

# Average and deviation for the stochastic FitzHugh–Nagumo system

W. Wang<sup>1</sup>

A. J. Roberts<sup>2</sup>

(Received 31 July 2008; revised 10 November 2008)

## Abstract

An averaged system for the slow-fast stochastic FitzHugh–Nagumo system is derived in this paper. The rate of convergence in probability is obtained as a byproduct. Moreover the deviation between the original system and the averaged system is studied. A martingale approach proves that the deviation is described by a Gaussian process. The deviation gives a more accurate asymptotic approximation than previous work.

## Contents

### 1 Introduction

C293

---

<http://anziamj.austms.org.au/ojs/index.php/ANZIAMJ/article/view/1391>  
gives this article, © Austral. Mathematical Soc. 2008. Published November 11, 2008.  
ISSN 1446-8735. (Print two pages per sheet of paper.)

1	<i>Introduction</i>	C293
2	<b>Preliminaries and result on fast component</b>	<b>C296</b>
3	<b>Averaged equation</b>	<b>C297</b>
4	<b>Deviation estimate</b>	<b>C300</b>
5	<b>Numerical simulations</b>	<b>C305</b>
6	<b>Discussion</b>	<b>C306</b>
	<b>References</b>	<b>C306</b>

## 1 Introduction

The FitzHugh–Nagumo system arises as a simplification of the Hodgkin–Huxley model describing signal transmission across axons in neurobiology [2, e.g.]. Modelling the integral input from surrounding cells to the given neuron as a random signal, the following stochastic FitzHugh–Nagumo system arises with the activity  $\mathbf{u}^\epsilon(\mathbf{x}, t)$  and recovery fields  $\mathbf{v}^\epsilon(\mathbf{x}, t)$  satisfying

$$d\mathbf{u}^\epsilon = [\partial_{xx}\mathbf{u}^\epsilon + \mathbf{u}^\epsilon - (\mathbf{u}^\epsilon)^3 + \mathbf{v}^\epsilon] dt + \sigma_1 d\mathbf{W}_1, \quad (1)$$

$$d\mathbf{v}^\epsilon = \frac{1}{\epsilon} [\partial_{xx}\mathbf{v}^\epsilon + \mathbf{u}^\epsilon - \mathbf{v}^\epsilon] dt + \frac{\sigma_2}{\sqrt{\epsilon}} d\mathbf{W}_2, \quad (2)$$

with zero Dirichlet boundary conditions, and initial conditions  $\mathbf{u}^\epsilon(0) = \mathbf{u}_0 \in L^2(-1, 1)$  and  $\mathbf{v}^\epsilon(0) = \mathbf{v}_0 \in L^2(-1, 1)$  where  $L^2(-1, 1)$  is the Lebesgue space of square integrable real valued functions on the nondimensional interval  $(-1, 1)$ .  $\mathbf{W}_1(\mathbf{x}, t)$  and  $\mathbf{W}_2(\mathbf{x}, t)$  are mutually independent  $L^2(-1, 1)$  valued Wiener processes. In the following we denote by  $\mathbf{f}(\mathbf{u}, \mathbf{v}) = \mathbf{u} - \mathbf{u}^3 + \mathbf{v}$ ,  $\mathbf{g}(\mathbf{u}, \mathbf{v}) = \mathbf{u} - \mathbf{v}$  and by  $\mathbf{A} = \partial_{xx}$  with zero Dirichlet boundary condition on  $(-1, 1)$ .

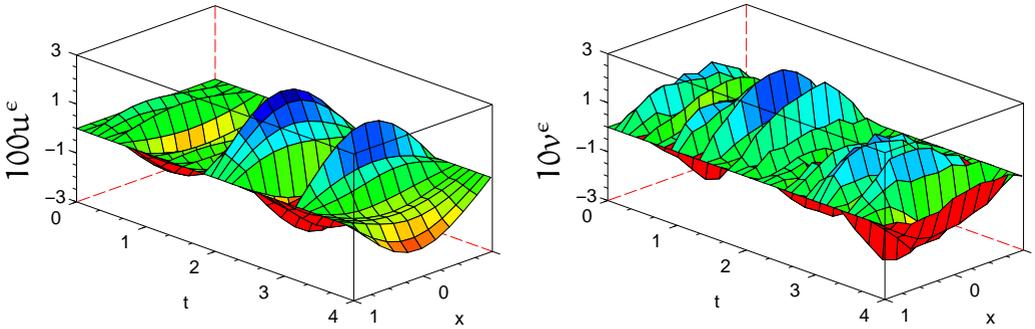


FIGURE 1: an example realisation of the FitzHugh–Nagumo system (1)–(2) with  $\sigma_1 = 0$ ,  $\sigma_2 = 3$ ,  $W_2 = (I - A)^{-1}Z$  for a cylindrical Wiener process  $Z(t)$ , and small parameter  $\epsilon = 0.1$ .

Figure 1 plots one example simulation of the FitzHugh–Nagumo system (1)–(2) showing that  $W_2$ , the noise forcing of  $v$ , feeds into the dynamics of  $u$ . Assume  $\epsilon > 0$  is small. Then the FitzHugh–Nagumo system (1)–(2) has two widely separated timescales. A simplified equation which governs the evolution of the system over the long time scale is highly desirable.

Such a simplified equation, capturing the dynamics of the system on the slow time scale, is often called an averaged equation. There is great deal of work on averaging principles for deterministic ordinary differential equations [9, e.g.] and for stochastic ordinary differential equations [4, e.g.]. But there are few results on averaging for stochastic partial differential equations (SPDEs). Recently an averaged equation for a class of reaction–diffusion equations with stochastic fast component was obtained by assuming that nonlinear terms are all Lipschitz [1]. The resultant averaged equation is deterministic. This article derives the averaged equation (3) modelling (1)–(2) and proves a square-root rate of convergence in probability. If for any fixed  $u$ , the fast system (2) has a unique invariant measure  $\mu_u$ , then as  $\epsilon \rightarrow 0$ , under a mixing assumption, the solution  $u^\epsilon$  of equation (1), converges in distribution

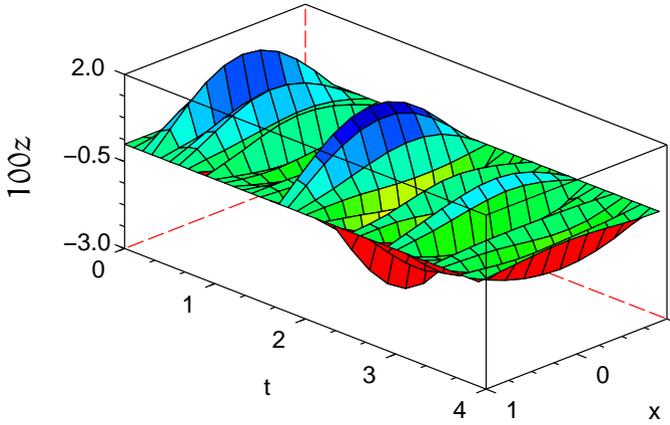


FIGURE 2: an example realisation of the stochastic deviation (22) for small parameter  $\epsilon = 0.1$ ,  $\sigma_2 = 3$  and  $Q_2 = (1 - \partial_{xx})^{-1}$ .

to the solution of

$$\begin{aligned} du &= [Au + u - u^3 + (I - A)^{-1}u] dt + \sigma_1 dW_1(t), & (3) \\ u(0) &= u_0. & (4) \end{aligned}$$

Section 3 proves the convergence rate is 1/2 in the sense that

$$\sup_{0 \leq t \leq T} \mathbb{E}|u^\epsilon(t) - u(t)| \leq C_T \epsilon^{1/2} \tag{5}$$

for some positive constant  $C_T$ .

Furthermore, the deviation between the original system and the averaged system is determined, the deviation SPDE (22), as shown in the example solution of Figure 2. By estimate (5), Section 4 proves that as  $\epsilon \rightarrow 0$  the limit of  $(u^\epsilon(t) - u(t))/\sqrt{\epsilon}$  is a Gaussian process and thus leads to the SPDE (22) for the deviation. Including the deviation SPDE (22) gives a much better approximation than the averaged equation (3). In particular, when the initial state  $u_0 = 0$  and there is no direct forcing of  $u$ ,  $\sigma_1 = 0$ , as used

in Figures 1 and 2, then the averaged solution is identically  $\mathbf{u}(t) = 0$ . In such a case, the dynamics of  $\mathbf{u}$  as seen in Figure 1 are modelled solely by deviations governed by the SPDE (22).

## 2 Preliminaries and result on fast component

Let  $H = L^2(-1, 1)$  with  $L^2$ -norm denoted by  $|\cdot|$  and inner product by  $\langle \cdot, \cdot \rangle$ . For any  $\alpha > 0$ ,  $\mathbf{u} \in H$  define  $|\mathbf{u}|_\alpha = |\mathbf{A}^{\alpha/2}\mathbf{u}|$  and for  $\alpha = 1$ , the norm is written as  $\|\cdot\|$ . Then let  $H_0^\alpha$  be the space of the closure of  $C_0^\infty(-1, 1)$ , the space of smooth functions with compact support on  $(-1, 1)$ , under the norm  $|\cdot|_\alpha$ . Furthermore, let  $H^{-\alpha}$  denote the dual space of  $H_0^\alpha$ . Denote by  $e^{\mathbf{A}t}$ ,  $t \geq 0$ , the compact analytic semigroup defined by  $\mathbf{A}$  on  $H$  and by  $\lambda_1 = -\pi^2/4$  the first eigenvalue of  $\mathbf{A}$ . Also we are given  $H$  valued  $Q$ -Wiener processes  $W_1(t)$  and  $W_2(t)$ ,  $t \geq 0$ , with covariance operator (spatial correlation)  $Q_1^2$  and  $Q_2^2$  respectively, which are mutually independent on a given probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  with  $\mathcal{F}$  a complete  $\sigma$ -algebra on  $\Omega$  [7]. Denote by  $\mathbb{E}$  the expectation operator with respect to  $\mathbb{P}$ . Here we assume  $\sigma_1 \in \mathbb{R}$ ,  $\sigma_2 > 0$  and make the hypothesis that

$$\text{tr}[(-\mathbf{A})^{1/2}Q_i^2] < \infty, \quad i = 1, 2. \tag{6}$$

By Theorem 11.7 of Prato and Zabczyk [7], we have the following result on the fast component.

**Theorem 1** *Assume hypothesis (6). For any fixed  $\mathbf{u} \in H$ , system (2) has a unique stationary solution,  $\eta_{\mathbf{u}}^\epsilon(t)$ , with distribution  $\mu_{\mathbf{u}} = \mathcal{N}(\bar{\eta}_{\mathbf{u}}, (\mathbf{I} - \mathbf{A})^{-1}\sigma_2^2Q_2^2/2 + (\mathbf{I} - \mathbf{A})^{-2}\mathbf{u} \otimes \mathbf{u})$  with  $\bar{\eta}_{\mathbf{u}} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{u}$ , which is exponential mixing in the following sense*

$$\mathbb{E}|v^\epsilon(t) - \eta_{\mathbf{u}}^\epsilon(t)|^2 \leq e^{-2(\lambda_1 - C_9)t/\epsilon} \mathbb{E}|v_0 - \eta_{\mathbf{u}}^\epsilon(0)|^2. \tag{7}$$

Moreover there is some  $\rho > 0$  such that

$$\left| \mathbb{E}f(\mathbf{u}, \mathbf{v}^\epsilon(t)) - \int_{\mathbb{H}} f(\mathbf{u}, \mathbf{x}) \mu_{\mathbf{u}}(d\mathbf{x}) \right| \leq C(1 + |\mathbf{v}_0|^2) e^{-\rho t/\epsilon}. \tag{8}$$

By the time scale transformation  $t \rightarrow \tau = t/\epsilon$ , (2) is transformed to

$$d\mathbf{v} = [\mathbf{A}\mathbf{v} + \mathbf{g}(\mathbf{u}, \mathbf{v})] d\tau + \sigma_2 d\tilde{W}_2(\tau), \quad \mathbf{v}(0) = \mathbf{v}_0, \tag{9}$$

where  $\tilde{W}_2$  is the scaled version of  $W_2$  and with same distribution. Then (9) has a unique stationary solution  $\eta_{\mathbf{u}}$  with distribution  $\mu_{\mathbf{u}}$ . Furthermore

$$\bar{f}(\mathbf{u}) := \int_{\mathbb{H}} f(\mathbf{u}, \mathbf{v}) \mu_{\mathbf{u}}(d\mathbf{v}) = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t f(\mathbf{u}, \eta_{\mathbf{u}}(s)) ds, \tag{10}$$

is well defined by the ergodic property of  $\mu_{\mathbf{u}}$ . It follows from the definition that  $\bar{f}(\mathbf{u}) = \mathbf{u} - \mathbf{u}^3 + \bar{\eta}_{\mathbf{u}}$ . Furthermore, by a generalized theorem on contracting maps depending on a parameter [1],  $\mathbf{v}_{\mathbf{u}}(t)$  is Fréchet differentiable with respect to  $\mathbf{u}$  and

$$\sup_{\mathbf{u}, \mathbf{v}_0 \in \mathbb{H}, 0 \leq t \leq T} |D_{\mathbf{u}} \mathbf{v}_{\mathbf{u}}|_{\mathcal{L}(\mathbb{H})} \leq C_T \tag{11}$$

where  $|\cdot|_{\mathcal{L}(\mathbb{H})}$  is the operator norm.

### 3 Averaged equation

This section gives the averaged equation and, as a byproduct, the convergence rate. First by the energy method [7], for any  $T > 0$ , there is a positive constant  $C_T$  which is independent of  $0 < \epsilon < 1$ , such that

$$\mathbb{E} \sup_{0 \leq t \leq T} \|\mathbf{u}^\epsilon(t)\|^2 + \mathbb{E} |\partial_t \mathbf{u}^\epsilon|_{L^2(0,T;H^{-1})} \leq C_T (\|\mathbf{u}_0\|^2 + \|\mathbf{v}_0\|^2) \tag{12}$$

and for any positive integer  $m$ ,

$$\mathbb{E} \int_0^T \|\mathbf{u}^\epsilon(s)\|^{2m} ds + \mathbb{E} \int_0^T \|\mathbf{v}^\epsilon(t)\|^2 ds \leq C_T(\|\mathbf{u}_0\|^{2m} + \|\mathbf{v}_0\|^{2m}). \quad (13)$$

Then a lemma by Simon [8] implies  $\{\mathcal{D}(\mathbf{u}^\epsilon)\}_{0 < \epsilon < 1}$ , the distributions of  $\mathbf{u}^\epsilon$ , are tight in space  $C(0, T; H)$ , that is, for any  $\kappa > 0$  there is a compact set  $K_\kappa$  in  $C(0, T; H)$  such that  $\mathbb{P}\{\mathbf{u}^\epsilon \in K_\kappa\} > 1 - \kappa$ . Here  $K_\kappa$  is chosen as a family of decreasing sets with respect to  $\kappa$ . Moreover, by the estimate (12) and the Markov inequality, we further choose the set  $K_\kappa$  such that there is some positive constant  $C_T^\kappa$ , for  $\mathbf{u}^\epsilon \in K_\kappa$ ,  $\|\mathbf{u}^\epsilon(t)\|^2 \leq C_T^\kappa$ ,  $t \in [0, T]$ . Then we have the following averaged result.

**Theorem 2** *Assume hypothesis (6). Given  $T > 0$ , for any  $\mathbf{u}_0 \in H$ , a solution  $\mathbf{u}^\epsilon(t, \mathbf{u}_0)$  of (1) converges in probability to  $\mathbf{u}$  in  $C(0, T; H)$  which solves (3)–(4). Moreover the convergence rate is  $1/2$ , that is, for some positive constant  $C_T > 0$*

$$\sup_{0 \leq t \leq T} \mathbb{E}|\mathbf{u}^\epsilon(t) - \mathbf{u}(t)| \leq C_T \sqrt{\epsilon}.$$

**Proof:** For any  $\kappa > 0$  we introduce a new probability space  $(\Omega_\kappa, \mathcal{F}_\kappa, \mathbb{P}_\kappa)$ , which is a subspace of  $(\Omega, \mathcal{F}, \mathbb{P})$  defined by

$$\Omega_\kappa = \{\omega \in \Omega : \mathbf{u}^\epsilon(\omega) \in K_\kappa\}, \quad \mathcal{F}_\kappa = \{S \cap \Omega_\kappa : S \in \mathcal{F}\}$$

and

$$\mathbb{P}_\kappa(S) = \frac{\mathbb{P}(S \cap \Omega_\kappa)}{\mathbb{P}(\Omega_\kappa)} \quad \text{for } S \in \mathcal{F}_\kappa.$$

Then  $\mathbb{P}(\Omega \setminus \Omega_\kappa) \leq \kappa$  by the choice of  $K_\kappa$ .

Now we restrict  $\omega \in \Omega_\kappa$  and introduce an auxiliary process. For any  $T > 0$ , partition the interval  $[0, T]$  into subintervals of length  $\delta = \sqrt{\epsilon}$ . Then

construct processes  $(\tilde{u}^\epsilon, \tilde{v}^\epsilon)$  such that for  $t \in [k\delta, (k+1)\delta)$ ,

$$\begin{aligned} \tilde{u}^\epsilon(t) &= e^{A(t-k\delta)}u^\epsilon(k\delta) + \int_{k\delta}^t e^{A(t-s)}f(u^\epsilon(k\delta), \tilde{v}^\epsilon(s)) ds \\ &\quad + \sigma_1 \int_{k\delta}^t e^{A(t-s)}dW_1(s), \quad \tilde{u}^\epsilon(0) = u_0, \end{aligned} \tag{14}$$

$$\begin{aligned} d\tilde{v}^\epsilon(t) &= \frac{1}{\epsilon} [A\tilde{v}^\epsilon(t) + g(u^\epsilon(k\delta), \tilde{v}^\epsilon(t))] dt + \frac{\sigma_2}{\sqrt{\epsilon}} dW_2(t), \\ \tilde{v}^\epsilon(k\delta) &= v^\epsilon(k\delta). \end{aligned} \tag{15}$$

By the definition of  $\Omega_\kappa$ ,  $\{u^\epsilon\}_\epsilon$  is compact in space  $C(0, T; H)$ , there is  $C_T > 0$ , such that

$$|u^\epsilon(t) - u^\epsilon(k\delta)|^2 \leq C_T\delta^2 \quad \text{for } t \in [k\delta, (k+1)\delta). \tag{16}$$

Then by the Itô formula and Gronwall lemma [7]

$$|v^\epsilon(t) - \tilde{v}^\epsilon(t)|^2 \leq C_T\delta^2, \quad t \in [0, T]. \tag{17}$$

Moreover, by the choice of  $\Omega_\kappa$  and the growth of  $f(\cdot, v)$ ,  $f(\cdot, v) : H \rightarrow H^{-\beta}$  is Lipschitz with  $-1/2 \leq \beta \leq -1/4$ . Then we have, by noticing (16),

$$|u^\epsilon(t) - \tilde{u}^\epsilon(t)| \leq C_T\delta, \quad t \in [0, T]. \tag{18}$$

On the other hand, in the mild sense

$$u(t) = e^{At}u_0 + \int_0^t e^{A(t-s)}\bar{f}(u(s)) ds + \sigma_1 \int_0^t e^{A(t-s)}dW_1(s).$$

Then, using  $[z]$  to denote the largest integer less than or equal to  $z$ ,

$$\begin{aligned} |\tilde{u}^\epsilon(t) - u(t)| &\leq \int_0^t e^{A(t-s)}|f(u^\epsilon(\lfloor s/\delta \rfloor \delta), \tilde{v}^\epsilon(s)) - \bar{f}(u^\epsilon(\lfloor s/\delta \rfloor \delta))| ds \\ &\quad + \int_0^t e^{A(t-s)}|\bar{f}(u^\epsilon(\lfloor s/\delta \rfloor \delta)) - \bar{f}(u^\epsilon(s))| ds \end{aligned}$$

$$+ \int_0^t e^{A(t-s)} |\bar{f}(\mathbf{u}^\epsilon(s)) - \bar{f}(\mathbf{u}(s))| ds.$$

Moreover, by estimate (11), we have

$$\begin{aligned} & \frac{1}{\tau} \left| \int_0^\tau [f(\mathbf{u}^\epsilon(s), \mathbf{v}_{\mathbf{u}^\epsilon(s)}(\delta)) - f(\mathbf{u}(s), \mathbf{v}_{\mathbf{u}(s)}(\delta))] d\delta \right|_{\mathbb{H}} \\ & \leq 2 \left[ \|\mathbf{u}^\epsilon(s)\|^2 + \|\mathbf{u}(s)\|^2 + \sup_{\mathbf{u}, \mathbf{v}_0 \in \mathbb{H}, 0 \leq t \leq T} |\mathbf{D}_{\mathbf{u}} \mathbf{v}(t)|_{\mathcal{L}(\mathbb{H})} \right] |\mathbf{u}^\epsilon(s) - \mathbf{u}(s)|. \end{aligned}$$

Then by (10) we have for  $t \in [0, T]$

$$|\tilde{\mathbf{u}}^\epsilon(t) - \mathbf{u}(t)| \leq C_T \left[ \delta + \int_0^T |\mathbf{u}^\epsilon(s) - \mathbf{u}(s)| ds \right]. \quad (19)$$

As  $|\mathbf{u}^\epsilon(t) - \mathbf{u}(t)| \leq |\mathbf{u}^\epsilon(t) - \tilde{\mathbf{u}}(t)| + |\tilde{\mathbf{u}}(t) - \mathbf{u}(t)|$ , by the Gronwall lemma and (16), (18) and (19) we have for  $t \in [0, T]$ ,

$$|\mathbf{u}^\epsilon(t) - \mathbf{u}(t)| \leq C_T \sqrt{\epsilon}. \quad (20)$$

Now by the arbitrariness of  $\kappa$ , the proof is complete. 

## 4 Deviation estimate

The previous section proved that, formally,  $\mathbf{u}^\epsilon(t) = \mathbf{u}(t) + \mathcal{O}(\epsilon^{1/2})$ . This section determines the coefficient of  $\epsilon^{1/2}$ , the so-called deviation:

$$\mathbf{z}^\epsilon(t) = \frac{1}{\sqrt{\epsilon}} (\mathbf{u}^\epsilon - \mathbf{u}). \quad (21)$$

**Theorem 3**  $z^\epsilon$  converges in distribution to a process  $z$  in space  $C(0, T; H)$  which solves

$$dz = [Az + (1 - 3u^2)z] dt + \sigma_2(I - A)^{-1}Q_2 d\bar{W}(t) \tag{22}$$

where  $\bar{W}(t)$  is a cylindrical Wiener process defined on an enlarged probability space.

**Proof:** The deviation  $z^\epsilon$  satisfies, by (1) and (3)

$$\dot{z}^\epsilon = Az^\epsilon + \frac{1}{\sqrt{\epsilon}} [f(u^\epsilon, v^\epsilon) - \bar{f}(u)] , \quad z^\epsilon(0) = 0. \tag{23}$$

Then by the assumptions on  $f$  and the compact property  $e^{At}$ , we have for some positive  $1 > \delta > 0$  and  $1 > \alpha > 0$  such that

$$\mathbb{E}|z^\epsilon|_{C^\delta(0, T; H)} \leq C_T(|u_0|^2 + |v_0|^2) \tag{24}$$

$$\text{and} \quad \mathbb{E} \sup_{0 \leq t \leq T} |z^\epsilon(t)|_\alpha \leq C_{T, \alpha}(|u_0|^2 + |v_0|^2). \tag{25}$$

Here  $C^\delta(0, T; H)$  is the Hölder space with exponent  $\delta$ . Then by the compact embedding of  $C^\delta(0, T; H) \cap C(0, T; H^\alpha) \subset C(0, T; H)$ , the distribution of  $\{z^\epsilon\}_\epsilon$ , denoted by  $\{\nu^\epsilon\}_\epsilon$ , is precompact.

Write  $z^\epsilon = z_1^\epsilon + z_2^\epsilon$  with

$$\dot{z}_1^\epsilon = Az_1^\epsilon + \frac{1}{\sqrt{\epsilon}}(\eta_u^\epsilon - \bar{\eta}_u), \quad z_1(0)^\epsilon = 0 \tag{26}$$

$$\text{and} \quad \dot{z}_2^\epsilon = Az_2^\epsilon + \frac{1}{\sqrt{\epsilon}}[f(u^\epsilon, v^\epsilon) - f(u, \eta_u^\epsilon)], \quad z_2^\epsilon(0) = 0, \tag{27}$$

respectively, and consider  $z_1^\epsilon$  and  $z_2^\epsilon$  separately. Theorem 1 defines  $\bar{\eta}_u$ . Denote by  $\nu_1^\epsilon$  be the probability measure of  $z_1^\epsilon$  induced on the space  $C(0, T; H)$ . For  $\gamma > 0$ , denote by  $UC^\gamma(H, \mathbb{R})$  the space of all functions from  $H$  to  $\mathbb{R}$

which, together with all the order till to  $\gamma$  order Fréchet derivatives, are uniformly continuous. For  $\mathbf{h} \in \mathbf{UC}^\gamma(\mathbf{H}, \mathbb{R})$ , denote by  $\mathbf{h}'$  and  $\mathbf{h}''$  the first and second order Fréchet derivative. We have the following lemma by a martingale approach [3].

**Lemma 4** Assume hypothesis (6). Any limiting measure of  $\nu_1^\epsilon$ , denote by  $\mathbb{P}^0$ , solves the following martingale problem on  $\mathbf{C}(0, T; \mathbf{H})$ :  $\mathbb{P}^0\{\mathbf{z}_1(0) = 0\} = 1$ ,

$$\mathbf{h}(\mathbf{z}_1(t)) - \mathbf{h}(\mathbf{z}_1(0)) - \int_0^t \langle \mathbf{h}'(\mathbf{z}_1(\theta)), \partial_{\mathbf{x}\mathbf{x}} \mathbf{z}_1(\theta) \rangle d\theta - \frac{1}{2} \int_0^t \text{tr} [\mathbf{h}''(\mathbf{z}_1(\theta)) \mathbf{B}(\mathbf{u})] d\theta$$

is a  $\mathbb{P}^0$ -martingale for any  $\mathbf{h} \in \mathbf{UC}^2(\mathbf{H}, \mathbb{R})$ . Here

$$\mathbf{B}(\mathbf{u}) = 2 \int_0^\infty \mathbb{E}[(\eta_{\mathbf{u}}(t) - \bar{\eta}_{\mathbf{u}}) \otimes (\eta_{\mathbf{u}}(0) - \bar{\eta}_{\mathbf{u}})] dt = \sigma_2^2 (\mathbf{I} - \mathbf{A})^{-2} \mathbf{Q}_2^2 \quad (28)$$

with  $\otimes$  being the tensor product.

**Proof:** For any  $0 < s \leq t < \infty$  and  $\mathbf{h} \in \mathbf{UC}^\infty(\mathbf{H}, \mathbb{R})$  we have

$$\begin{aligned} \mathbf{h}(\mathbf{z}_1^\epsilon(t)) - \mathbf{h}(\mathbf{z}_1^\epsilon(s)) &= \int_s^t \langle \mathbf{h}'(\mathbf{z}_1^\epsilon(\theta)), \partial_t \mathbf{z}_1^\epsilon \rangle d\tau \\ &= \int_s^t \langle \mathbf{h}'(\mathbf{z}_1^\epsilon(\theta)), \partial_{\mathbf{x}\mathbf{x}} \mathbf{z}_1^\epsilon(\theta) \rangle d\theta + \frac{1}{\sqrt{\epsilon}} \int_s^t \langle \mathbf{h}'(\mathbf{z}_1^\epsilon(\theta)), \eta_{\mathbf{u}}^\epsilon(\theta) - \bar{\eta}_{\mathbf{u}} \rangle d\theta. \end{aligned}$$

Rewrite the second term on the right-hand of above equation as

$$\begin{aligned} \mathbf{L}_1 + \mathbf{L}_2 + \mathbf{L}_3 &= \frac{1}{\sqrt{\epsilon}} \int_s^t \langle \mathbf{h}'(\mathbf{z}_1^\epsilon(t)), \eta_{\mathbf{u}}^\epsilon(\theta) - \bar{\eta}_{\mathbf{u}} \rangle d\theta \\ &\quad - \frac{1}{\sqrt{\epsilon}} \int_s^t \int_\theta^t \mathbf{h}''(\mathbf{z}_1^\epsilon(\delta)) (\eta_{\mathbf{u}}^\epsilon(\theta) - \bar{\eta}_{\mathbf{u}}, \mathbf{A} \mathbf{z}_1^\epsilon(\delta)) d\delta d\theta \\ &\quad - \frac{1}{\epsilon} \int_s^t \int_\theta^t \mathbf{h}''(\mathbf{z}_1^\epsilon(\delta)) (\eta_{\mathbf{u}}^\epsilon(\theta) - \bar{\eta}_{\mathbf{u}}, \eta_{\mathbf{u}}^\epsilon(\delta) - \bar{\eta}_{\mathbf{u}}) d\delta d\theta, \end{aligned}$$

where  $L_1$ ,  $L_2$  and  $L_3$  denote the separate lines of the right-hand of the above equation respectively. Denote by  $\Sigma^\epsilon(\delta, \theta) = (\eta_u^\epsilon(\theta) - \bar{\eta}_u) \otimes (\eta_u^\epsilon(\delta) - \bar{\eta}_u)$ . Then

$$\begin{aligned}
 L_3 &= -\frac{1}{\epsilon} \sum_{ij} \int_s^t \int_\theta^t \partial_{ij} h(z_1^\epsilon(\delta)) \langle \Sigma^\epsilon(\delta, \theta) e_i, e_j \rangle d\delta d\theta & (29) \\
 &= -\frac{1}{\epsilon} \sum_{ij} \int_s^t \int_\theta^t \int_\delta^t \langle \partial_{ij} h'(z_1^\epsilon(\lambda)), [z_1^\epsilon(\lambda) + \frac{1}{\sqrt{\epsilon}}(\eta_u^\epsilon(\lambda) - \bar{\eta}_u)] \rangle \\
 &\quad \times \langle \tilde{\Sigma}^\epsilon(\delta, \theta) e_i, e_j \rangle d\lambda d\delta d\theta \\
 &\quad + \frac{1}{\epsilon} \sum_{ij} \int_s^t \int_\tau^t \partial_{ij} h(z_1^\epsilon(t)) \langle \tilde{\Sigma}^\epsilon(\delta, \tau) e_i, e_j \rangle d\delta d\tau \\
 &\quad + \frac{1}{\epsilon} \sum_{ij} \int_s^t \int_\tau^t \partial_{ij} h(z_1^\epsilon(\delta)) \langle \mathbb{E}[\Sigma^\epsilon(\delta, \tau)] e_i, e_j \rangle d\delta d\tau & (30)
 \end{aligned}$$

where  $\tilde{\Sigma}^\epsilon(\delta, \theta) = \Sigma^\epsilon(\delta, \theta) - \mathbb{E}[\Sigma^\epsilon(\delta, \theta)]$ . Here  $\{e_i\}_{i=1}^\infty$  is one eigenbasis of  $H$  and  $\partial_{ij} = \partial_{e_i} \partial_{e_j}$  with  $\partial_{e_i}$  is the directional derivative in direction  $e_i$ . Now for any bounded continuous function  $\Phi$  on  $C(0, s; H)$ , let  $\Phi(\cdot, \omega) = \Phi(z_1^\epsilon(\cdot, \omega))$ . Denote by  $L_{31}$ ,  $L_{32}$  and  $L_{33}$  the separate lines of the right-hand of (30) respectively. Then by (8),  $|\mathbb{E}[(L_{31} + L_{32})\Phi]| \rightarrow 0$  as  $\epsilon \rightarrow 0$ . As  $\eta_u(t) - \bar{\eta}_u$  is stationary correlated and independent of  $u$ , by the exponential mixing property, if  $\epsilon_n \rightarrow 0$  as  $n \rightarrow \infty$ ,  $\nu^{\epsilon_n} \rightarrow P^0$ ,

$$\lim_{n \rightarrow \infty} \mathbb{E}[L_3\Phi] = \frac{1}{2} \int_s^t \mathbb{E}^{P^0} \left( \text{tr} [h''(z_1(\theta))B(u)] \Phi \right) d\theta,$$

where  $B(u)$  is defined by (28). Similarly by (8),  $\mathbb{E}[L_1\Phi + L_2\Phi] \rightarrow 0$  as  $\epsilon \rightarrow 0$ . By the tightness of  $z^\epsilon$  in  $C(0, T; H)$ , the sequence  $z_1^{\epsilon_n}$  has a limit process, denoted by  $z_1$ , in a weak sense. Then

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \int_s^t \langle h'(z_1^{\epsilon_n}(\theta)), Az_1^{\epsilon_n}(\theta) \rangle \Phi d\theta \right] = \mathbb{E} \left[ \int_s^t \langle h'(z_1(\theta)), Az_1(\theta) \rangle \Phi d\theta \right]$$

and

$$\lim_{n \rightarrow \infty} \mathbb{E}[(h(z_1^{\epsilon_n}(t)) - h(z_1^{\epsilon_n}(s)))\Phi] = \mathbb{E}[(h(z_1(t)) - h(z_1(s)))\Phi].$$

At last we have

$$\begin{aligned} \mathbb{E}^{P^0} [(h(z_1)(t) - h(z_1(s)))\Phi] = & \tag{31} \\ \mathbb{E}^{P^0} \left[ \int_s^t \langle h'(z_1(\theta)), Az_1(\theta) \rangle \Phi \, d\theta \right] + \frac{1}{2} \mathbb{E}^{P^0} \left\{ \int_s^t \text{tr} [h''(z_1(\theta))B(u)] \Phi \, d\theta \right\}. \end{aligned}$$

By an approximation argument we can prove (31) holds for all  $h \in UC^2(H)$ . This completes the proof. ♠

Denote by  $P^0$  the limit point of  $\nu_1^\epsilon$ . Then by the relation between the martingale problem and the weak solution of SPDES [6],  $P^0$  solves the martingale problem related to the following SPDE

$$dz_1 = Az_1 \, dt + \sqrt{B(u)} \, d\bar{W}(t). \tag{32}$$

where  $B(u)$  is defined in Lemma 4. Here  $\bar{W}(t)$  is a cylindrical Wiener process, with covariance operator being the identity operator on  $H$ , defined on a new probability space  $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$ . Then by the uniqueness of the solution to (32),  $z_1^\epsilon$  converges in distribution to  $z_1$  in  $C(0, T; H)$ .

On the other hand, let  $z_2$  be one limit point of  $z_2^\epsilon$  in  $C(0, T; H)$ . By the assumption on  $f$  and (7) we derive that  $z_2$  is the unique solution of the following equation

$$\dot{z}_2 = Az_2 + (1 - 3u^2)z, \quad z_2(0) = 0 \tag{33}$$

with  $z = z_1 + z_2$ . Then we have that  $z^\epsilon$  uniquely converges in distribution to  $z$  which solves (22). This proves Theorem 3. ♠

## 5 Numerical simulations

The numerical simulations plotted in Figures 1 and 2 employed standard routines for integrating stochastic ordinary differential equations (SDEs) [5]. The stochastic FitzHugh–Nagumo system (1)–(2) is discretised in space on just nine spatial grid points using simple centred differences to approximate spatial derivatives. Because of the dissipation of high wavenumber modes, no significant difference was observed with a more refined grid. The spatial noise was simply approximated by an independent white noise, Wiener process, at each grid point for the  $v^\epsilon$  equation (2). This spatial discretisation results in a coupled system of 18 stochastic ordinary differential equations with forcing by nine independent Wiener processes.

Analogous spatial discretisations modelled the stochastic deviation equation (22).

To integrate in time the spatially discretised systems, note that in both the stochastic FitzHugh–Nagumo system and the stochastic deviation equation the forcing noise appears linearly. Thus the Itô and Stratonovich version of a time integrator are identical. The time integration was performed with a standard explicit, strong, first order method for SDEs with multiple independent noises, the method chosen from the book by Kloeden and Platen [5]. Because the SDEs are sourced from SPDEs, the SDEs are stiff and a small time step of  $\Delta t = 0.002$  ensured stability. Since the simulations are to support a proof of principle, we did not seek more efficient implicit methods. All simulations executed quickly in the freely available package Scilab <sup>1</sup> that also drew Figures 1 and 2.

---

<sup>1</sup><http://www.scilab.org>

## 6 Discussion

Usually an averaged model, such as (3), excludes fluctuations caused by the fast component and is thus often not a good approximation. For example, as in Figure 1, when the initial  $\mathbf{u}_0 = \mathbf{0}$  the averaged model (3) predicts  $\mathbf{u} = \mathbf{0}$  and so misses the apparent stochastic fluctuations. However, the deviation estimate, the solution of (1) with  $\mathbf{u}_0 = \mathbf{0}$  is modelled appropriately by the deviation  $\sqrt{\epsilon}\mathbf{z}$  with  $\mathbf{z}$  solving (22), as in Figure 1. Our result gives a better model than the usual averaged equation. Furthermore, our approach shows that the fluctuation is Gaussian.

**Acknowledgements** This research is supported by the Australian Research Council grant DP0774311 and NSFC grant 10701072.

## References

- [1] S. Cerrai and M. Freidlin, Averaging principle for a class of stochastic reaction–diffusion equations, to appear in *Probab. Th. and Rel. Fields*. doi:10.1007/s00440-008-0144-z C294, C297
- [2] R. FitzHugh, Impulses and physiological states in theoretical models of nerve membrane. *Biophys J.*, **1**(1961). C293
- [3] H. Kesten and G. C. Papanicolaou, A limit theorem for turbulent diffusion, *Commun. Math. Phys.*, **65**(1979), 79–128. doi:10.1007/3-540-08853-9 C302
- [4] R. Z. Khasminskii, On the principle of averaging the Itô stochastic differential equations (Russian), *Kibernetika*, **4**(1968), 260–279. C294

- [5] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics*. Springer–Verlag, 1992. [C305](#)
- [6] M. Metivier, *Stochastic Partial Differential Equations in Infinite Dimensional Spaces*, Scuola Normale Superiore, Pisa, 1988. [C304](#)
- [7] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, 1992. [C296](#), [C297](#), [C299](#)
- [8] J. Simon, Compact sets in the space  $L^p(0, T; B)$ , *Ann. Mat. Pura Appl.*, **146** (1987), 65–96. [doi:10.1007/BF01762360](https://doi.org/10.1007/BF01762360) [C298](#)
- [9] V. M. Volosov, Averaging in systems of ordinary differential equations. *Russ. Math. Surv.*, **17**(1962), 1–126. [doi:10.1070/RM1962v017n06ABEH001130](https://doi.org/10.1070/RM1962v017n06ABEH001130) [C294](#)

## Author addresses

1. **W. Wang**, School of Mathematical Sciences, University of Adelaide, Adelaide, AUSTRALIA; and Department of Mathematics, Nanjing University, Nanjing, CHINA.  
<mailto:w.wang@adelaide.edu.au>, [wangweinju@yahoo.com.cn](mailto:wangweinju@yahoo.com.cn)
2. **A. J. Roberts**, School of Mathematical Sciences, University of Adelaide, Adelaide, AUSTRALIA.  
<mailto:anthony.roberts@adelaide.edu.au>