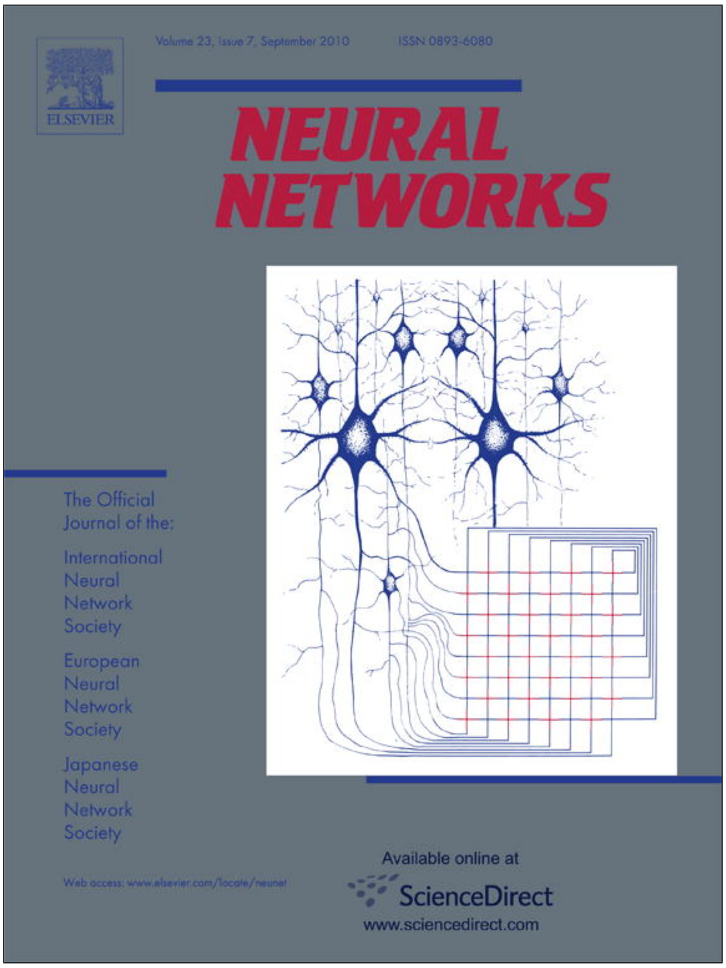
Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.


Volume 23, Issue 7, September 2010
ISSN 0893-6080
ELSEVIER
NEURAL NETWORKS
The Official Journal of the:
International Neural Network Society
European Neural Network Society
Japanese Neural Network Society
Web access: www.elsevier.com/locate/neunet
Available online at
ScienceDirect
www.sciencedirect.com

This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Neural Networks 23 (2010) 836–842

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

The role of chaotic resonance in cerebellar learning

Isao T. Tokuda [a,*], Cheol E. Han [b], Kazuyuki Aihara [c], Mitsuo Kawato [d], Nicolas Schweighofer [e]

[a] *School of Information Science, Japan Advanced Institute of Science and Technology, Isikawa, Japan*
[b] *Computer Science, University of Southern California, Los Angeles, USA*
[c] *Institute of Industrial Science, University of Tokyo, Tokyo, Japan*
[d] *Advanced Telecommunication Research, Kyoto, Japan*
[e] *Biokinesiology and Physical Therapy, University of Southern California, Los Angeles, USA*

ARTICLE INFO

Article history:
Received 16 November 2009
Received in revised form 9 April 2010
Accepted 27 April 2010

Keywords:
Chaotic resonance
Inferior olive
Cerebellar learning
Synchronization

ABSTRACT

According to the cerebellar learning hypothesis, the inferior olive neurons, despite their low firing rates, are thought to transmit high-fidelity error signals to the cerebellar cortex. "Chaotic resonance", via moderate electrical coupling between inferior olive neurons, has been proposed to realize efficient transmission of the error signal by desynchronizing spiking. Here, we first show that chaotic resonance is a robust phenomenon, as it does not depend upon the details of the inferior olive neuronal model. Second, we show that chaotic resonance enhances learning of a neural controller for fast arm movements. Furthermore, when both coupling and noise are considered simultaneously, we found that chaotic resonance widens the range of noise intensity within which efficient learning can be realized. We suggest that, from an energetic viewpoint, the spiking activity induced by chaos can be more economical than that induced by noise.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

The anatomy and the physiology of the cerebellum make it ideally suited to learn how to refine motor commands (Albus, 1971; Ito, 1970; Ito, Sakurai, & Tongroach, 1982; Kawato & Gomi, 1992; Marr, 1969; Schweighofer, Spoelstra, Arbib, & Kawato, 1998; Shidara, Kawano, Gomi, & Kawato, 1993) or to learn sensory predictions from motor commands (Miall, Christensen, Cain, & Stanley, 2007; Miall, Weir, Wolpert, & Stein, 1993; Tseng, Diedrichsen, Krakauer, Shadmehr, & Bastian, 2007). The Purkinje cells, the sole output neurons of the cerebellar cortex, receive two major types of synaptic inputs: (i) numerous parallel fibers that relay information from much of the cerebral cortex and spinal cord, and (ii) a single climbing fiber, which is an axon from an inferior olive (IO) neuron, that has been shown to transmit error signals (Gilbert & Thach, 1977; Kitazawa, Kimura, & Yin, 1998). When conjointly activated with parallel fibers, IO spikes modify cerebellar input–output transformations, in agreement with the known long-term depression (LTD) at the parallel fiber–Purkinje cell synapse (Ito et al., 1982).

Two apparently contradictory constraints must be met, however, for the cerebellum to realize efficient adaptive motor control or prediction. First, the IO must transmit error signals with high temporal resolution despite its low firing rate. Second, IO neurons must fire at a low firing rate so that complex spikes encoding error signals do not interfere with simple spikes carrying motor control commands or predictions (Kawato & Gomi, 1992; Kobayashi et al., 1998). We previously proposed that these two constraints are simultaneously met via low-rate IO chaotic spike firing (Schweighofer et al., 2004). Such chaotic behavior leads to the generation of IO spikes at different timings at each trial. Specifically, we showed that electrical coupling via gap junctions can provide the source of disorder that induced a "chaotic resonance" (Nishimura, Katada, & Aihara, 2000) in IO networks. Here, chaotic dynamics is not supplied externally but it originates internally from complex interaction among the neurons. This resonance leads to an increase in information transmission in IO neurons by distributing high-frequency components of the error inputs over the sporadic, irregular, and non-phase-locked spikes. Desynchronization is indeed necessary for scattering the spike timings of each neuron to increase the time resolution of the population rate coding (Masuda & Aihara, 2002, 2003). Purkinje cells can then reconstruct the complete error signal via spatio-temporal integration because functionally related Purkinje cells and IO cells are grouped in "microcomplexes" (Ito, 1990; Schweighofer, 1998).

The direct effect of electrical coupling in enhancing cerebellar learning has yet to be shown, however. Furthermore, the robustness of chaotic resonance is unclear for two reasons. First, chaos does not always imply destruction of synchrony, since synchronization between chaotic oscillators has been commonly observed in a variety of physical or biological systems (Pikovsky, Rosenblum, & Kurths, 2001). Second, in our original study, we used a rather

---

[*] Corresponding author. Tel.: +81 761 51 1226; fax: +81 761 51 1149.
E-mail address: isao@jaist.ac.jp (I.T. Tokuda).

0893-6080/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2010.04.006

complicated compartment model (Schweighofer, Doya, & Kawato, 1999), and many physiological parameters can be chosen rather arbitrarily in this model. Finally, it is unclear whether chaos is indispensable to desynchronize IO neurons and to realize efficient information transmission, since neural noise can also desynchronize IO neurons.

Here, we develop a simple model of IO neurons to test the hypothesis that chaotic spiking induced via electrical coupling in IO neurons robustly enhances the learning of complex motor commands compared to non-chaotic or noise-induced jittered spiking. In our simulations, the IO neurons provide error signals to an idealized model of the cerebellar cortex that learns, via feedback error learning (Kawato, Furukawa, & Suzuki, 1987; Kawato & Gomi, 1992), to control a simplified model of the human arm in rapid reaching movements.

2. Methods

2.1. Inferior olive model

The dynamical properties of the IO neuron can be summarized as follows. (i) Under an isolated condition, a single IO neuron generates a limit cycle oscillation (Manor, Rinzel, Segev, & Yarom, 1997). (ii) Through gap-junction connections with other neurons, the IO neuron gives rise to more complex spike patterns (Lang, Sugihara, & Llinas, 1996; Makarenko & Llinas, 1998; Schweighofer et al., 2004). The μ-model is a simplified two-dimensional neuronal model that satisfies these dynamical characteristics (Fujii & Tsuda, 2004; Tsuda, Fujii, Tadokoro, Yasuoka, & Yamaguti, 2004). In particular, when embedded in a one-dimensional chain, complex spiking patterns such as chaotic itinerancy can be generated (Tsuda et al., 2004). The dynamics of a one-dimensional chain of μ-neurons is given by

$$\begin{aligned} \eta_1 \frac{\mathrm{d}x_i}{\mathrm{d}t} &= -y_i - \mu_i x_i^2 \left(x_i - \frac{3}{2} \right) + I + J_i + \xi_i, \\ \eta_2 \frac{\mathrm{d}y_i}{\mathrm{d}t} &= -y_i + \mu_i x_i^2, \end{aligned} \tag{1}$$

where

$$J_i = \begin{cases} g(x_2 + x_N - 2x_1) & (i = 1) \\ g(x_{i+1} + x_{i-1} - 2x_i) & (i = 2, \ldots, N-1) \\ g(x_1 + x_{N-1} - 2x_N) & (i = N), \end{cases} \tag{2}$$

x_i and y_i represent the membrane potential and ion channel activity of the ith neuron ($i = 1, 2, \ldots, N$), N is the total number of the neurons, μ is a system parameter, η_1 and η_2 are time constants, g is the coupling strength of the gap junctions, and I is an external input. An advantage of using this model is its weak dependence on the parameter value, since μ is the only parameter that controls the qualitative dynamics of the neuron; the time constants, which are set equal in this study ($\eta_1 = \eta_2$), do not change the qualitative dynamics of the neuron. Furthermore, because the parameter dependence on the neural dynamics has been thoroughly analyzed (Fujii & Tsuda, 2004), the proper parameter value for μ to generate spiking dynamics is also well understood.

Real neurons are subject to various kinds of noise. Since noise can destroy synchronous firing activity in a similar way as chaos, it is natural to consider that noise can also enhance information transfer in the IO. To take into account such a noise effect, we added independent white *Gaussian* noise $\xi_i(t)$ to the original μ-model, with $E[\xi_i(t)] = 0$, $E[\xi_i(t)\xi_j(s)] = 2D\delta(t-s)\delta(i-j)$, where D is the noise intensity, as in Collins, Chow, and Imhoff (1995).

The spiking activity of the kth IO neuron is defined as a membrane potential that exceeds a threshold value of x_{th}. In the case of noise-free simulations ($D = 0$), Eq. (1) is integrated by the fourth-order Runge–Kutta algorithm started from a random initial condition. In the presence of noise, Eq. (1) becomes a stochastic differential equation, which is simulated by Euler's algorithm (Fox, Gatland, Roy, & Vemuri, 1988). In the following experiments, five simulations were run to compute the average quantities so that the dependence of the neural dynamics on the random initial conditions is weakened.

2.2. Mutual information

As a basic study to evaluate the information transmission of the IO network, we measured the mutual information (Rényi, 1970) between an input signal and the spike responses. As an input signal, we used chaotic signals from the Rössler equations ($\mathrm{d}x/\mathrm{d}t = -y - z$, $\mathrm{d}y/\mathrm{d}t = x + 0.36y$, $\mathrm{d}z/\mathrm{d}t = 0.4x - (4.5 - x)z$) (Rössler, 1979). The y-variable is injected to all neurons in the same manner as $I = I_0 + \beta \cdot y$ ($I_0 = 0.01$ and $\beta = 0.002$). The output $S(t)$ represents a time sequence of a number of spikes generated from the population of neurons within a time interval of 0.02. Then the mutual information between input $I(t)$ and output $S(t)$ is computed, where the signals are discretized into 25 bins for calculating the probability distributions.

It is noted that the chaotic input signal has been utilized merely as a typical example of complex signal in the brain. The same results can be obtained when a periodic or noisy signal is used as the input.

2.3. Synchrony

For the IO neurons with low firing frequency to transmit information efficiently, synchronous activity is not desired, because in this condition the network becomes equivalent to a single neuron. As an index to detect such synchronized activity of the neurons, the order parameter R (Kuramoto, 1984) has been utilized. The order parameter is defined as $R\exp(\mathrm{i}\Phi) = (1/N)\Sigma_{j=1}^{N}\exp(\mathrm{i}\phi_j)$, where ϕ_j represents phase of the jth neuron given by angle $\phi_j = \arctan\left(\frac{x_j(t-0.2)}{x_j(t)}\right)$. The order parameter takes a real value between 0 and 1, where a large value close to $R = 1$ implies strong mutual synchronization and a small value close to $R = 0$ implies desynchronization.

2.4. Chaos

We quantified the strength of chaotic activity of the IO neurons with the Lyapunov exponents, computed as in Shimada and Nagashima (1979). From the Lyapunov exponents ordered in a descending manner $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{2N}$, the Lyapunov dimension is defined as $D_L = k + \sum_{i=1}^{k}\lambda_i/|\lambda_{k+1}|$, where k is the maximal value of j such that $\sum_{i=1}^{j}\lambda_i \geq 0$ (Kaplan & Yorke, 1970). The Lyapunov dimension represents an effective dimension of the chaotic dynamics in the $2N$-dimensional state space. A larger Lyapunov dimension implies more complex dynamics of IO neurons. In the scenario of chaotic resonance, the information transmission is expected to be maximized in the regime where the Lyapunov dimension takes the largest value.

2.5. Feedback error learning

The IO neurons are supposed to provide error signals to an idealized model of the cerebellar cortex. Here, we assume that the cerebellum learns an inverse model of an arm via feedback error learning (Kawato et al., 1987; Kawato & Gomi, 1992; Schweighofer et al., 1998; Shidara et al., 1993). In feedback error learning, supervised learning of a feedforward controller occurs using a feedback control signal as the error signal. As the feedforward controller improves, the reliance on the feedback controller decreases. Note that our purpose here is to show that chaotic resonance of IOs with very low firing rates can enhance the learning of complex mappings, such as an inverse model for arm control; thus other complex mappings such as forward models would have been possible as well. We therefore do not model the cerebellum in great detail, but instead we model a simple network composed of granule cells that project to Purkinje cells with modifiable weights, as in Schweighofer, Doya, and Lay (2001). The granule cells receive a desired state in joint space with (fixed) random weights. Importantly, the Purkinje cells receive both granule cells and spiking IO inputs, the latter of which transmit the feedback error signals.

The feedback error learning controller controls a two-link human arm on a horizontal plane, with parameters adapted from Katayama and Kawato (1993). The arm dynamics is given by

$$\mathbf{M}(\theta)\ddot{\theta} + \mathbf{C}(\dot{\theta}, \theta)\dot{\theta} = \tau, \tag{3}$$

where θ is the vector of the arm joint angles and τ is the motor command (see below). The inertial and Coriolis matrices M and C are given by

$$\begin{aligned}
&M_{11} = I_1 + I_2 + 2W_2L_1\cos(\theta_e) + W_1L_1^2,\\
&M_{12} = M_{21} = I_2 + W_2L_1\cos(\theta_e), \qquad M_{22} = I_2,\\
&C_{11} = -2W_2L_1\sin(\theta_e)\dot{\theta}_s,\\
&C_{12} = -W_2L_1\sin(\theta_e)\dot{\theta}_e = -C_{21}, \qquad C_{22} = 0,
\end{aligned}$$

where θ_e is the elbow joint angle, θ_s is the shoulder joint angle, L_1 and L_2 are segment lengths, I_1 and I_2 are inertia parameters, and W_1 and W_2 are two other parameters.

In feedback error learning, the output of a crude feedback controller and the output of a feedforward controller are summed and form the motor command. The controller receives a desired minimum jerk trajectory (Flash & Hogan, 1985) in joint coordinates. The vectors of motor commands are given by the sum of the feedback and feedforward motor commands, *ufb* and *uff*, respectively:

$$\tau = ufb + uff. \tag{4}$$

We used a simple PD feedback controller.

$$ufb = K_P \cdot (\theta_d - \theta_{\text{sensed}}) + K_D \cdot (\dot{\theta}_d - \dot{\theta}_{\text{sensed}}),$$

where θ_d and θ_{sensed} are the vectors of the desired and sensed joint position, simply taken as the actual joint position; neither delay nor sensory noise is considered in this simple model. Note that *ufb* is an error signal, which is used both for feedback control and as input to the IO that will train the cerebellum/feedforward controller. This feedforward controller consists of a neural network with a "granule cell" layer, which sends its output to a "Purkinje cell" layer.

$$GC_j = \tanh\left(\sum_i v_{ji}s_i\right), \tag{5}$$

$$PC_k = \sum_j w_{kj}GC_j, \tag{6}$$

where GC_j is the jth granule cell activity, PC_k is the kth Purkinje cell activity, v represents the fixed weights from the inputs to the granule cells, w represents the modifiable weights from the granule cells to the Purkinje cells, and the input $s = [\theta_e, \theta_s, \dot{\theta}_e, \dot{\theta}_s, \ddot{\theta}_e, \ddot{\theta}_s]$ represents the desired state vector.

The weights from the granule cells to the Purkinje cell layer are updated based on a simplified model of plasticity at the granule cell–Purkinje cells synapses, as in Kawato and Gomi (1992):

$$w_{kj} = w_{kj} + \alpha \cdot (\text{IO}_k - \text{IO}_{\text{mean}}) \cdot GC_j \tag{7}$$

where IO_k is the spiking activity of the kth IO neuron and α is a learning rate. The error input to the IO cells is then given by $I = I_0 + \beta \cdot ufb$, where I_0 and β are the minimal input and the input gain, respectively. Before learning, the mean firing rate IO_{mean} is determined by averaging the mean firing rates over all IO neurons with constant input $I = I_0$.

2.6. Simulation parameters

The motor task is to reach successively four targets forming a square of 20 × 20 cm with a movement time of 0.5 s. The center location of the square is [0, 0.4] m, where the shoulder is located at [0, 0]. During 30 learning trials, the feedback command *ufb* is integrated as a learning error. We modeled 100 granule cells, which send their inputs to 50 Purkinje cells for each joint. We also have 50 IO cells per joint and one-to-one connections between each IO and Purkinje cell. The fixed weights from the desired state to the granule cells layer are initialized by random variables $N(0, 1)$ and the modifiable weights from the granule cell to the Purkinje cell layer are initialized to zero. The μ-values for the IO neurons are set inhomogeneously as $\mu \in [0.99 \cdot 1.65, 1.01 \cdot 1.65]$, where the mean value of $\mu = 1.65$ is considered as a physiologically plausible value for the spiking neuron (Fujii & Tsuda, 2004; Tsuda et al., 2004). To simulate the low firing frequency of the IO neuron, the time constants η_1, η_2 are set such that the mean firing rate becomes 2 Hz for a constant input. Other simulation parameters are given as $x_{\text{th}} = 0.75$, $\eta_1 = \eta_2 = 0.04$, $I_0 = 0.2$, $\beta = 0.05$, $\alpha = 0.02$, $L_1 = 0.33$ m, $L_2 = 0.34$, $I_1 = 0.067$ kg m^2, $I_2 = 0.97$ kg m^2, $W_1 = 1.52$ kg, and $W_2 = 0.34$ kg m, $K_P = 100$, $K_D = 1$, and the time step $\mathrm{d}t = 0.003$.

3. Results

3.1. Information transmission

The mutual information was calculated in the noise-free condition ($D = 0$) for the chaotic input signal. The coupling strength was varied in the range of $g \in [0, 0.3]$. The upper, middle, and lower graphs of Fig. 1 show the mutual information, the Lyapunov dimension, and the synchronization quantity, respectively. In the regime of weak coupling, a strong coherent activity was induced due to a common input signal (see Fig. 2(a)). Because of this synchrony, the efficiency of the information transmission was low, indicated by the small mutual information. Here, each neuron generated a simple limit cycle oscillation. As the coupling was increased, the network dynamics became chaotic, as indicated by the sudden increase in the Lyapunov dimension ($g \approx 0.03$). This chaos destroyed synchronous firings of the neurons as shown in Fig. 2(b). Accordingly, the synchronization quantity was decreased (see Fig. 1(c)) and the desynchronized firings enhanced the information transfer, as indicated in Fig. 1(a). Similarity between the input signal and the network output can be also confirmed in Fig. 3. By this mechanism, the peak of the Lyapunov dimension coincides with the peak of the mutual information. As the coupling was further increased, the Lyapunov dimension and efficiency of the information transmission were lowered (see also Fig. 3(c)).

To study how the three quantities (mutual information, chaos, and synchrony) are related, their correlations were computed. The correlation coefficient between the mutual information and the Lyapunov dimension was 0.84 ($p < 0.0001$), whereas the correlation coefficient between the mutual information and the synchrony was -0.69 ($p < 0.001$). Since the correlation coefficient between the mutual information and the largest Lyapunov exponent was merely 0.31 (significance level of $p = 0.12$), the Lyapunov dimension seems to be more appropriate than the largest Lyapunov exponent to measure the chaotic resonance.

3.2. Feedback error learning

We then inserted the network of IO neurons in the feedback error learning control scheme. Fig. 4(a) shows the dependence of the asymptotic learning error on the coupling strength. The solid line corresponds to the noise-free case ($D = 0$). The error was minimized at around $g \approx 0.05$. This is in good agreement with the maximal position of the mutual information of Fig. 1(a), implying that the optimal information transmission observed with the intermediate coupling is also important for efficient learning. Thus, chaos plays a key role in desynchronizing IO firings and enhancing the feedback error learning. It should be noted that

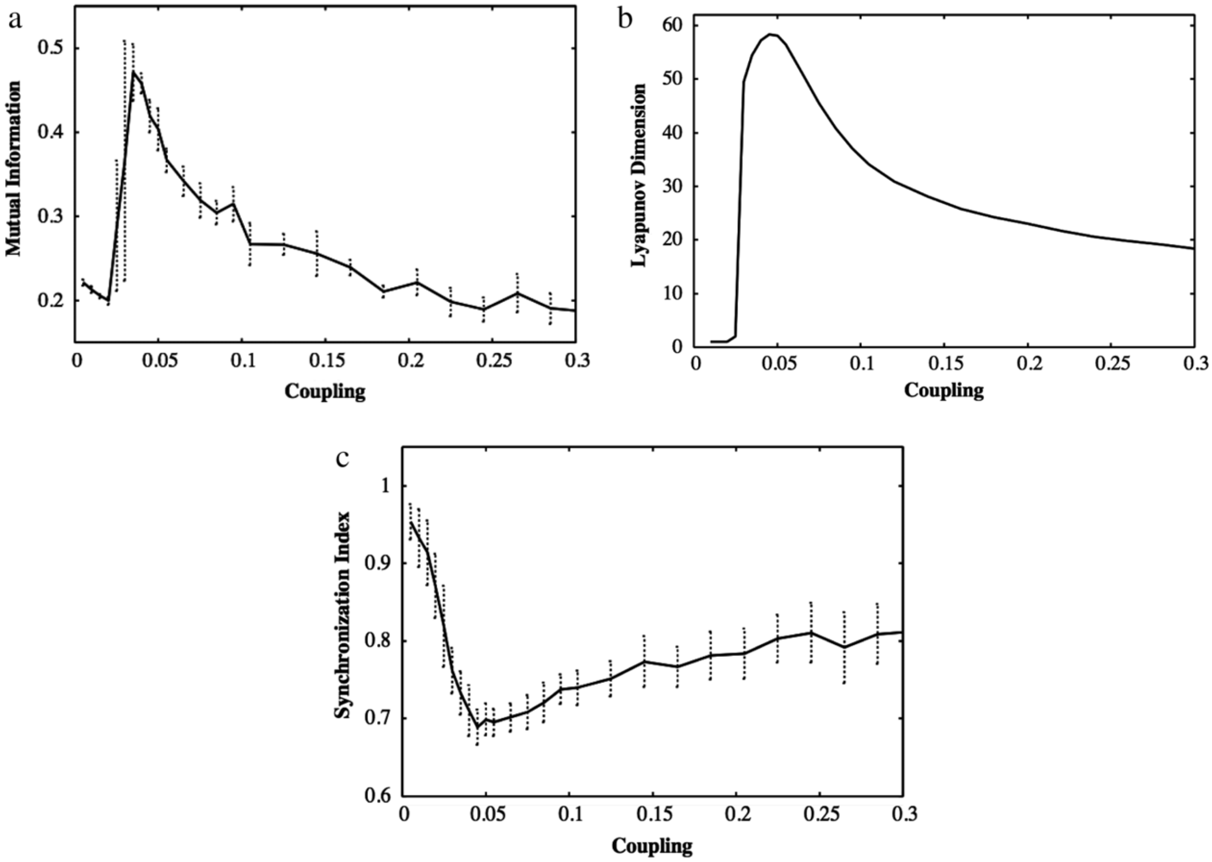


Fig. 1. Transmission of a chaotic signal via a network of coupled inferior olive neurons. The coupling strength is varied between $g \in [0, 0.3]$. (a) Mutual information between input and output. (b) Lyapunov dimension of the network dynamics of the inferior olive. (c) Synchronization index measured by computing the order parameter R.

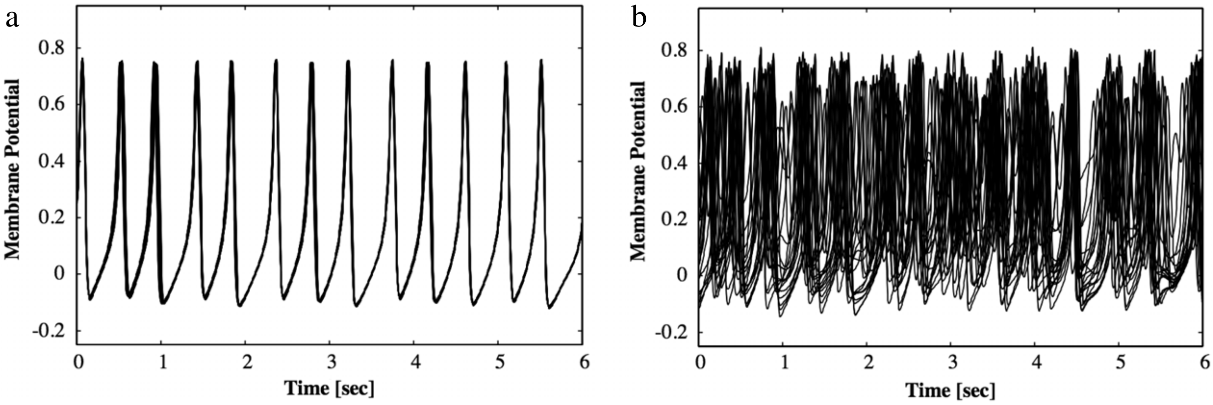


Fig. 2. Superimposed drawing of membrane potential activity of 20 neurons, which receive a common input signal. (a) No coupling ($g = 0$). (b) Intermediate coupling ($g = 0.05$).

a relatively small learning error is observed in a broad range of $g \in [0.02, 0.1]$ in Fig. 4(a).

Next, we studied the effect of noise. Fig. 4(b) shows the results of applying the dynamical noise in the range of $D \in [0, 1]$. Without coupling ($g = 0$; the solid line), the error was quite large for a weak noise ($D < 0.04$), because the noise was too weak to destroy the synchronous firings. As the noise level increased, the coherence disappeared and learning improved. Best learning was realized at the intermediate level of noise of around $D \approx 0.1$. As the noise was further increased, the efficiency of the error transmission decreased and learning deteriorated.

The dotted line of Fig. 4(b) shows the effect of noise in the case of intermediate coupling between the IO neurons ($g = 0.05$). The main difference with no coupling is that efficient learning was also realized in the range of weak noise ($D < 0.05$), where chaos enhances the learning process. Furthermore, worsening of the learning performance was to some extent suppressed in the range of large noise ($D > 0.4$).

The dotted line of Fig. 4(a) shows the dependence of the learning error on the coupling strength in the presence of intermediate noise ($D = 0.1$). Compared to the noise-free curve (the solid line), such intermediate noise realized good learning even for small coupling ($g < 0.05$), where the noise destroyed the coherence

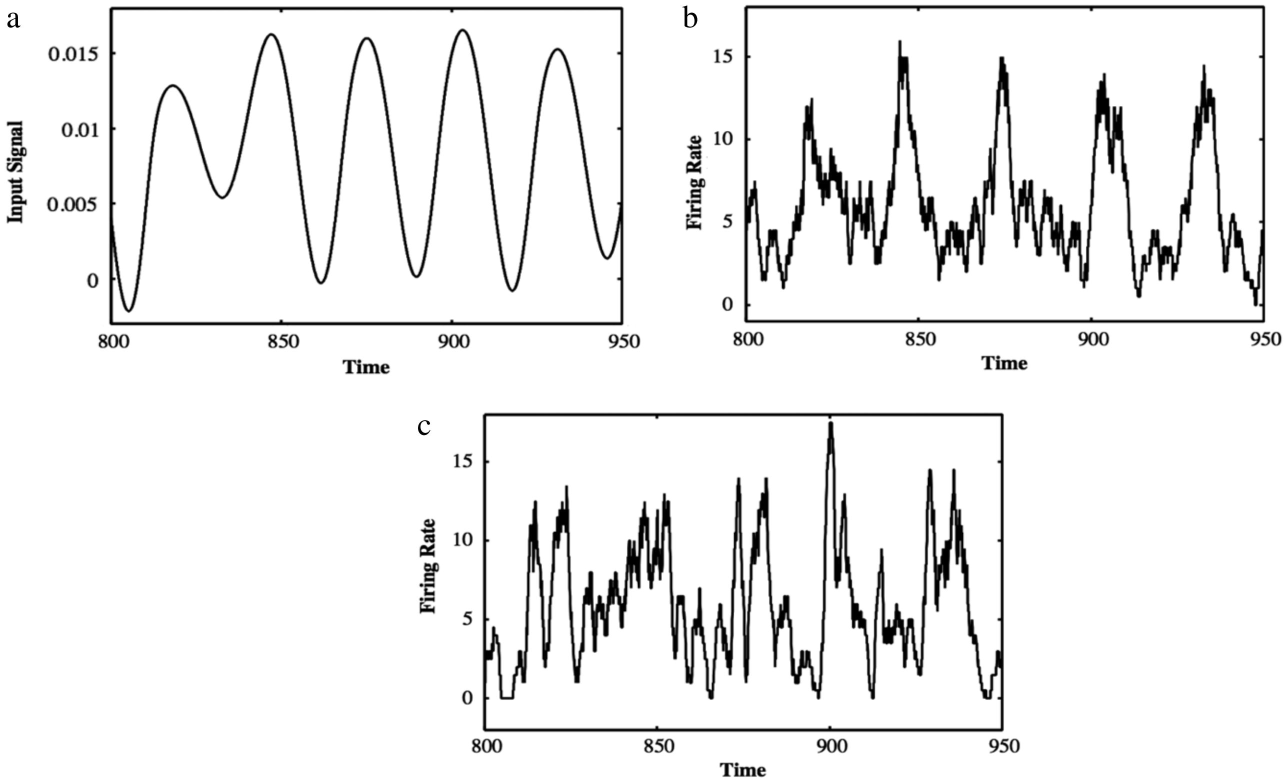


Fig. 3. (a) Input signal $I(t)$ to a network of inferior olive neurons. (b), (c) Firing rate $S(t)$ defined by the number of spikes within a time window of 0.02 normalized by the total neuron number and the time window. $g = 0.05$ for (b), $g = 0.3$ for (c).

of the neural activity instead of chaos. In the range of a strong coupling, which tends to induce synchronous activities of the neurons, the noise again destroyed the synchrony and maintained good learning.

Finally, as a simple quantity to measure the energy required for the neural activity, we computed the spike amplitudes with and without coupling (Fig. 4(c)). In the overall range of the noise, the spike amplitude was significantly lower in the presence of coupling than in its absence.

4. Discussion

To summarize, we used the μ-model as a simplified model for the electrically coupled IO neurons. We studied the capability of information transmission by a network of spatially coupled μ-neurons and found that there exists an intermediate strength of coupling that maximizes the complexity of chaotic firings and simultaneously optimizes the information transmission. The existence of such an optimal regime can be understood as follows. For weak coupling, a signal input common to all the neurons induces a strong coherent neuronal activity (Fig. 2(a)). This synchrony limits the efficient information transmission (Fig. 1(a)). An increase in the coupling strength induces chaotic dynamics, which destroys synchronous firings of the neurons (Fig. 1(b) and (c)). Since the desynchronized firings enhance the information transfer, the peak of the Lyapunov dimension coincides with the peak of the mutual information (Fig. 1(a) and (b)). As the coupling is further increased, the Lyapunov dimension is lowered rather quickly (Fig. 1(b)). As studied in detail by Tsuda et al. (2004), this is due to the characteristics of this particular neural system. Namely, after the system passes through the peak of the Lyapunov dimension, a fully developed chaos becomes localized in a relatively low-dimensional space with a strong coherent property. Such low-dimensional coexisting attractors are eventually destabilized and linked via high-dimensional dynamical paths to form intermittent switching between ruins of the attractors, known as chaotic itinerancy (Tsuda et al., 2004). Because of the low-dimensional coherent motion, the efficiency of the information transmission is lowered (Figs. 1(a) and 3(c)).

These observations lead to the first part of our conclusion that desynchronization of the neural firings is essential for realizing efficient information transfer (Masuda & Aihara, 2002, 2003; Schweighofer et al., 2004). With an intermediate coupling strength, chaos destroys synchronization and maximizes information transmission. This regime corresponds to the chaotic resonance, which has been predicted by Schweighofer et al. (2004). Our finding with the μ-model implies that the chaotic resonance is a general phenomenon, which does not depend upon the details of the neuronal models.

In the second part of our study, the IO network was embedded into the framework of feedback error learning. The simulation study demonstrated that chaotic resonance can greatly enhance the motor learning, where the optimal learning point is located closely to the point of chaotic resonance indicated with the maximal Lyapunov dimension (the solid line of Fig. 4(a)). Compared with the regime of efficient information transmission (Fig. 1(a)), a small learning error was observed in a relatively broad range of $g \in [0.02, 0.1]$. This implies that an efficient learning is realized also in a regime where the information transmission is not highly efficient. Two reasons can be considered for this broadening effect. The first reason might be due to the robustness of the learning. Even if the signals transmitted by the IO network contain certain amount of error or noise, the iterative learning procedure makes up such error to smooth out the noise effect (Holmström & Koistinen, 1992; Matsuoka, 1992). The second reason may come from the simplicity of the control problem. The two-link planar robotics is much simpler than any actual dynamics of human body parts. Especially for complicated movement control of manipulation of objects or locomotion on uneven terrains, different fidelities of the

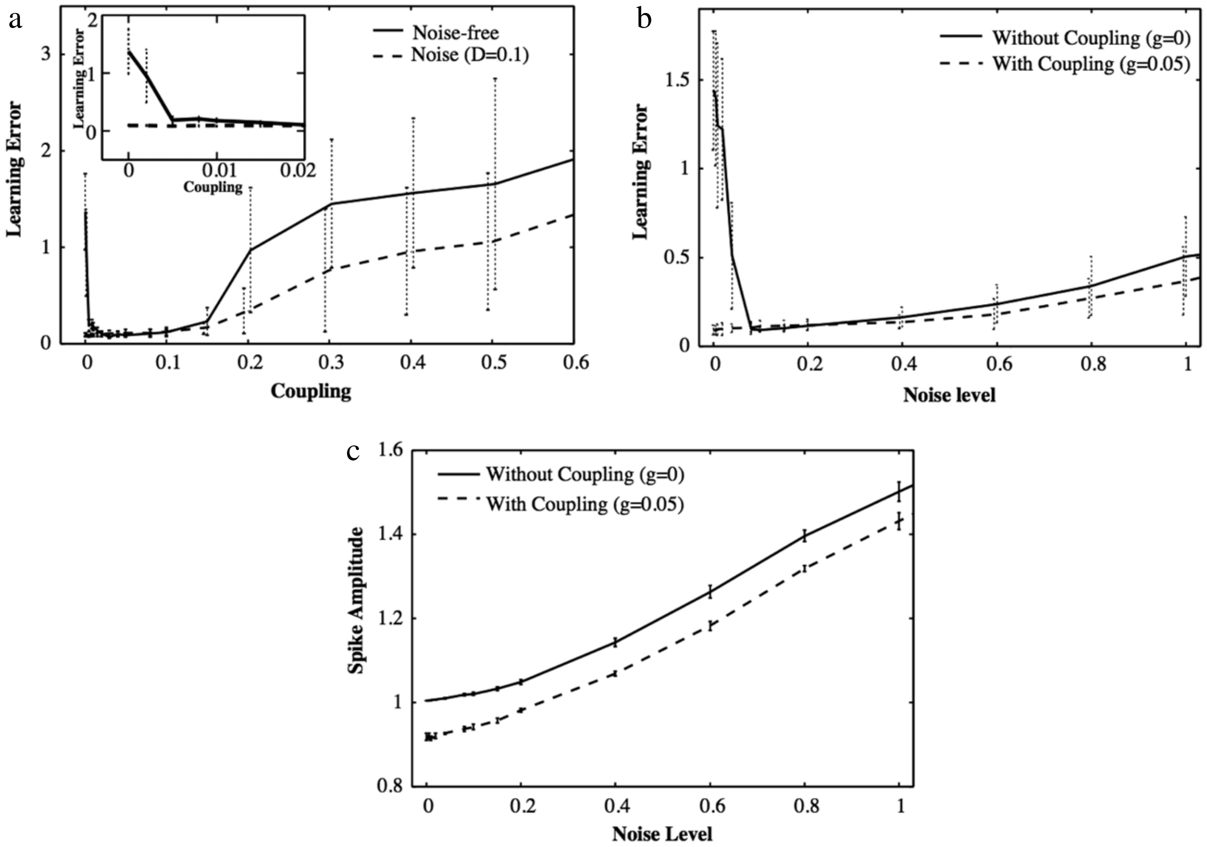


Fig. 4. Effect of coupling and noise in the inferior olive network on feedback error learning. (a) Dependence of the learning error on the coupling strength $g \in [0, 0.6]$. The solid line represents the case without noise ($D = 0$), whereas the dotted line represents the case with dynamical noise ($D = 0.1$). The small graph (inset) represents an enlargement of the small region $g \in [0, 0.02]$. (b) Dependence of the learning error on noise intensity $D \in [0, 1]$. The solid line represents the case without coupling ($g = 0$), whereas the dotted line shows the case with intermediate coupling ($g = 0.05$). (c) Dependence of the spiking amplitude on noise intensity $D \in [0, 1]$. The solid line represents the case without coupling ($g = 0$), whereas the dotted line shows the case with intermediate coupling ($g = 0.05$).

error signal should lead to quite different levels of learning errors. With more complicated problems, the learning curve may become more sensitive to the precision of the signal transmission.

Our key question on the feedback error learning was whether chaos is indispensable to enhance the motor learning. To compare the functionality of chaos with that of noise, a dynamical noise input was introduced to the μ-model. In the no-coupling case, we showed that a small amount of noise is not sufficient to destroy synchronous firings, which results in a large asymptotic error (see the solid line of Fig. 4(b)). Too strong noise, on the other hand, reduces the transmitted information and is again not adequate for learning. For an intermediate level of noise ($D \approx 0.1$ in Fig. 4(b)), the coherence of the neural dynamics disappears and learning is much enhanced, similarly to what is observed with intermediate coupling. The existence of this kind of optimal noise level is reminiscent of stochastic resonance in neural networks (e.g. Collins et al., 1995).

In the presence of the electrical coupling (the dotted line of Fig. 4(b)), efficient learning can be found even in the range of a weak noise ($D < 0.05$), where chaotic resonance enhances the learning process instead of noise. Moreover, in the range of a large noise ($D > 0.4$), learning is less degraded with coupling than without coupling. Here, electrical coupling weakens the noisy dynamics by cancelling the noise acting on opposite directions in the interconnected neurons. Thus, coupling is highly advantageous for widening the range of the noise intensity, which enhances the learning process.

Noise, on the other hand, can compensate for weak coupling. In the presence of an intermediate noise level (the dotted line of Fig. 4(a)), efficient learning is realized when coupling alone is too weak to induce chaos. The same effect can also be seen for a stronger coupling. These findings imply that coupling and noise can compensate each other. Namely, in the absence of noise, coupling enhances learning, whereas noise enhances learning if the coupling is too weak or too strong to induce chaos. This kind of interplay between coupling and noise enlarges the parameter ranges of both coupling strength and noise intensity that provide efficient learning.

Since the effects of noise and coupling in enhancing learning is similar, is there an advantage of chaos-induced learning over noise-induced learning? Our results on spike amplitude suggest that, from an energetic viewpoint, chaos may provide more economical desynchrony than noise for two reasons. First, we found that, in the overall range of noise, the spike amplitude is significantly lower in the presence of coupling than in its absence (Fig. 4(c)). Electrical coupling diminishes the spike amplitudes as follows: if a neuron spikes with relatively strong amplitude, the neighboring (desynchronized) neurons tend to pull the neuronal state back. Thus, chaotic spikes induced by electrical coupling give rise to relatively smaller spike amplitudes, which may require less energy expenditure. According to Lennie (2003), the recovery of ion imbalance associated with spike generation is the major energy consumption for spiking activity. Because the membrane potential is a temporal integration of the incoming ion flux, the spike amplitude may represent the total number of ions which are pumped out after the spiking, thus providing a good estimate for energy expenditure. Second, noise in the nervous system is thought to arise mainly from synaptic noise (Hubbard, Stenhouse, & Eccles, 1967). The generation of such noise is energetically costly because it requires spike generations in the presynaptic neurons and postsynaptic potentials in the postsynatic neurons. On the other hand, electrical coupling itself does not consume too much energy. Finally, our results show that the improvements in learning occur only for limited ranges of coupling strengths and noise amplitudes; thus, to be effective, the coupling or noise level should be relatively well controlled. It has been suggested that the coupling strength between IO neurons is modulated by inhibitory inputs from the cerebellar nucleus (Best & Regehr, 2009). It is unclear however how the noise level in neural systems could be controlled.

In summary, a comparison between our simulations with and without coupling showed that the presence of the electrical coupling can be advantageous in the sense that (1) chaotic resonance enlarges the range of noise level that realizes efficient learning and (2) the spiking activity is kept at relatively low amplitudes, leading to a reduction in the amount of energy necessary for spiking. Addition of the dynamical noise showed that noise can also improve learning with an enlarged range of the efficient coupling strength, where the interplay between noise and chaos works in a complementary manner. In future work, we will validate the hypothesis that physiological electrical coupling strengths induce chaotic firings in real inferior olive neurons.

Acknowledgements

This research was partially supported by Grants-in-Aid for Scientific Research (C) (No. 20560352) from MEXT of Japan and SCOPE (071705001) of the Ministry of Internal Affairs and Communications (MIC) of Japan to IT, by National Science Foundation Grant IIS 0535282 to NS, and by Grants-in-Aid for Scientific Research on Priority Areas (No. 17022012) from MEXT of Japan to KA.

References

Albus, J. S. (1971). The theory of cerebellar function. *Mathematical Biosciences*, *10*, 25–61.

Best, A. R., & Regehr, W. G. (2009). Inhibitory regulation of electrically coupled neurons in the inferior olive is mediated by asynchronous release of GABA. *Neuron*, *62*(4), 555–565.

Collins, J. J., Chow, C. C., & Imhoff, T. T. (1995). Stochastic resonance without tuning. *Nature*, *376*, 236–238.

Flash, T., & Hogan, N. (1985). The coordination of arm movements: an experimentally confirmed mathematical model. *Journal of Neuroscience*, *5*, 1688–1703.

Fox, R. F., Gatland, I. R., Roy, R., & Vemuri, G. (1988). Fast accurate algorithm for numerical simulation of exponentially correlated colored noise. *Physical Review A*, *38*, 5938–5940.

Fujii, H., & Tsuda, I. (2004). Itinerant dynamics of class I^* neurons coupled by gap junctions. *Lecture Notes in Computer Science*, *3146*, 140–160.

Gilbert, P. F., & Thach, W. T. (1977). Purkinje cell activity during motor learning. *Brain Research*, *128*, 309–328.

Holmström, L., & Koistinen, P. (1992). Using additive noise in back-propagation training. *IEEE Transactions on Neural Networks*, *3*, 24–38.

Hubbard, J. I., Stenhouse, D., & Eccles, R. M. (1967). Origin of synaptic noise. *Science*, *157*, 330–331.

Ito, M. (1970). Neurophysiological aspects of the cerebellar motor control system. *International Journal of Neuroscience*, *7*, 162–176.

Ito, M. (1990). A new physiological concept on cerebellum. *Revue Neurologique (Paris)*, *146*, 564–569.

Ito, M., Sakurai, M., & Tongroach, P. (1982). Climbing fibre induced long term depression of both mossy fibre responsiveness and glutamate sensitivity of cerebellar Purkinje cells. *The Journal of Physiology*, *324*, 113–134.

Kaplan, J. L., & Yorke, J. A. (1970). Chaotic behavior of multidimensional difference equations. In H. O. Walter, & H. Peitgen (Eds.), *Lecture notes in mathematics*: *Vol. 730*. *Functional differential equations and approximations of fixed points* (pp. 204–227). Berlin: Springer-Verlag.

Katayama, M., & Kawato, M. (1993). Virtual trajectory and stiffness ellipse during multijoint arm movement predicted by neural inverse models. *Biological Cybernetics*, *69*, 353–362.

Kawato, M., Furukawa, K., & Suzuki, R. (1987). A hierarchical neural-network model for control and learning of voluntary movement. *Biological Cybernetics*, *57*, 169–185.

Kawato, M., & Gomi, H. (1992). A computational model of four regions of the cerebellum based on feedback-error learning. *Biological Cybernetics*, *68*, 95–103.

Kitazawa, S., Kimura, T., & Yin, P. B. (1998). Cerebellar complex spikes encode both destinations and error in arm movements. *Nature*, *392*, 494–497.

Kobayashi, Y., Kawano, K., Takemura, A., Inoue, Y., Kitama, T., Gomi, H., & Kawato, M. (1998). Temporal firing patterns of Purkinje cells in the cerebellar ventral paraflocculus during ocular following responses in monkeys II. Complex spikes. *Journal of Neurophysiology*, *80*, 832–848.

Kuramoto, Y. (1984). *Chemical oscillations, waves and turbulence*. Berlin: Springer.

Lang, E. J., Sugihara, I., & Llinas, R. (1996). GABAergic modulation of complex spike activity by the cerebellar nucleoolivary pathway in rat. *Journal of Neurophysiology*, *76*, 255–275.

Lennie, P. (2003). The cost of cortical computation. *Current Biology*, *13*, 493–497.

Makarenko, V., & Llinas, R. (1998). Experimentally determined chaotic phase synchronization in a neuronal system. *Proceedings of the National Academy of Sciences of the United States of America*, *95*, 15747–15752.

Manor, Y., Rinzel, J., Segev, I., & Yarom, Y. (1997). Low-amplitude oscillations in the inferior olive: a model based on electrical coupling of neurons with heterogeneous channel densities. *Journal of Neurophysiology*, *77*, 2736–2752.

Marr, D. (1969). A theory of cerebellar cortex. *The Journal of Physiology*, *202*, 437–470.

Masuda, N., & Aihara, K. (2002). Briding rate coding and temporal spike coding by effect of noise. *Physical Review Letters*, *88*, 248101.

Masuda, N., & Aihara, K. (2003). Duality of rate coding and temporal coding in multilayered feedforward networks. *Neural Computation*, *15*, 103–125.

Matsuoka, K. (1992). Noise injection into inputs in back-propagation learning. *IEEE Transactions on Systems, Man, and Cybernetics*, *22*, 436–440.

Miall, R. C., Christensen, L. O., Cain, O., & Stanley, J. (2007). Disruption of state estimation in the human lateral cerebellum. *PLoS Biology*, *5*, e316.

Miall, R. C., Weir, D. J., Wolpert, D. M., & Stein, J. F. (1993). Is the cerebellum a Smith predictor? *Journal of Motor Behavior*, *25*, 203–216.

Nishimura, H., Katada, N., & Aihara, K. (2000). Coherent response in a chaotic neural network. *Neural Processing Letters*, *12*, 49–58.

Pikovsky, A., Rosenblum, M., & Kurths, J. (2001). *Synchronization: a universal concept in nonlinear sciences*. Cambridge: Cambridge University Press.

Rényi, A. (1970). *Probability theory*. Amsterdam: North-Holland.

Rössler, O. E. (1979). Continuous chaos. *Annals of the New York Academy of Sciences*, *31*, 376–392.

Schweighofer, N. (1998). A model of activity-dependent formation of cerebellar microzones. *Biological Cybernetics*, *79*, 97–107.

Schweighofer, N., Doya, K., Fukai, H., Chiron, J. V., Furukawa, T., & Kawato, M. (2004). Chaos may enhance information transmission in the inferior olive. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 4655–4660.

Schweighofer, N., Doya, K., & Kawato, M. (1999). Electrophysiological properties of inferior olive neurons: a compartmental model. *Journal of Neurophysiology*, *82*, 804–817.

Schweighofer, N., Doya, K., & Lay, F. (2001). Unsupervised learning of granule cell sparse codes enhances cerebellar adaptive control. *Neuroscience*, *103*, 35–50.

Schweighofer, N., Spoelstra, J., Arbib, M. A., & Kawato, M. (1998). Role of the cerebellum in reaching movements in humans. II. A neural model of the intermediate cerebellum. *European Journal of Neuroscience*, *10*, 95–105.

Shidara, M., Kawano, K., Gomi, H., & Kawato, M. (1993). Inverse-dynamics model eye movement control by Purkinje cells in the cerebellum. *Nature*, *365*, 50–52.

Shimada, I., & Nagashima, T. (1979). A numerical approach to ergodic problem of dissipative dynamical systems. *Progress of Theoretical Physics*, *61*, 1605–1616.

Tseng, Y. W., Diedrichsen, J., Krakauer, J. W., Shadmehr, R., & Bastian, A. J. (2007). Sensory prediction errors drive cerebellum-dependent adaptation of reaching. *Journal of Neurophysiology*, *98*, 54–62.

Tsuda, I., Fujii, H., Tadokoro, S., Yasuoka, T., & Yamaguti, Y. (2004). Chaotic itinerancy as a mechanism of irregular changes between synchronization and desynchronization in a neural network. *Journal of Integrative Neuroscience*, *3*, 159–182.