J. Math. Anal. Appl. 416 (2014) 609–628

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

Strong convergence rate in averaging principle for stochastic FitzHugh–Nagumo system with two time-scales ☆

Hongbo Fu [a], Li Wan [a], Youzhen Wang [a], Jicheng Liu [b,c,*]

[a] College of Mathematics and Computer Science, Wuhan Textile University, Wuhan, 430073, PR China
[b] School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan, 430074, PR China
[c] Department of Statistics, Purdue University, West Lafayette, IN 47907, USA

ARTICLE INFO

Article history:
Received 18 March 2013
Available online 6 March 2014
Submitted by U. Stadtmueller

Keywords:
Stochastic FitzHugh–Nagumo system
Averaging principle
Invariant measure
Strong convergence order

ABSTRACT

This article deals with averaging principle for stochastic FitzHugh–Nagumo system with different time-scales. Under suitable conditions, the existence of an averaging equation eliminating the fast variable for this coupled system is proved, and as a consequence, the system can be reduced to a single stochastic ordinary equation with a modified coefficient. Moreover, the rate of convergence for the slow component towards the solution of the averaging equation is of order 1/2.

1. Introduction

The FitzHugh–Nagumo (F–N) system [9,28] is a simplified version of the well-known Hodgkin–Huxley model [19], which describes the mechanism of the neural excitability, excitation phenomena for macro-receptors and other natural membranes. This system has attracted a lot of interest and there is extensive literature on mathematical analysis for it. The wellposedness and regularity of solutions for F–N system with inhomogeneous boundary conditions are considered in [20]. Boundedness and convergence to equilibrium for F–N system are studied in [24]. Also see Marion [27] and Shao [31] for results on long time behavior of the system and Deng [8], Jones [21] for existence and stability of traveling waves.

Quite recently much attention has also been brought to the dynamical property of F–N system with stochastic or random perturbations. The existence of random attractors for the stochastic F–N system

☆ This work was supported by China NSF Grant Nos. 11301403, 11271295, 11271013, 10901065 and Science and Technology Research Projects of Hubei Provincial Department of Education No. D20131602.

* Corresponding author at: School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan, 430074, PR China.

E-mail addresses: hbfuhust@gmail.com (H. Fu), wanlinju@yahoo.com.cn, wanlinju@aliyun.com (L. Wan), 1142207073@qq.com (Y. Wang), jcliu@hust.edu.cn (J. Liu).

http://dx.doi.org/10.1016/j.jmaa.2014.02.062
0022-247X/

defined on unbounded domain $\mathbb{R}^n$ has been studied in Wang [36]. The asymptotic limit for stochastic F–N system with small excitability has been derived by Lv and Wang [26]. The authors show that the solution of the stochastic F–N system subjected with additive noise converges in probability to that of the limit system. Other topics on stochastic F–N system, such as stochastic resonance, stochastic bifurcation and stochastic synchronization, can be found in [32,38,18,25], and the reference therein.

In this paper, we will consider the stochastic F–N system subjected with multiplicative noise that is white in time and homogeneous in space. Let $D = (0, l) \subset \mathbb{R}$ and consider the equations

$$\frac{\partial X_t^\epsilon(\xi)}{\partial t} = -\gamma X_t^\epsilon(\xi) + Y_t^\epsilon(\xi) + \sigma_1\big(X_t^\epsilon(\xi)\big)\dot{W}_t^{Q_1}, \tag{1.1}$$

$$\frac{\partial Y_t^\epsilon(\xi)}{\partial t} = \frac{1}{\epsilon}\frac{\partial^2 Y_t^\epsilon(\xi)}{\partial \xi^2} + \frac{1}{\epsilon}\big(g\big(Y_t^\epsilon(\xi)\big) - X_t^\epsilon(\xi)\big) + \frac{1}{\sqrt{\epsilon}}\sigma_2\big(X_t^\epsilon(\xi), Y_t^\epsilon(\xi)\big)\dot{W}_t^{Q_2} \tag{1.2}$$

for $t > 0$, $\xi \in D$, with initial conditions $X_0^\epsilon(\xi) = X_0(\xi)$, $Y_0^\epsilon(\xi) = Y_0(\xi)$, and zero Dirichlet boundary conditions, where the nonlinear function g satisfies certain dissipative conditions, σ_1 and σ_2 are real-valued functions which are assumed to be Lipschitz continuous and linear growth, the parameter γ is a positive number and ϵ is a small positive parameter describing the ratio of time scale between the process X^ϵ and Y^ϵ. With this time scale the variable X^ϵ is referred as *slow* component and Y^ϵ as the *fast* component. The driving processes W^{Q_1} and W^{Q_2} are mutually independent Wiener processes on the stochastic basis $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$, which will be specified later. In many cases, one is concerned with predicting time evolution of X^ϵ in case of ϵ is small or its asymptotic behavior under the limit $\epsilon \to 0$. Then a reduction system for X^ϵ, capturing the dynamics of the slow motion, is desirable. The main purpose of this article is to study the asymptotic limit ($\epsilon \to 0$) for slow variable X^ϵ in the context of an averaging method. Our results show that the asymptotic behavior can be characterized by a stochastic differential equation with averaged coefficient.

The theory of averaging principle serves as a tool in study of the qualitative behaviors for complex systems with multiscales. It was first studied by Bogoliubov [2], then by Gikhman [13], Volosov [35] and Besjes [1] for a non-linear ordinary differential equations. Subsequently, the theory of averaging was developed by Khasminskii [22] to the stochastic ordinary equations with different time-scales. The results in [22] showed that a stochastic averaging principle occurs for the slow component in a weak sense. Taking into account the generalized and refined results, it is worthy quoting the paper by Veretennikov [33,34], the work of Freidlin and Wentzell [10,11] with notably extensions to convergence in probability. The mean-square type convergence was treated in Golec and Ladde [17] and Givon and co-workers [15]. For the strong convergence, we refer to [16] and [14]. We also refer to the recent paper by Xu, Duan and Xu [39] which deals with averaging principle for stochastic differential equations (SDEs) with non-Gaussian Lévy noise. In particular, the convergence order is also estimated in terms of noise intensity.

However, there are few results on the averaging principle for stochastic systems in infinite dimension space. The recent work by Cerrai and Freidlin [6] presents an averaged result slow-fast stochastic partial differential equations (SPDEs) with additive noise and Cerrai [5] deals with the case of multiplicative noise. The two papers show that the averaging principle, in sense of convergence in probability without an explicit rate, holds for that stochastic systems.

In the case that noise is of additive type, Wang and Roberts [37] show that the strong convergence (approximation of trajectories) rate is to be 1/2 for stochastic partial differential equations with slow and fast components. Also, Bréhier [3] obtains the weak convergence (approximation of law) rate 1 under condition that the noise included only in the fast component. These order is the same as for SDEs (see [22, 23]).

As far as multiplicative noise is concerned, in [12] the Khasminskii technique [22] is used to prove the strong convergence principle for SPDEs with two time scales. But the explicit order of convergence is not presented. To this purpose, we will study the order for convergence principle uniformly in time for stochastic

F–N system (1.1)–(1.2), which is stronger than the convergence in [6,5,37]. To be more precise, we will prove that,

$$\mathbb{E}\sup_{0\leqslant t\leqslant T_0}\left\|X_t^\epsilon-\bar{X}_t\right\|^2=\mathcal{O}(\sqrt{\epsilon}\,)$$

for fixed $T_0>0$. Here, the norm $\|\cdot\|$ denotes the usual norm defined on Hilbert space $\mathcal{L}^2(D)$ consisting of square integral real valued function on the interval $D=(0,l)$, the $\bar{X}_t$ is the solution of a reduced equation which approximates the slow component X^ϵ in F–N system. To achieve this, a key step is to show the existence for an invariant measure with exponentially mixing property for the fast equation, where a dissipative condition is needed. As a result, an effective dynamics for slow equation can be derived by averaging its parameters in the fast variable. From a technical point of view, one of the main novelties of this paper is that we obtain an explicit rate of convergence to the averaged effective dynamics, which is crucial for numerical analysis.

This paper is organized as follows. In Section 2 we present the framework and some preliminary results. Section 3 deals with the ergodicity of the fast equation with frozen slow component. In Section 4, some priori estimates are presented. These results will be utilized in the subsequent discussion. In Section 5 we prove the strong convergence version for stochastic averaging principle with an explicit convergence order.

Throughout this paper, the letter C below with or without subscripts will denote positive constants whose value may change in different occasions. We will write the dependence of constant on parameters explicitly if it is essential.

2. Framework and preliminary results

Let $D=(0,l)\subset\mathbb{R}$ be an open bounded interval and H be the Hilbert space $\mathcal{L}^2(D)$ equipped with the inner product $(\cdot,\cdot)_H$ and corresponding norm $\|\cdot\|$. Given $T_0>0$, consider the following system of stochastic FitzHugh–Nagumo equation on the interval D with separated time scales

$$\frac{\partial X_t^\epsilon(\xi)}{\partial t}=-\gamma X_t^\epsilon(\xi)+Y_t^\epsilon(\xi)+\sigma_1\big(X_t^\epsilon(\xi)\big)\dot{W}_t^{Q_1},\tag{2.1}$$

$$\frac{\partial Y_t^\epsilon(\xi)}{\partial t}=\frac{1}{\epsilon}\frac{\partial^2 Y_t^\epsilon(\xi)}{\partial\xi^2}+\frac{1}{\epsilon}\big(g\big(Y_t^\epsilon(\xi)\big)-X_t^\epsilon(\xi)\big)+\frac{1}{\sqrt{\epsilon}}\sigma_2\big(X_t^\epsilon(\xi),Y_t^\epsilon(\xi)\big)\dot{W}_t^{Q_2}\tag{2.2}$$

for $t>0$ with the initial conditions

$$X_0^\epsilon(\xi)=X_0(\xi)\in H,\qquad Y_0^\epsilon(\xi)=Y_0(\xi)\in H\tag{2.3}$$

and boundary conditions

$$Y^\epsilon(t,0)=Y^\epsilon(t,l)=0,\quad 0\leqslant t\leqslant T_0,\tag{2.4}$$

where γ is a positive parameter, the small parameter ϵ is positive and represents the ratio of time scale in this system. The mapping $g:\mathbb{R}\to\mathbb{R}$ is a nonlinear function satisfying conditions for all $\xi\in\mathbb{R}$

$$g(\xi)\xi\leqslant -a\xi^{2m}+b,\tag{2.5}$$

$$\big|g(\xi)\big|\leqslant c|\xi|^{2m-1}+d\tag{2.6}$$

and

$$g'(\xi)\leqslant\lambda,\tag{2.7}$$

here a, b, c, d and λ are positive constants, m is a positive integer. Note that $g(\xi) = \xi(\xi - 2)(1/2 - \xi)$ is an example.

Define the abstract operator $A = \partial_{\xi\xi}$ with zero Dirichlet boundary condition. Let $\{e_k\}_{k\in\mathbb{N}}$ denote the complete orthonormal system of eigenvectors in H such that, for $k = 1, 2, \ldots,$

$$Ae_k = -\alpha_k e_k, \qquad e_k|_{\partial D} = 0,$$

with $0 < \alpha_1 \leqslant \alpha_2 \leqslant \cdots \leqslant \alpha_k \leqslant \cdots$. Let V be the Sobolev space H_0^1 of order 1 with Dirichlet boundary conditions, which is densely and continuously injected in the Hilbert space H. Identifying H with its dual space we get a Gelfand triple

$$V \subset H \subset V^*$$

and

$$A : V \to V^*.$$

Due to the Poincaré inequality, we have

$$\langle Av, v\rangle = -\|\nabla v\|^2 \leqslant -\alpha_1\|v\|^2, \tag{2.8}$$

where $\langle\cdot,\cdot\rangle$ denotes the dual pairs of (V^*, V).

Next we recall the definition of Wiener process on Hilbert space H. For more details, see [30]. For $i = 1, 2$, let $\{e_{i,k}\}_{k\in\mathbb{N}}$ be eigenvectors of a nonnegative, symmetric operator Q_i with corresponding eigenvalues $\{\lambda_{i,k}\}_{k\in\mathbb{N}}$, such that

$$Q_i e_{i,k} = \lambda_{i,k} e_{i,k}, \quad \lambda_{i,k} > 0,\ k \in \mathbb{N}.$$

For $i = 1, 2$, let $W_t^{Q_i}$ be an H valued Q_i-Wiener process with operator Q_i satisfying

$$Tr\, Q_i = \sum_{k=1}^{\infty} \lambda_{i,k} = \rho_i < \infty. \tag{2.9}$$

Then

$$W_t^{Q_i} = \sum_{k=1}^{\infty} \lambda_{i,k}^{\frac{1}{2}} \beta_{i,k}(t) e_{i,k}, \quad t \geqslant 0,$$

where $\{\beta_{i,k}(t)\}_{k\in\mathbb{N}}^{i=1,2}$ are independent real-valued Brownian motions on a probability base $(\Omega, \mathscr{F}, \mathscr{F}_t, \mathbb{P})$. For $i = 1, 2$, we recall that for an $\mathscr{F}_t$-adapted process Φ_t satisfying the condition

$$\mathbb{E}\int_0^T \big\|\Phi_t(\cdot)\big\|^2\, dt = \int_0^T \int_D \big|\Phi_t(\xi)\big|^2\, d\xi\, dt < \infty,$$

the stochastic integrals

$$\mathcal{I}_t^i(\Phi) = \int_0^t \Phi_t\, dW_t^{Q_i}, \quad t \in [0, T]$$

is well defined [30] and has the Itô isometry:

$$\mathbb{E}\left\|\mathcal{I}_t^i(\Phi)\right\|^2 = \mathbb{E}\int_0^t \|\Phi_s\|_{Q_i}^2\,ds,$$

where

$$\|\Phi_s\|_{Q_i}^2 = \sum_{k=1}^{\infty} \lambda_{i,k}\|\Phi_s\|^2, \quad s \in [0,T].$$

To give precise results, we rewrite the system (2.1)–(2.4) to abstract stochastic evolutionary equations

$$\frac{dX_t^\epsilon}{dt} = -\gamma X_t^\epsilon + Y_t^\epsilon + \sigma_1\left(X_t^\epsilon\right)\dot{W}_t^{Q_1}, \tag{2.10}$$

$$X^\epsilon(0) = X_0, \tag{2.11}$$

$$\frac{dY_t^\epsilon}{dt} = \frac{1}{\epsilon}\left(AY_t^\epsilon + g\left(Y_t^\epsilon\right) - X_t^\epsilon\right) + \frac{1}{\sqrt{\epsilon}}\sigma_2\left(X_t^\epsilon, Y_t^\epsilon\right)\dot{W}_t^{Q_2}, \tag{2.12}$$

$$Y^\epsilon(0) = Y_0 \tag{2.13}$$

with $(X_0, Y_0) \in H \times H$.

To ensure the existence and uniqueness for system (2.10)–(2.13), we impose the following conditions on diffusion coefficients:

Assumption 1. The real-valued functions σ_1 and σ_2 satisfy the global Lipschitz condition, specifically, there exist constants L_{σ_1} and L_{σ_2} such that

$$\left|\sigma_1(u_1) - \sigma_1(u_2)\right|^2 \leqslant L_{\sigma_1}|u_1 - u_2|^2, \tag{2.14}$$

$$\left|\sigma_2(u_1, v_1) - \sigma_2(u_2, v_2)\right|^2 \leqslant L_{\sigma_2}\left(|u_1 - u_2|^2 + |v_1 - v_2|^2\right) \tag{2.15}$$

for all $u_1, v_1, u_2, v_2 \in \mathbb{R}$.

Remark 2.1. With Assumption 1, it immediately follows

$$\left|\sigma_1(u)\right|^2 \leqslant 2L_{\sigma_1}|u|^2 + 2\left|\sigma_1(0)\right|^2, \tag{2.16}$$

$$\left|\sigma_2(u,v)\right|^2 \leqslant (1+\varrho)L_{\sigma_2}\left(|u|^2 + |v|^2\right) + C_\varrho\left|\sigma_2(0,0)\right|^2 \tag{2.17}$$

for $u, v \in \mathbb{R}$ and all $\varrho > 0$. Thus σ_1 and σ_2 satisfy the sublinear growth condition.

We now introduce some definitions of solutions of system (2.10)–(2.13) and the results of the existence and uniqueness. Denote by $\{e^{At}\}_{t\geqslant 0}$ the C_0-semigroup generated by A, then the *mild solution* [30] of (2.10)–(2.13) is given by

$$X_t^\epsilon = e^{-\gamma t}X_0 + \int_0^t e^{-\gamma(t-s)}Y_s^\epsilon\,ds + \int_0^t e^{-\gamma(t-s)}\sigma_1\left(X_s^\epsilon\right)dW_s^{Q_1},$$

$$Y_t^\epsilon = e^{At/\epsilon}Y_0 + \frac{1}{\epsilon}\int_0^t e^{A(t-s)/\epsilon}\left(g\left(Y_s^\epsilon\right) - X_s^\epsilon\right)ds + \frac{1}{\sqrt{\epsilon}}\int_0^t e^{A(t-s)/\epsilon}\sigma_2\left(X_s^\epsilon, Y_s^\epsilon\right)dW_s^{Q_2}.$$

We also recall the *strong solution* [7] for system (2.10)–(2.13) in the sense

$$(X_t^\epsilon, \varphi)_H = (X_0, \varphi)_H + \int_0^t (-\gamma X_s^\epsilon + Y_s^\epsilon, \varphi)_H \, ds + \int_0^t (\varphi, \sigma_1(X_s^\epsilon) \, dW_s^{Q_1})_H,$$

$$(Y_t^\epsilon, \varphi)_H = (Y_0, \varphi)_H + \frac{1}{\epsilon} \int_0^t \langle AY_s^\epsilon, \varphi \rangle \, ds + \frac{1}{\epsilon} \int_0^t (g(Y_s^\epsilon) - X_s^\epsilon, \varphi)_H \, ds + \frac{1}{\sqrt{\epsilon}} \int_0^t (\varphi, \sigma_2(X_s^\epsilon, Y_s^\epsilon) \, dW_s^{Q_2})_H$$

for any $\varphi \in V$. Since the function g is not global Lipschitz, the classical wellposedness result [30] for stochastic system (2.10)–(2.13) do not hold. By adapting the *cutoff technique* as in [4], we can show that for every $\epsilon > 0, T_0 > 0$, Assumption 1 guarantees a unique *strong solution* $(X^\epsilon, Y^\epsilon) \in (L^2(\Omega \times [0, T_0]; V) \cap L^2(\Omega; C([0, T_0]; H)))^2$, which is also a *mild solution*, for the system (2.10)–(2.13). Furthermore, we possess the energy identities

$$\begin{aligned} \|X_t^\epsilon\|^2 = \|X_0\|^2 &+ 2\int_0^t (-\gamma X_s^\epsilon + Y_s^\epsilon, X_s^\epsilon)_H \, ds + 2\int_0^t (X_s^\epsilon, \sigma_1(X_s^\epsilon) \, dW_s^{Q_1})_H \\ &+ \int_0^t \|\sigma_1(X_s^\epsilon)\|_{Q_1}^2 \, ds, \quad \text{a.s.} \end{aligned} \tag{2.18}$$

and

$$\begin{aligned} \|Y_t^\epsilon\|^2 = \|Y_0\|^2 &+ \frac{2}{\epsilon}\int_0^t \langle AY_s^\epsilon, Y_s^\epsilon \rangle \, ds + \frac{2}{\epsilon}\int_0^t (g(Y_s^\epsilon) - X_s^\epsilon, Y_s^\epsilon)_H \, ds \\ &+ \frac{2}{\sqrt{\epsilon}}\int_0^t (Y_s^\epsilon, \sigma_2(X_s^\epsilon, Y_s^\epsilon) \, dW_s^{Q_2})_H + \frac{1}{\epsilon}\int_0^t \|\sigma_2(X_s^\epsilon, Y_s^\epsilon)\|_{Q_2}^2 \, ds, \quad \text{a.s.} \end{aligned} \tag{2.19}$$

for $t > 0$.

For the process of solution for the fast motion equation, we introduce the following assumption:

Assumption 2. For given $T_0 > 0$ and some $\epsilon_0 > 0$ the solution process $\{Y^\epsilon\}_{\epsilon \in (0,\epsilon_0), t \in [0,T_0]}$ is uniform bounds in $L^2(\Omega, H)$. That is, there exists a constant C such that

$$\sup_{\epsilon \in (0,\epsilon_0), t \in [0,T_0]} \mathbb{E}\|Y_t^\epsilon\|^2 < C. \tag{2.20}$$

Remark 2.2. In the particular case, where σ_2 is bounded, one has

$$\sup_{\epsilon \in (0,1), t \in [0,T_0]} \mathbb{E}\|Y_t^\epsilon\|^2 \leqslant C < \infty. \tag{2.21}$$

In fact, by making use of (2.18) and (2.19) first, then taking (2.5), (2.8) and (2.9) into account, we can deduce that

$$\mathbb{E}\left(\frac{1}{\epsilon}\|X_t^\epsilon\|^2 + \|Y_t^\epsilon\|^2\right) \leqslant \frac{1}{\epsilon}\|X_0\|^2 + \|Y_0\|^2 + \frac{1}{\epsilon}Ct + \frac{1}{\epsilon}\mathbb{E}\int_0^t \|X_s^\epsilon\|^2 \, ds,$$

which implies

$$\sup_{0\leqslant t\leqslant T_0,0<\epsilon<1} \mathbb{E}\|X_t^\epsilon\|^2 \leqslant C_{T_0} < \infty. \tag{2.22}$$

Then taking expectation on both sides of the identity (2.19) we have

$$\mathbb{E}\|Y_t^\epsilon\|^2 = \|Y_0\|^2 + \frac{2}{\epsilon}\int_0^t \mathbb{E}\langle AY_s^\epsilon, Y_s^\epsilon\rangle\, ds + \frac{2}{\epsilon}\int_0^t \mathbb{E}\big(g(Y_s^\epsilon) - X_s^\epsilon, Y_s^\epsilon\big)_H\, ds + \frac{1}{\epsilon}\int_0^t \mathbb{E}\|\sigma_2(X_s^\epsilon, Y_s^\epsilon)\|_{Q_2}^2\, ds,$$

which means

$$\frac{d}{dt}\mathbb{E}\|Y_t^\epsilon\|^2 = \frac{2}{\epsilon}\mathbb{E}\langle AY_t^\epsilon, Y_t^\epsilon\rangle + \frac{2}{\epsilon}\mathbb{E}\big(g(Y_t^\epsilon), Y_t^\epsilon\big)_H - \frac{2}{\epsilon}\mathbb{E}\big(X_t^\epsilon, Y_t^\epsilon\big)_H + \frac{1}{\epsilon}\mathbb{E}\|\sigma_2(X_t^\epsilon, Y_t^\epsilon)\|_{Q_2}^2. \tag{2.23}$$

By invoking Poincaré inequality (2.8), condition (2.5) and the boundedness of σ_2, we get

$$2\langle AY_t^\epsilon, Y_t^\epsilon\rangle + 2\big(g(Y_t^\epsilon), Y_t^\epsilon\big)_H + \|\sigma_2(X_t^\epsilon, Y_t^\epsilon)\|_{Q_2}^2 \leqslant -2\alpha_1\|Y_t^\epsilon\|^2 + C. \tag{2.24}$$

Using Young's inequality in form of $|a_1a_2|^2 \leqslant \frac{\alpha_1}{2}|a_1|^2 + C_{\alpha_1}|a_2|^2$ for the third term on the right-hand side of (2.23), we obtain

$$2\big|\big(X_t^\epsilon, Y_t^\epsilon\big)_H\big| \leqslant \alpha_1\|Y_t^\epsilon\|^2 + C_{\alpha_1}\|X_t^\epsilon\|^2. \tag{2.25}$$

By (2.23), (2.24) and (2.25) we get that

$$\frac{d}{dt}\mathbb{E}\|Y_t^\epsilon\|^2 \leqslant -\frac{\alpha_1}{\epsilon}\mathbb{E}\|Y_t^\epsilon\|^2 + \frac{C}{\epsilon}\mathbb{E}\|X_t^\epsilon\|^2 + \frac{C}{\epsilon}.$$

According to (2.22), this implies that

$$\frac{d}{dt}\mathbb{E}\|Y_t^\epsilon\|^2 \leqslant -\frac{\alpha_1}{\epsilon}\mathbb{E}\|Y_t^\epsilon\|^2 + \frac{C}{\epsilon}. \tag{2.26}$$

With the aid of Gronwall's inequality introduced in [14], the estimate (2.21) can be derived from (2.26). This illustrates Assumption 2 is reasonable.

In order to obtain strong convergence results for averaging principle, we need to assume some dissipative condition on fast motion equation.

Assumption 3. Assume that the growth rate of nonlinear term g and diffusion term σ_2 in fast component equation is smaller than the decay rate α_1 of A, that is

$$\eta := 2\alpha_1 - 2\lambda - \rho_2 L_{\sigma_2} > 0. \tag{2.27}$$

Assumption (2.27) is interpreted as a strong dissipative condition which yields a unique invariant measure possessing exponentially mixing property for Markov semigroup associated to fast variable equation.

3. Ergodicity for the fast motion equation

For fixed $x \in H$ we consider the following problem associated with the fast equation with frozen slow component

$$\frac{dY_t}{dt} = AY_t + g(Y_t) - x + \sigma_2(x, Y_t)\dot{W}_t^{Q_2}, \tag{3.1}$$

$$Y_0 = y. \tag{3.2}$$

By arguing as before, for any fixed slow component $x \in H$ and any initial data $y \in H$, system (3.1)–(3.2) has a unique strong solution denoted by $Y_t^{x,y}$. The energy equality (2.19) reads

$$\begin{aligned}\|Y_t^{x,y}\|^2 = \|y\|^2 &+ 2\int_0^t \langle AY_s^{x,y}, Y_s^{x,y}\rangle\, ds + 2\int_0^t \big(g\big(Y_s^{x,y}\big), Y_s^{x,y}\big)_H\, ds \\ &- 2\int_0^t \big(x, Y_s^{x,y}\big)_H\, ds + 2\int_0^t \big(Y_s^{x,y}, \sigma_2\big(x, Y_s^{x,y}\big)\, dW_s^{Q_2}\big)_H \\ &+ \int_0^t \big\|\sigma_2\big(x, Y_s^{x,y}\big)\big\|_{Q_2}^2\, ds,\end{aligned}$$

which implies

$$\frac{d}{dt}\mathbb{E}\|Y_t^{x,y}\|^2 = 2\mathbb{E}\langle AY_t^{x,y}, Y_t^{x,y}\rangle + 2\mathbb{E}\big(g\big(Y_t^{x,y}\big), Y_t^{x,y}\big)_H - 2\mathbb{E}\big(x, Y_t^{x,y}\big)_H + \mathbb{E}\big\|\sigma_2\big(x, Y_t^{x,y}\big)\big\|_{Q_2}^2. \tag{3.3}$$

From Poincaré inequality (2.8) we have

$$\langle AY_s^{x,y}, Y_s^{x,y}\rangle \leqslant -\alpha_1\|Y_s^{x,y}\|^2. \tag{3.4}$$

By (2.5) and (2.17) in Remark 2.1 we have

$$2\big(g\big(Y_s^{x,y}\big), Y_s^{x,y}\big)_H + \big\|\sigma_2\big(x, Y_s^{x,y}\big)\big\|_{Q_2}^2 \leqslant (1+\varrho)\rho_2 L_{\sigma_2}\big(\|Y_s^{x,y}\|^2 + \|x\|^2\big) + C. \tag{3.5}$$

If we choose $\varrho > 0$ such that $\rho = 2\alpha_1 - (1+\varrho)\rho_2 L_{\sigma_2} > 0$, and then take $\rho > 0$ for Young's inequality in the form $|a_1 a_2| \leqslant \frac{1}{4}\rho|a_1|^2 + C_\rho|a_2|^2$, we obtain

$$\big|\big(x, Y_s^{x,y}\big)_H\big| \leqslant \frac{1}{4}\rho\|Y_s^{x,y}\|^2 + C_\rho\|x\|^2. \tag{3.6}$$

From (3.3)–(3.6), we immediately see that

$$\frac{d}{dt}\mathbb{E}\|Y_t^{x,y}\|^2 \leqslant -\frac{\rho}{2}\mathbb{E}\|Y_t^{x,y}\|^2 + C\big(1 + \|x\|^2\big).$$

According to Gronwall's inequality introduced in [14], this means

$$\mathbb{E}\|Y_t^{x,y}\|^2 \leqslant C\big(e^{-\frac{\rho}{2}t}\|y\|^2 + \|x\|^2 + 1\big), \quad t > 0 \tag{3.7}$$

for some constant $C > 0$. Let $Y_t^{x,y'}$ be the solution of system (3.1)–(3.2) with initial value $Y_0 = y'$. Thanks to (2.7), (2.8) and (2.15), it is immediate to check that for any $t \geqslant 0$,

$$\mathbb{E}\left\|Y_t^{x,y} - Y_t^{x,y'}\right\|^2 \leqslant \left\|y - y'\right\|^2 e^{-\eta t}. \tag{3.8}$$

For any $x \in H$ denote by P_t^x the Markov semigroup associated to Eqs. (3.1)–(3.2) defined by

$$P_t^x \psi(y) = \mathbb{E}\psi\left(Y_t^{x,y}\right), \quad t \geqslant 0,\ y \in H,$$

for any $\psi \in \mathcal{B}_b(H)$ the space of bounded functions on H. We also recall a probability μ^x on H which is called an invariant measure for $(P_t^x)_{t\geqslant 0}$ if

$$\int_H P_t^x \psi d\mu^x = \int_H \psi d\mu^x, \quad t \geqslant 0$$

for any bounded function $\psi \in \mathcal{B}_b(H)$. Thanks to (3.7), by adapting the arguments used in [4] and [6], it is possible to show that there exists a unique invariant measure μ^x for the semigroup P_t^x, which satisfies

$$\int_H \|y\|^2 \mu^x(dy) \leqslant C\left(1 + \|x\|^2\right). \tag{3.9}$$

Set $f(x, y) := -\gamma x + y$, $x, y \in H$ and define the averaging function

$$\bar{f}(x) := \int_H f(x, y)\mu^x(dy).$$

According to (3.9) we have

$$\left\|f(x, y) - \bar{f}(x)\right\|^2 \leqslant C\left(1 + \|x\|^2 + \|y\|^2\right). \tag{3.10}$$

By the invariant property of μ^x and (3.8) we have

$$\begin{aligned}
\left\|\mathbb{E}f\left(x, Y_t^{x,y}\right) - \bar{f}(x)\right\|^2 &= \left\|\int_H \mathbb{E}\left(f\left(x, Y_t^{x,y}\right) - f\left(x, Y_t^{x,z}\right)\right)\mu^x(dz)\right\|^2 \\
&\leqslant \int_H \mathbb{E}\left\|Y_t^{x,y} - Y_t^{x,z}\right\|^2 \mu^x(dz) \\
&\leqslant e^{-\eta t} \int_H \|y - z\|^2 \mu^x(dz) \\
&\leqslant Ce^{-\eta t}\left(1 + \|x\|^2 + \|y\|^2\right).
\end{aligned} \tag{3.11}$$

4. Some a priori estimates

In this section, we present some a priori estimates for solution processes X^ϵ and Y^ϵ.

Lemma 4.1. *For any $X_0, Y_0 \in H$ there exists a constant $C > 0$ such that*

$$\sup_{0<\epsilon<\epsilon_0, 0\leqslant t\leqslant T_0} \mathbb{E}\left\|X_t^\epsilon\right\|^2 \leqslant C\left(1 + \|X_0\|^2\right). \tag{4.1}$$

The above lemma is a direct result of energy equality (2.18) and (2.20) and so the proof is omitted.

Lemma 4.2. *For all $t \in [0, T_0]$, $h \in (0,1)$, the mean square displacement of the solution X^ϵ for slow equation satisfies*

$$\sup_{0<\epsilon<\epsilon_0} \mathbb{E}\left\|X_{t+h}^\epsilon - X_t^\epsilon\right\|^2 \leqslant Ch. \tag{4.2}$$

Proof. It is clear that for all $t \in [0, T_0]$, $h \in (0,1)$,

$$X_{t+h}^\epsilon - X_t^\epsilon = \left(e^{-\gamma h} - 1\right)X_t^\epsilon + \int_t^{t+h} e^{-\gamma(t+h-s)} Y_s^\epsilon \, ds + \int_t^{t+h} e^{-\gamma(t+h-s)} \sigma_1\left(X_s^\epsilon\right) dW_s^{Q_1}$$

and hence

$$\mathbb{E}\left\|X_{t+h}^\epsilon - X_t^\epsilon\right\|^2 \leqslant C\left|e^{-\gamma h} - 1\right|^2 \mathbb{E}\left\|X_t^\epsilon\right\|^2 + Ch \int_t^{t+h} e^{-2\gamma(t+h-s)} \mathbb{E}\left\|Y_s^\epsilon\right\|^2 ds$$
$$+ C \int_t^{t+h} e^{-2\gamma(t+h-s)} \left(1 + \mathbb{E}\left\|X_s^\epsilon\right\|^2\right) ds,$$

so that, from (4.1) and (2.20),

$$\sup_{0<\epsilon<\epsilon_0, 0\leqslant t\leqslant T_0} \mathbb{E}\left\|X_{t+h}^\epsilon - X_t^\epsilon\right\|^2 \leqslant Ch$$

is obtained. □

Next, we introduce an auxiliary process $(\hat{X}_t^\epsilon, \hat{Y}_t^\epsilon) \in H \times H$. Fix a positive number $\delta < 1$ and do a partition of time interval $[0, T_0]$ of size δ. We construct a process $\hat{Y}_t^\epsilon$, with initial datum $\hat{Y}_0^\epsilon = Y_0$, by means of the equations

$$d\hat{Y}_t^\epsilon = \frac{1}{\epsilon} A\hat{Y}_t^\epsilon \, dt + \frac{1}{\epsilon}\left(g\left(\hat{Y}_t^\epsilon\right) - X_{k\delta}^\epsilon\right) dt + \frac{1}{\sqrt{\epsilon}} \sigma_2\left(X_{k\delta}^\epsilon, \hat{Y}_t^\epsilon\right) dW_t^{Q_2}, \qquad \hat{Y}_{k\delta}^\epsilon = Y_{k\delta}^\epsilon$$

for $t \in [k\delta, \min\{(k+1)\delta, T_0\})$, $k \geqslant 0$, where $X_{k\delta}^\epsilon$ and $Y_{k\delta}^\epsilon$ are slow and fast solution processes at time $k\delta$, respectively. Denote $\lfloor\cdot\rfloor$ to be the integer function and define the process $\hat{X}_t^\epsilon$ by integral

$$\hat{X}_t^\epsilon = X_0 + \int_0^t \left(-\gamma X_{s(\delta)}^\epsilon + \hat{Y}_s^\epsilon\right) ds + \int_0^t \sigma_1\left(X_s^\epsilon\right) dW_s^{Q_1}$$

for $t \in [0, T_0]$, where $s(\delta) = \lfloor s/\delta \rfloor \delta$ is the nearest breakpoint preceding s. We will establish convergence of the auxiliary processes $\hat{Y}_t^\epsilon$ to the fast solution process Y_t^ϵ and $\hat{X}_t^\epsilon$ to the slow solution process X_t^ϵ, respectively.

Lemma 4.3. *There exists a constant $C > 0$ such that for any $t \in [0, T_0]$ it holds*

$$\mathbb{E}\left\|Y_t^\epsilon - \hat{Y}_t^\epsilon\right\|^2 \leqslant C\delta \tag{4.3}$$

and

$$\mathbb{E} \sup_{0\leqslant t\leqslant T_0} \left\|X_t^\epsilon - \hat{X}_t^\epsilon\right\|^2 \leqslant C\delta. \tag{4.4}$$

Proof. For $t \in [0, T_0]$ with $t \in [k\delta, (k+1)\delta)$ we have

$$\begin{aligned}
\mathbb{E}\|Y_t^\epsilon - \hat{Y}_t^\epsilon\|^2 &= \frac{2}{\epsilon}\int_{k\delta}^{t} \mathbb{E}\langle A(Y_s^\epsilon - \hat{Y}_s^\epsilon), Y_s^\epsilon - \hat{Y}_s^\epsilon\rangle\, ds \\
&\quad + \frac{2}{\epsilon}\int_{k\delta}^{t} \mathbb{E}(g(\hat{Y}_s^\epsilon) - g(Y_s^\epsilon), Y_s^\epsilon - \hat{Y}_s^\epsilon)_H\, ds \\
&\quad + \frac{2}{\epsilon}\int_{k\delta}^{t} \mathbb{E}(X_s^\epsilon - X_{k\delta}^\epsilon, Y_s^\epsilon - \hat{Y}_s^\epsilon)_H\, ds \\
&\quad + \frac{1}{\epsilon}\int_{k\delta}^{t} \mathbb{E}\|\sigma_2(X_{k\delta}^\epsilon, \hat{Y}_s^\epsilon) - \sigma_2(X_s^\epsilon, Y_s^\epsilon)\|_{Q_2}^2\, ds.
\end{aligned}$$

This show that

$$\begin{aligned}
\frac{d}{dt}\mathbb{E}\|Y_t^\epsilon - \hat{Y}_t^\epsilon\|^2 &= \frac{2}{\epsilon}\mathbb{E}\langle A(Y_t^\epsilon - \hat{Y}_t^\epsilon), Y_t^\epsilon - \hat{Y}_t^\epsilon\rangle \\
&\quad + \frac{2}{\epsilon}\mathbb{E}(g(\hat{Y}_t^\epsilon) - g(Y_t^\epsilon), Y_t^\epsilon - \hat{Y}_t^\epsilon)_H \\
&\quad + \frac{2}{\epsilon}\mathbb{E}(X_t^\epsilon - X_{k\delta}^\epsilon, Y_t^\epsilon - \hat{Y}_t^\epsilon)_H \\
&\quad + \frac{1}{\epsilon}\mathbb{E}\|\sigma_2(X_{k\delta}^\epsilon, \hat{Y}_t^\epsilon) - \sigma_2(X_t^\epsilon, Y_t^\epsilon)\|_{Q_2}^2.
\end{aligned} \tag{4.5}$$

From Poincaré inequality (2.8) we have

$$\mathbb{E}\langle A(Y_t^\epsilon - \hat{Y}_t^\epsilon), Y_t^\epsilon - \hat{Y}_t^\epsilon\rangle \leqslant -\alpha_1\mathbb{E}\|Y_t^\epsilon - \hat{Y}_t^\epsilon\|^2. \tag{4.6}$$

And by condition (2.7) we have

$$\mathbb{E}(g(\hat{Y}_t^\epsilon) - g(Y_t^\epsilon), Y_t^\epsilon - \hat{Y}_t^\epsilon)_H \leqslant \lambda\mathbb{E}\|Y_t^\epsilon - \hat{Y}_t^\epsilon\|^2. \tag{4.7}$$

Also by (2.15) and (2.9) we obtain

$$\begin{aligned}
\mathbb{E}\|\sigma_2(X_{k\delta}^\epsilon, \hat{Y}_t^\epsilon) - \sigma_2(X_t^\epsilon, Y_t^\epsilon)\|_{Q_2}^2 &\leqslant \rho_2 L_{\sigma_2}(\mathbb{E}\|X_{k\delta}^\epsilon - X_t^\epsilon\|^2 + \mathbb{E}\|\hat{Y}_t^\epsilon - Y_t^\epsilon\|^2) \\
&\leqslant C\delta + \rho_2 L_{\sigma_2}\mathbb{E}\|\hat{Y}_t^\epsilon - Y_t^\epsilon\|^2,
\end{aligned} \tag{4.8}$$

where we used (4.2) in the second inequality since $t \in [k\delta, (k+1)\delta)$. Then by taking $\eta = 2\alpha_1 - 2\lambda - \rho_2 L_{\sigma_2} > 0$ for Young's inequality in the form of $|ab| \leqslant \frac{\eta}{4}|a|^2 + C_\eta|b|^2$ we obtain

$$\begin{aligned}
\mathbb{E}|(X_t^\epsilon - X_{k\delta}^\epsilon, Y_t^\epsilon - \hat{Y}_t^\epsilon)_H| &\leqslant \frac{\eta}{4}\mathbb{E}\|Y_t^\epsilon - \hat{Y}_t^\epsilon\|^2 + C_\eta\mathbb{E}\|X_t^\epsilon - X_{k\delta}^\epsilon\|^2 \\
&\leqslant C\delta + \frac{\eta}{4}\mathbb{E}\|Y_t^\epsilon - \hat{Y}_t^\epsilon\|^2.
\end{aligned} \tag{4.9}$$

By taking (4.5), (4.6), (4.7), (4.8) and (4.9) into account, we get

$$\frac{d}{dt}\mathbb{E}\|Y_t^\epsilon - \hat{Y}_t^\epsilon\|^2 \leqslant -\frac{\eta}{2\epsilon}\mathbb{E}\|Y_t^\epsilon - \hat{Y}_t^\epsilon\|^2 + C\frac{\delta}{\epsilon}$$

and then, due to Gronwall's inequality presented in [14], we have

$$\mathbb{E}\big\|Y_t^\epsilon - \hat{Y}_t^\epsilon\big\|^2 \leqslant C\delta.$$

Therefore the first assertion follows. As for the second estimate, we have by Hölder's inequality and the Burkholder inequality that

$$\mathbb{E}\sup_{0\leqslant t\leqslant T_0}\big\|X_t^\epsilon - \hat{X}_t^\epsilon\big\|^2 \leqslant C\int_0^{T_0}\mathbb{E}\big\|X_s^\epsilon - X_{s(\delta)}^\epsilon\big\|^2\,ds + C\int_0^{T_0}\mathbb{E}\big\|Y_s^\epsilon - \hat{Y}_s^\epsilon\big\|^2\,ds.$$

Using (4.2) and the first estimate (4.3), we immediately get the desired second estimation. □

5. Averaging principle

In this section we will consider the effective dynamical system

$$\frac{\partial \bar{X}_t}{\partial t} = \bar{f}(\bar{X}_t) + \sigma_1(\bar{X}_t)\dot{W}_t^{Q_1}, \tag{5.1}$$

$$\bar{X}_0 = X_0 \tag{5.2}$$

with

$$\bar{f}(x) = \int_H f(x,y)\mu^x(dy), \quad x \in H,$$

where $f(x,y) = -\gamma x + y$ and μ^x denotes the unique invariant measure for system (3.1)–(3.2) introduced in Section 3. Moreover, due to [6], the mapping $\bar{f} : H \mapsto H$ is Lipschitz continuous. Our aim here is proving the main result of this paper. Namely, we are going to verify that the sequence $\{X_t^\epsilon:\ t \geqslant 0\}_{\epsilon>0}$ converges to the solution process $\{\bar{X}_t:\ t \geqslant 0\}$ of the averaged system (5.1)–(5.2) in space $L^2(\Omega, C([0,T_0];H))$ as ϵ goes to zero. To this end, we shall explore the difference between the solution of the averaged equation and the auxiliary process $\hat{X}_t^\epsilon$. Next, by the construction of $\hat{Y}_t^\epsilon$ and a time shift transformation, we have for any fixed k and $s \in [0,\delta)$ the equalities

$$\begin{aligned}
\hat{Y}_{s+k\delta}^\epsilon &= e^{As/\epsilon}Y_{k\delta}^\epsilon + \frac{1}{\epsilon}\int_{k\delta}^{k\delta+s} e^{A(k\delta+s-r)/\epsilon}\big(g\big(\hat{Y}_r^\epsilon\big) - X_{k\delta}^\epsilon\big)\,dr \\
&\quad + \frac{1}{\sqrt{\epsilon}}\int_{k\delta}^{k\delta+s} e^{A(k\delta+s-r)/\epsilon}\sigma_2\big(X_{k\delta}^\epsilon, \hat{Y}_r^\epsilon\big)\,dW_r^{Q_2} \\
&= e^{As/\epsilon}Y_{k\delta}^\epsilon + \frac{1}{\epsilon}\int_0^s e^{A(s-r)/\epsilon}\big(g\big(\hat{Y}_{r+k\delta}^\epsilon\big) - X_{k\delta}^\epsilon\big)\,dr \\
&\quad + \frac{1}{\sqrt{\epsilon}}\int_0^s e^{A(s-r)/\epsilon}\sigma_2\big(X_{k\delta}^\epsilon, \hat{Y}_{r+k\delta}^\epsilon\big)\,dW_r^{*Q_2},
\end{aligned} \tag{5.3}$$

where $W_t^{*Q_2} = W_{t+k\delta}^{Q_2} - W_{k\delta}^{Q_2}$ is the shift version of $W_t^{Q_2}$ and hence they have the same distribution. Let $\bar{W}_t$ be a Wiener process defined on the same stochastic basis and independent of $W_t^{Q_1}$ and $W_t^{Q_2}$. Construct a process $Y^{X_{k\delta}^\epsilon, Y_{k\delta}^\epsilon}$ by means of

$$
\begin{aligned}
Y_{s/\epsilon}^{X_{k\delta}^\epsilon, Y_{k\delta}^\epsilon} &= e^{As/\epsilon} Y_{k\delta}^\epsilon + \int_0^{s/\epsilon} e^{A(s/\epsilon - r)} \big(g\big(Y_r^{X_{k\delta}^\epsilon, Y_{k\delta}^\epsilon}\big) - X_{k\delta}^\epsilon\big)\, dr \\
&\quad + \int_0^{s/\epsilon} e^{A(s/\epsilon - r)} \sigma_2\big(X_{k\delta}^\epsilon, Y_r^{X_{k\delta}^\epsilon, Y_{k\delta}^\epsilon}\big)\, d\bar{W}_r \\
&= e^{As/\epsilon} Y_{k\delta}^\epsilon + \frac{1}{\epsilon} \int_0^{s} e^{A(s-r)/\epsilon} \big(g\big(Y_{r/\epsilon}^{X_{k\delta}^\epsilon, Y_{k\delta}^\epsilon}\big) - X_{k\delta}^\epsilon\big)\, dr \\
&\quad + \frac{1}{\sqrt{\epsilon}} \int_0^{s} e^{A(s-r)/\epsilon} \sigma_2\big(X_{k\delta}^\epsilon, Y_{r/\epsilon}^{X_{k\delta}^\epsilon, Y_{k\delta}^\epsilon}\big)\, d\bar{\bar{W}}_r^\epsilon,
\end{aligned}
\tag{5.4}
$$

where $\bar{\bar{W}}_t^\epsilon = \sqrt{\epsilon}\bar{W}_{t/\epsilon}$ is the scaled version of $\bar{W}_t$. By comparison, (5.3) and (5.4) yield

$$
\big(X_{k\delta}^\epsilon, \hat{Y}_{s+k\delta}\big) \sim \big(X_{k\delta}^\epsilon, Y_{s/\epsilon}^{X_{k\delta}^\epsilon, Y_{k\delta}^\epsilon}\big), \quad s \in [0, \delta),
$$

where $\sim$ denotes coincidence in distribution sense. Set

$$
\mathcal{J}_k^\epsilon := \mathbb{E}\left\| \int_0^\delta \big(f\big(X_{k\delta}^\epsilon, Y_{s/\epsilon}^{X_{k\delta}^\epsilon, Y_{k\delta}^\epsilon}\big) - \bar{f}\big(X_{k\delta}^\epsilon\big)\big)\, ds \right\|^2, \quad 0 \leqslant k \leqslant \lfloor T_0/\delta \rfloor - 1,
$$

then we state the critical lemma, which will be used later.

Lemma 5.1. *Suppose that Assumptions 1–3 hold, then there exists a constant $C > 0$ such that*

$$
\mathcal{J}_k^\epsilon \leqslant C\delta\epsilon, \quad 0 \leqslant k \leqslant \lfloor T_0/\delta \rfloor - 1. \tag{5.5}
$$

Proof. Let $\mathbb{Q}^y$ denote the probability law of the diffusion process $\{Y_t^x\colon t \geqslant 0\}$ which is governed by differential equation

$$
dY_t^x = AY_t^x\, dt + \big(g\big(Y_t^x\big) - x\big)\, dt + \sigma_2\big(x, Y_t^x\big)\, d\bar{W}_t.
$$

When its initial value is $Y_0^x = y$ and we denote the solution by $Y_t^{x,y}$. The expectation with respect to $\mathbb{Q}^y$ is denoted by $\mathbb{E}^y$. Hence we have

$$
\mathbb{E}^y\big(\psi\big(Y_t^x\big)\big) = \mathbb{E}\big(\psi\big(Y_t^{x,y}\big)\big)
$$

for all bounded function ψ. For more details on $\mathbb{Q}^y$ the readers are referred to [29]. First we note that it is easy to show that $\mathcal{J}_k^\epsilon < \infty$, $k = 0, 1, \ldots, \lfloor T_0/\delta \rfloor - 1$. Let $\{e_k\}_{k\in\mathbb{N}}$ be an orthonormal basis of H. Then by the Fourier expansion and the Fubini theorem, we see that

$$
\begin{aligned}
\mathcal{J}_k^\epsilon &= \sum_{i=1}^\infty \mathbb{E}\left| \int_0^\delta \big(f\big(X_{k\delta}^\epsilon, Y_{s/\epsilon}^{X_{k\delta}^\epsilon, Y_{k\delta}^\epsilon}\big) - \bar{f}\big(X_{k\delta}^\epsilon\big), e_i\big)_H\, ds \right|^2 \\
&= 2\int_0^\delta \int_\tau^\delta \sum_{i=1}^\infty \mathbb{E}\Big[\big(f\big(X_{k\delta}^\epsilon, Y_{s/\epsilon}^{X_{k\delta}^\epsilon, Y_{k\delta}^\epsilon}\big) - \bar{f}\big(X_{k\delta}^\epsilon\big), e_i\big)_H \big(f\big(X_{k\delta}^\epsilon, Y_{\tau/\epsilon}^{X_{k\delta}^\epsilon, Y_{k\delta}^\epsilon}\big) - \bar{f}\big(X_{k\delta}^\epsilon\big), e_i\big)_H\Big]\, ds\, d\tau,
\end{aligned}
$$

$k = 0, 1, \dots, \lfloor T_0/\delta \rfloor - 1$. For $i = 1, 2, \dots,$ define

$$J_i(\tau, s, x, y) := \mathbb{E}\big[\big(f\big(x, Y_s^{x,y}\big) - \bar{f}(x), e_i\big)_H \big(f\big(x, Y_\tau^{x,y}\big) - \bar{f}(x), e_i\big)_H\big].$$

It follows from the Markov property of $Y_t^{x,y}$ that

$$\begin{aligned} J_i(\tau, s, x, y) &= \mathbb{E}^y\big\{\mathbb{E}^y\big[\big(f\big(x, Y_s^x\big) - \bar{f}(x), e_i\big)_H \big(f\big(x, Y_\tau^x\big) - \bar{f}(x), e_i\big)_H \big| \mathscr{M}_\tau^x\big]\big\} \\ &= \mathbb{E}^y\big\{\big(f\big(x, Y_\tau^x\big) - \bar{f}(x), e_i\big)_H \times \mathbb{E}^{Y_\tau^{x,y}}\big[\big(f\big(x, Y_{s-\tau}^x\big) - \bar{f}(x), e_i\big)_H\big]\big\} \end{aligned}$$

for $i = 1, 2, \dots,$ where $\mathscr{M}_t^x$ denotes the σ-field generated by $\{Y_r^x;\ r \leqslant t\}$, $\mathbb{E}^{Y_\tau^{x,y}}[(f(x, Y_{s-\tau}^x) - \bar{f}(x), e_i)]$ means the function $\mathbb{E}^{\tilde{y}}[(f(x, Y_{s-\tau}^x) - \bar{f}(x), e_i)]$ evaluated at $\tilde{y} = Y_\tau^{x,y}$. Therefore the Hölder inequality yields

$$\sum_{i=1}^{\infty} J_i(\tau, s, x, y) \leqslant \big\{\mathbb{E}^y \big\|f\big(x, Y_\tau^x\big) - \bar{f}(x)\big\|^2\big\}^{\frac{1}{2}} \big\{\mathbb{E}^y\big(\big\|\mathbb{E}^{\tilde{y}}\big(f\big(x, Y_{s-\tau}^x\big)\big) - \bar{f}(x)\big\|^2_{\tilde{y}=Y_\tau^{x,y}}\big)\big\}^{\frac{1}{2}},$$

which, with aid of (3.7), (3.10) and (3.11), implies that

$$\begin{aligned} \sum_{i=1}^{\infty} J_i(\tau, s, x, y) &\leqslant C\big\{\mathbb{E}^y \big\|f\big(x, Y_\tau^x\big) - \bar{f}(x)\big\|^2\big\}^{\frac{1}{2}} \big\{\mathbb{E}^y\big(1 + \|x\|^2 + \big\|Y_{s-\tau}^x\big\|^2\big) \cdot e^{-\eta(s-\tau)}\big\}^{\frac{1}{2}} \\ &\leqslant C\big(1 + \|x\|^2 + \|y\|^2\big) e^{-\frac{\eta}{2}(s-\tau)}. \end{aligned} \tag{5.6}$$

Let $\mathcal{M}_{k\delta}^\epsilon$ be the σ-field generated by $X_{k\delta}^\epsilon$ and $Y_{k\delta}^\epsilon$, which is independent of $\{Y_r^{x,y}\colon\ r \geqslant 0\}$. By adapting the approach in [29] (Theorem 7.1.2) we can deduce from (5.6) and Lemma 4.1 that

$$\begin{aligned} \mathcal{J}_k^\epsilon &= 2\int_0^\delta \int_\tau^\delta \sum_{i=1}^{\infty} \mathbb{E}\big(\mathbb{E}\big[\big(f\big(X_{k\delta}^\epsilon, Y_{s/\epsilon}^{X_{k\delta}^\epsilon, Y_{k\delta}^\epsilon}\big) - \bar{f}\big(X_{k\delta}^\epsilon\big), e_i\big)_H \big(f\big(X_{k\delta}^\epsilon, Y_{\tau/\epsilon}^{X_{k\delta}^\epsilon, Y_{k\delta}^\epsilon}\big) - \bar{f}\big(X_{k\delta}^\epsilon\big), e_i\big)_H \big| \mathcal{M}_{k\delta}^\epsilon\big]\big)\, ds\, d\tau \\ &= 2\int_0^\delta \int_\tau^\delta \mathbb{E}\left[\left(\sum_{i=1}^{\infty} J_i(\tau/\epsilon, s/\epsilon, x, y)\right)\bigg|_{(x,y)=(X_{k\delta}^\epsilon, Y_{k\delta}^\epsilon)}\right] ds\, d\tau \\ &\leqslant C\int_0^\delta \int_\tau^\delta e^{-\frac{\eta}{2}(s-\tau)/\epsilon}\, ds\, d\tau \\ &= \frac{2\epsilon}{\eta}\left[\delta - \frac{2\epsilon}{\eta}\big(1 - e^{\frac{\delta\eta}{2\epsilon}}\big)\right], \end{aligned}$$

which completes the proof. $\square$

Lemma 5.2. *Suppose that Assumptions 1–3 are satisfied. Then for any $T_0 > 0$, we have*

$$\mathbb{E} \sup_{0 \leqslant t \leqslant T_0} \left\| \int_0^t \big(f\big(X_{s(\delta)}^\epsilon, \hat{Y}_s^\epsilon\big) - \bar{f}\big(X_s^\epsilon\big)\big)\, ds \right\|^2 \leqslant C\left(\delta + \frac{\epsilon}{\delta}\right), \tag{5.7}$$

where C is a constant independent of (ϵ, δ).

Proof. Notice that for any $t \in [0, T_0)$, there exists an $n_t = \lfloor t/\delta \rfloor$ such that $t \in [n_t\delta, (n_t+1)\delta \wedge T_0)$. Therefore, we have representation in the form

$$\int_0^t \left(f\left(X^\epsilon_{s(\delta)}, \hat{Y}^\epsilon_s\right) - \bar{f}\left(X^\epsilon_s\right)\right) ds := I_1(t,\epsilon) + I_2(t,\epsilon) + I_3(t,\epsilon), \tag{5.8}$$

where

$$I_1(t,\epsilon) = \sum_{k=0}^{n_t-1} \int_{k\delta}^{(k+1)\delta} \left(f\left(X^\epsilon_{k\delta}, \hat{Y}^\epsilon_s\right) - \bar{f}\left(X^\epsilon_{k\delta}\right)\right) ds,$$

$$I_2(t,\epsilon) = \sum_{k=0}^{n_t-1} \int_{k\delta}^{(k+1)\delta} \left(\bar{f}\left(X^\epsilon_{k\delta}\right) - \bar{f}\left(X^\epsilon_s\right)\right) ds = \int_0^{n_t\delta} \left(\bar{f}\left(X^\epsilon_{s(\delta)}\right) - \bar{f}\left(X^\epsilon_s\right)\right) ds,$$

$$I_3(t,\epsilon) = \int_{n_t\delta}^{t} \left(f\left(X^\epsilon_{n_t\delta}, \hat{Y}^\epsilon_s\right) - \bar{f}\left(X^\epsilon_s\right)\right) ds.$$

Concerning $I_2(t,\epsilon)$, we have the inequalities

$$\begin{aligned}
\mathbb{E} \sup_{0\leqslant t\leqslant T_0} \left\|I_2(t,\epsilon)\right\|^2 &\leqslant C\mathbb{E} \sup_{0\leqslant t\leqslant T_0} \int_0^{n_t\delta} \left\|X^\epsilon_{s(\delta)} - X^\epsilon_s\right\|^2 ds \\
&\leqslant C \int_0^{T_0} \mathbb{E}\left\|X^\epsilon_{s(\delta)} - X^\epsilon_s\right\|^2 ds \\
&\leqslant C\delta
\end{aligned} \tag{5.9}$$

for $T_0 > 0$, where the last inequality is due to (4.2). We proceed next to the estimation of $I_3(t,\epsilon)$. Set

$$\tilde{f}(x) = \int_H y\mu^x(dy), \quad x \in H,$$

which, according to (3.9), possesses property

$$\left\|\tilde{f}(x)\right\|^2 \leqslant C\left(1 + \|x\|^2\right), \quad x \in H, \tag{5.10}$$

then

$$\bar{f}(x) = -\gamma x + \tilde{f}(x). \tag{5.11}$$

For all $T_0 > 0$, we have by (5.11) and Hölder's inequality the estimate

$$\begin{aligned}
\mathbb{E} \sup_{0\leqslant t\leqslant T_0} \left\|I_3(t,\epsilon)\right\|^2 \leqslant{}& \delta C\mathbb{E} \sup_{0\leqslant t\leqslant T_0} \int_{\lfloor t/\delta\rfloor\delta}^{t} \left\|X^\epsilon_{\lfloor t/\delta\rfloor\delta} - X^\epsilon_s\right\|^2 ds + \delta C\mathbb{E} \sup_{0\leqslant t\leqslant T_0} \int_{\lfloor t/\delta\rfloor\delta}^{t} \left\|\hat{Y}^\epsilon_s\right\|^2 ds \\
&+ \delta C\mathbb{E} \sup_{0\leqslant t\leqslant T_0} \int_{\lfloor t/\delta\rfloor\delta}^{t} \left\|\tilde{f}\left(X^\epsilon_s\right)\right\|^2 ds.
\end{aligned} \tag{5.12}$$

Let us denote by $I_3^1(T_0,\epsilon)$, $I_3^2(T_0,\epsilon)$, $I_3^3(T_0,\epsilon)$ the three terms in the right hand side of (5.12), respectively. We observe that

$$\begin{aligned}
I_3^1(T_0,\epsilon) &\leqslant \delta C\mathbb{E}\Bigg\{\max_{\substack{k\delta\leqslant t\leqslant(k+1)\delta,\\ 0\leqslant k\leqslant\lfloor T_0/\delta\rfloor-1}}\int_{k\delta}^{t}\big\|X_{k\delta}^\epsilon-X_s^\epsilon\big\|^2\,ds\Bigg\}\\
&\quad+\delta C\mathbb{E}\Bigg\{\max_{\lfloor T_0/\delta\rfloor\delta\leqslant t\leqslant T_0}\int_{\lfloor T_0/\delta\rfloor\delta}^{t}\big\|X_{\lfloor T_0/\delta\rfloor\delta}^\epsilon-X_s^\epsilon\big\|^2\,ds\Bigg\}\\
&\leqslant \delta C\sum_{k=0}^{\lfloor T_0/\delta\rfloor-1}\int_{k\delta}^{(k+1)\delta}\mathbb{E}\big\|X_{k\delta}^\epsilon-X_s^\epsilon\big\|^2\,ds+\delta C\int_{\lfloor T_0/\delta\rfloor\delta}^{T_0}\mathbb{E}\big\|X_{\lfloor T/\delta\rfloor\delta}^\epsilon-X_s^\epsilon\big\|^2\,ds.
\end{aligned}$$

By (4.2), the above inequality yields

$$\begin{aligned}
I_3^1(T_0,\epsilon) &\leqslant \delta C\sum_{k=0}^{\lfloor T_0/\delta\rfloor-1}\int_{k\delta}^{(k+1)\delta}\delta\,ds+\delta C\int_{\lfloor T_0/\delta\rfloor\delta}^{T_0}\delta\,ds\\
&\leqslant C\delta.
\end{aligned}\tag{5.13}$$

Repeating the same argument used in the estimation of $I_3^1(T_0,\epsilon)$, it follows that

$$I_3^2(T_0,\epsilon)\leqslant \delta C\sum_{k=0}^{\lfloor T_0/\delta\rfloor-1}\int_{k\delta}^{(k+1)\delta}\mathbb{E}\big\|\hat{Y}_s^\epsilon\big\|^2\,ds+\delta C\int_{\lfloor T_0/\delta\rfloor}^{T_0}\mathbb{E}\big\|\hat{Y}_s^\epsilon\big\|^2\,ds.\tag{5.14}$$

Note that we can easily deduce the auxiliary process $\hat{Y}^\epsilon$ possess the property

$$\sup_{0<\epsilon<\epsilon_0}\ \sup_{0\leqslant t\leqslant T_0}\mathbb{E}\big\|\hat{Y}_t^\epsilon\big\|^2\leqslant C$$

for some constant $C>0$. Hence (5.14) yields

$$I_3^2(T_0,\epsilon)\leqslant C\delta.\tag{5.15}$$

Using (5.10) and (4.1) we can deduce

$$\begin{aligned}
I_3^3(T_0,\epsilon) &\leqslant \delta C\mathbb{E}\sup_{0\leqslant t\leqslant T_0}\int_{\lfloor t/\delta\rfloor\delta}^{t}\big(1+\big\|X_s^\epsilon\big\|^2\big)\,ds\\
&\leqslant C\delta^2+\delta C\int_0^{T_0}\mathbb{E}\big\|X_s^\epsilon\big\|^2\,ds\\
&\leqslant C\delta.
\end{aligned}\tag{5.16}$$

Collecting together (5.12), (5.13), (5.15) and (5.16) we obtain

$$\mathbb{E}\sup_{0\leqslant t\leqslant T_0}\big\|I_3(t,\epsilon)\big\|^2\leqslant C\delta.\tag{5.17}$$

Concerning $I_1(t,\epsilon)$, we can deduce

$$\begin{aligned}
\mathbb{E}\sup_{0\leqslant t\leqslant T_0}\left\|I_1(t,\epsilon)\right\|^2 &= \mathbb{E}\sup_{0\leqslant t\leqslant T_0}\left\|\sum_{k=0}^{\lfloor t/\delta\rfloor-1}\int_{k\delta}^{(k+1)\delta}\left(f\left(X^\epsilon_{k\delta},\hat{Y}^\epsilon_s\right)-\bar{f}\left(X^\epsilon_{k\delta}\right)\right)ds\right\|^2 \\
&\leqslant \mathbb{E}\sup_{0\leqslant t\leqslant T_0}\left\{\lfloor t/\delta\rfloor\sum_{k=0}^{\lfloor t/\delta\rfloor-1}\left\|\int_{k\delta}^{(k+1)\delta}\left(f\left(X^\epsilon_{k\delta},\hat{Y}^\epsilon_s\right)-\bar{f}\left(X^\epsilon_{k\delta}\right)\right)ds\right\|^2\right\} \\
&\leqslant \frac{T_0}{\delta}\sum_{k=0}^{\lfloor T_0/\delta\rfloor-1}\mathbb{E}\left\|\int_{k\delta}^{(k+1)\delta}\left(f\left(X^\epsilon_{k\delta},\hat{Y}^\epsilon_s\right)-\bar{f}\left(X^\epsilon_{k\delta}\right)\right)ds\right\|^2 \\
&\leqslant \frac{T_0^2}{\delta^2}\max_{0\leqslant k\leqslant\lfloor T_0/\delta\rfloor-1}\mathbb{E}\left\|\int_{k\delta}^{(k+1)\delta}\left(f\left(X^\epsilon_{k\delta},\hat{Y}^\epsilon_s\right)-\bar{f}\left(X^\epsilon_{k\delta}\right)\right)ds\right\|^2 \\
&= \frac{T_0^2}{\delta^2}\max_{0\leqslant k\leqslant\lfloor T_0/\delta\rfloor-1}\mathbb{E}\left\|\int_0^{\delta}\left(f\left(X^\epsilon_{k\delta},\hat{Y}^\epsilon_{s+k\delta}\right)-\bar{f}\left(X^\epsilon_{k\delta}\right)\right)ds\right\|^2 \\
&= \frac{C}{\delta^2}\max_{0\leqslant k\leqslant\lfloor T_0/\delta\rfloor-1}\mathbb{E}\left\|\int_0^{\delta}\left(f\left(X^\epsilon_{k\delta},Y^{X^\epsilon_{k\delta},Y^\epsilon_{k\delta}}_{s/\epsilon}\right)-\bar{f}\left(X^\epsilon_{k\delta}\right)\right)ds\right\|^2 \\
&= \frac{C}{\delta^2}\max_{0\leqslant k\leqslant\lfloor T_0/\delta\rfloor-1}\mathcal{J}^\epsilon_k.
\end{aligned}$$

Taking (5.5) into account, we can deduce that

$$\mathbb{E}\sup_{0\leqslant t\leqslant T_0}\left\|I_1(t,\epsilon)\right\|^2\leqslant C\frac{\epsilon}{\delta},$$

which, taking into account (5.8), (5.9) and (5.17), gives

$$\mathbb{E}\sup_{0\leqslant t\leqslant T_0}\left\|\int_0^t\left(f\left(X^\epsilon_{s(\delta)},\hat{Y}^\epsilon_s\right)-\bar{f}\left(X^\epsilon_s\right)\right)ds\right\|^2\leqslant C\left(\delta+\frac{\epsilon}{\delta}\right).$$

The proof is completed. □

Lemma 5.3. *Suppose that Assumptions 1–3 hold, then there exists a constant C such that*

$$\mathbb{E}\sup_{0\leqslant t\leqslant T_0}\left\|\hat{X}^\epsilon_t-\bar{X}_t\right\|^2\leqslant C\left(\delta+\frac{\epsilon}{\delta}\right). \tag{5.18}$$

Proof. For any $t\in[0,T_0]$ we write

$$\hat{X}^\epsilon_t-\bar{X}_t=\sum_{i=1}^5 A_i(t), \tag{5.19}$$

where

$$A_1(t) = \int_0^t \left(f\left(X^\epsilon_{s(\delta)}, \hat{Y}^\epsilon_s\right) - \bar{f}\left(X^\epsilon_s\right)\right) ds,$$

$$A_2(t) = \int_0^t \left(\bar{f}\left(X^\epsilon_s\right) - \bar{f}\left(\hat{X}^\epsilon_s\right)\right) ds,$$

$$A_3(t) = \int_0^t \left(\bar{f}\left(\hat{X}^\epsilon_s\right) - \bar{f}(\bar{X}_s)\right) ds,$$

$$A_4(t) = \int_0^t \left(\sigma_1\left(X^\epsilon_s\right) - \sigma_1\left(\hat{X}^\epsilon_s\right)\right) dW^{Q_1}_s,$$

$$A_5(t) = \int_0^t \left(\sigma_1\left(\hat{X}^\epsilon_s\right) - \sigma_1(\bar{X}_s)\right) dW^{Q_1}_s.$$

Using Hölder's inequality, the Burkholder–Davis–Gundy inequality and (4.4) we obtain

$$\begin{aligned} \mathbb{E} \sup_{0\leqslant t\leqslant T_0} \left\|A_2(t)\right\|^2 + \mathbb{E} \sup_{0\leqslant t\leqslant T_0} \left\|A_4(t)\right\|^2 &\leqslant C\mathbb{E} \sup_{0\leqslant t\leqslant T_0} \int_0^t \left\|X^\epsilon_s - \hat{X}^\epsilon_s\right\|^2 ds \\ &\leqslant C\mathbb{E} \sup_{0\leqslant t\leqslant T_0} \left\|X^\epsilon_t - \hat{X}^\epsilon_t\right\|^2 \\ &\leqslant C\delta. \end{aligned} \tag{5.20}$$

It is also easy to see that, for any $u \in [0, T_0]$, the estimation

$$\begin{aligned} \mathbb{E} \sup_{0\leqslant t\leqslant u} \left\|A_3(t)\right\|^2 + \mathbb{E} \sup_{0\leqslant t\leqslant u} \left\|A_5(t)\right\|^2 &\leqslant C\mathbb{E} \sup_{0\leqslant t\leqslant u} \int_0^t \left\|\hat{X}^\epsilon_s - \bar{X}_s\right\|^2 ds \\ &\leqslant C \int_0^u \mathbb{E} \sup_{0\leqslant r\leqslant s} \left\|\hat{X}^\epsilon_r - \bar{X}_r\right\|^2 ds. \end{aligned} \tag{5.21}$$

Form (5.7), (5.19), (5.20) and (5.21), we have for all $u \in [0, T_0]$,

$$\mathbb{E} \sup_{0\leqslant t\leqslant u} \left\|\hat{X}^\epsilon_t - \bar{X}_t\right\|^2 \leqslant C\left(\delta + \frac{\epsilon}{\delta}\right) + C \int_0^u \mathbb{E} \sup_{0\leqslant r\leqslant s} \left\|\hat{X}^\epsilon_r - \bar{X}_r\right\|^2 ds,$$

which, with aid of Gronwall's inequality, yields

$$\mathbb{E} \sup_{0\leqslant t\leqslant T_0} \left\|\hat{X}^\epsilon_t - \bar{X}_t\right\|^2 \leqslant C\left(\delta + \frac{\epsilon}{\delta}\right).$$

The proof is completed. □

With Lemma 4.3 and Lemma 5.3 in hand, we now formulate the main result in this paper.

Theorem 5.1. *Suppose that Assumptions 1–3 hold and $X_0, Y_0 \in H$, then we have the stochastic averaging principle with convergence rate* $1/2$, *that is,*

$$\mathbb{E} \sup_{0 \leqslant t \leqslant T_0} \left\| X_t^\epsilon - \bar{X}_t \right\|^2 = \mathcal{O}(\sqrt{\epsilon}).$$

Proof. Thanks to (4.4) and (5.18) we obtain

$$\mathbb{E} \sup_{0 \leqslant t \leqslant T_0} \left\| X_t^\epsilon - \bar{X}_t \right\|^2 \leqslant C\left(\delta + \frac{\epsilon}{\delta}\right).$$

Taking $\delta = \sqrt{\epsilon}$ in above inequality, we get

$$\mathbb{E} \sup_{0 \leqslant t \leqslant T_0} \left\| X_t^\epsilon - \bar{X}_t \right\|^2 \leqslant C\sqrt{\epsilon},$$

which completes the proof. □

Remark 5.1. It should be stressed that we have confined ourselves to the case the diffusion coefficient of the slow variable do not depend on the fast variable, that is, $\sigma_1(x, y) = \sigma_1(x)$. In fact, a simple example (see [14]) can show that for slow-fast system, strong convergence does not hold where the noise coefficient of the slow variable depends on fast variable.

Acknowledgments

The authors thank the editor and the anonymous reviewers for their detailed comments and valuable suggestions. We would also like to thank Professor Jinqiao Duan and Professor Xu Sun for helpful discussions and comments.

References

[1] J.G. Besjes, On the asymptotic methods for non-linear differential equations, J. Mecanique 8 (1969) 357–373.
[2] N.N. Bogoliubov, Y.A. Mitropolsky, Asymptotic Methods in the Theory of Non-Linear Oscillations, Gordon & Breach Science Publishers, New York, 1961.
[3] C.E. Bréhier, Strong and weak orders in averaging for SPDEs, Stochastic Process. Appl. 122 (2012) 2553–2593.
[4] S. Cerrai, Stochastic reaction–diffusion systems with multiplicative noise and non-Lipschitz reaction term, Probab. Theory Related Fields 125 (2003) 271–304.
[5] S. Cerrai, A Khasminkii type averaging principle for stochastic reaction–diffusion equations, Ann. Appl. Probab. 19 (3) (2009) 899–948.
[6] S. Cerrai, M.I. Freidlin, Averaging principle for a class of stochastic reaction–diffusion equations, Probab. Theory Related Fields 144 (2009) 137–177.
[7] P.L. Chow, Stochastic Partial Differential Equations, Chapman & Hall/CRC, New York, 2007.
[8] B. Deng, The existence of infinite many travelling front and backwaves in the FitzHugh–Nagumo equations, SIAM J. Math. Anal. 22 (1991) 1631–1650.
[9] R. Fitzhugh, Impulse and physiological states in models of nerve membrane, Biophys. J. 1 (1961) 445–466.
[10] M.I. Freidlin, A.D. Wentzell, Random Perturbation of Dynamical Systems, second ed., Springer, New York, 1998.
[11] M.I. Freidlin, A.D. Wentzell, Long-time behavior of weakly coupled oscillators, J. Stat. Phys. 123 (2006) 1311–1337.
[12] H. Fu, J. Liu, Strong convergence in stochastic averaging for two time-scales stochastic partial differential equations, J. Math. Anal. Appl. 384 (2011) 70–86.
[13] I.I. Gikhman, On a theorem of N. N. Bogoliubov, Ukrain. Mat. Zh. 4 (1952) 215–219 (in Russian).
[14] D. Givon, Strong convergence rate for two-time-scale jump-diffusion stochastic differential systems, Multiscale Model. Simul. 6 (2) (2007) 577–594.
[15] D. Givon, I.G. Kevrekidis, R. Kupferman, Strong convergence of projective integration schemes for singular perturbed stochastic differential systems, Commun. Math. Sci. 4 (2006) 707–729.
[16] J. Golec, Stochastic averaging principle for systems with pathwise uniqueness, Stoch. Anal. Appl. 13 (3) (1995) 307–322.
[17] J. Golec, G. Ladde, Averaging principle and systems of singularly perturbed stochastic differential equations, J. Math. Phys. 31 (1990) 1116–1123.

[18] H. Hasegawa, Stochastic bifurcation in FitzHugh–Nagumo ensembles subjected to additive and/or multiplicative noises, Phys. D 237 (2) (2008) 137–155.
[19] A.L. Hodgkin, A.F. Huxley, A quantitative description of membrane current and its application to conduction and excitation in nerve, J. Physiol. 117 (1952) 500–544.
[20] D.E. Jackson, Existence and regularity for the Fitzhugh–Nagumo equations with inhomogeneous boundary conditions, Nonlinear Anal. 14 (3) (1990) 201–216.
[21] C. Jones, Stability of the travelling waves solution of the FitzHugh–Nagumo system, Trans. Amer. Math. Soc. 286 (1984) 431–469.
[22] R.Z. Khasminskii, On the principle of averaging the Itô stochastic differential equations, Kibernetika 4 (1968) 260–279 (in Russian).
[23] D. Liu, Strong convergence of principle of averaging for multiscale stochastic dynamical systems, Commun. Math. Sci. 8 (4) (2010) 999–1020.
[24] O. Lopes, FitzHugh–Nagumo system: Boundedness and convergence to equilibrium, J. Differential Equations 44 (3) (1982) 400–413.
[25] D.G. Luchinsky, P.V.E. McClintock, A.V. Polovinkin, G.V. Osipov, Stochastic excitation and synchronization in coupled FitzHugh–Nagumo elements, in: L. Schimansky-Geier, D. Abbott, A. Neiman, C. Van den Broeck (Eds.), Noise in Complex Systems and Stochastic Dynamics, SPIE, Washington, 2003, pp. 301–308.
[26] Y. Lv, W. Wang, Limit dynamics for the stochastic FitzHugh–Nagumo system, Nonlinear Anal. Real World Appl. 11 (2010) 3091–3105.
[27] M. Marion, Finite-dimensional attractors associated with partly dissipative reaction–diffusion systems, SIAM J. Math. Anal. 20 (1989) 816–844.
[28] J.S. Nagumo, S. Arimoto, S. Yoshizawa, An active pulse transmission line simulating nerve axon, Proc. Inst. Radio Eng. 50 (1962) 2061–2071.
[29] B. Øksendal, Stochastic Differential Equations, sixth ed., Springer, Berlin, 2003.
[30] G. Da Prato, J. Zabczyk, Stochastic Equations in Infinite Dimensions, Cambridge University Press, Cambridge, 1992.
[31] Z. Shao, Existence of inertial manifold for partly dissipative reaction–diffusion systems in higher space dimensions, J. Differential Equations 144 (1998) 1–43.
[32] J. Tessonea, H.S. Wiob, Stochastic resonance in an extended FitzHugh–Nagumo system: The role of selective coupling, Phys. A 374 (1) (2007) 46–54.
[33] A.Y. Veretennikov, On the averaging principle for systems of stochastic differential equations, Math. USSR Sb. 69 (1991) 271–284.
[34] A.Y. Veretennikov, On large deviations in the averaging principle for SDEs with full dependence, Ann. Probab. 27 (1999) 284–296.
[35] V.M. Volosov, Averaging in systems of ordinary differential equations, Russian Math. Surveys 17 (1962) 1–126.
[36] B.X. Wang, Random attractors for the stochastic FitzHugh–Nagumo system on unbounded domains, Nonlinear Anal. 71 (2009) 2811–2828.
[37] W. Wang, A.J. Roberts, Average and deviation for slow-fast stochastic partial differential equations, J. Differential Equations 253 (2012) 1265–1286.
[38] D. Wu, S.Q. Zhu, Stochastic resonance in FitzHugh–Nagumo system with time-delayed feedback, Phys. Lett. A 372 (32) (2008) 5299–5304.
[39] Y. Xu, J. Duan, W. Xu, An averaging principle for stochastic dynamical systems with Lévy noise, Phys. D (2011) 1395–1401.