NOLTA, IEICE

Paper

Selective averaging with application to phase reduction and neural control

Gerd S. Schmidt [1], *Dan Wilson* [2], *Frank Allgöwer* [1], *and Jeff Moehlis* [2a)]

[1] *Institute of Systems Theory and Automatic Control, University of Stuttgart, Germany*

[2] *Department of Mechanical Engineering, University of California, Santa Barbara, CA, 93106, USA*

[a)] *moehlis@engineering.ucsb.edu*

Received January 23, 2014; Published October 1, 2014

Abstract: For a class of vector fields, we show that one can selectively average terms which are of the same order in a small parameter, giving an extension of standard averaging results. Such selective averaging is illustrated for the phase reduction of a system of oscillators with both coupling and external input, for which the coupling can be averaged to give a term which only depends on phase differences, while the external input term is not averaged. For a coupled two-neuron system, we use selectively averaged equations to find the optimal input which takes the in-phase state to the anti-phase state.

Key Words: averaging, phase reduction, optimal control, mathematical neuroscience

1. Introduction

Averaging is a powerful analysis technique in the dynamical systems toolbox, particularly for nonlinear oscillations subjected to small perturbations [5, 13]. Typically, the averaged system retains important information about the solutions for the original system, but is easier to analyze.

In this paper, we will extend averaging theorems from [13] to the case that we call *selective averaging*, in which certain terms for a vector field are averaged while others are not. In

$$\dot{x} = \varepsilon f^0(x,t) + \varepsilon f^1(x,t), \qquad x(0) = x_0, \tag{1}$$

where $f^i : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ is continuous in x and t for $i = 0, 1$, this corresponds to averaging f^0 but not f^1. The use of selective averaging will be illustrated for the phase reduction of a system of oscillators with both coupling and external input. Moreover, for a coupled two-neuron system, we use selectively averaged equations to find the optimal input which takes the in-phase state to the anti-phase state.

2. The selective averaging theorem

We assume that

$$M_i \equiv \sup_{x \in D} \sup_{0 \le \varepsilon t \le L} ||f^i(x,t)|| < \infty, \qquad i = 0, 1,$$

Nonlinear Theory and Its Applications, IEICE, vol. 5, no. 4, pp. 424–435 ©IEICE 2014 DOI: 10.1588/nolta.5.424

where $D \subset \mathbb{R}^n$ and L is chosen so that $x(t) \in D$ for all $0 \le t \le L/\varepsilon$. Moreover, we assume that f^0 and f^1 satisfy

$$||f^i(x,t) - f^i(y,t)|| \le \lambda_{f^i}||x-y||, \qquad i = 0, 1$$

for $x, y \in D$; here λ_{f^i} is called a Lipschitz constant for f^i.

We first define the *local average* f_T of a function $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ to be

$$f_T(x,t) := \frac{1}{T}\int_0^T f(x, t+s)\mathrm{d}s.$$

The following lemmas will allow us to prove the Selective Averaging Theorem.

Lemma 2.1 (Lemma 4.2.3 from [13]) If the continuous vector field $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ is T-periodic in t, then

$$f_T(x,t) = \overline{f}(x) = \frac{1}{T}\int_0^T f(x,s)ds.$$

Lemma 2.2 Consider the initial value problem (1). With t on the time scale $\frac{1}{\varepsilon}$, the solution $x(t)$ satisfies

$$\left\| x_T(t) - x_0 - \varepsilon\int_0^t f_T^0(x(\sigma),\sigma)\mathrm{d}\sigma - \frac{\varepsilon}{T}\int_0^T\int_0^{t+s} f^1(x(\sigma),\sigma)\mathrm{d}\sigma\mathrm{d}s \right\| \le \frac{1}{2}\varepsilon T((1+\lambda_{f^0}L)M_0 + \lambda_{f^0}LM_1).$$

Proof The proof is similar to [13, Lemma 4.2.7]. We express the solution to (1) as

$$x(t) = x_0 + \varepsilon\int_0^t f^0(x(\sigma),\sigma)\mathrm{d}\sigma + \varepsilon\int_0^t f^1(x(\sigma),\sigma)\mathrm{d}\sigma.$$

Calculating the local average of the solution, we obtain

$$x_T(t) = x_0 + \frac{\varepsilon}{T}\int_0^T\int_0^{t+s} f^0(x(\sigma),\sigma)\mathrm{d}\sigma\mathrm{d}s + \frac{\varepsilon}{T}\int_0^T\int_0^{t+s} f^1(x(\sigma),\sigma)\mathrm{d}\sigma\mathrm{d}s.$$

Now, the term

$$\begin{aligned}\frac{\varepsilon}{T}\int_0^T\int_0^{t+s} f^0(x(\sigma),\sigma)\mathrm{d}\sigma\mathrm{d}s &= \frac{\varepsilon}{T}\int_0^T\int_0^{t} f^0(x(\sigma+s),\sigma+s)\mathrm{d}\sigma\mathrm{d}s + \varepsilon R_1\\ &= \frac{\varepsilon}{T}\int_0^t\int_0^{T} f^0(x(\sigma),\sigma+s)\mathrm{d}s\mathrm{d}\sigma + \varepsilon R_1 + \varepsilon R_2,\end{aligned}$$

where

$$\|R_1\| = \left\|\frac{1}{T}\int_0^T\int_0^s f^0(x(\sigma),\sigma)\mathrm{d}\sigma\mathrm{d}s\right\| \le \frac{1}{T}\int_0^T\int_0^s M_0\mathrm{d}\sigma\mathrm{d}s = \frac{1}{2}M_0T,$$

and

$$\begin{aligned}\|R_2\| &= \left\|\frac{1}{T}\int_0^t\int_0^T [f^0(x(\sigma+s),\sigma+s) - f^0(x(\sigma),\sigma+s)]\mathrm{d}s\mathrm{d}\sigma\right\|\\ &\le \frac{\lambda_{f^0}}{T}\int_0^t\int_0^T \|x(\sigma+s) - x(\sigma)\|\mathrm{d}s\mathrm{d}\sigma\\ &= \varepsilon\frac{\lambda_{f^0}}{T}\int_0^t\int_0^T \left\|\int_\sigma^{\sigma+s}(f^0(x(\zeta),\zeta) + f^1(x(\zeta),\zeta))\mathrm{d}\zeta\right\|\mathrm{d}s\mathrm{d}\sigma\\ &\le \varepsilon\frac{\lambda_{f^0}}{T}\int_0^t\int_0^T\int_\sigma^{\sigma+s} \|f^0(x(\zeta),\zeta) + f^1(x(\zeta),\zeta)\|\mathrm{d}\zeta\mathrm{d}s\mathrm{d}\sigma\\ &\le \varepsilon\frac{\lambda_{f^0}}{T}\int_0^t\int_0^T\int_\sigma^{\sigma+s} (\|f^0(x(\zeta),\zeta)\| + \|f^1(x(\zeta),\zeta)\|)\mathrm{d}\zeta\mathrm{d}s\mathrm{d}\sigma\\ &\le \varepsilon\frac{\lambda_{f^0}}{T}\int_0^t\int_0^T (M_0s + M_1s)\mathrm{d}s\mathrm{d}\sigma = \frac{1}{2}\varepsilon\lambda_{f^0}t(M_0+M_1)T\\ &\le \frac{1}{2}\lambda_{f^0}L(M_0+M_1)T.\end{aligned}$$

Putting these expressions together gives the result. ∎

Lemma 2.3 Consider the initial value problem (1). If y is the solution of the initial value problem

$$\dot{y} = \varepsilon f_T^0(y,t) + \varepsilon f^1(y,t), \qquad y(0) = x_0,$$

then $x(t) = y(t) + \mathcal{O}(\varepsilon T)$ on the time scale $\frac{1}{\varepsilon}$.

Proof The proof is similar to [13, Lemma 4.2.8].

$$\begin{aligned}
&\left\| x(t) - x_0 - \varepsilon \int_0^t f_T^0(x(\sigma),\sigma)\mathrm{d}\sigma - \varepsilon \int_0^t f^1(x(\sigma),\sigma)\mathrm{d}\sigma \right\| \\
&\leq \left\| x(t) - x_0 - \varepsilon \int_0^t f_T^0(x(\sigma),\sigma)\mathrm{d}s - \frac{\varepsilon}{T}\int_0^T \int_0^{t+s} f^1(x(\sigma),\sigma)\mathrm{d}\sigma\mathrm{d}s \right\| \\
&\quad + \left\| \frac{\varepsilon}{T}\int_0^T \int_0^{t+s} f^1(x(\sigma),\sigma)\mathrm{d}\sigma\mathrm{d}s - \varepsilon \int_0^t f^1(x(\sigma),\sigma)\mathrm{d}\sigma \right\|.
\end{aligned}$$

Now, the term

$$\begin{aligned}
&\left\| \frac{\varepsilon}{T}\int_0^T \int_0^{t+s} f^1(x(\sigma),\sigma)\mathrm{d}\sigma\mathrm{d}s - \varepsilon \int_0^t f^1(x(\sigma),\sigma)\mathrm{d}\sigma \right\| \\
&= \left\| \frac{\varepsilon}{T}\int_0^T \int_0^{t+s} f^1(x(\sigma),\sigma)\mathrm{d}\sigma\mathrm{d}s - \frac{\varepsilon}{T}\int_0^T \int_0^{t} f^1(x(\sigma),\sigma)\mathrm{d}\sigma\mathrm{d}s \right\| \\
&= \left\| \frac{\varepsilon}{T}\int_0^T \left(\int_0^{t+s} f^1(x(\sigma),\sigma)\mathrm{d}\sigma - \int_0^{t} f^1(x(\sigma),\sigma)\mathrm{d}\sigma \right) \mathrm{d}s \right\| \\
&= \left\| \frac{\varepsilon}{T}\int_0^T \left(\int_t^{t+s} f^1(x(\sigma),\sigma)\mathrm{d}\sigma \right) \mathrm{d}s \right\| \leq \frac{\varepsilon}{T}\int_0^T sM_1\mathrm{d}s = \frac{1}{2}\varepsilon M_1 T.
\end{aligned}$$

Putting this together with Lemma 2.2, we obtain

$$\left\| x(t) - x_0 - \varepsilon \int_0^t f_T^0(x(\sigma),\sigma)\mathrm{d}\sigma - \varepsilon \int_0^t f^1(x(\sigma),\sigma)\mathrm{d}\sigma \right\| \leq \frac{1}{2}\varepsilon T((1+\lambda_{f^0}L)M_0 + \lambda_{f^0}LM_1) + \frac{1}{2}\varepsilon M_1 T.$$

Thus, we have

$$x(t) = x_0 + \varepsilon \int_0^t f_T^0(x(\sigma),\sigma)\mathrm{d}\sigma + \varepsilon \int_0^t f^1(x(\sigma),\sigma)\mathrm{d}\sigma + \mathcal{O}(\varepsilon T).$$

Now,

$$y(t) = x_0 + \varepsilon \int_0^t f_T^0(y(\sigma),\sigma)\mathrm{d}\sigma + \varepsilon \int_0^t f^1(y(\sigma),\sigma)\mathrm{d}\sigma,$$

so

$$x(t) - y(t) = \varepsilon \int_0^t [f_T^0(x(\sigma),\sigma) - f_T^0(y(\sigma),\sigma)]\mathrm{d}\sigma + \varepsilon \int_0^t [f^1(x(\sigma),\sigma) - f^1(y(\sigma),\sigma)]\mathrm{d}\sigma + \mathcal{O}(\varepsilon T).$$

Therefore,

$$\|x(t) - y(t)\| \leq \varepsilon \int_0^t (\lambda_{f^0} + \lambda_{f^1})\|x(\sigma) - y(\sigma)\|\mathrm{d}\sigma + \mathcal{O}(\varepsilon T).$$

Then, applying Gronwall's Lemma [13, Lemma 1.3.1],

$$\|x(t) - y(t)\| = \mathcal{O}\left(\varepsilon T e^{\varepsilon(\lambda_{f^0} + \lambda_{f^1})t}\right). \qquad \blacksquare$$

We can now prove the following.

Theorem 2.4 (Selective Averaging Theorem) Let $x(t)$ be the solution to

$$\dot{x} = \varepsilon f^0(x,t) + \varepsilon f^1(x,t), \qquad x(0) = x_0, \tag{2}$$

and let $y(t)$ be the solution to

$$\dot{y} = \varepsilon \overline{f^0}(y) + \varepsilon f^1(y,t), \qquad y(0) = x_0, \tag{3}$$

where f^0 is T-periodic, and f^0 and f^1 satisfy the assumptions given in Section 1. Then

$$\|x(t) - y(t)\| = \mathcal{O}(\varepsilon)$$

on the time scale $1/\varepsilon$.

Proof This follows from Lemmas 2.1 and 2.3. ∎

3. Application to phase reduction

A powerful technique for analyzing biological oscillators is the rigorous reduction to a phase model, with a single variable for each oscillator describing the phase of the oscillation with respect to some reference state [6, 7, 10, 15]. This tremendous reduction in the dimensionality and complexity of a system often retains enough information to yield a useful understanding of its dynamics, and can allow for the implementation of phase-based control algorithms.

Phase reduction is commonly applied to systems of coupled oscillators, where in the limit of weak coupling one can use averaging to obtain terms which only depend on the phase differences of the oscillators; see, for example, [4]. Phase reduction has also been applied to systems of uncoupled oscillators which receive an external input, for example in [2]. Here we consider phase reduction for coupled oscillators with an external input; by averaging only the coupling term, we provide justification for models that are sometimes useful for neural control problems, e.g., [11, 14].

Suppose that the system

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x})$$

has a periodic orbit $\mathbf{x}_\gamma(t)$ with period $T = \frac{2\pi}{\omega}$. Now consider

$$\frac{d\mathbf{x}_i}{dt} = \mathbf{F}(\mathbf{x}_i) + \varepsilon \sum_j \mathbf{p}(\mathbf{x}_i, \mathbf{x}_j) + \varepsilon u(t)\hat{e}_1, \qquad i = 1, \cdots, N,$$

where $\mathbf{x}_i$ is the state of the i^{th} oscillator, $\mathbf{p}$ represents coupling between oscillators, $u(t)$ is the external input, and $\hat{e}_1$ is a unit vector in the x_1-direction. (For a neuron, this could correspond to an input $u(t)$ in the voltage equation.) Here, for simplicity we have assumed that all oscillators are identical and have identical coupling to all other oscillators. Moreover, we have assumed that each receives the same input $u(t)$, but it is straightforward to generalize this to the case of different inputs as in Section 3.1. We transform to phase variables as follows, cf. [10]:

$$\begin{aligned} \frac{d\theta_i}{dt} &= \frac{\partial \theta_i}{\partial \mathbf{x}_i} \cdot \frac{d\mathbf{x}_i}{dt} = \frac{\partial \theta_i}{\partial \mathbf{x}_i} \cdot \left(\mathbf{F}(\mathbf{x}_i) + \varepsilon \sum_j \mathbf{p}(\mathbf{x}_i, \mathbf{x}_j) + \varepsilon u(t)\hat{e}_1 \right) \\ &= \omega + \varepsilon \frac{\partial \theta_i}{\partial \mathbf{x}_i} \cdot \sum_j \mathbf{p}(\mathbf{x}_i, \mathbf{x}_j) + \varepsilon \frac{\partial \theta_i}{\partial \mathbf{x}_i} \cdot (u(t)\hat{e}_1). \end{aligned}$$

To lowest order in ε,

$$\frac{d\theta_i}{dt} = \omega + \varepsilon \mathbf{Z}(\theta_i) \cdot \sum_j \mathbf{p}(\theta_i, \theta_j) + \varepsilon \mathbf{Z}(\theta_i) \cdot (u(t)\hat{e}_1), \tag{4}$$

$$\mathbf{Z}(\theta_i) = \left. \frac{\partial \theta_i}{\partial \mathbf{x}_i} \right|_{\mathbf{x}_\gamma(\theta_i)}, \qquad \mathbf{p}(\theta_i, \theta_j) = \mathbf{p}(\mathbf{x}_\gamma(\theta_i), \mathbf{x}_\gamma(\theta_j)).$$

Here, $\mathbf{Z}(\theta)$ is known as the phase response curve [15]. Let $\theta_i = \phi_i + \omega t$; then

$$\frac{d\phi_i}{dt} = \varepsilon \mathbf{Z}(\phi_i + \omega t) \cdot \sum_j \mathbf{p}(\phi_i + \omega t, \phi_j + \omega t) + \varepsilon \mathbf{Z}(\phi_i + \omega t) \cdot (u(t)\hat{e}_1).$$

Now, apply the Selective Averaging Theorem to average the coupling term (to use this theorem, we can consider the lift of ϕ_i to $\mathbb{R}$):

$$\frac{d\varphi_i}{dt} = \frac{\varepsilon}{T} \int_0^T \mathbf{Z}(\varphi_i + \omega t) \cdot \sum_j \mathbf{p}(\varphi_i + \omega t, \underbrace{\varphi_j + \omega t}_{\varphi_j - \varphi_i + \varphi_i + \omega t})dt + \varepsilon \mathbf{Z}(\varphi_i + \omega t) \cdot (u(t)\hat{e}_1).$$

Let $s = \varphi_i + \omega t$, which gives

$$\frac{d\varphi_i}{dt} = \frac{\varepsilon}{2\pi} \sum_j \int_0^{2\pi} \mathbf{Z}(s) \cdot \mathbf{p}(s, \varphi_j - \varphi_i + s)ds + \varepsilon \mathbf{Z}(\varphi_i + \omega t) \cdot (u(t)\hat{e}_1).$$

Then, letting $\vartheta_i = \varphi_i + \omega t$,

$$\frac{d\vartheta_i}{dt} = \omega + \frac{\varepsilon}{2\pi} \sum_j \int_0^{2\pi} \mathbf{Z}(s) \cdot \mathbf{p}(s, \vartheta_j - \vartheta_i + s)ds + \varepsilon \mathbf{Z}(\vartheta_i) \cdot (u(t)\hat{e}_1).$$

That is,

$$\frac{d\vartheta_i}{dt} = \omega + \varepsilon \sum_j h(\vartheta_j - \vartheta_i) + \varepsilon \mathbf{Z}(\vartheta_i) \cdot (u(t)\hat{e}_1), \tag{5}$$

where

$$h(\psi) = \frac{1}{2\pi} \int_0^{2\pi} \mathbf{Z}(s) \cdot \mathbf{p}(s, \psi + s)ds.$$

From the Selective Averaging Theorem, we expect $\theta_i(t) - \vartheta_i(t) = \phi_i(t) - \varphi_i(t) = \mathcal{O}(\varepsilon)$ on the time scale $1/\varepsilon$.

We note that a similar phase reduction result has been obtained using different means in [9].

3.1 Thalamic neural model

As mentioned previously, selective averaging can be useful for phase reduced neural systems. Here, we formulate and solve an optimal control problem for a selectively averaged model of thalamic neurons to compare the errors made in the phase reduction to errors made by using selective averaging. We illustrate the utility of selective averaging for two periodically spiking neurons [12]:

$$\begin{aligned} C\dot{V}_1 &= I_{m,1} - g_{G\to G}(V_1 - V_{G\to G})s_2 + I_{ext}(t), \\ C\dot{V}_2 &= I_{m,2} - g_{G\to G}(V_2 - V_{G\to G})s_1, \end{aligned} \tag{6}$$

where

$$\begin{aligned} I_{m,i} &= -I_L(V_i) - I_{Na}(V_i, h_i) - I_K(V_i, h_i) - I_T(V_i, r_i) + I_{SM}, \\ \dot{h}_i &= (h_\infty(V_i) - h_i)/\tau_h(V_i), \\ \dot{r}_i &= (r_\infty(V_i) - r_i)/\tau_r(V_i), \\ \dot{s}_i &= \alpha H_\infty(V_i - \theta_g)(1 - s_i) - \beta s_i, \quad i = 1, 2. \end{aligned} \tag{7}$$

Here

$$H_\infty(V) = 1/(1 + \exp(-(V - \theta_g^H)/\sigma_g^H)),$$

$C = 1\mu\text{F/cm}^2, g_{G\to G} = 0.02\text{nS}/\mu\text{m}^2, V_{G\to G} = -100\text{mV}, \alpha = 2\text{msec}^{-1}, \theta_g = 20.0, \beta = 0.08\text{msec}^{-1}, \theta_g^H = -57.0$, and $\sigma_g^H = 2.0$. The term I_{SM} represents a baseline current which we take to be $5\mu\text{A/cm}^2$ causing the neuron to fire with a period $T = 8.395$ ms. This model reproduces the firing patterns of two synaptically coupled thalamic neurons. The synaptic coupling is inhibitory (inhibits spiking and slows down the periodic orbit), and we assume neuron 1 can be controlled through direct current injection,

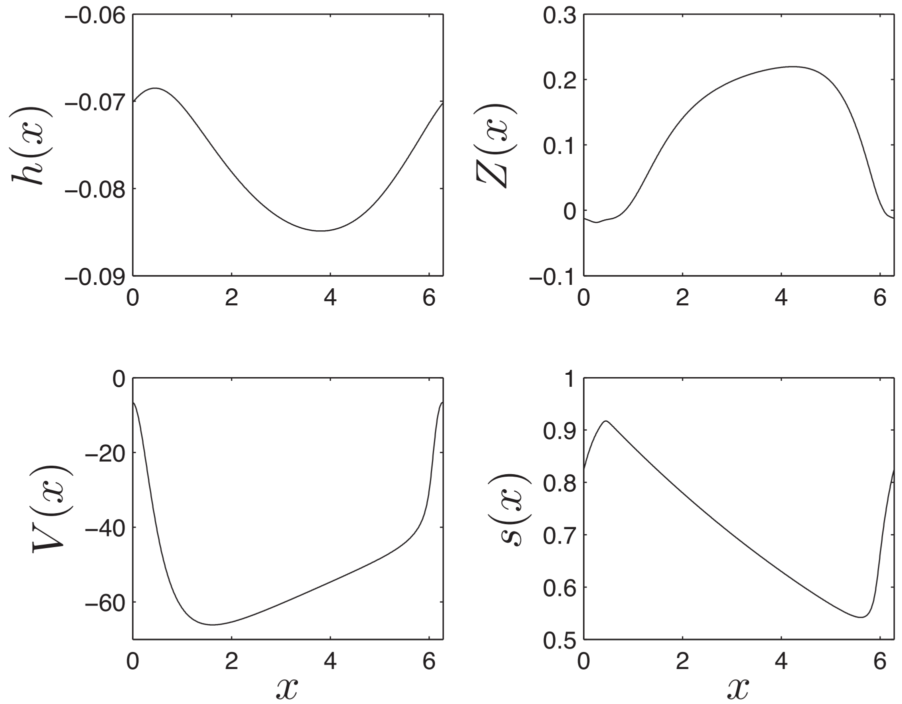

Fig. 1. Plots of relevant functions from equations (8) and (9).

$I_{ext}(t)$. For a full explanation of the functions $I_L, I_{Na}, I_K, I_t, h_\infty, \tau_h, r_\infty$ and τ_r, we refer the reader to [12].

In the absence of external stimuli and coupling, individual neurons from equations (6) and (7) eventually settle to a periodic orbit. Employing phase reduction techniques, we reduce the model equations to (cf. (4))

$$\begin{aligned} \theta(t) &= [\theta_1(t), \theta_2(t)]^T \ \in \mathbb{S}^2, \\ \dot{\theta}_1 &= \omega + Z(\theta_1)\left[u(t) - g_{G\to G}(V(\theta_1) - V_{G\to G})s(\theta_2)\right], \\ \dot{\theta}_2 &= \omega - Z(\theta_2)\left[g_{G\to G}(V(\theta_2) - V_{G\to G})s(\theta_1)\right]. \end{aligned} \tag{8}$$

Here $\theta_i \in [0, 2\pi)$ for $i = 1, 2$ with $\theta = 2\pi$ corresponding to when the neuron spikes, $u(t) = I_{ext}(t)/C$ is the control input to neuron 1, $\omega = 2\pi/T$, $Z(x)$ is each neuron's phase response curve, and $V(x)$ and $s(x)$ describe the transmembrane voltage and coupling variables, respectively, as a function of the phase, evaluated on the periodic orbit which exists in the absence of coupling and external input. As we show in Section 2, provided the neural coupling is small enough, equation (8) can be further simplified by selectively averaging the coupling between the neurons to get (cf. (5))

$$\begin{aligned} \vartheta(t) &= [\vartheta_1(t), \vartheta_2(t)]^T \ \in \mathbb{S}^2, \\ \dot{\vartheta}_1 &= \omega + h(\vartheta_2 - \vartheta_1) + Z(\vartheta_1)u(t), \\ \dot{\vartheta}_2 &= \omega + h(\vartheta_1 - \vartheta_2), \end{aligned} \tag{9}$$

where $h(x)$ is the selectively averaged coupling function,

$$h(x) = \frac{1}{2\pi}\int_0^{2\pi} \left(-Z(a)\left[g_{G\to G}(V(a) - V_{G\to G})s(x+a)\right]\right) da,$$

and $\vartheta_i \in [0, 2\pi)$ for $i = 1, 2$. Relevant functions are shown in Fig. 1. Notice $h(x) < 0$ which is consistent with inhibitory synaptic coupling; moreover, since $h'(0) > 0$ and $h'(\pi) < 0$, the in-phase solution is stable and the anti-phase solution is unstable in the absence of input [1].

3.2 Minimum energy control of the reduced systems

Suppose we want to control the selectively averaged phase reduced system (9) from an in-phase state to a particular anti-phase state while minimizing the overall energy used. To this end, we first consider a domain $\mathcal{D}$ comprised of continuous external stimuli $u(t)$ that bring the system $\vartheta_1(0) = \vartheta_2(0) = 0$ to $\vartheta_1(t_1) = 0$ and $\vartheta_2(t_1) = \pi$. Of these stimuli, we seek to find the one which minimizes the energy functional, $\mathcal{G}[u(t)] = \int_0^{t_1} u(t)^2 dt$.

Adopting the framework of Calculus of Variations [8], this amounts to finding all continuous relative minimum functions u^*, which are defined such that there exists an $\epsilon > 0$ such that for all other continuous functions u that satisfy $||u - u^*|| < \epsilon$, $\mathcal{G}(u) - \mathcal{G}(u^*) \geq 0$. The minimum of all relative minimum functions will be the absolute minimum. These minimum functions can be found by first defining cost functionals for each system. For the selectively averaged system (9), the cost functional is $\mathcal{A}[u(t)] = \int_0^{t_1} \mathcal{B}[u(t)]dt$ where

$$\mathcal{B}[u(t)] = u(t)^2 + \lambda_1 \left(\frac{d\vartheta_1}{dt} - \omega - Z(\vartheta_1)u(t) - h(\vartheta_2 - \vartheta_1) \right) + \lambda_2 \left(\frac{d\vartheta_2}{dt} - \omega - h(\vartheta_1 - \vartheta_2) \right),$$

and λ_1 and λ_2 are Lagrange multipliers that force the dynamics to obey (9). Following the formalism of Calculus of Variations, by defining $\mathbf{\Phi} = [\vartheta_1, \vartheta_2, \lambda_1, \lambda_2]$, extremal functions (both minimizing and maximizing) for the averaged system must satisfy the following Euler-Lagrange equations:

$$\frac{\partial \mathcal{B}}{\partial u} = \frac{d}{dt}\left(\frac{\partial \mathcal{B}}{\partial \dot{u}}\right), \qquad \frac{\partial \mathcal{B}}{\partial \mathbf{\Phi}} = \frac{d}{dt}\left(\frac{\partial \mathcal{B}}{\partial \dot{\mathbf{\Phi}}}\right).$$

These equations are

$$\begin{aligned}
u(t) &= \lambda_1 Z(\vartheta_1)/2, \\
\dot{\lambda_1} &= \lambda_1(h'(\vartheta_2 - \vartheta_1) - Z'(\vartheta_1)u(t)) - \lambda_2 h'(\vartheta_1 - \vartheta_2), \\
\dot{\lambda_2} &= -\lambda_1 h'(\vartheta_2 - \vartheta_1) + \lambda_2 h'(\vartheta_1 - \vartheta_2), \\
\dot{\vartheta}_1 &= \omega + Z(\vartheta_1)u(t) + h(\vartheta_2 - \vartheta_1), \\
\dot{\vartheta}_2 &= \omega + h(\vartheta_1 - \vartheta_2),
\end{aligned} \tag{10}$$

where $' = d/d\vartheta$.

For the non-averaged, phase reduced system, we can also calculate an analogous optimal control. In this case the cost functional is $\mathcal{N}[u(t)] = \int_0^{t_1} \mathcal{M}[u(t)]dt$, where

$$\begin{aligned}
\mathcal{M}[u(t)] = u(t)^2 &+ \psi_1 \left(\frac{d\theta_1}{dt} - \omega - Z(\theta_1)\left[u(t) - g_{G\to G}(V(\theta_1) - V_{G\to G})s(\theta_2)\right] \right) \\
&+ \psi_2 \left(\frac{d\theta_2}{dt} - \omega + Z(\theta_2) g_{G\to G}(V(\theta_2) - V_{G\to G})s(\theta_1) \right),
\end{aligned}$$

and ψ_1 and ψ_2 are Lagrange multipliers that force the dynamics to obey (8). Extremal functions for the nonaveraged system can be found from the following Euler-Lagrange equations:

$$\begin{aligned}
u(t) &= \psi_1 Z(\theta_1)/2, \\
\dot{\psi_1} &= \psi_1 Z'(\theta_1) g_{G\to G}(V(\theta_1) - V_{G\to G})s(\theta_2) + \psi_1[Z(\theta_1) g_{G\to G} V'(\theta_1)s(\theta_2) - Z'(\theta_1)u(t)] \\
&\quad + \psi_2 Z(\theta_2) g_{G\to G}(V(\theta_2) - V_{G\to G})s'(\theta_1), \\
\dot{\psi_2} &= \psi_1 Z(\theta_1) g_{G\to G}(V(\theta_1) - V_{G\to G})s'(\theta_2) + \psi_2 Z'(\theta_2) g_{G\to G}(V(\theta_2) - V_{G\to G})s(\theta_1) \\
&\quad + \psi_2 Z(\theta_2) g_{G\to G} V'(\theta_2)s(\theta_1), \\
\dot{\theta}_1 &= \omega + Z(\theta_1)\left[u(t) - g_{G\to G}(V(\theta_1) - V_{G\to G})s(\theta_2)\right], \\
\dot{\theta}_2 &= \omega - Z(\theta_2)\left[g_{G\to G}(V(\theta_2) - V_{G\to G})s(\theta_1)\right],
\end{aligned} \tag{11}$$

where $' = d/d\theta$.

We choose $t_1 = 41.98$, corresponding to $5T$, or 5 spikes of an unperturbed, uncoupled neuron. For the equations (10), obtained for the selectively averaged system, we solve a two point boundary value problem using a double bisection algorithm described in [3] to determine inital values of Lagrange multipliers $\lambda_1(0)$ and $\lambda_2(0)$ which give resulting stimuli $u(t) \in \mathcal{D}$. We note that the above methodology

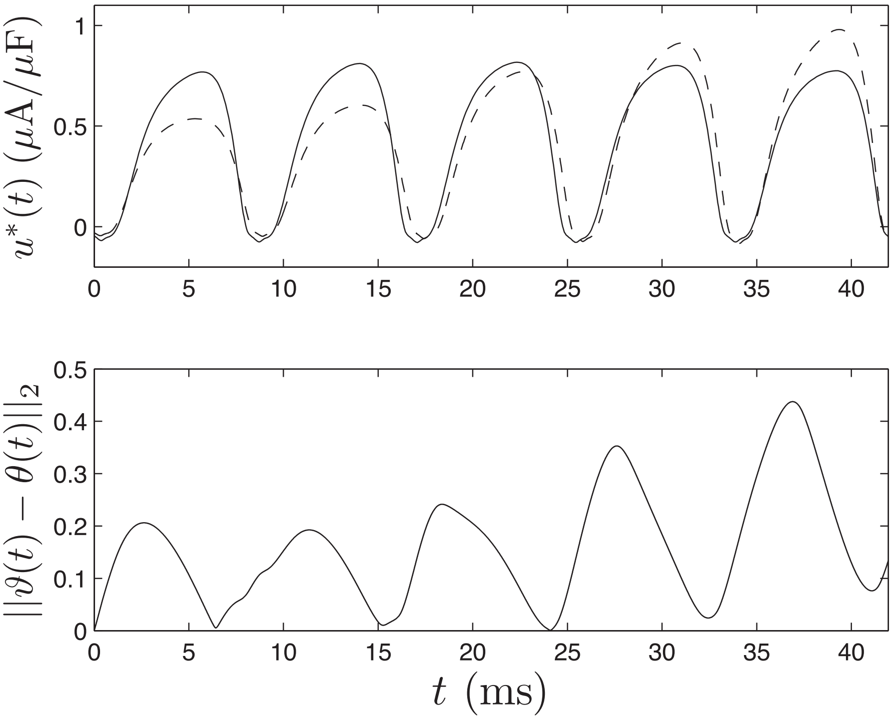

Fig. 2. Comparison of the results for the selectively averaged and nonaveraged systems. The top panel shows the optimal control using the selectively averaged equations and nonaveraged equations as a solid and dashed line, respectively. The bottom panel compares the selectively averaged and nonaveraged solutions, with the optimal control for the selectively averaged system applied to both systems.

only guarantees a locally optimal solution, but we can limit our search for $\lambda_1(0)$ and $\lambda_2(0)$ so that the resulting solutions yield optimal solutions that are small enough so that the phase reduction (8), which requires sufficiently small inputs and coupling, is still valid.

We repeat the procedure for finding locally optimal inputs for the equations (11), which correspond to the nonaveraged system, to compare the results. The top panel of Fig. 2 shows the optimal control for the averaged (9) and nonaveraged (8) equations as solid and dashed lines, respectively. We see both strategies give solutions which are not far from each other. Next, we simulate (9) and (8) using the optimal control obtained through selective averaging. We find that the solution for the nonaveraged system, $\theta(t)$, remains close to the solution for the averaged system, $\vartheta(t)$, throughout the simulation, with the difference in magnitude growing as the simulation time increases, as expected from the selective averaging theorem. Using the optimal control for the selectively averaged system, $\vartheta(5T) = [0, \pi]^T$ as expected, and applying the same control on the nonaveraged system (which is *not* optimal for this system) yields $\theta(5T) = [0.10, 3.06]^T$, and is quite close to the antiphase target.

Finally, we compare the solution to the full equations given by (6) and (7) to the solution of the phase reduced equations (8) and to the averaged phase reduced equations (9). For the full equations, we infer the phase of each neuron $\Theta(t) = [\Theta_1(t), \Theta_2(t)]^T \in \mathbb{S}^2$ at each time step by simulating each neuron separately in the absence of input and coupling to determine when it spikes next. Results are shown in Fig. 3. We find that the error between the averaged and nonaveraged phase reductions and the full equations is much larger than the error between the phase reductions themselves. The ending condition of the full system is $\Theta(5T) = [0.11, 3.90]^T$ and the error between the phase of the target and the phase of the second neuron is much greater than for either of the phase reduced systems. This larger error comes from the open loop nature of the solutions, as small errors made at the beginning of the simulation become larger as time progresses.

In this context, selective averaging may be useful in reducing the complexity of an optimal control problem. Comparing the equations of optimality for the averaged equations (10) and the nonaveraged equations (11), we see that averaging makes the optimal equations less complicated, and less compu-

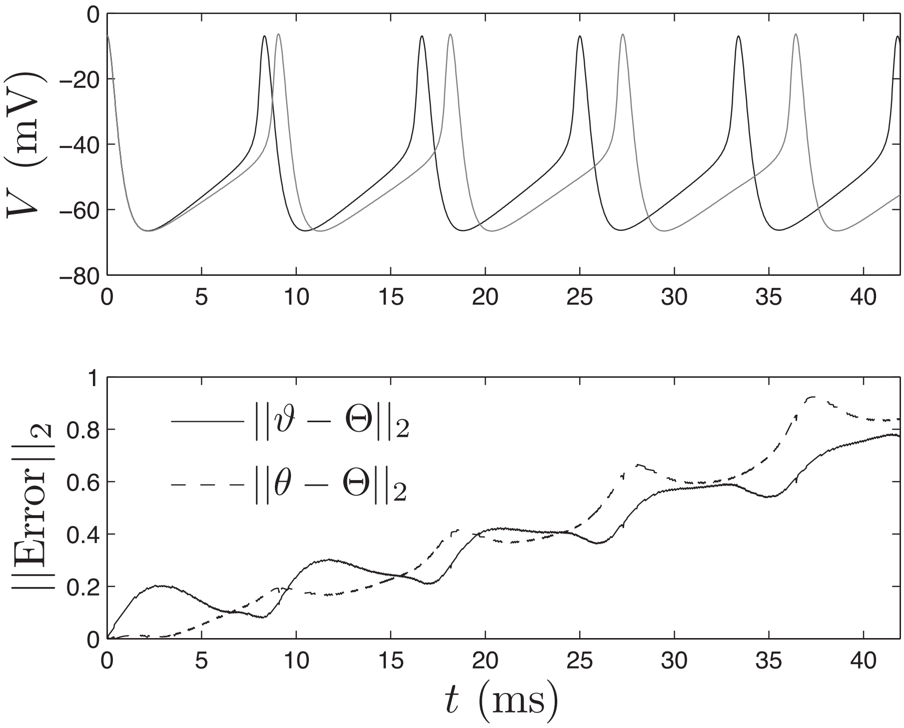

Fig. 3. The top panel shows each neuron's transmembrane voltage as a function of time found from (6), and the bottom panel shows the relative error between solutions as a function of time.

tationally intensive. For a more complicated form of coupling, selective averaging may dramatically reduce time required to iteratively calculate an optimal solution, which may be attractive if the optimal control must be calculated in real time and applied in an *in vitro* experiment. Furthermore, phase reduction is necessary to decrease the dimensionality of this optimal control problem from 8 to 2, and we see from Fig. 3 that the additional error made through selective averaging is small. Finally, selective averaging may also be attractive because it allows us to connect our intuition and results for systems with phase difference coupling. For instance, averaging the coupling function to get $h(x)$ allows us to see that the synaptic coupling will be inhibitory (negative sign), meaning that the optimal control will most likely need to be predominantly excitatory (positive sign), in order reach the particular target we have prescribed. This intuition would be more difficult to obtain by using the nonaveraged equations (8).

4. The general selective averaging theorem

Following [13], we can also prove a general selective averaging theorem. Consider a vector field $f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ which is continuous in x and t, and Lipschitz continuous in x on $D \subset \mathbb{R}^n$. If the average

$$\overline{f}(x) = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(x,s)\mathrm{d}s$$

exists and the limit is uniform in x on compact sets $K \subset D$, then we call f a *Krylov-Bogoliubov-Mitropolsky (KBM) vector field*. The following lemmas will allow us to prove the General Selective Averaging Theorem.

Lemma 4.1 (Lemma 4.3.1 from [13]) If f^0 is a KBM vector field, and assuming that $\varepsilon T = o(1)$ as $\varepsilon \downarrow 0$ (that is, $\lim_{\varepsilon\downarrow 0} \varepsilon T = 0$), then on a time scale $\frac{1}{\varepsilon}$ one has

$$f_T^0(x,t) = \overline{f^0}(x) + \mathcal{O}\left(\frac{\delta_0(\varepsilon)}{\varepsilon T}\right)$$

where

$$\delta_0(\varepsilon) = \sup_{x\in D} \sup_{t\in[0,\frac{L}{\varepsilon})} \varepsilon \left\| \int_0^t \left[f^0(x,s) - \overline{f^0}(x) \right] \mathrm{d}s \right\|.$$

Here $\delta_0(\varepsilon)$ quantifies the behavior of f^0 relative to its average.

Lemma 4.2 Let y be the solution of the initial value problem

$$\dot{y} = \varepsilon f_T^0(y,t) + \varepsilon f^1(y,t), \qquad y(0) = x_0, \tag{12}$$

and suppose f^0 is a KBM vector field with order function $\delta_0(\varepsilon)$. Then the solution of

$$\dot{z} = \varepsilon \overline{f^0}(z) + \varepsilon f^1(z,t), \qquad z(0) = x_0 \tag{13}$$

satisfies

$$y(t) = z(t) + \mathcal{O}\left(\frac{\delta_0(\varepsilon)}{\varepsilon T} \right)$$

with t on a time scale $\frac{1}{\varepsilon}$.

Proof The proof is similar to ([13, Lemma 4.3.5]). We express the solutions to (12) and (13) respectively as:

$$y(t) = x_0 + \varepsilon \int_0^t f_T^0(y(s),s)ds + \varepsilon \int_0^t f^1(y(s),s)ds,$$

$$z(t) = x_0 + \varepsilon \int_0^t \overline{f^0}(z(s))ds + \varepsilon \int_0^t f^1(z(s),s)ds.$$

Using Lemma 4.1,

$$y(t) - z(t) = \varepsilon \int_0^t \left(\overline{f^0}(y(s)) - \overline{f^0}(z(s)) \right) \mathrm{d}s + \mathcal{O}\left(\frac{\delta_0(\varepsilon)t}{T} \right) + \varepsilon \int_0^t (f^1(y(s),s) - f^1(z(s),s))\mathrm{d}s.$$

The Lipschitz-continuity of f^0 implies

$$\begin{aligned}
&\left\| \int_0^t \left[\overline{f^0}(y(s)) - \overline{f^0}(z(s)) \right] ds \right\| \\
&\quad = \left\| \int_0^t \left(\lim_{T\to\infty} \frac{1}{T} \int_0^T \left[f^0(y(s),\rho) - f^0(z(s),\rho) \right] d\rho \right) ds \right\| \\
&\quad \leq \int_0^t \left(\lim_{T\to\infty} \frac{1}{T} \int_0^T \left\| f^0(y(s),\rho) - f^0(z(s),\rho) \right\| d\rho \right) ds \\
&\quad \leq \int_0^t \left(\lim_{T\to\infty} \frac{1}{T} \int_0^T \lambda_{f^0} \|y(s) - z(s)\| d\rho \right) ds \\
&\quad = \int_0^t \lambda_{f^0} \|y(s) - z(s)\| ds.
\end{aligned}$$

Thus,

$$\|y(t) - z(t)\| \leq \varepsilon \int_0^t (\lambda_{f^0} + \lambda_{f^1}) \|y(s) - z(s)\| ds + \mathcal{O}\left(\frac{\delta_0(\varepsilon)t}{T} \right).$$

Applying Gronwall's Lemma [13, Lemma 1.3.1] we obtain

$$\|y(t) - z(t)\| = \mathcal{O}\left(\frac{\delta_0(\varepsilon)t}{T} e^{\varepsilon(\lambda_{f^0} + \lambda_{f^1})t} \right).$$

The result follows by taking t on the time scale $1/\varepsilon$. ∎

We can now prove the following.

Theorem 4.3 (General Selective Averaging Theorem) Let x be a solution of the initial value problem

$$\dot{x} = \varepsilon f^0(x,t) + \varepsilon f^1(x,t), \qquad x(0) = x_0.$$

We assume that f^0 is a KBM-vector field with order function $\delta_0(\varepsilon)$. Let z be the solution of the initial value problem

$$\dot{z} = \varepsilon \overline{f^0}(z) + \varepsilon f^1(z,t), \qquad z(0) = x_0.$$

Then

$$x(t) = z(t) + \mathcal{O}(\sqrt{\delta_0(\epsilon)}).$$

Proof The proof similar to [13, Theorem 4.3.6]. By Lemma 2.3 we know that the solution y of

$$\dot{y} = \varepsilon f^0_T(y,t) + \varepsilon f^1(y,t)$$

satisfies

$$x(t) = y(t) + \mathcal{O}(\varepsilon T)$$

on a time scale $\frac{1}{\varepsilon}$. Also, from Lemma 4.2,

$$y(t) = z(t) + \mathcal{O}\left(\frac{\delta_0(\varepsilon)}{\varepsilon T}\right).$$

Then, from the triangle inequality

$$\|x(t) - z(t)\| \leq \|x(t) - y(t)\| + \|y(t) - z(t)\|,$$

we have

$$x(t) = z(t) + \mathcal{O}(\varepsilon T) + \mathcal{O}\left(\frac{\delta_0(\varepsilon)}{\varepsilon T}\right).$$

If we let $T = \sqrt{\delta_0(\varepsilon)}/\varepsilon$, then the result follows. ∎

We note that f^0 does not need to be periodic for this theorem.

5. Conclusion

In this paper, we extended averaging theorems from [13] to the case in which terms of the same order are selectively averaged. We applied such selective averaging to the phase reduction of a system of oscillators with both coupling and external input; here the coupling term was averaged to give phase difference coupling, while the external input term was not averaged. Our results give rigorous justification for models considered in papers such as [11, 14], which are motivated by neural control. As a new illustration that selective averaging can be a useful tool for problems in neural control, we applied it to obtain a simpler system for calculating the optimal input which takes the in-phase state to the anti-phase state for a coupled two-neuron system.

Acknowledgments

The authors would like to thank the German Research Foundation (DFG) for financial support of the project within the Cluster of Excellence in Simulation Technology (EXC 310/1) at the University of Stuttgart, and the National Science Foundation for support through Grant NSF-1000678.

References

[1] E. Brown, P. Holmes, and J. Moehlis, "Globally coupled oscillator networks," In E. Kaplan, J. Marsden, and K.R. Sreenivasan, editors, *Perspectives and Problems in Nonlinear Science: A Celebratory Volume in Honor of Larry Sirovich*, pp. 183–215, Springer-Verlag, New York, 2003.

[2] E. Brown, J. Moehlis, and P. Holmes, "On the phase reduction and response dynamics of neural oscillator populations," *Neural Comp.*, vol. 16, pp. 673–715, 2004.

[3] P. Danzl, A. Nabi, and J. Moehlis, "Charge-balanced spike timing control for phase models of spiking neurons," *Discrete and Continuous Dynamical Systems*, vol. 28, pp. 1413–1435, 2010.

[4] G.B. Ermentrout and N. Kopell, "Multiple pulse interactions and averaging in systems of coupled neural oscillators," *J. Math. Bio.*, vol. 29, pp. 195–217, 1991.

[5] J. Guckenheimer and P.J. Holmes, *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields*, Springer-Verlag, New York, 1983.

[6] F.C. Hoppensteadt and E.M. Izhikevich, *Weakly Connected Neural Networks*, Springer-Verlag, New York, 1997.

[7] E.M. Izhikevich, *Dynamical Systems in Neuroscience*, MIT Press, Cambridge, Massachusetts, 2007.

[8] D. Kirk, *Optimal Control Theory*, Dover Publications, New York, 1998.

[9] H. Kori, Y. Kawamura, H. Nakao, K. Arai, and Y. Kuramoto, "Collective-phase description of coupled oscillators with general network structure," *Phys. Rev. E*, 80:036207, 2009.

[10] Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence*, Springer-Verlag, New York, 1984.

[11] A. Nabi and J. Moehlis, "Single input optimal control for globally coupled neuron networks," *J. Neural Eng.*, 8, Article No. 065008, 2011.

[12] J. Rubin and D. Terman, "High frequency stimulation of the subthalamic nucleus eliminates pathological thalamic rhythmicity in a computational model," *Journal of Computational Neuroscience*, vol. 16, pp. 211–235, 2004.

[13] J.A. Sanders, F. Verhulst, and J. Murdock, *Averaging Methods in Nonlinear Dynamical Systems, Second Edition*, Springer-Verlag, New York, 2007.

[14] P.A. Tass, "Synergetics of the nervous system: from basic principles to therapy," *Nonlin. Phenom. Complex Systems*, vol. 5, pp. 470–478, 2002.

[15] A. Winfree, *The Geometry of Biological Time, Second Edition*, Springer, New York, 2001.