

# Multi-Layer Perceptron with Pulse Glial Chain Having Oscillatory Excitation Threshold

Chihiro Ikuta

Dept. of Electrical and Electronic Eng.,  
Tokushima University  
2-1 Minami-Josanjima,  
Tokushima Japan  
Email: ikuta@ee.tokushima-u.ac.jp

Yoko Uwate

Dept. of Electrical and Electronic Eng.,  
Tokushima University  
2-1 Minami-Josanjima,  
Tokushima Japan  
Email: uwate@ee.tokushima-u.ac.jp

Yoshifumi Nishio

Dept. of Electrical and Electronic Eng.,  
Tokushima University  
2-1 Minami-Josanjima,  
Tokushima Japan  
Email: nishio@ee.tokushima-u.ac.jp

**Abstract**—A brain has neuron and glial cells. In the brain, these cells correlate each other and make a higher brain function. In this study, we propose a Multi-Layer Perceptron (MLP) with pulse glial chain having oscillatory excitation threshold. We connect artificial glia units with the neurons in a hidden-layer. When the connecting neuron output is larger than an excitation threshold of the connected glia, the glia excites and generates the pulse. This pulse transmits to the neighboring glia and the connecting neuron. The pulse increases a threshold of the connecting neuron, thus the glia gives energy for solving tasks. In this model, the excitation threshold is oscillating within a defined value. Even if the connecting neuron output does not change, the pulse generation occurs by the oscillation of the excitation threshold. The oscillation of the excitation threshold gives more energy to the network and improves a learning performance of the MLP. By computer simulation, we confirm that the oscillation of the excitation threshold improves a learning performance.

## I. INTRODUCTION

A glia and a neuron are cells existing in a brain. We already have many investigations and many applications about the neuron, because the neuron transmits an electric signal each other. Thus, the neuron closely relates to brain works. On the other hand, a main function of the glia were considered to a support of the neuron. Actually, the glia has many support functions for the neuron. However, several researchers discovered that the glia has important functions for a higher brain function [1][3]. The glia has many receptors to ions such as glutamate acid, adenosine triphosphate, calcium [4][6]. Among them, we focus on the calcium ion. The glia transmits a signal by using the change of calcium ion concentration [7][8]. This change of ion concentration also influences a membrane potential of the neuron. Thus, the glia closely relates to the neurons and makes the brain works. We consider that the glia function can apply to an artificial neural network.

For the application, we use a Multi-Layer Perceptron (MLP). The MLP is a famous artificial neural network. In this model, the neurons connect with the neurons in neighboring layers. Each neuron has weights of connections between the neurons. We can change the output value of the MLP by tuning of the weights of connection. In general, a Back Propagation (BP) algorithm is used for determination of a value of the weights of connections [9]. By this learning, the MLP can be applied to several tasks such as pattern recognition, data mining. However, the BP algorithm has a local minimum

problem, because the BP algorithm uses the steepest decent method. If the MLP falls into local minimum, the MLP cannot escape out from there. For this problem, many methods have proposed such as a noise inputting, introduction of a momentum coefficient, and so on.

In this study, we propose a Multi-Layer Perceptron (MLP) with pulse glial chain having oscillatory excitation threshold which is inspired by the biological glia functions. We one-by-one connect the glia with the neuron in the hidden-layer. Each glia receives the connecting neuron output. When the connecting neuron has a large output than an excitation threshold of the glia, the glia is excited. The excited glia generates a pulse. This pulse transmits to the neighboring glia and the connecting neuron. If the neighboring glia receives other glia pulse, this glia is also excited. Therefore, the glia pulse transmits to all glia. In the case of the neuron, the glia pulse increases the threshold of the inner state of the neuron. The neuron output is changed by the glia pulse and obtains an energy from the glia. In this model, we introduce the oscillatory excitation threshold of the glia. The excitation threshold oscillates within defined value during a learning. When the learning of the network converges, the output of the neuron is fixity. Thereby, the pattern of the glia pulse generation is also fixed. In the proposed model, the glia can change between a generation and a stopping generation of the pulse by the oscillatory excitation threshold of the glia. By the change of the pulse generation, the pulse propagation is also changed. We consider that the oscillatory excitation threshold influences the convergent of the MLP learning. We confirm that the glia improves the MLP learning performance by using Two-Spiral Problem (TSP).

## II. PROPOSED MLP

The MLP is one of feed forward neural network. The neurons make layers and have connections with the neurons in the neighboring layers. In this study, we propose the MLP with pulse glial chain having oscillatory excitation threshold shown as Fig. 1.

### A. Pulse glial chain having oscillatory excitation threshold

The glia are one-by-one connected with the neurons in the hidden-layer. In this model, the glia and the neurons correlate each other. The glia receives the connecting neuron output

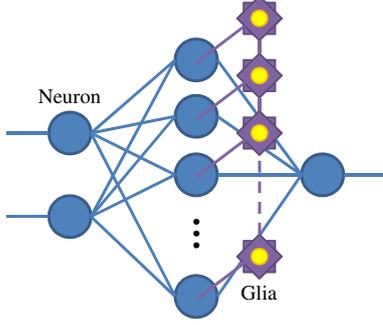


Fig. 1. Proposed MLP.

and is excited by the connecting neuron output. The excitation conditions of the glia are described by Eq. (1).

$$\psi_i(t+1) = \begin{cases} 1, & \{\theta_n(t) < y_i \cup \psi_{i+1}(t-D) = 1 \\ & \cup \psi_{i-1}(t-D) = 1\} \cap (t - \tau_i > \theta_g), \\ -1, & \{1 - \theta_n(t) > y_i \cup \psi_{i+1}(t-D) = 1 \\ & \cup \psi_{i-1}(t-D) = 1\} \cap (t - \tau_i > \theta_g) \\ \gamma\psi_i(t), & \text{else,} \end{cases}, (1)$$

where  $\psi$  is an output of a glia,  $i$  is a position number of the glia in the hidden-layer,  $\gamma$  is an attenuation parameter ( $0 < \gamma < 1$ ),  $y$  is an output of a connecting neuron,  $\theta_n$  is a excitation threshold of the glia,  $\tau$  is a time of a previous pulse generation,  $\theta_g$  is a period of inactivity, and  $D$  is a delay time of a glial effect. When the connecting neuron output is larger than the excitation threshold ( $\theta_n(t)$ ) of the glia, the glia is excited and generates the positive pulse. And also, when the connecting neuron output is smaller than  $1 - \theta_n(t)$  of the glia, the glia is excited and generates the negative pulse. The generated pulse transmits to the neighboring glia and the connecting neuron. The pulse excites the neighboring glia and increases the inner state of the connecting neuron. The neighboring glia also generate the pulse, thereby the pulse transmits to all glia. When the glia generates the pulse, the glia starts the period of inactivity. If this glia receives the neuron output and/or the neighboring glial pulse, the glia cannot be excited again during period of inactivity.

In this study, we introduce the oscillatory excitation threshold of the glia to the pulse glial chain. The oscillatory excitation threshold of the glia is described by Eq. (2).

$$\theta_n(t) = r(t) \quad (a < r < b), \quad (2)$$

where  $r$  is random function, and  $a$  and  $b$  is constant value. The value of the excitation threshold oscillates between  $a$  and  $b$ . We use Mersenne Twister pseudo-random number [10] for the random function. Figure 2 shows example of the relationships between the oscillatory excitation threshold of the glia and the output of the neurons. We consider three cases of the output of the neuron in the simulation. The output of the neuron (A) does not reach the excitation threshold of the glia, thereby the connected glia is not excited by this neuron. The output of the neuron (B) is into the amplitude of the oscillatory excitation threshold of the glia when the output of this neuron is converged. In this case, the connected glia with neuron (B) is stochastically excited. The output of the neuron (C) completely

overcomes the oscillatory excitation threshold of the glia when the output of this neuron is converged. The connected glia with the neuron (C) is excited independent from the oscillation of the excitation threshold of the glia. In the proposed method, several glia are stochastically excited by the oscillation of the excitation threshold of the glia. We consider that the oscillatory excitation threshold breaks the periodic pulse generation of the glia, thereby the glia can effectively give the energy to the MLP.

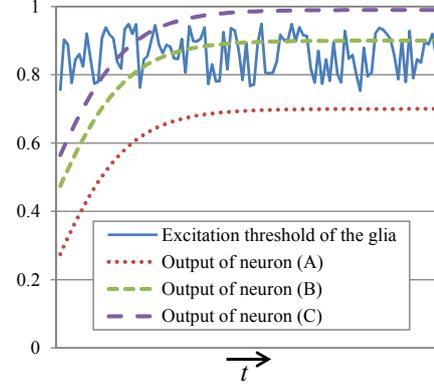


Fig. 2. Relationships between oscillatory excitation threshold of the glia and the output of the neurons.

### B. Updating rule of neuron

The neuron has multi-input and single-output, and we can change the neuron output by tuning the weights of connections between the neurons. The standard updating rule of the neuron is defined as shown in Eq. (3).

$$y_i(t+1) = f \left( \sum_{j=1}^n w_{ij}(t)x_j(t) - \theta_i(t) \right), \quad (3)$$

where  $y$  is an output of the neuron,  $w$  is a weight of the connection,  $x$  is an input of the neuron, and  $\theta$  is an excitation threshold of the neuron. In this equation, the weight of the connection and the threshold of the neuron are learned based on the BP algorithm. Next, we show the proposed updating rule of the neuron. We add the generated pulse of the glia  $\psi$  to the excitation threshold of the neuron. In this study, this updating rule is only used for the neurons in the hidden-layer. The updating rule is described by Eq. (4).

$$y_i(t+1) = f \left( \sum_{j=1}^n w_{ij}(t)x_j(t) - \theta_i(t) + \alpha\psi_i(t) \right), \quad (4)$$

where  $\alpha$  is a weight of the glial effect. A peek of the generated pulse is changed according to  $\alpha$ . We choose an optimal value of  $\alpha$  for solving a task by a heuristic search. The generated pulse is independent from the learning of the MLP, thus this pulse can gives an energy to the network and helps for escaping out from the local minimum.  $\psi$  is updated by Eq. (1). Equations (3) and (4) are used as a sigmoidal function to an activating function of neuron and is described by Eq. (5).

$$f(a) = \frac{1}{1 + e^{-a}}, \quad (5)$$

where  $a$  is an inner state of the neuron.

### III. SIMULATION

We use five kinds of the MLPs for comparison of the performance.

- (1) The standard MLP
- (2) The MLP with random noise
- (3) The MLP with random timing pulses
- (4) The MLP with pulse glial chain
- (5) The MLP with pulse glial chain having oscillatory excitation threshold

The standard MLP (1) does not have the external unit, thus this MLP often falls into the local minimum. The MLP with random noise (2) has a uniformed random noise in the excitation threshold of the neurons in the hidden-layer. The MLP with random timing pulses (3) has a pulse oscillation in the excitation threshold of the neurons in the hidden-layer, and this pulse is generated at random. The MLP with pulse glial chain has a glial pulse in the excitation threshold of the neurons in the hidden-layer. The MLP with pulse glial chain having oscillatory excitation threshold is the proposed MLP. In this model, the glias generate and transmit pulse, and the glia excitation threshold oscillates within a defined value. We use a Two-Spiral Problem (TSP) for a task of the MLPs. The TSP is a famous task for the artificial neural network and has a high nonlinearity [11][12]. In this task, we input the coordinates of the spiral points, and the MLP learns the classification of each spiral coordinate. We obtain ideal classification from norms between the each coordinates and the learning spiral points. The learning spiral points and the ideal classification are shown in Fig. 3.

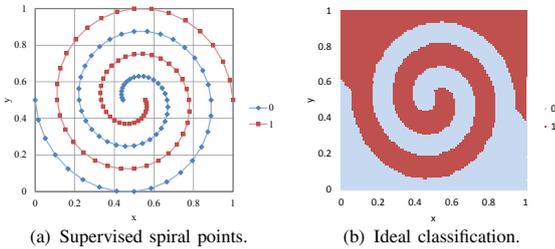


Fig. 3. Target points.

We obtain the experimental results from 100 trials about each MLP. In each trial, we give the different initial weights of connections. The MLP has 50000 iterations during one trial and learns 130 data sets of spiral points. We use the three layers MLP, and the number of neurons in each layer is 2-40-1. We use Mean Square Error (MSE) for the error function. The MSE is described by Eq. (6).

$$MSE = \frac{1}{N} \sum_{n=1}^N (T_n - O_n)^2, \quad (6)$$

where  $N$  is the number of learning datum,  $T$  is a target value, and  $O$  is an output of MLP. We obtain results which are an average error, a minimum error, a maximum error, and a standard deviation of the results.

#### A. Learning performance

Firstly, we obtain a learning performance. The learning performance means that the accuracy of the classification for

the learning data sets. We compare the output of the MLP and supervised classification of the spiral points. The learning performance of each MLP is shown in Table I. The standard MLP (1) is the worst of all in the average error, the minimum error and the maximum error. We can say that the standard MLP (1) falls into local minimum. The MLP with random noise (2) and the MLP with random timing pulse (3) reduce the error than the standard MLP (1), however these methods do not have a large difference from the standard MLP (1). On the other hand, the MLP with pulse glial chain (4) and the MLP with pulse glial chain having oscillatory excitation threshold (5) reduce the error more than the others. In the case of the MLP with pulse glial chain having oscillatory excitation threshold (5), the maximum error is smaller than the MLP with pulse glial chain (4). From this result, we consider that the oscillatory excitation threshold of the glia helps escaping out from the local minimum. Thereby, the maximum error of the proposed MLP (5) is reduced than that of the MLP with pulse glial chain.

TABLE I. LEARNING PERFORMANCE.

	Average	Minimum	Maximum	Std. Dev.
(1)	0.12269	0.00831	0.23857	0.05554
(2)	0.10847	0.00047	0.24278	0.05742
(3)	0.11439	0.00740	0.26349	0.05742
(4)	0.01990	0.00067	0.11664	0.02226
(5)	0.01436	0.00069	0.08139	0.01688

Next, we show the difference of the learning curves between the MLP with pulse glial chain (4) and the MLP with pulse glial chain having oscillatory excitation threshold (5) as shown in Fig. 4. The learning curves are obtained from the average error in each iteration. The error reductions are similar from start of the learning to 10000 iterations. After that, the error reduction of the MLP with pulse glial chain having oscillatory excitation threshold (5) is faster than the MLP with pulse glial chain (4). We consider that the pulse generations both methods are similar to each other at start of the learning, because the neurons' outputs are often changed by the learning. When the neuron output becomes stable in each neuron, the oscillatory excitation threshold influences the pulse generation. Hence, the MLP with pulse glial chain having oscillatory excitation threshold (5) can reduce the error than the MLP with pulse glial chain (4).

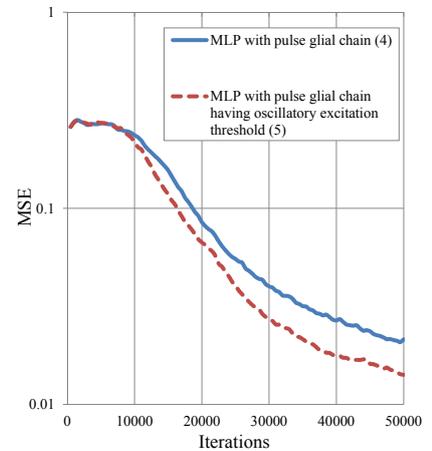


Fig. 4. Learning curves.

## B. Classification performance

We show the classification performance of the MLPs. The classification performance shows a generalization capability. We give the unlearning coordinates to the learned MLPs from 0 to 1 in x-y plane. And also, we compare the output of the MLP and the ideal classification in each coordinate. The classification performance is shown in Table II. The standard MLP (1), the MLP with random noise (2) and the MLP with random timing pulse (3) have similar classification performance. The MLP with pulse glial chain (4) and the MLP with pulse glial chain having oscillatory excitation threshold (5) have a better classification performance than the others. Especially, the MLP with pulse glial chain having oscillatory excitation threshold (5) obtains the smallest of all in the minimum error. In general, the MLP becomes over fitting when the MLP is excessively learned. In the MLP with pulse glial chain having oscillatory excitation threshold (5) has a high learning performance and a high classification performance. From this result, we can say that this MLP has a high searching ability of the solutions.

TABLE II. CLASSIFICATION PERFORMANCE.

	Average	Minimum	Maximum	Std. Dev.
(1)	0.21782	0.10565	0.29477	0.03858
(2)	0.19278	0.10460	0.33065	0.04434
(3)	0.20432	0.12082	0.31958	0.03851
(4)	0.12538	0.08027	0.19639	0.02625
(5)	0.11884	0.06693	0.19198	0.02344

Finally, we show examples of the classification of the unlearning coordinates of the MLPs in Fig. 5. We obtain the classification image about average results in the MLPs. The standard MLP (1), the MLP with random noise (2) and the MLP with random timing pulses (3) cannot draw the two spirals. On the other hand, the MLP with pulse glial chain (4) and the MLP with pulse glial chain having oscillatory excitation threshold (5) can draw the two spirals. From this result, we consider that the MLP with pulse glial chain having oscillatory excitation threshold can solve the TSP.

## IV. CONCLUSION

In this study, we have proposed the MLP with pulse glial chain having oscillatory excitation threshold which is inspired from the glia functions. The glia are one-by-one connected with the neurons in the hidden-layer. The glia generates the pulse according to the connecting neuron output. This pulse transmits to the neighboring glia and the neurons. In the proposed model, we introduce the oscillatory excitation threshold of the glia. By the oscillatory excitation threshold, the excited glia are sometimes randomly chosen. Thereby, the pulse glial chain having oscillatory excitation threshold gives more energy to the MLP performance than the previous method. We confirmed that the proposed MLP has the better performance than the conventional MLPs.

## ACKNOWLEDGMENT

This work was partly supported by MEXT/JSPS Grant-in-Aid for JSPS Fellows (24-10018).

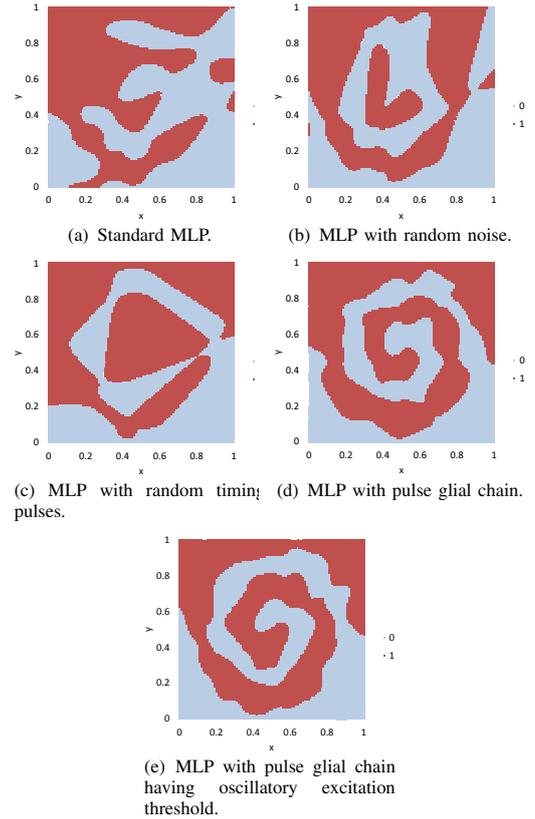


Fig. 5. Examples of classification results.

## REFERENCES

- [1] P.G. Haydon, "Glia: Listening and Talking to the Synapse," *Nature Reviews Neuroscience*, vol.2, pp.844-847, 2001.
- [2] R.D. Fields, B. Stevens-Graham, "New Insights into Neuron-Glia Communication," *Science*, vol.298, pp.556-562, 2002.
- [3] G.I. Hatton, V. Parpura, "Glia Neuronal Signaling," 2004, Kluwer Academic Publishers.
- [4] S. Koizumi, M. Tsuda, Y. Shigemoto-Nogami and K. Inoue, "Dynamic Inhibition of Excitatory Synaptic Transmission by Astrocyte-Derived ATP in Hippocampal Cultures," *Proc. National Academy of Science of U.S.A.*, vol.100, pp.11023-11028, Mar. 2003.
- [5] S. Ozawa, "Role of Glutamate Transporters in Excitatory Synapses in Cerebellar Purkinje Cells," *Brain and Nerve*, vol.59, pp.669-676, 2007.
- [6] G. Perea and A. Araque, "Glial Calcium Signaling and Neuro-Glia Communication," *Cell Calcium*, vol.38, pp.375-382, 2005.
- [7] S. Kriegler and S.Y. Chiu, "Calcium Signaling of Glial Cells along Mammalian Axons," *The Journal of Neuroscience*, vol.13, pp.4229-4245, 1993.
- [8] M.P. Mattoson and S.L. Chan, "Neuronal and Glial Calcium Signaling in Alzheimer's Disease," *Cell Calcium*, vol.34, pp.385-397, 2003.
- [9] D.E. Rumelhart, G.E. Hinton and R.J. Williams, "Learning Representations by Back-Propagating Errors," *Nature*, vol.323-9, pp.533-536, 1986.
- [10] M. Matsumoto and T. Nishimura, "Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator," *ACM Transactions on Modeling and Computer Simulation*, vol.8 (1), pp.3-30, 1998.
- [11] J.R. Alvarez-Sanchez, "Injecting Knowledge into the Solution of the Two-Spiral Problem," *Neural Computing & Applications*, vol.8, pp.265-272, 1999.
- [12] H. Sasaki, T. Shiraishi and S. Morishita, "High Precision Learning for Neural Networks by Dynamic Modification of Their Network Structure," *Dynamics & Design Conference*, pp.411-1-411-6, 2004.