/\epsilon}^Y\right] - E[a(x,Y_u)]\right| du \\ &\leq \epsilon \int_{\frac{t}{\epsilon}}^{\frac{\tau_\delta^\epsilon}{\epsilon}} c_3|x| \cdot k_\delta|x| \cdot \phi\left(u - \frac{t}{\epsilon}\right) du \\ &\leq \epsilon c_3 c_6 k_\delta |x|^2 \int_{\frac{t}{\epsilon}}^{\frac{\tau_\delta^\epsilon}{\epsilon}} e^{-\beta(u-\frac{t}{\epsilon})} du \leq \epsilon \frac{c_3 c_6 k_\delta}{\beta}|x|^2,\end{aligned} \tag{4.35}$$

and for $t \geq \tau_\delta^\epsilon$,

$$\epsilon \left|V_{1,1}^\epsilon(x,t)\right| = \epsilon \left| \int_{\frac{\tau_\delta^\epsilon}{\epsilon}}^{\frac{\tau_\delta^\epsilon}{\epsilon}} \left(\frac{\partial V(x)}{\partial x}\right)^T (E[a(x,Y_u)|\mathcal{F}_t^\epsilon] - E[a(x,Y_u)])\, du\right| = 0. \tag{4.36}$$

Thus for any $t \geq 0$,

$$\epsilon \left|V_{1,1}^\epsilon(x,t)\right| \leq \epsilon \frac{c_3 c_6 k_\delta}{\beta}|x|^2. \tag{4.37}$$

By Hölder's inequality, (3.11), (4.33), and (4.34), we get that

$$\begin{aligned}\epsilon \left|V_{1,2}^\epsilon(x,t)\right| &\leq \epsilon \int_{\frac{\tau_\delta^\epsilon(t)}{\epsilon}}^{\frac{\tau_\delta^\epsilon}{\epsilon}} \left|\int_{S_Y} \left(\frac{\partial V(x)}{\partial x}\right)^T a(x,y)\, (P_u(dy) - \mu(dy))\right| du \\ &\leq \epsilon \int_{\frac{\tau_\delta^\epsilon(t)}{\epsilon}}^{\frac{\tau_\delta^\epsilon}{\epsilon}} \left(\int_{S_Y} \left|\left(\frac{\partial V(x)}{\partial x}\right)^T a(x,y)\right|^2 [P_u(dy) + \mu(dy)]\right)^{\frac{1}{2}} \\ &\qquad \cdot \left(\int_{S_Y} |P_u - \mu|(dy)\right)^{\frac{1}{2}} du \\ &\leq \epsilon \int_{\frac{\tau_\delta^\epsilon(t)}{\epsilon}}^{\frac{\tau_\delta^\epsilon}{\epsilon}} \left(\int_{S_Y} (k_\delta c_3)^2 |x|^4 [P_u(dy) + \mu(dy)]\right)^{\frac{1}{2}} (\mathrm{var}(P_u - \mu))^{\frac{1}{2}} du \\ &\leq \epsilon c_3 k_\delta |x|^2 \int_{\frac{\tau_\delta^\epsilon(t)}{\epsilon}}^{\frac{\tau_\delta^\epsilon}{\epsilon}} \left(\int_{S_Y} [P_u(dy) + \mu(dy)]\right)^{\frac{1}{2}} \left(c_5 e^{-\alpha u}\right)^{\frac{1}{2}} du \\ &= \epsilon\sqrt{2c_5}\, c_3 k_\delta |x|^2 \int_{\frac{\tau_\delta^\epsilon(t)}{\epsilon}}^{\frac{\tau_\delta^\epsilon}{\epsilon}} e^{-\frac{\alpha}{2}u} du \leq \epsilon \frac{2\sqrt{2c_5}\, c_3 k_\delta}{\alpha}|x|^2.\end{aligned} \tag{4.38}$$

Therefore, by (4.31), (4.37), and (4.38), for any $x \in \mathbb{R}^n$ with $|x| \le \delta$, and $t \ge 0$,

$$-\epsilon C_1(\delta)|x|^2 \le V_1^\epsilon(x,t) \le \epsilon C_1(\delta)|x|^2, \tag{4.39}$$

where $C_1(\delta) = \frac{2\sqrt{2c_5}c_3k_\delta}{\alpha} + \frac{c_3c_6k_\delta}{\beta}$. By (3.10), (4.30), and (4.39), there exists an $\epsilon_1 > 0$ such that $\frac{\epsilon_1 C_1}{c_1} < 1$, and for $0 < \epsilon \le \epsilon_1$, $x \in \mathbb{R}^n$ with $|x| \le \delta$, and $t \ge 0$,

$$k_1(\delta)V(x) \le V^\epsilon(x,t) \le k_2(\delta)V(x), \tag{4.40}$$

where $k_1(\delta) = 1 - \frac{\epsilon_1 C_1(\delta)}{c_1} > 0$, $k_2(\delta) = 1 + \frac{\epsilon_1 C_1(\delta)}{c_1} > 0$.

Step 2 *(action of the p-infinitesimal operator on a Lyapunov function in the case with local conditions).* We discuss the action of the p-infinitesimal operator $\hat{\mathcal{A}}_\delta^\epsilon$ of the vector process $(X_t^{\epsilon,\delta}, Y_{t/\epsilon})$ on the perturbed Lyapunov function $V^\epsilon(x,t)$.

Recall that $\tau_\delta^\epsilon(t)$ is defined by (4.27). By the continuity of the process X_t^ϵ, we know that for any $t \ge 0$, $X^\epsilon_{\tau_\delta^\epsilon(t)} \in D_\delta = \{x' \in \mathbb{R}^n : |x'| \le \delta\}$. Define

$$\begin{aligned} G(x,y) &= \left(\frac{\partial V(x)}{\partial x}\right)^T a(x,y), \quad \bar{G}(x) = \left(\frac{\partial V(x)}{\partial x}\right)^T \bar{a}(x), \\ \tilde{G}(x,y) &= G(x,y) - \bar{G}(x). \end{aligned} \tag{4.41}$$

Notice that $X^\epsilon_{\tau_\delta^\epsilon(t)}$ is measurable with respect to the σ-field $\mathcal{F}_t^\epsilon$. Then by the definition in (4.30), $V^\epsilon(X^\epsilon_{\tau_\delta^\epsilon(t)}, t) = V(X^\epsilon_{\tau_\delta^\epsilon(t)}) + V_1^\epsilon(X^\epsilon_{\tau_\delta^\epsilon(t)}, t)$. Now we prove that for $0 < \epsilon \le \epsilon_1$, $V^\epsilon(X^\epsilon_{\tau_\delta^\epsilon(t)}, t) \in \mathcal{D}(\hat{\mathcal{A}}_\delta^\epsilon)$, the domain of p-infinitesimal operator $\hat{\mathcal{A}}_\delta^\epsilon$ (for definitions of p-limit and p-infinitesimal operator, see Appendix A), and

$$\begin{aligned} &\hat{\mathcal{A}}_\delta^\epsilon V^\epsilon(X^\epsilon_{\tau_\delta^\epsilon(t)}, t) \\ &= I_{\{t<\tau_\delta^\epsilon\}} \cdot \left\{ \bar{G}(X_t^\epsilon) + \int_{\tau_\delta^\epsilon(t)}^{\tau_\delta^\epsilon} \left[\left.\frac{\partial E_t^\epsilon[\tilde{G}(x, Y_{s/\epsilon})]}{\partial x}\right|_{x=X_t^\epsilon} \right]^T a(X_t^\epsilon, Y_{t/\epsilon}) ds \right\} \triangleq g_\delta^\epsilon(t), \end{aligned} \tag{4.42}$$

where $E_t^\epsilon[\,\cdot\,]$ stands for the conditional expectation $E[\,\cdot\,|\mathcal{F}_t^\epsilon]$, i.e., $E[\,\cdot\,|\mathcal{F}^Y_{t/\epsilon}]$.

Since X_t^ϵ and Y_t are both continuous processes, we know that $V^\epsilon(X^\epsilon_{\tau_\delta^\epsilon(t)}, t)$ and $g_\delta^\epsilon(t)$ are progressively measurable with respect to $\{\mathcal{F}_t^\epsilon\}$. In order to prove (4.42), we need only prove the following three claims for $0 < \epsilon \le \epsilon_1$:

(i) $V^\epsilon(X^\epsilon_{\tau_\delta^\epsilon(t)}, t) \in \overline{\mathcal{M}}^\epsilon_\delta$, where $\overline{\mathcal{M}}^\epsilon_\delta$ is defined with respect to the vector process $(X_t^{\epsilon,\delta}, Y_{t/\epsilon})$ similarly as $\overline{\mathcal{M}}^\epsilon$ is defined in Appendix A.

(ii) $g_\delta^\epsilon(t) \in \overline{\mathcal{M}}^\epsilon_\delta$.

(iii) $p\text{-}\lim_{\delta'\downarrow 0} \frac{E_t^\epsilon[V^\epsilon(X^\epsilon_{\tau_\delta^\epsilon(t+\delta')}, t+\delta')] - V^\epsilon(X^\epsilon_{\tau_\delta^\epsilon(t)}, t)}{\delta'} = g_\delta^\epsilon(t)$.

By (4.40) and the definition of $\tau_\delta^\epsilon(t)$, we get that for $0 < \epsilon \le \epsilon_1$,

$$\sup_{t\ge 0} E\left[\left|V^\epsilon(X^\epsilon_{\tau_\delta^\epsilon(t)}, t)\right|\right] \le \sup_{t\ge 0} E\left[k_2(\delta)V(X^\epsilon_{\tau_\delta^\epsilon(t)})\right] \le k_2(\delta) \cdot \sup_{x\in D_\delta} V(x) < \infty.$$

Thus (i) holds. For the proofs of (ii) and (iii), see Lemmas B.2 and B.3.

Hence, by (4.42), (3.13), (B.9), and (3.10), for any $t \ge 0$ and $0 < \epsilon \le \epsilon_1$,

$$\begin{aligned} \hat{\mathcal{A}}_\delta^\epsilon V^\epsilon(X^\epsilon_{\tau_\delta^\epsilon(t)}, t) &\le I_{\{t<\tau_\delta^\epsilon\}} \left(-\gamma V(X^\epsilon_{\tau_\delta^\epsilon(t)}) + \epsilon \frac{C_2(\delta)}{c_1} V(X^\epsilon_{\tau_\delta^\epsilon(t)})\right) \\ &= -\left(\gamma - \epsilon\frac{C_2(\delta)}{c_1}\right) V(X^\epsilon_{\tau_\delta^\epsilon(t)}) \cdot I_{\{t<\tau_\delta^\epsilon\}}. \end{aligned} \tag{4.43}$$

Take $\epsilon_1' > 0$ such that $\gamma - \epsilon_1' \frac{C_2(\delta)}{c_1} > 0$. Let $\epsilon_2 = \min\{\epsilon_1, \epsilon_1'\}$. Then for $0 < \epsilon \le \epsilon_2$ and any $t \ge 0$,

$$\hat{\mathcal{A}}_\delta^\epsilon V^\epsilon(X^\epsilon_{\tau^\epsilon_\delta(t)}, t) \le 0. \tag{4.44}$$

Step 3 *(proof of stability in probability* (3.14)*).* Suppose $\epsilon \in (0, \epsilon_2]$, $r \in (0, \delta)$, and $X_0^\epsilon = x$ satisfying that $|x| \le r$. For $t \ge 0$, define two stopping times τ_r^ϵ and $\tau_r^\epsilon(t)$ by $\tau_r^\epsilon = \inf\{s \ge 0 : |X_s^\epsilon| > r\}$ and $\tau_r^\epsilon(t) = \tau_r^\epsilon \wedge t$. Then for any $t \ge 0$, $|X^\epsilon_{\tau_r^\epsilon(t)}| \le r < \delta$, $\tau_r^\epsilon(t) \le \tau_\delta^\epsilon(t)$, and $\tau_\delta^\epsilon(\tau_r^\epsilon(t)) = \tau_\delta^\epsilon \wedge \tau_r^\epsilon(t) = \tau_\delta^\epsilon \wedge (\tau_r^\epsilon \wedge t) = \tau_\delta^\epsilon(t) \wedge \tau_r^\epsilon(t) = \tau_r^\epsilon(t)$. Thus by Theorem A.1, the property of conditional expectation, and (4.44),

$$\begin{aligned} &E\left[V^\epsilon(X^\epsilon_{\tau_r^\epsilon(t)}, \tau_r^\epsilon(t)) - V^\epsilon(x,0)\right] = E\left[V^\epsilon(X^\epsilon_{\tau_\delta^\epsilon(\tau_r^\epsilon(t))}, \tau_r^\epsilon(t)) - V^\epsilon(x,0)\right] \\ &= E\left[E_0^\epsilon\left[V^\epsilon(X^\epsilon_{\tau_\delta^\epsilon(\tau_r^\epsilon(t))}, \tau_r^\epsilon(t))\right] - V^\epsilon(x,0)\right] = E\left[E_0^\epsilon\left[\int_0^{\tau_r^\epsilon(t)} \hat{\mathcal{A}}_\delta^\epsilon V^\epsilon(X^\epsilon_{\tau_\delta^\epsilon(u)}, u)du\right]\right] \\ &= E\left[\int_0^{\tau_r^\epsilon(t)} \hat{\mathcal{A}}_\delta^\epsilon V^\epsilon(X^\epsilon_{\tau_\delta^\epsilon(u)}, u)du\right] \le 0. \end{aligned} \tag{4.45}$$

By (4.40) and (4.45),

$$E\left[k_1(\delta)V(X^\epsilon_{\tau_r^\epsilon(t)})\right] \le E\left[V^\epsilon(X^\epsilon_{\tau_r^\epsilon(t)}, \tau_r^\epsilon(t))\right] \le E[V^\epsilon(x,0)] \le k_2(\delta)V(x). \tag{4.46}$$

Denote $V_r = \inf_{r \le |x| \le \delta} V(x)$. Then for any $T > 0$, we have

$$\begin{aligned} E[V(X^\epsilon_{\tau_r^\epsilon(T)})] &= \int_{\{\tau_r^\epsilon < T\}} V(X^\epsilon_{\tau_r^\epsilon(T)})dP + \int_{\{\tau_r^\epsilon \ge T\}} V(X^\epsilon_{\tau_r^\epsilon(T)})dP \\ &\ge \int_{\{\tau_r^\epsilon < T\}} V(X^\epsilon_{\tau_r^\epsilon(T)})dP \ge \int_{\{\sup_{0 \le t \le T} |X_t^\epsilon| > r\}} V(X^\epsilon_{\tau_r^\epsilon(T)})dP \\ &\ge V_r \cdot P\left\{\sup_{0 \le t \le T} |X_t^\epsilon| > r\right\}, \end{aligned}$$

which, together with (4.46), implies

$$P\left\{\sup_{0 \le t \le T} |X_t^\epsilon| > r\right\} \le \frac{E[V(X^\epsilon_{\tau_r^\epsilon(T)})]}{V_r} \le \frac{k_2(\delta)V(x)}{k_1(\delta)V_r}.$$

Letting $T \to \infty$, we get $P\left\{\sup_{t \ge 0} |X_t^\epsilon| > r\right\} \le \frac{k_2(\delta)V(x)}{k_1(\delta)V_r}$. Hence $P\left\{\sup_{t \ge 0} |X_t^\epsilon| \le r\right\} > 1 - \frac{k_2(\delta)V(x)}{k_1(\delta)V_r}$. Since $V(0) = 0$ and $V(x)$ is continuous, for any $\varsigma > 0$, there exists $\delta_1(r, \varsigma) \in (0, \delta)$ such that $V(x) < \frac{k_1(\delta)V_r}{k_2(\delta)}\varsigma$ for all $|x| < \delta_1(r, \varsigma)$. Thus we obtain that for any $0 < \epsilon \le \epsilon^*$ with $\epsilon^* = \min\{\epsilon_1, \epsilon_2\} = \epsilon_2$, for any given $r > 0$, $\varsigma > 0$, there exists $\delta_0 = \delta_1(\min(r, \delta/2), \varsigma) \in (0, \delta)$ such that for all $|x| < \delta_0$,

$$P\left\{\sup_{t \ge 0} |X_t^\epsilon| \le r\right\} \ge P\left\{\sup_{t \ge 0} |X_t^\epsilon| \le \min(r, \delta/2)\right\} > 1 - \varsigma;$$

equivalently, for any $0 < \epsilon \le \epsilon^*$, and any given $r > 0$,

$$\lim_{x \to 0} P\left\{\sup_{t \ge 0} |X_t^\epsilon| > r\right\} = 0. \tag{4.47}$$

Step 4 *(proof of asymptotic convergence property* (3.15)*).* Let $0 < \epsilon < \epsilon^*$ $(= \epsilon_2)$. By Theorem A.1, for any $0 \leq s \leq t$,

$$E\left[V^\epsilon(X^\epsilon_{\tau^\epsilon_\delta(t)}, t)|\mathcal{F}^\epsilon_s\right] = V^\epsilon(X^\epsilon_{\tau^\epsilon_\delta(s)}, s) + \int_s^t E\left[\hat{\mathcal{A}}^\epsilon_\delta V^\epsilon(X^\epsilon_{\tau^\epsilon_\delta(u)}, u)|\mathcal{F}^\epsilon_s\right] du \quad \text{a.s.,} \tag{4.48}$$

where $\mathcal{F}^\epsilon_s$ is defined by (4.29). By (4.40), we know that for any $t \geq 0$, $V^\epsilon(X^\epsilon_{\tau^\epsilon_\delta(t)}, t)$ is integrable. By (4.44) and (4.48), we obtain that for any $0 \leq s \leq t$, $E\left[V^\epsilon(X^\epsilon_{\tau^\epsilon_\delta(t)}, t)|\mathcal{F}^\epsilon_s\right] \leq V^\epsilon(X^\epsilon_{\tau^\epsilon_\delta(s)}, s)$ a.s. Hence by definition $\{V^\epsilon(X^\epsilon_{\tau^\epsilon_\delta(t)}, t) : t \geq 0\}$ is a nonnegative supermartingale with respect to $\{\mathcal{F}^\epsilon_t\}$. By Doob's theorem,

$$\lim_{t\to\infty} V^\epsilon(X^\epsilon_{\tau^\epsilon_\delta(t)}, t) = \xi \quad \text{a.s.,} \tag{4.49}$$

and ξ is finite almost surely. Let B^ϵ_x denote the set of sample paths of $(X^\epsilon_t : t \geq 0)$ with $X^\epsilon_0 = x$ such that $\tau^\epsilon_\delta = \infty$. Since $X^\epsilon_t \equiv 0$ is stable in probability, by (4.47),

$$\lim_{x\to 0} P(B^\epsilon_x) = 1. \tag{4.50}$$

Note that $\epsilon^* = \epsilon_2 = \min\{\epsilon_1, \epsilon_1'\}$, and $\epsilon_1' > 0$ satisfies $\gamma - \epsilon_1' \frac{C_2(\delta)}{c_1} > 0$. Then by (4.43), we get that for any $0 < \epsilon \leq \epsilon^*$,

$$\hat{\mathcal{A}}^\epsilon_\delta V^\epsilon(X^\epsilon_{\tau^\epsilon_\delta(t)}, t) \leq -c_\epsilon V(X^\epsilon_{\tau^\epsilon_\delta(t)}) \cdot I_{\{t<\tau^\epsilon_\delta\}}, \tag{4.51}$$

where $c_\epsilon = \gamma - \epsilon \frac{C_2(\delta)}{c_1} > 0$. For any $0 < \varsigma < \delta$, let $c^\varsigma_\epsilon = c_\epsilon c_1 \varsigma^2$. Notice that for any $t \geq 0$, $|X^\epsilon_{\tau^\epsilon_\delta(t)}| \leq \delta$. Then by (3.10) and (4.51), we obtain that if $0 < \epsilon \leq \epsilon^*$ and $|X^\epsilon_{\tau^\epsilon_\delta(t)}| \geq \varsigma$, then

$$\hat{\mathcal{A}}^\epsilon_\delta V^\epsilon(X^\epsilon_{\tau^\epsilon_\delta(t)}, t) \leq -c^\varsigma_\epsilon \cdot I_{\{t<\tau^\epsilon_\delta\}}. \tag{4.52}$$

For $0 < \epsilon \leq \epsilon^*$, $0 < \varsigma < \delta$, and any $t \geq 0$, define two stopping times $\tau^\epsilon_{\varsigma,\delta}$ and $\tau^\epsilon_{\varsigma,\delta}(t)$ by $\tau^\epsilon_{\varsigma,\delta} = \inf\{t : |X^\epsilon_t| \notin [\varsigma, \delta]\} = \inf\{t : |X^\epsilon_t| < \varsigma \text{ or } |X^\epsilon_t| > \delta\}$ and $\tau^\epsilon_{\varsigma,\delta}(t) = \tau^\epsilon_{\varsigma,\delta} \wedge t$. Then for any $t \geq 0$, we have that $\tau^\epsilon_{\varsigma,\delta}(t) \leq \tau^\epsilon_\delta(t)$. Suppose that $X^\epsilon_0 = x$ and $|x| \in (\varsigma, \delta)$. Then for any $t \in [0, \tau^\epsilon_{\varsigma,\delta}]$, $|X^\epsilon_t| \in [\varsigma, \delta]$. If $u \in [0, \tau^\epsilon_{\varsigma,\delta}(t)]$, then $0 \leq \tau^\epsilon_\delta(u) = \tau^\epsilon_\delta \wedge u \leq u \leq \tau^\epsilon_{\varsigma,\delta}(t) \leq \tau^\epsilon_{\varsigma,\delta}$, and thus $|X^\epsilon_{\tau^\epsilon_\delta(u)}| \in [\varsigma, \delta]$. Hence by Theorem A.1, the property of conditional expectation, and (4.52), we obtain that

(4.53)

$$\begin{aligned}
&E\left[V^\epsilon(X^\epsilon_{\tau^\epsilon_\delta(\tau^\epsilon_{\varsigma,\delta}(t))}, \tau^\epsilon_{\varsigma,\delta}(t))\right] - E[V^\epsilon(x, 0)] = E\left[V^\epsilon(X^\epsilon_{\tau^\epsilon_\delta(\tau^\epsilon_{\varsigma,\delta}(t))}, \tau^\epsilon_{\varsigma,\delta}(t)) - V^\epsilon(x, 0)\right] \\
&= E\left[E^\epsilon_0\left[V^\epsilon(X^\epsilon_{\tau^\epsilon_\delta(\tau^\epsilon_{\varsigma,\delta}(t))}, \tau^\epsilon_{\varsigma,\delta}(t))\right] - V^\epsilon(x, 0)\right] = E\left[\int_0^{\tau^\epsilon_{\varsigma,\delta}(t)} \hat{\mathcal{A}}^\epsilon_\delta V^\epsilon(X^\epsilon_{\tau^\epsilon_\delta(u)}, u) du\right] \\
&\leq E\left[\int_0^{\tau^\epsilon_{\varsigma,\delta}(t)} (-c^\varsigma_\epsilon \cdot I_{\{t<\tau^\epsilon_\delta\}}) du\right] = -c^\varsigma_\epsilon E\left[\tau^\epsilon_{\varsigma,\delta}(t) \cdot I_{\{t<\tau^\epsilon_\delta\}}\right].
\end{aligned}$$

Thus by (4.53) and (4.40),

$$E\left[\tau^\epsilon_{\varsigma,\delta}(t) \cdot I_{\{t<\tau^\epsilon_\delta\}}\right] \leq \frac{E[V^\epsilon(x, 0)]}{c^\varsigma_\epsilon} \leq \frac{k_2(\delta)V(x)}{c^\varsigma_\epsilon}. \tag{4.54}$$

By the definitions of $\tau^\epsilon_{\varsigma,\delta}$ and τ^ϵ_δ, we have that $\tau^\epsilon_{\varsigma,\delta} \le \tau^\epsilon_\delta$. Thus by the property of expectation and (4.54), we have

$$P\left\{t < \tau^\epsilon_{\varsigma,\delta}\right\} = P\left\{t < \tau^\epsilon_{\varsigma,\delta},\, t < \tau^\epsilon_\delta\right\} \le \frac{E\left[\tau^\epsilon_{\varsigma,\delta}(t)\cdot I_{\{t<\tau^\epsilon_\delta\}}\right]}{t} \le \frac{k_2(\delta)V(x)}{c^\varsigma_\epsilon t},$$

which means that the solution process X^ϵ_t beginning in the domain $\varsigma < |x| < \delta$ a.s. reaches the boundary of this domain in a finite time. Then by the definition of the set B^ϵ_x, for all paths contained in the set B^ϵ_x, except for a set of paths of probability zero, we have $\inf_{t>0}|X^\epsilon_t| = 0$. Since $a(0,y)\equiv 0$, if $X^\epsilon_s = 0$ for some $s\ge 0$, then $X^\epsilon_t = 0$ for all $t \ge s$. Hence we obtain $\liminf_{t\to\infty}|X^\epsilon_t| = 0$, and then by (3.10) and (4.40), for any $0 < \epsilon \le \epsilon^*$, we have $\liminf_{t\to\infty} V^\epsilon(X^\epsilon_t, t) = 0$. But by (4.49) and the definition of the set B^ϵ_x, the limit $\lim_{t\to\infty} V^\epsilon(X^\epsilon_{\tau^\epsilon_\delta(t)}, t) = \lim_{t\to\infty} V^\epsilon(X^\epsilon_t, t)$ exists for almost all paths in B^ϵ_x. By the above discussion this limit is equal to zero. Thus by (4.40) and (4.50), we obtain $\lim_{x\to 0} P\{\lim_{t\to\infty}|X^\epsilon_t| = 0\} = 1$.

4.2.2. Proof of Theorem 3.12. For brevity and to avoid overlap, we refer to parts of the proof of Theorem 3.10 that are adapted in the proof of Theorem 3.12.

Step 1 *(action of the p-infinitesimal operator on a Lyapunov function in the case with global conditions).* In the proof of Theorem 3.10, take $\delta = M$ for some positive integer M. Then similar to (4.40) and (4.43), we obtain that there exists an $\epsilon_1 > 0$ such that for any $0<\epsilon<\epsilon_1$, $x\in\mathbb{R}^n$ with $|x|\le M$, and $t\ge 0$,

$$k_1 V(x) \le V^\epsilon(x,t) \le k_2 V(x), \tag{4.55}$$

$$\hat{\mathcal{A}}^\epsilon_M V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)}, t) \le -\left(\gamma - \epsilon\frac{C_2}{c_1}\right) V(X^\epsilon_{\tau^\epsilon_M(t)})\cdot I_{\{t<\tau^\epsilon_M\}}, \tag{4.56}$$

where $k_1 = 1 - \frac{\epsilon_1 C_1}{c_1} > 0$, $k_2 = 1 + \frac{\epsilon_1 C_1}{c_1}$, $C_1 = \frac{2\sqrt{2c_5}c_3 k}{\alpha} + \frac{c_3 c_6 k}{\beta}$, $C_2 = \frac{c_6(c_3+c_4)k^2}{\beta} + \frac{2\sqrt{2c_5}(c_3+c_4)k^2}{\alpha}$ (independent of M used in the truncation).

Step 2 *(proof of global asymptotical stability in probability).* Let $0 < \epsilon'_0 < \min\{\frac{c_1}{C_2}\gamma, \epsilon_1\}$, and denote $\hat\gamma = \frac{1}{2k_2}\left(\gamma - \epsilon'_0\frac{C_2}{c_1}\right)$. Then by (4.55), (4.56), we get that for any $\epsilon\in(0,\epsilon'_0]$,

$$\hat{\mathcal{A}}^\epsilon_M V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)}, t) \le -2\hat\gamma\, k_2 V(X^\epsilon_{\tau^\epsilon_M(t)})\cdot I_{\{t<\tau^\epsilon_M\}} \le -2\hat\gamma\, V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)}, t)\cdot I_{\{t<\tau^\epsilon_M\}}. \tag{4.57}$$

By Lemma B.4, $\hat{\mathcal{A}}^\epsilon_M\left(V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)}, t)\cdot I_{\{t<\tau^\epsilon_M\}}\right) = \hat{\mathcal{A}}^\epsilon_M V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)}, t)$, which together with (4.57) implies that

$$\left(\hat{\mathcal{A}}^\epsilon_M + 2\hat\gamma\right)\left(V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)}, t)\cdot I_{\{t<\tau^\epsilon_M\}}\right) \le 0. \tag{4.58}$$

For $t\ge 0$, define

$$\begin{aligned} M^\epsilon_t = {}& e^{2\hat\gamma t} V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)}, t)\cdot I_{\{t<\tau^\epsilon_M\}} + e^{2\hat\gamma\tau^\epsilon_M} V(X^\epsilon_{\tau^\epsilon_M})\cdot I_{\{\tau^\epsilon_M\le t\}} - V^\epsilon(x,0) \\ & - \int_0^t e^{2\hat\gamma s}(\hat{\mathcal{A}}^\epsilon_M + 2\hat\gamma)\left(V^\epsilon(X^\epsilon_{\tau^\epsilon_M(s)}, s)\cdot I_{\{s<\tau^\epsilon_M\}}\right) ds. \end{aligned} \tag{4.59}$$

Then by the fact that $\tau^\epsilon_M > 0$ a.s., we know that $M^\epsilon_0 = 0$ a.s. By the definition of $V^\epsilon(x,t)$, we can verify that $e^{2\hat\gamma t} V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)}, t)\cdot I_{\{t<\tau^\epsilon_M\}} + e^{2\hat\gamma\tau^\epsilon_M} V(X^\epsilon_{\tau^\epsilon_M})\cdot I_{\{\tau^\epsilon_M\le t\}}$ is continuous in t, and thus M^ϵ_t is continuous. By (4.55), (4.41), the definition of

$\hat{\mathcal{A}}^\epsilon_M V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)}, t)$ (replace δ by M in (4.42)), (B.10) with δ replaced by M, and the fact that $|X^\epsilon_{\tau^\epsilon_M}| \le M$, we know that for any $t \ge 0$, M^ϵ_t is integrable. By Lemma B.5, we know that M^ϵ_t is a martingale relative to $\{\mathcal{F}^\epsilon_t\}$, and thus it is a zero-mean, continuous martingale relative to $\{\mathcal{F}^\epsilon_t\}$.

By (4.55), (4.58), and (4.59), we get that

$$\begin{aligned} 0 &\le k_1\, e^{2\hat{\gamma}t} V(X^\epsilon_{\tau^\epsilon_M(t)}) \cdot I_{\{t<\tau^\epsilon_M\}} \le e^{2\hat{\gamma}t} V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)}, t) \cdot I_{\{t<\tau^\epsilon_M\}} \\ &\le e^{2\hat{\gamma}t} V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)}, t) \cdot I_{\{t<\tau^\epsilon_M\}} + e^{2\hat{\gamma}\tau^\epsilon_M} V(X^\epsilon_{\tau^\epsilon_M}) \cdot I_{\{\tau^\epsilon_M \le t\}} \qquad (\text{since } V(x) \ge 0) \\ &= V^\epsilon(x,0) + M^\epsilon_t + \int_0^t e^{2\hat{\gamma}s} (\hat{\mathcal{A}}^\epsilon_M + 2\hat{\gamma}) \left(V^\epsilon(X^\epsilon_{\tau^\epsilon_M(s)}, s) \cdot I_{\{s<\tau^\epsilon_M\}} \right) ds \\ &\le V^\epsilon(x,0) + M^\epsilon_t \le k_2 V(x) + M^\epsilon_t, \end{aligned} \tag{4.60}$$

which means $k_2V(x) + M^\epsilon_t$ is a nonnegative continuous martingale relative to $\{\mathcal{F}^\epsilon_t\}$. By (4.60) and Doob's inequality (cf. section 2.III.9 of [5]), we have that for any $\eta > 0$, and $T > 0$,

$$\begin{aligned} P\left\{ \sup_{0\le t\le T} k_1 e^{2\hat{\gamma}t} V(X^\epsilon_{\tau^\epsilon_M(t)}) \cdot I_{\{t<\tau^\epsilon_M\}} > \eta \right\} &\le P\left\{ \sup_{0\le t\le T} \{k_2V(x) + M^\epsilon_t\} > \eta \right\} \\ &\le \frac{k_2V(x)}{\eta}. \end{aligned} \tag{4.61}$$

Letting $T \uparrow \infty$ in (4.61) yields

$$P\left\{ \sup_{t\ge 0} k_1 e^{2\hat{\gamma}t} V(X^\epsilon_{\tau^\epsilon_M(t)}) \cdot I_{\{t<\tau^\epsilon_M\}} > \eta \right\} \le \frac{k_2V(x)}{\eta}. \tag{4.62}$$

Notice that under Assumption 7, the original system (2.1) is globally Lipschitz. Then we know that the solution process X^ϵ_t is regular (cf. section 7.2 of [7]), i.e.,

$$\lim_{M\to\infty} \tau^\epsilon_M = \infty \quad \text{a.s.} \tag{4.63}$$

Notice that k_1, k_2, and $\hat{\gamma}$ are independent of M, and $\tau^\epsilon_M(t) = t \wedge \tau^\epsilon_M$. Then by (4.63),

$$\sup_{t\ge 0} k_1 e^{2\hat{\gamma}t} V(X^\epsilon_t) \le \liminf_{M\to\infty} \sup_{t\ge 0} k_1 e^{2\hat{\gamma}t} V(X^\epsilon_{\tau^\epsilon_M(t)}) \cdot I_{\{t<\tau^\epsilon_M\}} \quad \text{a.s.} \tag{4.64}$$

Now, by (4.64), Fatou's lemma, and (4.62), we obtain

$$\begin{aligned} P\left\{ \sup_{t\ge 0} k_1 e^{2\hat{\gamma}t} V(X^\epsilon_t) > \eta \right\} &= E\left[I_{(\eta,\infty]} \left(\sup_{t\ge 0} k_1 e^{2\hat{\gamma}t} V(X^\epsilon_t) \right) \right] \\ &\le E\left[\liminf_{M\to\infty} I_{(\eta,\infty]} \left(\sup_{t\ge 0} k_1 e^{2\hat{\gamma}t} V(X^\epsilon_{\tau^\epsilon_M(t)}) \cdot I_{\{t<\tau^\epsilon_M\}} \right) \right] \\ &\le \liminf_{M\to\infty} E\left[I_{(\eta,\infty]} \left(\sup_{t\ge 0} k_1 e^{2\hat{\gamma}t} V(X^\epsilon_{\tau^\epsilon_M(t)}) \cdot I_{\{t<\tau^\epsilon_M\}} \right) \right] \\ &= \liminf_{M\to\infty} P\left\{ \sup_{t\ge 0} k_1 e^{2\hat{\gamma}t} V(X^\epsilon_{\tau^\epsilon_M(t)}) \cdot I_{\{t<\tau^\epsilon_M\}} > \eta \right\} \le \frac{k_2V(x)}{\eta}. \end{aligned} \tag{4.65}$$

By Assumption 6, we have $\left\{ c_1|X^\epsilon_t|^2 \le e^{-2\hat{\gamma}t} \frac{\eta}{k_1},\ t \ge 0 \right\} \supseteq \left\{ V(X^\epsilon_t) \le e^{-2\hat{\gamma}t} \frac{\eta}{k_1},\ t \ge 0 \right\}$, which together with (4.65) implies $P\left\{ |X^\epsilon_t| \le e^{-\hat{\gamma}t} \left(\frac{\eta}{k_1c_1} \right)^{\frac{1}{2}},\ t \ge 0 \right\} \ge 1 - \frac{k_2V(x)}{\eta}$. Let

$\eta_1 > 0$ and $\eta_2 > 0$ be given. Choose η such that $\left(\frac{\eta}{k_1 c_1}\right)^{\frac{1}{2}} \le \eta_2$, and then choose $\delta_0 > 0$ such that if $|x| < \delta_0$, then $\frac{k_2 V(x)}{\eta} \le \eta_1$. Thus we have $P\left\{|X_t^\epsilon| \le \eta_2 e^{-\hat{\gamma}t},\ t \ge 0\right\} \ge 1 - \eta_1$.

Now, we prove that for any $x \in \mathbb{R}^n$, $P\{\lim_{t\to\infty} |X_t^\epsilon| = 0\} = 1$. Notice that for any $H > 0$, $\{\lim_{t\to\infty} |X_t^\epsilon| = 0\} = \{\lim_{t\to\infty} V(X_t^\epsilon) = 0\} \supseteq \left\{\sup_{t\ge 0} k_1 e^{2\hat{\gamma}t} V(X_t^\epsilon) \le H\right\}$. Then by (4.65), we obtain $P\{\lim_{t\to\infty} |X_t^\epsilon| = 0\} \ge 1 - \frac{k_2 V(x)}{H}$, and letting $H \uparrow \infty$ yields $P\{\lim_{t\to\infty} |X_t^\epsilon| = 0\} = 1$. The proof is completed.

4.2.3. Proof of Theorem 3.13. The only condition of Theorem 3.13 that is different from the conditions in Theorem 3.12 is $a(0, y) \equiv 0$ replaced with $\sup_{y\in S_Y} |a(0, y)| < \infty$. Thus here we use the same approach as in the proof of Theorem 3.12.

Step 1 *(Lyapunov estimates for Theorem* 3.13*).* Let $c = \left(\sup_{y\in S_Y} |a(0, y)|\right) \vee 1$. Then by Assumption 8 (assume $k \ge 1$; otherwise, replace k by $k \vee 1$), we get that for any $x \in \mathbb{R}^n$ and $y \in S_Y$,

$$|a(x, y)| \le c + k|x| \le k(c + |x|). \tag{4.66}$$

By (3.9) and (4.66), we get that for any $x \in \mathbb{R}^n$, $|\bar{a}(x)| \le k(c + |x|)$. Then following the proofs of Theorem 3.10, we obtain that for $x \in \mathbb{R}^n$ with $|x| \le M$, and $t \ge 0$,

$$-\epsilon C_1 |x|(c + |x|) \le V_1^\epsilon(x, t) \le \epsilon C_1 |x|(c + |x|), \tag{4.67}$$

where $C_1 = \frac{2\sqrt{2c_5}c_3 k}{\alpha} + \frac{c_3 c_6 k}{\beta}$ (the same with the one in the proof of Theorem 3.12).

By Assumption 6, the definition of $V^\epsilon(x, t)$, and (4.67), we get that for any $\epsilon > 0$, $x \in \mathbb{R}^n$ with $|x| \le M$, and $t \ge 0$,

$$V(x) - \epsilon C_1 |x|(c + |x|) \le V^\epsilon(x, t) \le V(x) + \epsilon C_1 |x|(c + |x|). \tag{4.68}$$

It follows from (4.68) and $c \ge 1$ that if $|x| \le 1$, then

$$V(x) - 2\epsilon c C_1 \le V^\epsilon(x, t) \le V(x) + 2\epsilon c C_1. \tag{4.69}$$

By Assumption 6 and $c \ge 1$, we have that if $|x| \ge 1$, then $|x|(c + |x|) \le 2c|x|^2 \le \frac{2c}{c_1} V(x)$, and thus by (4.68), if $|x| \ge 1$, then

$$\left(1 - \frac{2\epsilon c C_1}{c_1}\right) V(x) \le V^\epsilon(x, t) \le \left(1 + \frac{2\epsilon c C_1}{c_1}\right) V(x). \tag{4.70}$$

Take a positive constant $\epsilon_1' < \frac{c_1}{2cC_1}$, and define $k_1' = 1 - \frac{2cC_1}{c_1}\epsilon_1'$, $k_2' = 1 + \frac{2cC_1}{c_1}\epsilon_1'$. Then by (4.70), we get that for any $0 < \epsilon \le \epsilon_1'$ and $|x| \ge 1$,

$$k_1' V(x) \le V^\epsilon(x, t) \le k_2' V(x). \tag{4.71}$$

Step 2 *(action of the p-infinitesimal operator on a Lyapunov function without an equilibrium condition).* By (4.66) and Assumptions 6 and 8, we get that for any $x \in \mathbb{R}^n$, $y \in S_Y$,

$$|Q(x, y)| \le c_4 k(c + |x|) + c_3 k|x|, \tag{4.72}$$

where $Q(x, y)$ is given by (B.6). Then by (4.66) and (4.72), following the proof of (B.9), we obtain that

$$\begin{aligned} &\left| \int_{\tau_M^\epsilon(t)}^{\tau_M^\epsilon} \left[\frac{\partial E_t^\epsilon[\tilde{G}(x, Y_{s/\epsilon})]}{\partial x}\right]^T a(x, Y_{t/\epsilon}) ds \right| \\ &\quad \le \epsilon\, k^2 \left[\frac{c_6}{\beta} + \frac{2\sqrt{2c_5}}{\alpha}\right] \cdot \left[c_4 c^2 + (c_3 + 2c_4)c|x| + (c_3 + c_4)|x|^2\right]. \end{aligned} \tag{4.73}$$

By Assumption 6 and $c \geq 1$, we have that if $|x| \geq 1$, then $c_4c^2 + (c_3 + 2c_4)c|x| + (c_3 + c_4)|x|^2 \leq \frac{c_4c^2+(c_3+2c_4)c+(c_3+c_4)}{c_1}V(x)$. Denote $C_2' = k^2[\frac{c_6}{\beta} + \frac{2\sqrt{2c_5}}{\alpha}]\frac{c_4c^2+(c_3+2c_4)c+(c_3+c_4)}{c_1}$. Then by (4.73), we obtain that if $|x| \geq 1$, then

$$\left| \int_{\tau_M^\epsilon(t)}^{\tau_M^\epsilon} \left[\frac{\partial E_t^\epsilon[\tilde{G}(x, Y_{s/\epsilon})]}{\partial x} \right]^T a(x, Y_{t/\epsilon}) ds \right| \leq \epsilon C_2' V(x); \tag{4.74}$$

if $|x| < 1$, then

$$\left| \int_{\tau_M^\epsilon(t)}^{\tau_M^\epsilon} \left[\frac{\partial E_t^\epsilon[\tilde{G}(x, Y_{s/\epsilon})]}{\partial x} \right]^T a(x, Y_{t/\epsilon}) ds \right| \leq \epsilon c_1 C_2'. \tag{4.75}$$

By the definition of $\hat{\mathcal{A}}_M^\epsilon V^\epsilon(X_{\tau_M^\epsilon(t)}^\epsilon, t)$, Assumption 6, (4.74), and (4.75), for any $t \geq 0$,

$$\begin{aligned} &\hat{\mathcal{A}}_M^\epsilon V^\epsilon(X_{\tau_M^\epsilon(t)}^\epsilon, t) \\ &\leq \begin{cases} \left(-\gamma V(X_{\tau_M^\epsilon(t)}^\epsilon) + \epsilon c_1 C_2'\right) \cdot I_{\{t<\tau_M^\epsilon\}} & \text{if } |X_{\tau_M^\epsilon(t)}^\epsilon| < 1; \\ -(\gamma - \epsilon C_2') V(X_{\tau_M^\epsilon(t)}^\epsilon) \cdot I_{\{t<\tau_M^\epsilon\}} & \text{if } |X_{\tau_M^\epsilon(t)}^\epsilon| \geq 1. \end{cases} \end{aligned} \tag{4.76}$$

Step 3 *(proof of boundedness in probability).* Let $0 < \epsilon^* < \min\{\frac{\gamma}{C_2'}, \epsilon_1'\}$, and denote $\hat{\gamma} = \frac{\gamma - \epsilon^* C_2'}{k_2'}$. Then by (4.76) and (4.71), we get that for any $\epsilon \in (0, \epsilon^*]$, if $|X_{\tau_M^\epsilon(t)}^\epsilon| \geq 1$, then

$$\hat{\mathcal{A}}_M^\epsilon V^\epsilon(X_{\tau_M^\epsilon(t)}^\epsilon, t) \leq -\hat{\gamma} k_2' V(X_{\tau_M^\epsilon(t)}^\epsilon) \cdot I_{\{t<\tau_M^\epsilon\}} \leq -\hat{\gamma} V^\epsilon(X_{\tau_M^\epsilon(t)}^\epsilon, t) \cdot I_{\{t<\tau_M^\epsilon\}}. \tag{4.77}$$

By $\hat{\mathcal{A}}_M^\epsilon \left(V^\epsilon(X_{\tau_M^\epsilon(t)}^\epsilon, t) \cdot I_{\{t<\tau_M^\epsilon\}}\right) = \hat{\mathcal{A}}_M^\epsilon V^\epsilon(X_{\tau_M^\epsilon(t)}^\epsilon, t)$ (see Lemma B.4) and (4.77), we get that if $|X_{\tau_M^\epsilon(t)}^\epsilon| \geq 1$, then

$$\left(\hat{\mathcal{A}}_M^\epsilon + \hat{\gamma}\right) \left(V^\epsilon(X_{\tau_M^\epsilon(t)}^\epsilon, t) \cdot I_{\{t<\tau_M^\epsilon\}}\right) \leq 0. \tag{4.78}$$

For $t \geq 0$, define

$$\begin{aligned} M_t^\epsilon = {} & e^{\hat{\gamma} t} V^\epsilon(X_{\tau_M^\epsilon(t)}^\epsilon, t) \cdot I_{\{t<\tau_M^\epsilon\}} + e^{\hat{\gamma}\tau_M^\epsilon} V(X_{\tau_M^\epsilon}^\epsilon) \cdot I_{\{\tau_M^\epsilon \leq t\}} - V^\epsilon(x, 0) \\ & - \int_0^t e^{\hat{\gamma} s} (\hat{\mathcal{A}}_M^\epsilon + \hat{\gamma}) \left(V^\epsilon(X_{\tau_M^\epsilon(s)}^\epsilon, s) \cdot I_{\{s<\tau_M^\epsilon\}}\right) ds. \end{aligned}$$

As in the proof of Theorem 3.12, we can prove that M_t^ϵ is a zero-mean, continuous martingale relative to $\{\mathcal{F}_t^\epsilon\}$. Thus by $\hat{\mathcal{A}}_M^\epsilon \left(V^\epsilon(X_{\tau_M^\epsilon(t)}^\epsilon, t) \cdot I_{\{t<\tau_M^\epsilon\}}\right) = \hat{\mathcal{A}}_M^\epsilon V^\epsilon(X_{\tau_M^\epsilon(t)}^\epsilon, t)$, (4.69), (4.76), (4.78), and the fact that $\gamma > \hat{\gamma}$, we have for any $0 < \epsilon \leq \epsilon^*$,

$$\begin{aligned} & E\left[e^{\hat{\gamma} t} V^\epsilon(X_{\tau_M^\epsilon(t)}^\epsilon, t) I_{\{t<\tau_M^\epsilon\}}\right] \leq E\left[e^{\hat{\gamma} t} V^\epsilon(X_{\tau_M^\epsilon(t)}^\epsilon, t) \cdot I_{\{t<\tau_M^\epsilon\}} + e^{\hat{\gamma}\tau_M^\epsilon} V(X_{\tau_M^\epsilon}^\epsilon) I_{\{\tau_M^\epsilon \leq t\}}\right] \\ & = V^\epsilon(x, 0) + \int_0^t E\left[e^{\hat{\gamma} s} (\hat{\mathcal{A}}_M^\epsilon + \hat{\gamma}) \left(V^\epsilon(X_{\tau_M^\epsilon(s)}^\epsilon, s) \cdot I_{\{s<\tau_M^\epsilon\}}\right)\right] ds \\ & = V^\epsilon(x, 0) + \int_0^t E\left[e^{\hat{\gamma} s} (\hat{\mathcal{A}}_M^\epsilon + \hat{\gamma}) \left(V^\epsilon(X_{\tau_M^\epsilon(s)}^\epsilon, s) \cdot I_{\{s<\tau_M^\epsilon\}}\right) I_{\{|X_{\tau_M^\epsilon(s)}^\epsilon| < 1\}}\right] ds \\ & \quad + \int_0^t E\left[e^{\hat{\gamma} s} (\hat{\mathcal{A}}_M^\epsilon + \hat{\gamma}) \left(V^\epsilon(X_{\tau_M^\epsilon(s)}^\epsilon, s) \cdot I_{\{s<\tau_M^\epsilon\}}\right) I_{\{|X_{\tau_M^\epsilon(s)}^\epsilon| \geq 1\}}\right] ds \end{aligned}$$

$$
\begin{aligned}
&\le V^\epsilon(x,0) + \int_0^t E\left[e^{\hat\gamma s}(\hat{\mathcal{A}}_M^\epsilon + \hat\gamma)\left(V^\epsilon(X^\epsilon_{\tau^\epsilon_M(s)}, s)\cdot I_{\{s<\tau^\epsilon_M\}}\right) I_{\{|X^\epsilon_{\tau^\epsilon_M(s)}|<1\}}\right] ds \\
&\le V^\epsilon(x,0) + \int_0^t E\left[e^{\hat\gamma s}\left(-\gamma V(X^\epsilon_{\tau^\epsilon_M(s)}) + \epsilon c_1 C_2' \right.\right. \\
&\qquad\qquad \left.\left. + \hat\gamma V^\epsilon(X^\epsilon_{\tau^\epsilon_M(s)}, s)\right) I_{\{s<\tau^\epsilon_M\}} I_{\{|X^\epsilon_{\tau^\epsilon_M(s)}|<1\}}\right] ds \\
&\le V^\epsilon(x,0) + \int_0^t E\left[e^{\hat\gamma s}\left(-\gamma V(X^\epsilon_{\tau^\epsilon_M(s)}) + \epsilon c_1 C_2' \right.\right. \\
&\qquad\qquad \left.\left. + \hat\gamma\left(V(X^\epsilon_{\tau^\epsilon_M(s)}) + 2\epsilon c C_1\right)\right) I_{\{s<\tau^\epsilon_M\}} I_{\{|X^\epsilon_{\tau^\epsilon_M(s)}|<1\}}\right] ds \\
&\le V^\epsilon(x,0) + \int_0^t E\left[e^{\hat\gamma s}\left(\epsilon c_1 C_2' + 2\hat\gamma\epsilon c C_1\right)\right] ds \\
&= V^\epsilon(x,0) + \frac{\epsilon c_1 C_2' + 2\hat\gamma\epsilon c C_1}{\hat\gamma}\left(e^{\hat\gamma t} - 1\right) \le V^\epsilon(x,0) + \frac{\epsilon c_1 C_2' + 2\hat\gamma\epsilon c C_1}{\hat\gamma} e^{\hat\gamma t}.
\end{aligned}
$$

Thus we have that

$$
E\left[V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)}, t)\cdot I_{\{t<\tau^\epsilon_M\}}\right] \le e^{-\hat\gamma t} V^\epsilon(x,0) + \frac{\epsilon c_1 C_2' + 2\hat\gamma\epsilon c C_1}{\hat\gamma}. \tag{4.79}
$$

By (4.71), Assumption 6, and the property of expectation, we get that for any $r > 1$,

$$
\begin{aligned}
&P\left\{|X^\epsilon_{\tau^\epsilon_M(t)}| > r,\, t < \tau^\epsilon_M\right\} \\
&= P\left\{|X^\epsilon_{\tau^\epsilon_M(t)}| > r,\, k_1' V(X^\epsilon_{\tau^\epsilon_M(t)}) \le V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)}, t) \le k_2' V(X^\epsilon_{\tau^\epsilon_M(t)}),\, t < \tau^\epsilon_M\right\} \\
&= P\left\{|X^\epsilon_{\tau^\epsilon_M(t)}| > r,\, V(X^\epsilon_{\tau^\epsilon_M(t)}) > c_1 r^2,\, k_1' V(X^\epsilon_{\tau^\epsilon_M(t)}) \le V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)}, t) \right. \\
&\qquad\qquad \left. \le k_2' V(X^\epsilon_{\tau^\epsilon_M(t)}),\, t < \tau^\epsilon_M\right\} \\
&\le P\left\{|X^\epsilon_{\tau^\epsilon_M(t)}| > 1,\, V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)}, t) > c_1 k_1' r^2,\, t < \tau^\epsilon_M\right\} \\
&\le \frac{1}{c_1 k_1' r^2} E\left[V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)}, t)\cdot I_{\{t<\tau^\epsilon_M\}}\cdot I_{\{|X^\epsilon_{\tau^\epsilon_M(t)}|>1\}}\right].
\end{aligned} \tag{4.80}
$$

Thus by (4.79), (4.69), and (4.71), we obtain for any $0 < \epsilon \le \epsilon^*$, and any $t \ge 0$,

$$
\begin{aligned}
&E\left[V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)}, t)\cdot I_{\{t<\tau^\epsilon_M\}}\cdot I_{\{|X^\epsilon_{\tau^\epsilon_M(t)}|>1\}}\right] \\
&= E\left[V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)}, t)\cdot I_{\{t<\tau^\epsilon_M\}}\right] - E\left[V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)}, t)\cdot I_{\{t<\tau^\epsilon_M\}}\cdot I_{\{|X^\epsilon_{\tau^\epsilon_M(t)}|\le 1\}}\right] \\
&\le e^{-\hat\gamma t} V^\epsilon(x,0) + \frac{\epsilon c_1 C_2' + 2\hat\gamma\epsilon c C_1}{\hat\gamma} \\
&\qquad\qquad - E\left[\left(V(X^\epsilon_{\tau^\epsilon_M(t)}) - 2\epsilon c C_1\right)\cdot I_{\{t<\tau^\epsilon_M\}}\cdot I_{\{|X^\epsilon_{\tau^\epsilon_M(t)}|\le 1\}}\right] \\
&\le e^{-\hat\gamma t} V^\epsilon(x,0) + \frac{\epsilon c_1 C_2' + 2\hat\gamma\epsilon c C_1}{\hat\gamma} + 2\epsilon c C_1 \\
&\le \max\left\{V(x) + 2\epsilon^* c C_1, k_2' V(x)\right\} + \frac{\epsilon^* c_1 C_2'}{\hat\gamma} + 4\epsilon^* c C_1 \triangleq C,
\end{aligned} \tag{4.81}
$$

where C is a positive constant dependent on x, ϵ^*, c, c_1, C_1, C_2', k_2', and $\hat{\gamma}$. Thus by (4.80) and (4.81), we get that for any $0 < \epsilon \le \epsilon^*$, any $r > 1$, and any $t \ge 0$,

$$P\left\{|X^{\epsilon}_{\tau^{\epsilon}_M(t)}| > r,\, t < \tau^{\epsilon}_M\right\} \le \frac{C}{c_1 k_1' r^2}. \tag{4.82}$$

By the fact that $\lim_{M\to\infty} \tau^{\epsilon}_M = \infty$ a.s. (see (4.63)), the dominated convergence theorem, and (4.82), we get that for any $0 < \epsilon \le \epsilon^*$ and any $r > 1$,

$$\begin{aligned}
&\sup_{t\ge 0} P\{|X^{\epsilon}_t| > r\} = \sup_{t\ge 0} E\left[I_{(r,\infty]}(|X^{\epsilon}_t|)\right] = \sup_{t\ge 0} E\left[\lim_{M\to\infty} I_{(r,\infty]}(|X^{\epsilon}_{\tau^{\epsilon}_M(t)}|)\cdot I_{\{t<\tau^{\epsilon}_M\}}\right]\\
&= \sup_{t\ge 0}\left(\lim_{M\to\infty} E\left[I_{(r,\infty]}(|X^{\epsilon}_{\tau^{\epsilon}_M(t)}|)\cdot I_{\{t<\tau^{\epsilon}_M\}}\right]\right)\\
&= \sup_{t\ge 0}\left(\lim_{M\to\infty} P\left\{|X^{\epsilon}_{\tau^{\epsilon}_M(t)}| > r,\, t < \tau^{\epsilon}_M\right\}\right) \le \frac{C}{c_1 k_1' r^2},
\end{aligned}$$

which implies that $\lim_{r\to\infty}\sup_{t\ge 0} P\{|X^{\epsilon}_t| > r\} = 0$; i.e., the solution process X^{ϵ}_t is bounded in probability. The proof is completed.

5. Examples.

5.1. Perturbation process is asymptotically periodic. Consider the following system:

$$\frac{dx^{\epsilon}_t}{dt} = \xi^2_{t/\epsilon}(x^{\epsilon}_t + 1) - \frac{1}{2}(x^{\epsilon}_t + 1)^2, \tag{5.1}$$

where the perturbation process is

$$dY_t = -pY_t dt + q dw_t, \qquad \xi_t = \sin t + e^{-at}\sin Y_t \tag{5.2}$$

with $p, q, a > 0$. Noticing that for any $t \ge 0$

$$\begin{aligned}
\lim_{T\to\infty}\frac{1}{T}\int_t^{t+T}\xi^2_s ds &= \lim_{T\to\infty}\frac{1}{T}\int_t^{t+T}\left(\sin^2 s + 2\sin s e^{-as}\sin Y_s + e^{-2as}\sin^2 Y_s\right)ds\\
&= \lim_{T\to\infty}\frac{1}{T}\int_t^{t+T}\sin^2 s ds = \frac{1}{2\pi}\int_0^{2\pi}\sin^2 s ds = \frac{1}{2} \quad \text{a.s.},
\end{aligned}$$

we obtain the average system of (5.1) as $\dot{\bar{x}}_t = -\left(\bar{x}_t + \bar{x}^2_t\right)/2$, which is locally exponentially stable at $\bar{x}_t = 0$. Figure 1 shows the simulation results with $\bar{x}_0 = x^{\epsilon}_0 = 0.5$, $p = 1$, $q = 2$, $a = 0.01$, $\epsilon = 0.09$, from which we can see that the solution of the original system (5.1) converges (in probability) to the solution of the average system $\dot{\bar{x}}_t = -\left(\bar{x}_t + \bar{x}^2_t\right)/2$ (see (3.3) in Theorem 3.2) and the solution of system (5.1) is exponentially practically stable in probability (Theorem 3.5).

5.2. Perturbation process is a.s. exponentially stable. Consider the following system:

$$\frac{dx^{\epsilon}_t}{dt} = -\sin^2(\xi_{t/\epsilon}) + \left(\sin(\xi_{t/\epsilon}) - \frac{1}{2}\right)(x^{\epsilon}_t)^2 - x^{\epsilon}_t, \qquad d\xi_t = p\xi_t dt + q\xi_t dw_t, \tag{5.3}$$

where $p < q^2/2$. We know that $\xi_t = \xi_0 e^{(p-\frac{q^2}{2})t + qw_t}$. By the law of the iterated logarithm of Brownian motion (see Theorem 2.9.23 of [9]), we know that $\xi_t \to 0$

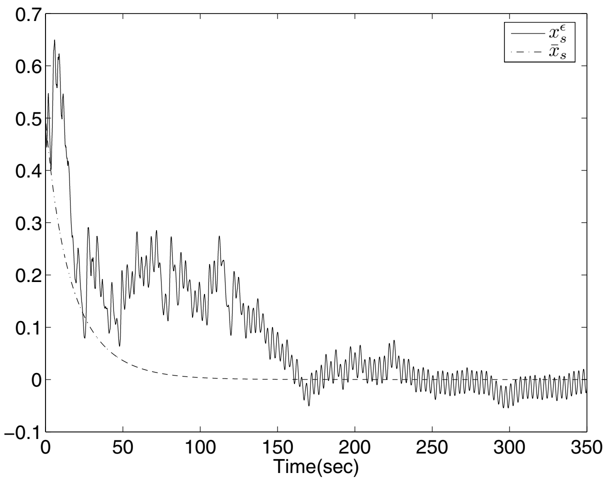

FIG. 1. *States of the original and average systems for system* (5.1)–(5.2) *to illustrate Theorems* 3.2 *and* 3.5.

a.s. as $t \to \infty$. Noticing that $\lim_{T\to\infty} \frac{1}{T}\int_t^{t+T} f(s)ds = \lim_{s\to\infty} f(s)$ for continuous function f when the latter limit exists, we have that for any $t \geq 0$,

$$\lim_{T\to\infty} \frac{1}{T}\int_t^{t+T} \sin^2(\xi_s)ds = 0 \quad \text{a.s.},$$
$$\lim_{T\to\infty} \frac{1}{T}\int_t^{t+T} \left(\left(\sin(\xi_s) - \frac{1}{2}\right)x^2 - x\right)ds = -\frac{1}{2}x^2 - x \quad \text{a.s.}$$

Thus we obtain the average system of (5.3) as $\dot{\bar{x}}_t = -\bar{x}_t - \bar{x}_t^2/2$, which is locally exponentially stable at $\bar{x}_t = 0$. Figure 2 shows the simulation results with $\bar{x}_0 = x_0^\epsilon = 0.2$, $\xi_0 = 1$, $p = 0.4$, $q = 1$, from which we can see that the solution of the original system (5.3) converges (in probability) to the solution of the average system $\dot{\bar{x}}_t = -\bar{x}_t - \bar{x}_t^2/2$ (see (3.4) in Corollary 3.4) and the solution of system (5.3) is exponentially practically stable in probability (Theorem 3.5).

5.3. Perturbation process is Brownian motion on the unit circle. While in sections 5.1 and 5.2 we illustrated the theorems in section 3.1 for uniform strong ergodic perturbation processes, in this section we illustrate the theorems in section 3.2 for ϕ-mixing perturbation processes. Consider the system

$$\begin{aligned}\frac{dx_t^\epsilon}{dt} = &-\begin{bmatrix} 0 & 1 \end{bmatrix}\begin{bmatrix} Y_1^2\left(\frac{t}{\epsilon}\right) & Y_2^2\left(\frac{t}{\epsilon}\right)\end{bmatrix}^T x_t^\epsilon \\ &+\left(\begin{bmatrix} 0 & 1 \end{bmatrix}\begin{bmatrix} Y_1\left(\frac{t}{\epsilon}\right) & Y_2\left(\frac{t}{\epsilon}\right)\end{bmatrix}^T - \frac{1}{2}\right)(x_t^\epsilon)^2,\end{aligned} \tag{5.4}$$

where the perturbation process $Y(t) = [Y_1(t), Y_2(t)]^T$ is Brownian motion on the unit circle, $dY_t = -\frac{1}{2}Y_t dt + BY_t dW_t$, and $Y_0 = [\cos(\vartheta), \sin(\vartheta)]^T$ for all $\vartheta \in \mathbb{R}$, with $B = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$. In fact, we have the simple expression [25, Example 5.4, p. 63]

$$Y(t) = [\cos(\vartheta + W_t), \sin(\vartheta + W_t)]^T = e^{i(\vartheta+W_t)}. \tag{5.5}$$

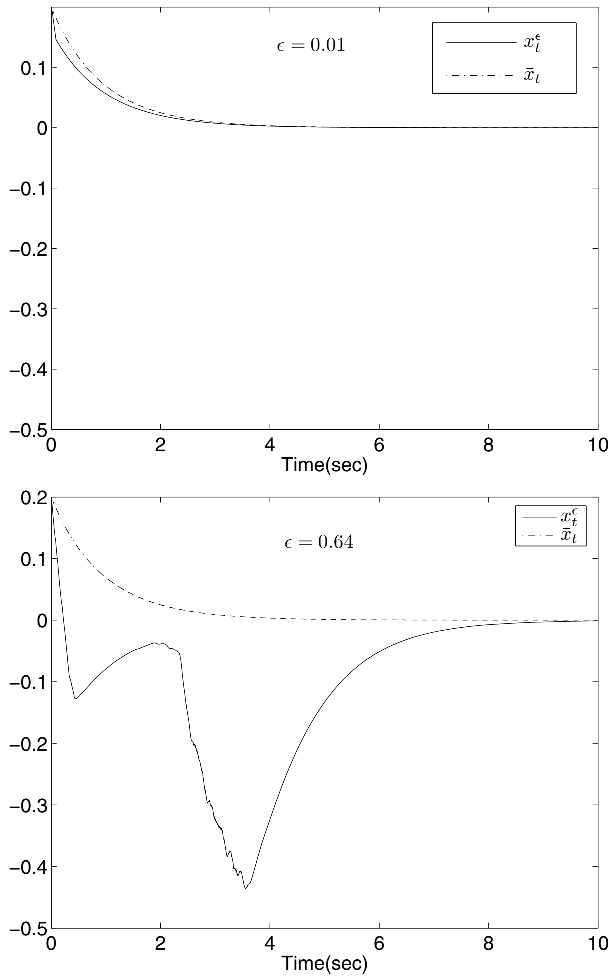

FIG. 2. *States of the original and average systems for system* (5.3). *Top: for* $\epsilon = 0.01$, *which is small (the average approximation is tight). Bottom: for* $\epsilon = 0.64$, *which is large (the average approximation is qualitatively correct, but it is not very accurate since the condition on the smallness of* ϵ *in Corollary* 3.4 *and Theorem* 3.5 *is not met).*

We know that the stochastic process $(Y(t), t \geq 0)$ is ϕ-mixing with an exponential mixing rate and exponentially ergodic with invariant distribution $\mu(dS) = \frac{l(S)}{2\pi}$ for any set $S \subset T$, where $T = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$, and $l(S)$ denotes the length (Lebesgue measure) of S. Corresponding to system (5.4), we have the function $a(x, y_1, y_2) = -y_2^2 x + \left(y_2 - \frac{1}{2}\right) x^2$. Noticing that $\int_T -y_2^2 \mu(dy_1, dy_2) = -\int_0^{2\pi} \sin^2(\theta) \frac{1}{2\pi} d\theta = -\frac{1}{2}$, and $\int_T \left(y_2 - \frac{1}{2}\right) \mu(dy_1, dy_2) = \int_0^{2\pi} \left(\sin\theta - \frac{1}{2}\right) \frac{1}{2\pi} d\theta = -\frac{1}{2}$, we obtain the average system of (5.4) as $\dot{\bar{x}}_t = -\left(\bar{x}_t + \bar{x}_t^2\right)/2$, which is locally exponentially stable at $\bar{x}_t = 0$. Figure 3 shows the simulation results with $\bar{x}_0 = x_0^\epsilon = 0.1$, $\epsilon = 0.64$, $Y_0 = [1, 0]^T$, from which we can see that the solution $x_t^\epsilon \equiv 0$ of the system (5.4) is asymptotically stable (in probability) (see (3.14) and (3.15) in Theorem 3.10).

6. Conclusions. We developed several basic theorems of stochastic infinite-time averaging for a class of nonlinear systems with uniform strong ergodic stochastic per-

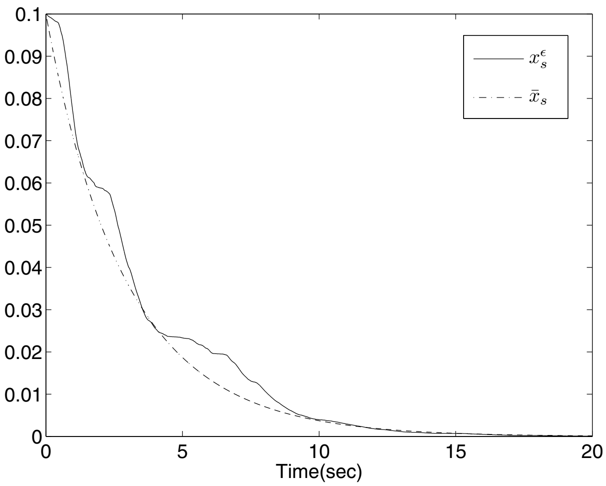

FIG. 3. *States of the original and average systems for system* (5.4), (5.5) *to illustrate Theorem* 3.10.

turbations and ϕ-mixing perturbations. For the former class, under the condition of exponential stability of average equilibrium, the original system is exponentially practically stable in probability. For the latter class, under the condition of exponential stability of average equilibrium, which is also an equilibrium of the original system, the original system is asymptotically stable in probability. This is the first work on infinite-time stochastic averaging for locally (rather than globally) Lipschitz systems and represents an extension of the deterministic general averaging for systems with aperiodic vector fields.

Appendix A. Some properties of *p*-limit and *p*-infinitesimal operator. Let $\mathcal{F}_t^\epsilon = \sigma\{X_s^\epsilon, Y_{s/\epsilon}, 0 \le s \le t\} = \sigma\{Y_{s/\epsilon}, 0 \le s \le t\} = \sigma\{Y_s, 0 \le s \le \frac{t}{\epsilon}\}$, and let E_t^ϵ denote the expectation conditioning on $\mathcal{F}_t^\epsilon$. Let $\mathcal{M}^\epsilon$ be the linear space of real-valued processes $f(t,\omega) \triangleq f(t)$ progressively measurable with respect to $\{\mathcal{F}_t^\epsilon\}$ such that $f(t)$ has a finite expectation for all t, and let $\overline{\mathcal{M}}^\epsilon$ be one subspace of $\mathcal{M}^\epsilon$ defined by $\overline{\mathcal{M}}^\epsilon = \left\{f \in \mathcal{M}^\epsilon : \sup_{t\ge 0} E|f(t)| < \infty\right\}$. A function f is said to be p-right continuous (or right continuous in the mean) if for each t, $E|f(t+\delta) - f(t)| \to 0$ as $\delta \downarrow 0$ and $\sup_{t\ge 0} E|f(t)| < \infty$. Following [17, 27], we define the *p-limit* and the *p-infinitesimal operator* $\hat{\mathcal{A}}^\epsilon$ as follows. Let $f, f^\delta \in \overline{\mathcal{M}}^\epsilon$ for each $\delta > 0$. Then we say that $f = p\text{-}\lim_{\delta\to 0} f^\delta$ if $\sup_{t,\delta} E|f^\delta(t)| < \infty$ and $\lim_{\delta\to 0} E|f^\delta(t) - f(t)| = 0$ for each t. We say that $f \in \mathcal{D}(\hat{\mathcal{A}}^\epsilon)$, the domain of $\hat{\mathcal{A}}^\epsilon$, and $\hat{\mathcal{A}}^\epsilon f = g$ if f and g are in $\overline{\mathcal{M}}^\epsilon$, and

$$p\text{-}\lim_{\delta\to 0} \frac{E_t^\epsilon[f(t+\delta)] - f(t)}{\delta} = g(t).$$

For our need, the most useful properties of $\hat{\mathcal{A}}^\epsilon$ are given by the following theorem.

THEOREM A.1 (Kurtz [17]). *Let $f(\cdot) \in \mathcal{D}(\hat{\mathcal{A}}^\epsilon)$. Then $M_\epsilon^f(t) = f(t) - f(0) - \int_0^t \hat{\mathcal{A}}^\epsilon f(u)du$ is a zero-mean martingale with respect to $\{\mathcal{F}_t^\epsilon\}$, and $E_t^\epsilon[f(t+s)] - f(t) = \int_t^{t+s} E_t^\epsilon\left[\hat{\mathcal{A}}^\epsilon f(u)\right] du$ a.s. Furthermore, if τ and σ are bounded $\{\mathcal{F}_t^\epsilon\}$ stopping times and each takes only countably many values and $\sigma \ge \tau$, then $E_\tau^\epsilon[f(\sigma)] - f(\tau) = E_\tau^\epsilon\left[\int_\tau^\sigma \hat{\mathcal{A}}^\epsilon f(u)du\right]$. If $f(\cdot)$ is right continuous a.s., we can drop the "countability" requirement.*

Appendix B. Auxiliary proofs for section 4.2.

LEMMA B.1 (Kushner [18, Lemma 4.4]). *Let $\xi(\cdot)$ be a ϕ-mixing process. Let $\mathcal{F}_0^t = \sigma\{\xi(s) : 0 \le s \le t\}$, $\mathcal{F}_t^\infty = \sigma\{\xi(s) : s \ge t\}$. Suppose that $h(t)$ is bounded with bound $K > 0$ and measurable on $\mathcal{F}_t^\infty$. Then $|E[h(t+s)|\mathcal{F}_0^t] - E[h(t+s)]| \le K\phi(s)$.*

LEMMA B.2. $g_\delta^\epsilon(t) \in \overline{\mathcal{M}}_\delta^\epsilon$.

Proof. By (4.41),

$$\tilde{G}(x,y) = G(x,y) - \bar{G}(x) = \left(\frac{\partial V(x)}{\partial x}\right)^T (a(x,y) - \bar{a}(x)). \tag{B.1}$$

Then we have that

$$\frac{\partial \tilde{G}(x,y)}{\partial x} = \left(\frac{\partial^2 V(x)}{\partial x^2}\right)^T (a(x,y) - \bar{a}(x)) + \left(\frac{\partial a(x,y)}{\partial x} - \frac{\partial \bar{a}(x)}{\partial x}\right)^T \frac{\partial V(x)}{\partial x}. \tag{B.2}$$

By (B.2), (3.12), (4.33), (4.33), (4.32), (3.9), and (3.11), we get that there exists $C_\delta > 0$ such that for any $x \in D_{\delta+1} = \{x' \in \mathbb{R}^n : |x'| \le \delta + 1\}$ and any $y \in S_Y$,

$$\left|\frac{\partial \tilde{G}(x,y)}{\partial x}\right| \le C_\delta. \tag{B.3}$$

First, we prove that for any $x = [x_1, \ldots, x_n] \in D_\delta$, $t \ge 0$, and $s \ge 0$,

$$\frac{\partial E_t^\epsilon[\tilde{G}(x, Y_{s/\epsilon})]}{\partial x} = E_t^\epsilon\left[\frac{\partial \tilde{G}(x, Y_{s/\epsilon})}{\partial x}\right]. \tag{B.4}$$

Without loss of generality, we need only prove that $\frac{\partial E_t^\epsilon[\tilde{G}(x,Y_{s/\epsilon})]}{\partial x_1} = E_t^\epsilon\left[\frac{\partial \tilde{G}(x,Y_{s/\epsilon})}{\partial x_1}\right]$. The proofs about the partial derivatives with respect to $x_2, \ldots, x_n$ are similar. By linearity of conditional expectation, the differential mean value theorem, and the dominated convergence theorem for conditional expectation (cf. (B.3)), we obtain

$$\begin{aligned}
&\frac{\partial E_t^\epsilon[\tilde{G}(x, Y_{s/\epsilon})]}{\partial x_1} \\
&= \lim_{\Delta x_1 \to 0} \frac{E_t^\epsilon[\tilde{G}(x_1 + \Delta x_1, x_2, \ldots, x_n, Y_{s/\epsilon})] - E_t^\epsilon[\tilde{G}(x_1, x_2, \ldots, x_n, Y_{s/\epsilon})]}{\Delta x_1} \\
&= \lim_{\Delta x_1 \to 0} E_t^\epsilon\left[\frac{\partial \tilde{G}}{\partial x_1}(x_1 + \theta\Delta x_1, x_2, \ldots, x_n, Y_{s/\epsilon})\right] \qquad (\text{where } 0 < \theta < 1) \\
&= E_t^\epsilon\left[\lim_{\Delta x_1 \to 0} \frac{\partial \tilde{G}}{\partial x_1}(x_1 + \theta\Delta x_1, x_2, \ldots, x_n, Y_{s/\epsilon})\right] \\
&= E_t^\epsilon\left[\frac{\partial \tilde{G}}{\partial x_1}(x_1, x_2, \ldots, x_n, Y_{s/\epsilon})\right];
\end{aligned} \tag{B.5}$$

i.e., $\frac{\partial E_t^\epsilon[\tilde{G}(x,Y_{s/\epsilon})]}{\partial x_1} = E_t^\epsilon\left[\frac{\partial \tilde{G}(x,Y_{s/\epsilon})}{\partial x_1}\right]$ holds. For simplicity, we denote

$$Q(x,y) = \left(\frac{\partial^2 V(x)}{\partial x^2}\right)^T a(x,y) + \left(\frac{\partial a(x,y)}{\partial x}\right)^T \frac{\partial V(x)}{\partial x}. \tag{B.6}$$

Then we have that

(B.7)

$$\int_{S_Y} Q(x,y)\mu(dy) = \left(\frac{\partial^2 V(x)}{\partial x^2}\right)^T \int_{S_Y} a(x,y)\mu(dy) + \left(\int_{S_Y} \frac{\partial a(x,y)}{\partial x}\mu(dy)\right)^T \frac{\partial V(x)}{\partial x}$$
$$= \left(\frac{\partial^2 V(x)}{\partial x^2}\right)^T \bar{a}(x) + \left(\frac{\partial \bar{a}(x)}{\partial x}\right)^T \frac{\partial V(x)}{\partial x},$$

where in the last equality we used $\int_{S_Y} \frac{\partial a(x,y)}{\partial x}\mu(dy) = \frac{\partial}{\partial x}\int_{S_Y} a(x,y)\mu(dy)$, which can be proved by following the deduction in (B.5). By (B.6), (3.12), (4.33), (4.32), and (3.11), we get that for any $x \in \mathbb{R}^n$ with $|x| \le \delta$, and $y \in S_Y$,

$$|Q(x,y)| \le (c_3 + c_4)k_\delta |x|. \tag{B.8}$$

By (B.4), (B.2), (B.6), (B.7), the fact that $\mathcal{F}_t^\epsilon = \mathcal{F}_{t/\epsilon}^Y$, (B.8), Lemma B.1, (4.33), and (4.34), we obtain that for any $x \in D_\delta$,

(B.9)

$$\left|\int_{\tau_\delta^\epsilon(t)}^{\tau_\delta^\epsilon} \left[\frac{\partial E_t^\epsilon[\tilde{G}(x,Y_{s/\epsilon})]}{\partial x}\right]^T a(x,Y_{t/\epsilon})ds\right| \le \int_{\tau_\delta^\epsilon(t)}^{\tau_\delta^\epsilon} \left|\left[\frac{\partial E_t^\epsilon[\tilde{G}(x,Y_{s/\epsilon})]}{\partial x}\right]^T a(x,Y_{t/\epsilon})\right| ds$$
$$= \int_{\tau_\delta^\epsilon(t)}^{\tau_\delta^\epsilon} \left|E_t^\epsilon \left[\frac{\partial \tilde{G}(x,Y_{s/\epsilon})}{\partial x}\right]^T a(x,Y_{t/\epsilon})\right| ds = \epsilon \int_{\frac{\tau_\delta^\epsilon(t)}{\epsilon}}^{\frac{\tau_\delta^\epsilon}{\epsilon}} \left|E_t^\epsilon \left[\frac{\partial \tilde{G}(x,Y_u)}{\partial x}\right]^T a(x,Y_{t/\epsilon})\right| du$$
$$= \epsilon \int_{\frac{\tau_\delta^\epsilon(t)}{\epsilon}}^{\frac{\tau_\delta^\epsilon}{\epsilon}} \left|E_t^\epsilon \left[Q(x,Y_u) - \int_{S_Y} Q(x,y)\mu(dy)\right]\right| \left|a(x,Y_{t/\epsilon})\right| du$$
$$= \epsilon \int_{\frac{\tau_\delta^\epsilon(t)}{\epsilon}}^{\frac{\tau_\delta^\epsilon}{\epsilon}} \left|E_t^\epsilon \left[Q(x,Y_u) - \int_{S_Y} Q(x,y)(P_u(dy) - P_u(dy) + \mu(dy))\right]\right| \left|a(x,Y_{t/\epsilon})\right| du$$
$$\le \epsilon \int_{\frac{\tau_\delta^\epsilon(t)}{\epsilon}}^{\frac{\tau_\delta^\epsilon}{\epsilon}} \left|E\left[Q(x,Y_u)|\mathcal{F}_{t/\epsilon}^Y\right] - E[Q(x,Y_u)]\right| \left|a(x,Y_{t/\epsilon})\right| du$$
$$+ \epsilon \int_{\frac{\tau_\delta^\epsilon(t)}{\epsilon}}^{\frac{\tau_\delta^\epsilon}{\epsilon}} \left|\int_{S_Y} Q(x,y)(P_u(dy) - \mu(dy))\right| \left|a(x,Y_{t/\epsilon})\right| du$$
$$\le \epsilon(c_3+c_4)k_\delta|x| \cdot k_\delta|x| \int_{\frac{\tau_\delta^\epsilon(t)}{\epsilon}}^{\frac{\tau_\delta^\epsilon}{\epsilon}} \phi\left(u - \frac{\tau_\delta^\epsilon(t)}{\epsilon}\right) du$$
$$+ \epsilon\sqrt{2c_5}(c_3+c_4)k_\delta|x| \cdot k_\delta|x| \int_{\frac{\tau_M^\epsilon(t)}{\epsilon}}^{\tau_M^\epsilon} e^{-\frac{\alpha}{2}u} du$$
$$\le \epsilon C_2(\delta)|x|^2 \qquad \text{(see (4.35), (4.36), (4.37), (4.38))},$$

where $C_2(\delta) = \frac{c_6(c_3+c_4)k_\delta^2}{\beta} + \frac{2\sqrt{2c_5}(c_3+c_4)k_\delta^2}{\alpha}$. Hence, by (4.42), (B.9), (4.41), (3.11), (4.33),

$$\sup_{t\ge 0} E[|g_\delta^\epsilon(t)|] \le \sup_{t\ge 0} E\left[I_{\{t<\tau_\delta^\epsilon\}} \cdot \left(|\bar{G}(X_t^\epsilon)| + \epsilon C_2(\delta)|X_t^\epsilon|^2\right)\right] \tag{B.10}$$
$$\le \sup_{t\ge 0} E\left[\sup_{|x|\le\delta} \left\{\left|\left(\frac{\partial V(x)}{\partial x}\right)^T \bar{a}(x)\right| + \epsilon C_2(\delta)|x|^2\right\}\right]$$

$$\leq \sup_{|x|\leq\delta}\left\{c_3k_\delta|x|^2+\epsilon C_2(\delta)|x|^2\right\}\leq (c_3k_\delta+\epsilon C_2(\delta))\delta^2<\infty,$$

and thus $g_\delta^\epsilon(t)\in\overline{\mathcal{M}}_\delta^\epsilon$. ∎

LEMMA B.3. $p\text{-}\lim_{\delta'\downarrow 0}\frac{E_t^\epsilon[V^\epsilon(X^\epsilon_{\tau_\delta^\epsilon(t+\delta')},t+\delta')]-V^\epsilon(X^\epsilon_{\tau_\delta^\epsilon(t)},t)}{\delta'}=g_\delta^\epsilon(t)$.

Proof. We prove a stronger result:

$$\lim_{\delta'\downarrow 0}\frac{E_t^\epsilon[V^\epsilon(X^\epsilon_{\tau_\delta^\epsilon(t+\delta')},t+\delta')]-V^\epsilon(X^\epsilon_{\tau_\delta^\epsilon(t)},t)}{\delta'}=g_\delta^\epsilon(t)\quad\text{a.s.,}\tag{B.11}$$

from which the statement of the lemma follows. Denote $(\frac{\partial V(x)}{\partial x})^T|_{x=X^\epsilon_{\tau_\delta^\epsilon(t)}}$ by $V_x^T(X^\epsilon_{\tau_\delta^\epsilon(t)})$. By (4.30), (4.31), (B.1), and the definition of $V^\epsilon(X^\epsilon_{\tau_\delta^\epsilon(t)},t)$, the property of conditional expectation, we have that

$$\begin{aligned}
&\frac{E_t^\epsilon[V^\epsilon(X^\epsilon_{\tau_\delta^\epsilon(t+\delta')},t+\delta')]-V^\epsilon(X^\epsilon_{\tau_\delta^\epsilon(t)},t)}{\delta'}\\
&=\frac{1}{\delta'}\Bigg\{E_t^\epsilon\Bigg[V(X^\epsilon_{\tau_\delta^\epsilon(t+\delta')})\\
&\qquad+\int_{\tau_\delta^\epsilon(t+\delta')}^{\tau_\delta^\epsilon}V_x(X^\epsilon_{\tau_\delta^\epsilon(t+\delta')})E^\epsilon_{t+\delta'}\left[a(X^\epsilon_{\tau_\delta^\epsilon(t+\delta')},Y_{s/\epsilon})-\bar a(X^\epsilon_{\tau_\delta^\epsilon(t+\delta')})\right]ds\Bigg]\\
&\qquad-\left[V(X^\epsilon_{\tau_\delta^\epsilon(t)})+\int_{\tau_\delta^\epsilon(t)}^{\tau_\delta^\epsilon}V_x(X^\epsilon_{\tau_\delta^\epsilon(t)})E_t^\epsilon\left[a(X^\epsilon_{\tau_\delta^\epsilon(t)},Y_{s/\epsilon})-\bar a(X^\epsilon_{\tau_\delta^\epsilon(t)})\right]ds\right]\Bigg\}\\
&=\frac{1}{\delta'}\left\{E_t^\epsilon[V(X^\epsilon_{\tau_\delta^\epsilon(t+\delta')})]-V(X^\epsilon_{\tau_\delta^\epsilon(t)})\right\}\\
&\quad-\frac{1}{\delta'}\int_{\tau_\delta^\epsilon(t)}^{\tau_\delta^\epsilon(t+\delta')}V_x(X^\epsilon_{\tau_\delta^\epsilon(t)})E_t^\epsilon\left[a(X^\epsilon_{\tau_\delta^\epsilon(t)},Y_{s/\epsilon})-\bar a(X^\epsilon_{\tau_\delta^\epsilon(t)})\right]ds\\
&\quad+\frac{1}{\delta'}\int_{\tau_\delta^\epsilon(t+\delta')}^{\tau_\delta^\epsilon}\Bigg\{E_t^\epsilon\Bigg[V_x(X^\epsilon_{\tau_\delta^\epsilon(t+\delta')})\left(a(X^\epsilon_{\tau_\delta^\epsilon(t+\delta')},Y_{s/\epsilon})-\bar a(X^\epsilon_{\tau_\delta^\epsilon(t+\delta')})\right)\\
&\qquad\qquad-V_x(X^\epsilon_{\tau_\delta^\epsilon(t)})\left(a(X^\epsilon_{\tau_\delta^\epsilon(t)},Y_{s/\epsilon})-\bar a(X^\epsilon_{\tau_\delta^\epsilon(t)})\right)\Bigg]\Bigg\}ds\\
&=\frac{1}{\delta'}\left\{E_t^\epsilon[V(X^\epsilon_{\tau_\delta^\epsilon(t+\delta')})]-V(X^\epsilon_{\tau_\delta^\epsilon(t)})\right\}-\frac{1}{\delta'}\int_{\tau_\delta^\epsilon(t)}^{\tau_\delta^\epsilon(t+\delta')}E_t^\epsilon\left[\tilde G(X^\epsilon_{\tau_\delta^\epsilon(t)},Y_{s/\epsilon})\right]ds\\
&\quad+\frac{1}{\delta'}\int_{\tau_\delta^\epsilon(t+\delta')}^{\tau_\delta^\epsilon}E_t^\epsilon\left[\tilde G(X^\epsilon_{\tau_\delta^\epsilon(t+\delta')},Y_{s/\epsilon})-\tilde G(X^\epsilon_{\tau_\delta^\epsilon(t)},Y_{s/\epsilon})\right]ds\\
&\triangleq g_1^{\epsilon,\delta}(t,\delta')-g_2^{\epsilon,\delta}(t,\delta')+g_3^{\epsilon,\delta}(t,\delta').
\end{aligned}\tag{B.12}$$

Following the proof of (B.5), we get

$$\begin{aligned}
\lim_{\delta'\downarrow 0}g_1^{\epsilon,\delta}(t,\delta')&=\lim_{\delta'\downarrow 0}\frac{1}{\delta'}\left\{E_t^\epsilon[V(X^\epsilon_{\tau_\delta^\epsilon(t+\delta')})]-V(X^\epsilon_{\tau_\delta^\epsilon(t)})\right\}\\
&=\lim_{\delta'\downarrow 0}E_t^\epsilon\left[\frac{V_x^T\left(X^\epsilon_{\tau_\delta^\epsilon(t)}+\theta\left(X^\epsilon_{\tau_\delta^\epsilon(t+\delta')}-X^\epsilon_{\tau_\delta^\epsilon(t)}\right)\right)\left(X^\epsilon_{\tau_\delta^\epsilon(t+\delta')}-X^\epsilon_{\tau_\delta^\epsilon(t)}\right)}{\delta}\right]
\end{aligned}\tag{B.13}$$

$$= \lim_{\delta'\downarrow 0} E_t^\epsilon \left[\frac{V_x^T \left(X^\epsilon_{\tau^\epsilon_\delta(t)} + \theta \left(X^\epsilon_{\tau^\epsilon_\delta(t+\delta')} - X^\epsilon_{\tau^\epsilon_\delta(t)} \right) \right) \int_{\tau^\epsilon_\delta(t)}^{\tau^\epsilon_\delta(t+\delta')} a(X_u^\epsilon, Y_{u/\epsilon}) du}{\delta} \right]$$

$$= \lim_{\delta'\downarrow 0} E_t^\epsilon \left[\frac{V_x^T \left(X^\epsilon_{\tau^\epsilon_\delta(t)} + \theta \left(X^\epsilon_{\tau^\epsilon_\delta(t+\delta')} - X^\epsilon_{\tau^\epsilon_\delta(t)} \right) \right) \int_t^{t+\delta'} a(X_u^\epsilon, Y_{u/\epsilon}) I_{\{u<\tau^\epsilon_\delta\}} du}{\delta} \right]$$

$$= V_x^T(X^\epsilon_{\tau^\epsilon_\delta(t)}) a(X_t^\epsilon, Y_{t/\epsilon}) \cdot I_{\{t<\tau^\epsilon_\delta\}} = V_x^T(X_t^\epsilon) a(X_t^\epsilon, Y_{t/\epsilon}) \cdot I_{\{t<\tau^\epsilon_\delta\}} \quad \text{a.s.},$$

$$\text{(B.14)} \qquad \lim_{\delta'\downarrow 0} g_2^{\epsilon,\delta}(t,\delta') = \lim_{\delta'\downarrow 0} \frac{1}{\delta'} \int_{\tau^\epsilon_\delta(t)}^{\tau^\epsilon_\delta(t+\delta')} E_t^\epsilon \left[\tilde{G}(X^\epsilon_{\tau^\epsilon_\delta(t)}, Y_{s/\epsilon}) \right] ds$$

$$= \lim_{\delta'\downarrow 0} \frac{1}{\delta'} \int_{\tau^\epsilon_\delta \wedge t}^{\tau^\epsilon_\delta \wedge (t+\delta')} E_t^\epsilon \left[\tilde{G}(X^\epsilon_{\tau^\epsilon_\delta(t)}, Y_{s/\epsilon}) \right] ds$$

$$= \lim_{\delta'\downarrow 0} \frac{1}{\delta'} \int_t^{t+\delta'} E_t^\epsilon \left[\tilde{G}(X^\epsilon_{\tau^\epsilon_\delta(t)}, Y_{s/\epsilon}) \right] I_{\{s<\tau^\epsilon_\delta\}} ds$$

$$= \tilde{G}(X^\epsilon_{\tau^\epsilon_\delta(t)}, Y_{t/\epsilon}) I_{\{t<\tau^\epsilon_\delta\}} = \tilde{G}(X_t^\epsilon, Y_{t/\epsilon}) I_{\{t<\tau^\epsilon_\delta\}} \quad \text{a.s.}$$

Following the proof of (B.13) and by (B.4), we get that

$$\lim_{\delta'\downarrow 0} g_3^{\epsilon,\delta}(t,\delta) = \lim_{\delta'\downarrow 0} \frac{1}{\delta'} \int_{\tau^\epsilon_\delta(t+\delta')}^{\tau^\epsilon_\delta} E_t^\epsilon \left[\tilde{G}(X^\epsilon_{\tau^\epsilon_\delta(t+\delta')}, Y_{s/\epsilon}) - \tilde{G}(X^\epsilon_{\tau^\epsilon_\delta(t)}, Y_{s/\epsilon}) \right] ds$$

$$= \lim_{\delta'\downarrow 0} \int_{\tau^\epsilon_\delta(t+\delta')}^{\tau^\epsilon_\delta} E_t^\epsilon \left[\frac{\tilde{G}(X^\epsilon_{\tau^\epsilon_\delta(t+\delta')}, Y_{s/\epsilon}) - \tilde{G}(X^\epsilon_{\tau^\epsilon_\delta(t)}, Y_{s/\epsilon})}{\delta'} \right] ds$$

$$= \lim_{\delta'\downarrow 0} \int_{\tau^\epsilon_\delta(t+\delta')}^{\tau^\epsilon_\delta} E_t^\epsilon \left[\frac{\tilde{G}_x^T \left(X^\epsilon_{\tau^\epsilon_\delta(t)} + \theta(X^\epsilon_{\tau^\epsilon_\delta(t+\delta')} - X^\epsilon_{\tau^\epsilon_\delta(t)}), Y_{s/\epsilon} \right) (X^\epsilon_{\tau^\epsilon_\delta(t+\delta')} - X^\epsilon_{\tau^\epsilon_\delta(t)})}{\delta'} \right] ds$$

$$= \lim_{\delta'\downarrow 0} \int_{\tau^\epsilon_\delta(t+\delta')}^{\tau^\epsilon_\delta} E_t^\epsilon \left[\frac{\tilde{G}_x^T \left(X^\epsilon_{\tau^\epsilon_\delta(t)} + \theta(X^\epsilon_{\tau^\epsilon_\delta(t+\delta')} - X^\epsilon_{\tau^\epsilon_\delta(t)}), Y_{s/\epsilon} \right)}{\delta'} \cdot \int_t^{t+\delta'} a(X_u^\epsilon, Y_{u/\epsilon}) I_{\{u<\tau^\epsilon_\delta\}} du \right] ds$$

$$= \int_{\tau^\epsilon_\delta(t)}^{\tau^\epsilon_\delta} E_t^\epsilon \left[\tilde{G}_x^T \left(X_t^\epsilon, Y_{s/\epsilon} \right) a(X_t^\epsilon, Y_{t/\epsilon}) I_{\{t<\tau^\epsilon_\delta\}} \right] ds$$

$$= I_{\{t<\tau^\epsilon_\delta\}} \int_{\tau^\epsilon_\delta(t)}^{\tau^\epsilon_\delta} \left[\left. \frac{\partial E_t^\epsilon[\tilde{G}(x, Y_{s/\epsilon})]}{\partial x} \right|_{x=X_t^\epsilon} \right]^T a(X_t^\epsilon, Y_{t/\epsilon}) ds \quad \text{a.s.},$$

which together with (B.12)–(B.14), (4.41), and (4.42) implies that (B.11) holds. ∎

LEMMA B.4. $\hat{\mathcal{A}}_M^\epsilon \left(V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)}, t) \cdot I_{\{t<\tau^\epsilon_M\}} \right) = g_M^\epsilon(t)$, *i.e.*,

$$p\text{-}\lim_{\delta\downarrow 0} \frac{E_t^\epsilon[V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t+\delta)}, t+\delta) \cdot I_{\{t+\delta<\tau^\epsilon_M\}}] - V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)}, t) \cdot I_{\{t<\tau^\epsilon_M\}}}{\delta} = g_M^\epsilon(t).$$

Proof. As in the proof of Lemma B.3, we prove

$$\lim_{\delta\downarrow 0}\frac{E_t^\epsilon[V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t+\delta)},t+\delta)I_{\{t+\delta<\tau^\epsilon_M\}}]-V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)},t)I_{\{t<\tau^\epsilon_M\}}}{\delta}=g^\epsilon_M(t)\quad\text{a.s.}\tag{B.15}$$

Denote $(\frac{\partial V(x)}{\partial x})^T|_{x=X^\epsilon_{\tau^\epsilon_M(t)}}$ by $V_x^T(X^\epsilon_{\tau^\epsilon_M(t)})$. By the definition of $V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)},t)$, following the proof of Lemma B.3, we get that

$$\begin{aligned}
&\frac{E_t^\epsilon[V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t+\delta)},t+\delta)I_{\{t+\delta<\tau^\epsilon_M\}}]-V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)},t)I_{\{t<\tau^\epsilon_M\}}}{\delta}\\
&=\frac{1}{\delta}\left\{E_t^\epsilon[V(X^\epsilon_{\tau^\epsilon_M(t+\delta)})\cdot I_{\{t+\delta<\tau^\epsilon_M\}}]-V(X^\epsilon_{\tau^\epsilon_M(t)})\cdot I_{\{t<\tau^\epsilon_M\}}\right\}\\
&\quad-\frac{1}{\delta}\int_{\tau^\epsilon_M(t)}^{\tau^\epsilon_M(t+\delta)}E_t^\epsilon\left[\tilde{G}(X^\epsilon_{\tau^\epsilon_M(t)},Y_{s/\epsilon})\right]ds\\
&\quad+\frac{1}{\delta}\int_{\tau^\epsilon_M(t+\delta)}^{\tau^\epsilon_M}E_t^\epsilon\left[\tilde{G}(X^\epsilon_{\tau^\epsilon_M(t+\delta)},Y_{s/\epsilon})-\tilde{G}(X^\epsilon_{\tau^\epsilon_M(t)},Y_{s/\epsilon})\right]ds\\
&\triangleq\bar{g}_1^{\epsilon,M}(t,\delta)-g_2^{\epsilon,M}(t,\delta)+g_3^{\epsilon,M}(t,\delta),
\end{aligned}$$

where the functions $g_2^{\epsilon,M}(\cdot,\cdot)$ and $g_3^{\epsilon,M}(\cdot,\cdot)$ are the same as the corresponding ones in (B.12) with δ replaced by M. And so we need only to consider $\bar{g}_1^{\epsilon,M}(t,\delta)$. Following the proof of (B.13), we get that

$$\begin{aligned}
&\lim_{\delta\downarrow 0}\bar{g}_1^{\epsilon,M}(t,\delta)=\lim_{\delta\downarrow 0}\frac{1}{\delta}\left\{E_t^\epsilon[V(X^\epsilon_{\tau^\epsilon_M(t+\delta)})I_{\{t+\delta<\tau^\epsilon_M\}}]-V(X^\epsilon_{\tau^\epsilon_M(t)})I_{\{t<\tau^\epsilon_M\}}\right\}\\
&=\lim_{\delta\downarrow 0}E_t^\epsilon\left[\frac{V(X^\epsilon_{\tau^\epsilon_M(t+\delta)})I_{\{t+\delta<\tau^\epsilon_M\}}-V(X^\epsilon_{\tau^\epsilon_M(t)})I_{\{t<\tau^\epsilon_M\}}}{\delta}\right]\\
&=\lim_{\delta\downarrow 0}E_t^\epsilon\left[\frac{V(X^\epsilon_{\tau^\epsilon_M(t+\delta)})\left(I_{\{t+\delta<\tau^\epsilon_M\}}-I_{\{t<\tau^\epsilon_M\}}\right)}{\delta}\right]\\
&\quad+\lim_{\delta\downarrow 0}E_t^\epsilon\left[\frac{\left(V(X^\epsilon_{\tau^\epsilon_M(t+\delta)})-V(X^\epsilon_{\tau^\epsilon_M(t)})\right)I_{\{t<\tau^\epsilon_M\}}}{\delta}\right]\\
&=0+\lim_{\delta\downarrow 0}E_t^\epsilon\left[\frac{\left(V(X^\epsilon_{\tau^\epsilon_M(t+\delta)})-V(X^\epsilon_{\tau^\epsilon_M(t)})\right)I_{\{t<\tau^\epsilon_M\}}}{\delta}\right]\\
&=V_x^T(X^\epsilon_{\tau^\epsilon_M(t)})a(X_t^\epsilon,Y_{t/\epsilon})\cdot I_{\{t<\tau^\epsilon_M\}}I_{\{t<\tau^\epsilon_M\}}\\
&=V_x^T(X_t^\epsilon)a(X_t^\epsilon,Y_{t/\epsilon})I_{\{t<\tau^\epsilon_M\}}=\lim_{\delta\downarrow 0}g_1^{\epsilon,M}(t,\delta).
\end{aligned}$$

Hence by the proof of Lemma B.3, we get that (B.15) holds. ∎

LEMMA B.5. *M_t^ϵ is a martingale relative to $\{\mathcal{F}_t^\epsilon\}$.*

Proof. For any $s,t\geq 0$, by (4.59), the property of conditional expectation, and $\hat{\mathcal{A}}^\epsilon_M\left(V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)},t)\cdot I_{\{t<\tau^\epsilon_M\}}\right)=\hat{\mathcal{A}}^\epsilon_M V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)},t)$ (see Lemma B.4), we have that

$$
\begin{aligned}
\text{(B.16)}\quad & E[M^\epsilon_{t+s}-M^\epsilon_t|\mathcal{F}^\epsilon_t] \\
&= E\Bigg[e^{2\hat\gamma(t+s)}V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t+s)},t+s)I_{\{t+s<\tau^\epsilon_M\}}-e^{2\hat\gamma t}V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)},t)I_{\{t<\tau^\epsilon_M\}} \\
&\qquad -\int_t^{t+s}e^{2\hat\gamma u}(\hat{\mathcal{A}}^\epsilon_M+2\hat\gamma)\left(V^\epsilon(X^\epsilon_{\tau^\epsilon_M(u)},u)I_{\{u<\tau^\epsilon_M\}}\right)du\,\Bigg|\,\mathcal{F}^\epsilon_t\Bigg] \\
&\quad + E\left[e^{2\hat\gamma\tau^\epsilon_M}V(X^\epsilon_{\tau^\epsilon_M})I_{\{\tau^\epsilon_M\le t+s\}}-e^{2\hat\gamma\tau^\epsilon_M}V(X^\epsilon_{\tau^\epsilon_M})I_{\{\tau^\epsilon_M\le t\}}\,\middle|\,\mathcal{F}^\epsilon_t\right] \\
&= \Bigg\{E\left[e^{2\hat\gamma(t+s)}V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t+s)},t+s)\,\middle|\,\mathcal{F}^\epsilon_t\right]-e^{2\hat\gamma t}V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)},t) \\
&\qquad\qquad -\int_t^{t+s}E\left[e^{2\hat\gamma u}(\hat{\mathcal{A}}^\epsilon_M+2\hat\gamma)\left(V^\epsilon(X^\epsilon_{\tau^\epsilon_M(u)},u)\right)\,\middle|\,\mathcal{F}^\epsilon_t\right]du\Bigg\} \\
&\quad -\Bigg\{E\left[e^{2\hat\gamma(t+s)}V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t+s)},t+s)I_{\{t+s\ge\tau^\epsilon_M\}}\,\middle|\,\mathcal{F}^\epsilon_t\right]-e^{2\hat\gamma t}V^\epsilon(X^\epsilon_{\tau^\epsilon_M(t)},t)I_{\{t\ge\tau^\epsilon_M\}} \\
&\qquad\qquad -\int_t^{t+s}E\left[2\hat\gamma e^{2\hat\gamma u}V^\epsilon(X^\epsilon_{\tau^\epsilon_M(u)},u)I_{\{u\ge\tau^\epsilon_M\}}\,\middle|\,\mathcal{F}^\epsilon_t\right]du\Bigg\} \\
&\quad + E\left[e^{2\hat\gamma\tau^\epsilon_M}V(X^\epsilon_{\tau^\epsilon_M})I_{\{\tau^\epsilon_M\le t+s\}}-e^{2\hat\gamma\tau^\epsilon_M}V(X^\epsilon_{\tau^\epsilon_M})I_{\{\tau^\epsilon_M\le t\}}\,\middle|\,\mathcal{F}^\epsilon_t\right] \\
&\triangleq g_1(t,s,\omega)-g_2(t,s,\omega)+g_3(t,s,\omega).
\end{aligned}
$$

For $u\ge t$, define $f(u,\omega)=E\left[e^{2\hat\gamma(u)}V^\epsilon(X^\epsilon_{\tau^\epsilon_M(u)},u)|\mathcal{F}^\epsilon_t\right](\omega)$. Then for any $u\ge t$, we have (if $u=t$, we consider the right derivative)

$$
\begin{aligned}
f'(u,\omega) &= \lim_{s\to0}\frac{f(u+s,\omega)-f(u,\omega)}{s} \\
&= \lim_{s\to0}\frac{E\left[e^{2\hat\gamma(u+s)}V^\epsilon(X^\epsilon_{\tau^\epsilon_M(u+s)},u+s)|\mathcal{F}^\epsilon_t\right]-E\left[e^{2\hat\gamma(u)}V^\epsilon(X^\epsilon_{\tau^\epsilon_M(u)},u)|\mathcal{F}^\epsilon_t\right]}{s} \\
&= \lim_{s\to0}E\left[\frac{e^{2\hat\gamma(u+s)}V^\epsilon(X^\epsilon_{\tau^\epsilon_M(u+s)},u+s)-e^{2\hat\gamma u}V^\epsilon(X^\epsilon_{\tau^\epsilon_M(u)},u)}{s}\,\middle|\,\mathcal{F}^\epsilon_t\right] \\
&= \lim_{s\to0}E\left[\frac{\left(e^{2\hat\gamma(u+s)}-e^{2\hat\gamma u}\right)V^\epsilon(X^\epsilon_{\tau^\epsilon_M(u+s)},u+s)}{s}\,\middle|\,\mathcal{F}^\epsilon_t\right] \\
&\quad +\lim_{s\to0}E\left[\frac{e^{2\hat\gamma u}\left(V^\epsilon(X^\epsilon_{\tau^\epsilon_M(u+s)},u+s)-V^\epsilon(X^\epsilon_{\tau^\epsilon_M(u)},u)\right)}{s}\,\middle|\,\mathcal{F}^\epsilon_t\right] \\
&= E\left[e^{2\hat\gamma u}(\hat{\mathcal{A}}^\epsilon_M+2\hat\gamma)\left(V^\epsilon(X^\epsilon_{\tau^\epsilon_M(u)},u)\right)\,\middle|\,\mathcal{F}^\epsilon_t\right],
\end{aligned}
$$

and thus

$$
g_1(t,s,\omega)=f(t+s,\omega)-f(t,\omega)-\int_t^{t+s}f'(u,\omega)du=0\quad\text{a.s.}\tag{B.17}
$$

By the definitions of τ^ϵ_M and $V^\epsilon(x,t)$, we have

$$
\begin{aligned}
\text{(B.18)}\quad g_2(t,s,\omega) &= E\left[e^{2\hat\gamma(t+s)}V(X^\epsilon_{\tau^\epsilon_M})I_{\{t+s\ge\tau^\epsilon_M\}}\,\middle|\,\mathcal{F}^\epsilon_t\right]-e^{2\hat\gamma t}V(X^\epsilon_{\tau^\epsilon_M})I_{\{t\ge\tau^\epsilon_M\}} \\
&\qquad -\int_t^{t+s}E\left[2\hat\gamma e^{2\hat\gamma u}V(X^\epsilon_{\tau^\epsilon_M})I_{\{u\ge\tau^\epsilon_M\}}\,\middle|\,\mathcal{F}^\epsilon_t\right]du
\end{aligned}
$$

$$= E\left[e^{2\hat{\gamma}(t+s)}V(X^{\epsilon}_{\tau^{\epsilon}_M})I_{\{t+s\geq\tau^{\epsilon}_M\}} - e^{2\hat{\gamma}t}V(X^{\epsilon}_{\tau^{\epsilon}_M})I_{\{t\geq\tau^{\epsilon}_M\}} - \int_t^{t+s} 2\hat{\gamma}e^{2\hat{\gamma}u}V(X^{\epsilon}_{\tau^{\epsilon}_M})I_{\{u\geq\tau^{\epsilon}_M\}}du \,\middle|\, \mathcal{F}^{\epsilon}_t\right].$$

Now, we analyze the item within the conditional expectation on the right-hand side of (B.18). For simplicity, let

$$h(t,s,\omega) = e^{2\hat{\gamma}(t+s)}V(X^{\epsilon}_{\tau^{\epsilon}_M})\cdot I_{\{t+s\geq\tau^{\epsilon}_M\}} - e^{2\hat{\gamma}t}V(X^{\epsilon}_{\tau^{\epsilon}_M})\cdot I_{\{t\geq\tau^{\epsilon}_M\}} - \int_t^{t+s} 2\hat{\gamma}e^{2\hat{\gamma}u}V(X^{\epsilon}_{\tau^{\epsilon}_M})\cdot I_{\{u\geq\tau^{\epsilon}_M\}}du. \tag{B.19}$$

Case 1: $t+s<\tau^{\epsilon}_M(\omega)$. Then $h(t,s,\omega)=0$.

Case 2: $t\geq\tau^{\epsilon}_M(\omega)$. Then we have $h(t,s,\omega) = e^{2\hat{\gamma}(t+s)}V(X^{\epsilon}_{\tau^{\epsilon}_M}) - e^{2\hat{\gamma}t}V(X^{\epsilon}_{\tau^{\epsilon}_M}) - \int_t^{t+s} 2\hat{\gamma}e^{2\hat{\gamma}u}V(X^{\epsilon}_{\tau^{\epsilon}_M})du = 0$, since

$$\frac{d\left(e^{2\hat{\gamma}u}V(X^{\epsilon}_{\tau^{\epsilon}_M})\right)}{du} = 2\hat{\gamma}e^{2\hat{\gamma}u}V(X^{\epsilon}_{\tau^{\epsilon}_M}). \tag{B.20}$$

Case 3: $t<\tau^{\epsilon}_M(\omega)\leq t+s$. Then by (B.19) and (B.20), we have $h(t,s,\omega) = e^{2\hat{\gamma}(t+s)}V(X^{\epsilon}_{\tau^{\epsilon}_M}) - \int_{\tau^{\epsilon}_M}^{t+s} 2\hat{\gamma}e^{2\hat{\gamma}u}V(X^{\epsilon}_{\tau^{\epsilon}_M})du = e^{2\hat{\gamma}\tau^{\epsilon}_M}V(X^{\epsilon}_{\tau^{\epsilon}_M})$.

Hence we have $-g_2(t,s,\omega) = -E\left[h(t,s,\omega)|\mathcal{F}^{\epsilon}_t\right] = -E\left[e^{2\hat{\gamma}\tau^{\epsilon}_M}V(X^{\epsilon}_{\tau^{\epsilon}_M})I_{\{t<\tau^{\epsilon}_M\leq t+s\}} \middle| \mathcal{F}^{\epsilon}_t\right] = -E\left[e^{2\hat{\gamma}\tau^{\epsilon}_M}V(X^{\epsilon}_{\tau^{\epsilon}_M})\,I_{\{\tau^{\epsilon}_M\leq t+s\}} - e^{2\hat{\gamma}\tau^{\epsilon}_M}V(X^{\epsilon}_{\tau^{\epsilon}_M})\,I_{\{\tau^{\epsilon}_M\leq t\}}\,\middle|\,\mathcal{F}^{\epsilon}_t\right]$, which implies that $-g_2(t,s,\omega)+g_3(t,s,\omega) = E[0|\mathcal{F}^{\epsilon}_t]=0$ a.s. This together with (B.16) and (B.17) provides that $E[M^{\epsilon}_{t+s}-M^{\epsilon}_t|\mathcal{F}^{\epsilon}_t]=0$ a.s. ∎

REFERENCES

[1] K. B. Ariyur and M. Krstić, *Real-Time Optimization by Extremum Seeking Control*, John Wiley & Sons, New York, 2003.

[2] A. Benveniste, M. Métivier, and P. Priouret, *Adaptive Algorithms and Stochastic Approximations*, Springer-Verlag, Berlin, 1990.

[3] G. Blankenship and G. C. Papanicolaou, *Stability and control of stochastic systems with wide-band noise disturbances*. I, SIAM J. Appl. Math., 34 (1978), pp. 437–476.

[4] N. N. Bogoliubov and Y. A. Mitropolsky, *Asymptotic Methods in the Theory of Nonlinear Oscillation*, Gordon and Breach Science Publishers, New York, 1961.

[5] J. L. Doob, *Classical Potential Theory and Its Probabilistic Counterpart*, Springer-Verlag, New York, 1984.

[6] P. Doukhan, *Mixing: Properties and Examples*, Springer-Verlag, New York, 1994.

[7] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, Springer-Verlag, New York, 1984.

[8] S. N. Hashemi and A. J. Heunis, *Averaging principle for diffusion processes*, Stochastics Stochastics Rep., 62 (1998), pp. 201–216.

[9] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, New York, 2005.

[10] H. K. Khalil, *Nonlinear Systems*, 3rd ed., Prentice–Hall, Upper Saddle River, NJ, 2002.

[11] R. Z. Khas'minskii, *A limit theorem for the solutions of differential equations with random right-hand sides*, Theory Probab. Appl., 11 (1966), pp. 390–406.

[12] R. Z. Khas'minskii and G. Yin, *On averaging principles: An asymptotic expansion approach*, SIAM J. Math. Anal., 35 (2004), pp. 1534–1560.

[13] Y. Kifer, *Averaging in dynamical systems and large deviations*, Invent. Math., 110 (1992), pp. 337–370.

[14] V. S. Korolyuk, *Average and stability of dynamical system with rapid stochastic switching*, in Exploring Stochastic Laws, A. V. Skorokhod and Yu. V. Borovskikh, eds., Brill Academic Publishers, Leiden, The Netherlands, 1995, pp. 219–232.

[15] V. S. Korolyuk, *Stability of stochastic systems in the scheme of diffusion approximation*, Ukraïn. Mat. Zh., 50 (1998), pp. 36–47.

[16] V. S. Korolyuk and Y. M. Chabanyuk, *Stability of dynamical system with semi-Markov switches under conditions of stability of the averaged system*, Ukraïn. Mat. Zh., 54 (2002), pp. 239–252.

[17] T. G. Kurtz, *Semigroups of conditioned shifts and approximation of Markov processes*, Ann. Probability, 3 (1975), pp. 618–642.

[18] H. J. Kushner, *Approximation and Weak Convergence Method for Random Processes, with Applications to Stochastic Systems Theory*, MIT Press, Cambridge, MA, 1984.

[19] H. J. Kushner and K. M. Ramachandran, *Nearly optimal singular controls for wideband noise driven systems*, SIAM J. Control Optim., 26 (1988), pp. 569–591.

[20] H. J. Kushner and G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications*, 2nd ed., Spinger-Verlag, New York, 2003.

[21] R. S. Liptser and A. N. Shiryaev, *Theory of Martingales*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1989.

[22] R. S. Liptser and J. Stoyanov, *Stochastic version of the averaging principle for diffusion type processes*, Stochastics Stochastics Rep., 32 (1990), pp. 145–163.

[23] L. Ljung, *Analysis of recursive stochastic algorithms*, IEEE Trans. Automat. Control, 22 (1977), pp. 551–575.

[24] C. Manzie and M. Krstic, *Extremum seeking with stochastic perturbations*, IEEE Trans. Automat. Control, 54 (2009), pp. 580–585.

[25] B. Øksendal, *Stochastic Differential Equations*, 4th ed., Springer-Verlag, Berlin, 1995.

[26] G. C. Papanicolaou and W. Kohler, *Asymptotic theory of mixing stochastic ordinary differential equations*, Comm. Pure Appl. Math., 27 (1974), pp. 641–668.

[27] R. Rishel, *Necessary and sufficient dynamic programming conditions for continuous time stochastic optimal control*, SIAM J. Control, 8 (1970), pp. 559–571.

[28] J. B. Roberts and P. D. Spanos, *Stochastic averaging: An approximate method of solving random vibration problems*, Internat. J. Non-Linear Mech., 21 (1986), pp. 111–134.

[29] J. A. Sanders, F. Verhulst, and J. Murdock, *Averaging Methods in Nonlinear Dynamical Systems*, 2nd ed., Springer-Verlag, New York, 2007.

[30] S. Sastry and M. Bodson, *Adaptive Control (Stability, Convergence, and Robustness)*, Prentice–Hall, Englewood Cliffs, NJ, 1989.

[31] A. V. Skorokhod, *Asymptotic Methods in the Theory of Stochastic Differential Equations*, Transl. Math. Monogr. 78, American Mathematical Society, Providence, RI, 1989.

[32] A. V. Skorokhod, F. C. Hoppensteadt, and H. Salehi, *Random Perturbation Methods with Applications in Science and Engineering*, Springer-Verlag, New York, 2002.

[33] V. Solo and X. Kong, *Adaptive Signal Processing Algorithms: Stability and Performance*, Prentice–Hall, Englewood Cliffs, NJ, 1994.

[34] J. C. Spall, *Introduction to Stochastic Search and Optimization: Estimation, Simulation, and Control*, Wiley-Interscience, Hoboken, NJ, 2003.

[35] M. S. Stankovic and D. M. Stipanovic, *Stochastic extremum seeking with applications to mobile sensor networks*, in Proceedings of the 2009 American Control Conference, St. Louis, MO, 2009.

[36] Y. Tan, D. Nesic, and I. M. Y. Mareels, *On non-local stability properties of extremum seeking controllers*, Automatica J. IFAC, 42 (2006), pp. 889–903.

[37] G. Yin and Q. Zhang, *Near optimality of stochastic control in systems with unknown parameter processes*, Appl. Math. Optim., 29 (1994), pp. 263–284.

[38] W. Q. Zhu and Y. Q. Yang, *Stochastic averaging of quasi-non integrable-Hamiltonian systems*, Trans. ASME J. Appl. Mech., 64 (1997), pp. 157–164.