Available online at www.sciencedirect.com
ScienceDirect
Neural Networks 20 (2007) 290–297
Neural Networks
www.elsevier.com/locate/neunet

2007 Special Issue

The cerebellum as a liquid state machine

Tadashi Yamazaki[1], Shigeru Tanaka*

Laboratory for Visual Neurocomputing, RIKEN Brain Science Institute, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

Abstract

We examined closely the cerebellar circuit model that we have proposed previously. The model granular layer generates a finite but very long sequence of active neuron populations without recurrence, which is able to represent the passage of time. For all the possible binary patterns fed into mossy fibres, the circuit generates the same number of different sequences of active neuron populations. Model Purkinje cells that receive parallel fiber inputs from neurons in the granular layer learn to stop eliciting spikes at the timing instructed by the arrival of signals from the inferior olive. These functional roles of the granular layer and Purkinje cells are regarded as a liquid state generator and readout neurons, respectively. Thus, the cerebellum that has been considered to date as a biological counterpart of a perceptron is reinterpreted to be a liquid state machine that possesses powerful information processing capability more than a perceptron.

Keywords: Cerebellum; Perceptron; Liquid state machine; Spatiotemporal activity patterns; Sparse coding; Recurrent inhibitory network; Long-term depression; Hybrid network

1. Introduction

In the Marr–Albus–Ito theory of cerebellar computation (Albus, 1971; Ito, 1984; Marr, 1969), the cerebellum is considered as a biological counterpart of a simple perceptron, which is a two-layer neural network with learning capability (Rosenblatt, 1958). Specifically, granule cells and Purkinje cells in the cerebellar cortex constitute the input and output layers, respectively, and connections between them by granule cell axons called parallel fibres are modifiable by an instruction signal coming from the inferior olive to Purkinje cells through climbing fibres, which is well known as long-term depression (LTD) (Ito, 1989, 2002b). While a simple perceptron is able to compute a function that describes only linear separation of input signals (Haykin, 1999), the cerebellum plays an essential role in motor control for coordinating movements of different body parts into a harmoniously integrated body movement, in which an enormous number of muscles must be activated precisely in a correct order and timing. Moreover, recent studies have suggested that the cerebellum is involved in higher cognitive functions including time perception and language processing (Ito (2002a) for review). How does the cerebellum as a simple perceptron perform such a complex task?

The granule cell layer comprises granule cells in a recurrent inhibitory network with Golgi cells (Ito, 1984), the major granule cell layer interneuron, suggesting that the input layer of the cerebellum represents a recurrent circuit. According to this observation, several groups have proposed cerebellar models in which the input layer is a recurrent network (Buonomano & Mauk, 1994; Hofstötter, Mitz, & Verschure, 2002; Medina, Garcia, Nores, Taylor, & Mauk, 2000; Yamazaki & Tanaka, 2005a). Yet, the computational power of the entire network has not been clarified. We have studied the dynamics of the recurrent circuit theoretically with a simplified rate-coding model (Yamazaki & Tanaka, 2005a) and a realistic model composed of spiking neuron units (Yamazaki & Tanaka, 2005b). These studies have shown that model granule cells exhibit a random repetition of transitions between active and inactive states. The sparse population of active cells changes with time, and there is no recurrence of active cell populations. Therefore, one population in a sequence of active cell populations is able to represent exclusively a specific time interval. Namely, the model cerebellum represents the passage of time in a sparse-population coding scheme. We have also demonstrated (1) resetability of activity pattern generation,

* Corresponding author. Tel.: +81 48 467 9667; fax: +81 48 467 9684.
E-mail address: shigeru@riken.jp (S. Tanaka).
[1] Current address: Laboratory for Motor Learning Control, RIKEN Brain Science Institute, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan.

0893-6080/$ - see front matter © 2007 Published by Elsevier Ltd
doi:10.1016/j.neunet.2007.04.004

(2) robust generation of an activity pattern against noise, (3) representation of different time passages associated with different input signals.

In the present study, we examined in more detail the information representation capability of the recurrent circuit in the model granular layer and the ability of model Purkinje cells to read out information. We considered to learn and represent a Boolean function, which has K bit inputs and M bit outputs. Model neurons in the recurrent circuit were able to distinguish all combinations of 2^K binary input signal patterns by a sparse-population coding, suggesting spatial discrimination capability of the network. Furthermore, the neurons could distinguish elapsed time from the onset of an input signal, indicating temporal discrimination capability as well. Taken together, the recurrent circuit was a spatiotemporal discriminator of input signals. On the other hand, we verified that our model cerebellar circuit is able to learn and represent any given Boolean function.

The spatiotemporal activity patterns of neurons generated by the recurrent circuit in the model granular layer have no fixed point attractors, because the same pattern does not appear more than once. On the other hand, model Purkinje cells, which receive the activity patterns and instruction signals, learn to generate desired output signals. In terms of the liquid state machine (Maass, Natschläger, & Markram, 2002), the granular layer in the cerebellum corresponds to a liquid state generator, and Purkinje cells work as readout neurons. Such functional resemblance of the cerebellum to a liquid state machine explains the huge computational power of the previous cerebellar models having a recurrent network as an input layer (Buonomano & Mauk, 1994; Hofstötter et al., 2002; Medina et al., 2000; Yamazaki & Tanaka, 2005a). Moreover, our liquid state machine hypothesis will supersede the classical perceptron hypothesis of the cerebellum.

2. Model description

Fig. 1 shows a schematic diagram of cell types and synaptic connections in the cerebellum (Eccles, Ito, & Szentágothai, 1967; Ito, 1984). The entire network computes a function $f : \left(\mathcal{B}^K, \mathcal{N}\right) \to \mathcal{B}^M$, where $\mathcal{B}$ and $\mathcal{N}$ denote the Boolean and natural integers, respectively. That is,

$$\mathbf{o} = f(\mathbf{x}, t), \tag{1}$$

where $\mathbf{x}$ and $\mathbf{o}$ denote $(x_1, x_2, \ldots, x_K)$ and $(o_1(t), o_2(t), \ldots, o_M(t))$, respectively, and t represents the discrete elapsed time from the onset of an input signal.

The model granular layer transforms static K inputs $(x_1, x_2, \ldots, x_K)$ into N spatiotemporal activity patterns $(z_1(t), z_2(t), \ldots, z_N(t))$ as shown below. Model Purkinje cells receive the spatiotemporal activity patterns and generate M outputs $(r_1(t), r_2(t), \ldots, r_M(t))$. For any i and t, $r_i(t)$ is given by

$$r_i(t) = \sum_j J_{ij} z_j(t), \tag{2}$$

where J_{ij} is the synaptic weight of the connection from neuron j in the granular layer to Purkinje cell i. The final output of the network $(\hat{o}_1(t), \hat{o}_2(t), \ldots, \hat{o}_M(t))$ is calculated as

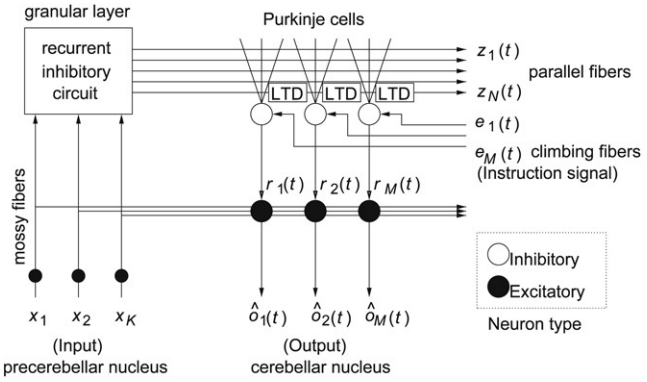

Fig. 1. Schematic diagram of cell types and synaptic connections in cerebellum.

$$\hat{o}_i(t) = \Theta\left[1 - r_i(t) - \theta\right], \quad \text{for } i = 1, \ldots, M, \tag{3}$$

where 1 in the argument of Θ represents the normalized amplitude of excitatory input signals through mossy fibres, $\Theta[x] = 1$ for $x \geq 0$ and 0 otherwise, and θ is a threshold constant. The weights of J_{ij} of parallel fibre inputs to Purkinje cells, are modified via a climbing fibre dependent LTD rule (Ito, 1989). That is, the conjunctive stimulation of parallel fibres (i.e. $z_j(t) > 0$) and the climbing fibre (i.e. $e_i(t) = 1$) depresses synapses of the parallel fiber terminals on the Purkinje cell dendrites. The instruction signal pattern is set at a desired output pattern as

$$e_i(t) = o_i(t), \quad \text{for } i = 1, \ldots, M. \tag{4}$$

Accordingly, J_{ij} is set as follows:

$$J_{ij} = \begin{cases} 0 & z_j(t) > 0 \text{ and } e_i(t) = 1, \\ 1 & \text{otherwise.} \end{cases} \tag{5}$$

Next, the model granular layer consists of N neurons (model granule cells). Let $z_i(t)$ be the activity of neuron i at time t, which is given by

$$z_i(t) = [u_i(t)]^+, \tag{6}$$

Here $[x]^+ = x$ for $x > 0$ and 0 otherwise. $u_i(t)$ is the internal state of neuron i at time t. $u_i(t)$ is defined by

$$u_i(t) = I_i - \sum_j w_{ij} \sum_{s=1}^{t} \exp\left(-(t-s)/\tau\right) z_j(s-1), \tag{7}$$

where I_i and w_{ij} denote the afferent input signal to neuron i through mossy fibres and the synaptic weight of recurrent inhibition from neuron j to neuron i. The summation with respect to s in the second term on the right-hand side represents the temporal integration of activities over a long time. This indicates that the activity of neuron j is integrated through time by the convolution with an exponential decay factor, and τ determines the integration range. Derivation of Eq. (7) is found in our previous paper (Yamazaki & Tanaka, 2005a).

We determine the value of w_{ij} randomly, indicating that the model granular layer is a random recurrent inhibitory network.

In the cerebellar granular layer, granule cells functionally inhibit other granule cells via Golgi cells, indicating that the granular layer is a recurrent inhibitory circuit. We assumed that connections between granule and Golgi cells are random. So long as we do not know the specific connectivity rule between these cells, it would be reasonable to assume the connections random. This is the same logic as adopted by "random matrix theory" for studying the energy level statistics of high-energy nuclear states, where the matrix elements of the Hamiltonian are assumed to be random variables due to the lack of knowledge on the matrix elements. As derived in our previous paper, effective granule–Golgi–granule connections consequently become random.

We assumed that neurons and afferent mossy fibres are connected randomly, where I_i is set as follows:

$$I_i = \frac{\sum_j Q_{ij} x_j}{\sum_j x_j}, \tag{8}$$

where $\Pr\left(Q_{ij} = 1/4\right) = 4/K$ and $\Pr\left(Q_{ij} = 0\right) = 1 - 4/K$, suggesting that each granule cell receives, on average, four mossy fibres (Mugnaini, Atluri, & Houk, 1974).

We hypothesize that when an input signal pattern is fed into mossy fibres, a spatiotemporal activity pattern is generated using Eqs. (6) and (7), and the population of active neurons change with time without recurrence for finite but sufficiently long time. To examine the nonrecurrence, we use the following correlation function, which we call the similarity index:

$$C(t_1, t_2) = \frac{\sum_i z_i(t_1) z_i(t_2)}{\sqrt{\sum_i z_i^2(t_1)} \sqrt{\sum_i z_i^2(t_2)}}. \tag{9}$$

The numerator calculates the inner product of population vectors of active neurons at times t_1 and t_2, and the denominator normalizes the vector length. The index takes a value from 0 to 1. It takes 0 if the two vectors are complement, 0.5 if they are random, and 1 if they are identical. We let $C(t_1, t_2) = 0$ if $\sum_i z_i^2(t_1) = 0$ or $\sum_i z_i^2(t_2) = 0$.

We also hypothesize that different spatiotemporal activity patterns are generated for different input signals. To demonstrate this hypothesis, we use the following similarity index.

$$C(t_1, t_2) = \frac{\sum_i z_i^{(1)}(t_1) z_i^{(2)}(t_2)}{\sqrt{\sum_i z_i^{(1)2}(t_1)} \sqrt{\sum_i z_i^{(2)2}(t_2)}}, \tag{10}$$

where $z_i^{(1)}(t)$ and $z_i^{(2)}(t)$ denote the activities of neuron i at time t under different input signals. We let $C(t_1, t_2) = 0$ if $\sum_i z_i^{(1)2}(t_1) = 0$ or $\sum_i z_i^{(2)2}(t_2) = 0$.

We conduct computer simulations in T steps for a given input signal, and evaluate similarity indices using the following parameters: $N = 1000$, $T = 1000$, $\tau = 100$, and w_{ij} were given under binomial distribution $\Pr(w_{ij} = 0) = \Pr(w_{ij} = 2\kappa/N) = 0.5$, where $\kappa = 2.0$. In our previous paper (Yamazaki & Tanaka, 2005a), we studied how the quality of a spatiotemporal activity pattern depends on these parameters, and the same parameters described above were used as the default setting in the paper except κ.

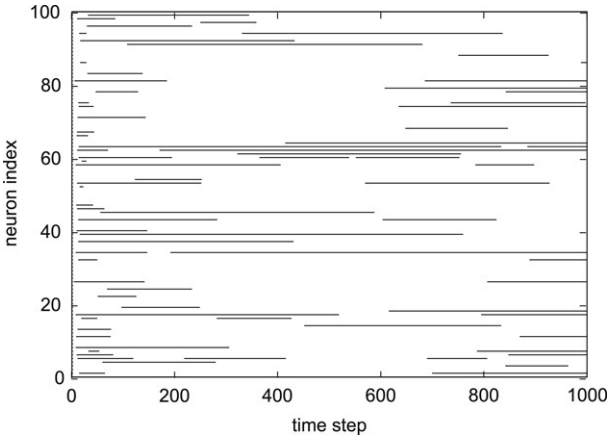

Fig. 2. Raster plot of active states ($z_i(t) > 0$) of 100 neurons that were chosen randomly from 500 neurons which received nonzero input signals. Neurons exhibit random repetition of transitions between active (denoted by horizontal solid lines) and inactive states. Different neurons show different activity patterns.

3. Results

3.1. Generation of spatiotemporal activity pattern by recurrent inhibitory network

First of all, we examined the basic dynamics of the recurrent inhibitory network by setting $K = 8$ and $\mathbf{x} = (0, 0, 0, 0, 0, 0, 0, 1)$, where one-half of neurons (=500) randomly chosen received input signals and the value I_i was $2/8 = 0.25$.

Fig. 2 shows state transitions of 100 neurons that were randomly chosen out of 500 neurons which received nonzero input signals. At $t = 0$, all neurons were active because $z_i(0) = I_i$. Then, neurons started to exhibit random repetition of transitions between active and inactive states. The number of active neurons rapidly converged to a constant value of about 70. That is, only about 14% of neurons were active at each time step. This indicates that an active neuronal population at each time step is sparse. The possible number of active neuron populations given by $\binom{0.5N}{0.07N}$ is found to be of the exponential order of N ($O(1.22^N/\sqrt{N})$). We also found that the total activity of neurons ($\sum_i z_i(t)$) rapidly converges to a constant value.

The left panel in Fig. 3 shows the similarity index between active neuron populations at t_1 and t_2 in the form of a $T \times T$ matrix, where the row and the column are specified by t_1 and t_2, respectively. Each value of the index was plotted in a grey scale in which black indicates 0 and white 1. Because the index at identical steps ($t_2 = t_1$) takes one for any given t_1, the diagonal line appeared white. A broad white band around the diagonal line indicates that the index decreased monotonically as the

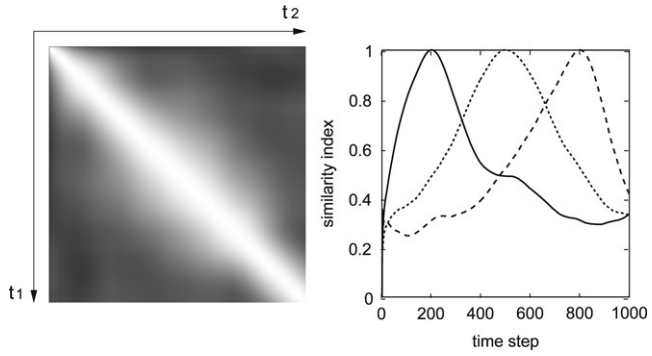

Fig. 3. Left: Plot of similarity index ($C(t_1, t_2)$) in a grey scale calculated using Eq. (9). Black indicates 0 and white 1. The index at the identical time step ($t_1 = t_2$) takes 1, and decreases as the difference between t_1 and t_2 increases. This is observed by the appearance of the white diagonal band. Right: Plots of $C(200, t)$ (solid line), $C(500, t)$ (dotted line) and $C(800, t)$ (dashed line), namely, the 200th, 500th and 800th rows in the matrix. The similarity index decreases almost monotonically as time elapses.

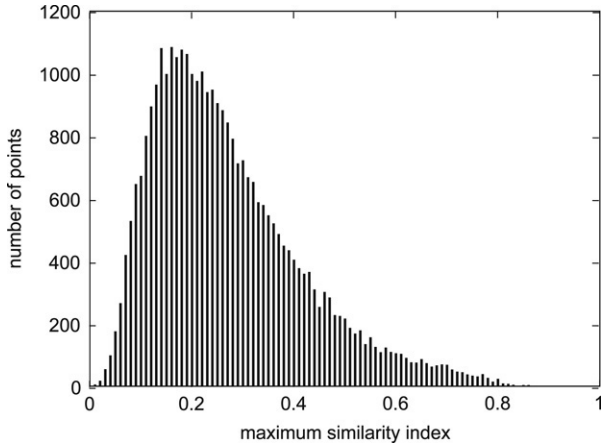

Fig. 5. Histogram of maximum values of similarity indices. The histogram plots the number of pairs of activity patterns with respect to the maximum value of the similarity index for the pair. For most pairs the similarity index takes the value of 0.2, and for very few pairs the index exceeds 0.8. Bin size is 0.01.

interval between t_1 and t_2 became longer. This is confirmed as shown in the right panel of Fig. 3, where $C(200, t)$, $C(500, t)$ and $C(800, t)$ are values of similarity indices along the 200th, 500th and 800th rows in the matrix. These features of the similarity index illustrated that the activity pattern of neurons changed gradually with time and did not recur for at least 1000 steps. We have examined longer time steps up to 100 000 steps ($T = 100\,000$) and observed that a sequence of active neuron populations was generated without recurrence (Yamazaki & Tanaka, 2005a). This implies a one-to-one correspondence between the active neuron population and time since signal input onset. Therefore, a dynamically changing population of active neurons is able to represent the passage of time in the rate-coding scheme.

3.2. Generation of different activity patterns for different input signals

Next, we tested if different input signals can be discriminated by different activity patterns. First, we conducted simulations for two different input signal patterns: $\mathbf{x}_1 = (0, 0, 0, 0, 0, 0, 0, 1)$ and $\mathbf{x}_2 = (0, 0, 0, 0, 0, 0, 1, 0)$. Let $z_i^{(1)}(t)$ and $z_i^{(2)}(t)$ be the activity patterns generated by $\mathbf{x}_1$ and $\mathbf{x}_2$, respectively. The left and middle panels in Fig. 4 show the similarity indices for input signal patterns $\mathbf{x}_1$ and $\mathbf{x}_2$, respectively. The passage of time is represented successfully in both cases as indicated by the appearance of the white diagonal band. The right panel in Fig. 4 shows the similarity index between $z_i^{(1)}(t)$ and $z_i^{(2)}(t)$. Although one fourth of the simulated neurons likely participated in the representation of time commonly for input signals $\mathbf{x}_1$ and $\mathbf{x}_2$, the values of the similarity index were small, where the maximum value was 0.34. This result indicates that the two activity patterns are independent of each other.

Then, we investigated how many input signal patterns are separately represented by spatiotemporal activity patterns of neurons in the model granular layer. For each pattern of K bit binary input signal patterns, we calculated the spatiotemporal activity pattern. Then, for all pairs of 2^K activity patterns, we calculated $2^K(2^K - 1)/2$ similarity indices and obtained the same number of maximum values for respective indices. We used 8 bit binary input signal patterns ($K = 8$), which was the upper limit due to the limitation of computational resources for calculating $2^K(2^K - 1)/2$ similarity indices. We discarded the case of $\mathbf{x} = (0, 0, 0, 0, 0, 0, 0, 0)$ that does not activate any neurons.

The distribution of $(255 \times 254)/2 (=32\,385)$ maximum values of the similarity indices is shown in Fig. 5. The largest value of the similarity index was 0.87, which was obtained by the activity pattern at $t = t_1$ with the input

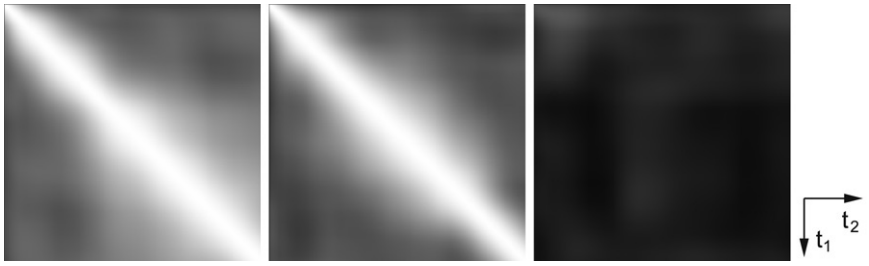

Fig. 4. Left and middle: Similarity indices for first and second simulations calculated using Eq. (9). In both cases, the white diagonal band appears in the matrix, indicating the successful generation of a time-varying neuron population. Right: Similarity index between two activity patterns in the simulations calculated using Eq. (10). The matrix looks totally black, suggesting that the two activity patterns are mutually uncorrelated. In all panels, the values of indices are represented in a grey scale. Black indicates 0 and white 1.

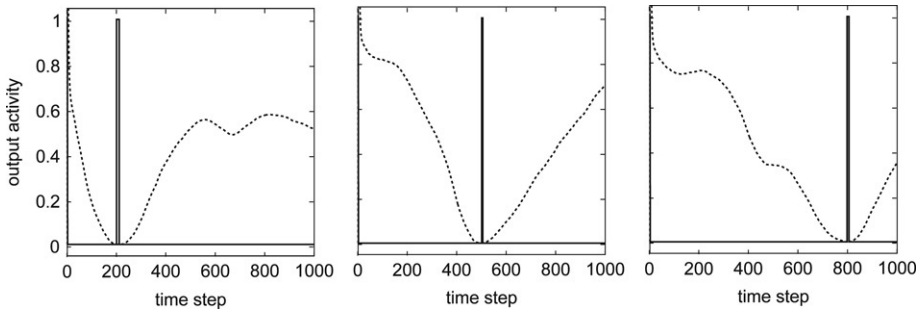

Fig. 6. Plots of $r(t)$ (dotted line) and $\hat{o}(t)$ (solid line) for representation of function defined by Eq. (11), in which t_* is set at 200 (left), 500 (middle) and 800 (right). The value of $r(t)$ decreases as t increases and reaches 0 at $t \approx t_*$. Then the value gradually increases. The value of $\hat{o}(t)$ is 1 at $t \approx t_*$ and 0 otherwise.

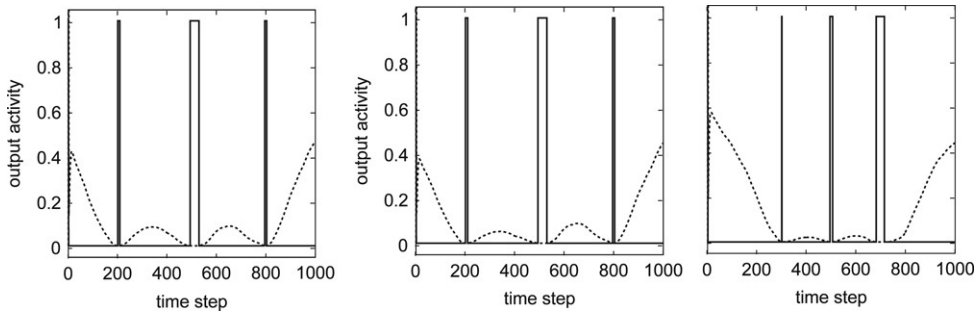

Fig. 7. Plots of $r(t)$ (dotted line) and $\hat{o}(t)$ (solid line) when the function defined by Eq. (13) is learned (left). Likewise, $r(t)$ (dotted line) and $\hat{o}(t)$ (solid line) for input signals $\mathbf{x}_1$ (middle) and $\mathbf{x}_2$ (right) when the function defined by Eq. (14) is learned. $r(t)$ converges to 0 at trained time steps while $\hat{o}(t)$ takes the value of 1 at the same time steps.

$\mathbf{x} = \mathbf{x}_1$ and that at $t = t_2$ with the input $\mathbf{x} = \mathbf{x}_2$, where $\mathbf{x}_1 = (1, 0, 0, 1, 0, 1, 1, 1)$ (151 in decimal), $\mathbf{x}_2 = (1, 1, 0, 1, 0, 1, 1, 1)$ (215 in decimal), $t_1 = 600$ and $t_2 = 449$. There were 8.5% of the pairs of activity patterns whose maximum values of the similarity index exceeded 0.5; only 0.15% of the pairs (49 pairs) exceeded 0.8. The peak of the distribution was located at 0.2. This distribution demonstrated that spatiotemporal populations of neurons evoked by different binary input signal patterns are sufficiently different. Based on the above analyses, it turns out that a recurrent inhibitory network with fixed random connections generates nonrecurrent spatiotemporal activity patterns which are different for different input signal patterns. A common feature between this inhibitory recurrent network and a recently proposed liquid state machine is the absence of fixed point attractors in the state space. Because the absence of fixed point attractors is a necessary condition of a liquid state machine, we are able to regard the network as a liquid state generator, and a spatiotemporal activity pattern as a liquid state. Therefore, the granular layer in the cerebellum can be a biological counterpart of a liquid state generator.

3.3. Learning and representation of Boolean functions by Purkinje cells and output neurons

Next, we examined the dynamics of model Purkinje cells and output neurons defined by Eqs. (2) and (3). Without loss of generality, we considered here a single Purkinje cell and a single output neuron ($M = 1$), because we assumed that there were no interactions among Purkinje cells or among output neurons. We denote $r_1(t)$, $e_1(t)$ and $\hat{o}_1(t)$ by simply $r(t)$, $e(t)$ and $\hat{o}(t)$, respectively.

First, we studied the temporal discriminability. We tested if the network can learn the following function:

$$f\left(\overbrace{0, \ldots, 0, 1}^{K}, t\right) = \begin{cases} 1 & t = t_* \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

To this end, we set the instruction signal $e(t)$ as follows:

$$e(t) = \begin{cases} 1 & t = t_* \\ 0 & \text{otherwise,} \end{cases} \tag{12}$$

and set J_{ij} according to Eq. (5).

The three panels in Fig. 6 show the changes of $r(t)$ and $\hat{o}(t)$ for $t_* = 200$ (left), 500 (middle) and 800 (right). For all cases, $r(t)$ became 0 at $t \approx t_*$. When we set $\theta = 1$, we obtained $\hat{o}(t) = 1$ at $t \approx t_*$ and $\hat{o}(t) = 0$ otherwise. This indicates that the function was successfully learned. We note that after T steps has elapsed, $r(t)$ converged to a constant level and stayed there, whereas $\hat{o}(t)$ was kept to be 0.

The model was also able to learn the following function that represents multiple timings.

$$f\left(\overbrace{0, \ldots, 0, 1}^{K}, t\right) = \begin{cases} 1 & t \in \{200, 500, 800\} \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

The left panel in Fig. 7 shows $r(t)$ and $\hat{o}(t)$ when this function was learned. $r(t)$ became 0 at $t \approx 200$, 500 and 800, and thereby $\hat{o}(t) = 1$ at these time steps whereas $\hat{o}(t) = 0$ at the other time

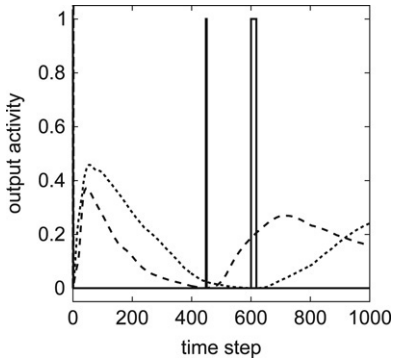

Fig. 8. Plots of $r(t)$ when either $\mathbf{x}_1$ is given (dotted line) or $\mathbf{x}_2$ is given (dashed line) after the function defined by Eq. (15) is learned. Plots of $\hat{o}(t)$ are also shown in solid lines.

steps. In this case, $\hat{o}(t)$ became 1 at 200 and 800 only for a few steps, whereas $\hat{o}(t)$ became 1 for 40 time steps at $t = 500$.

Next we attempted to embed the following function that switches different output signals representing multiple timings for different input signal patterns fed into the network:

$$f(\mathbf{x}, t) = \begin{cases} 1 & \mathbf{x} = \mathbf{x}_1, t \in \{200, 500, 800\}, \\ 1 & \mathbf{x} = \mathbf{x}_2, t \in \{300, 500, 700\}, \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

where $\mathbf{x}_1$ and $\mathbf{x}_2$ are the same as those described in Section 3.2. The middle and right panels in Fig. 7 show $r(t)$ and $\hat{o}(t)$ for $\mathbf{x} = \mathbf{x}_1$ and $\mathbf{x} = \mathbf{x}_2$, respectively. In both cases, because $r(t)$ was 0 at the desired timings, $\hat{o}(t)$ became 1 at the same timings and thereby able to represent the pertinent function. This result indicates that the network is able to learn to represent different timings for different input patterns.

We then studied whether the two activity patterns, which yielded the largest similarity index as shown in Section 3.2, are discriminable. We tested the following function for learning:

$$f(\mathbf{x}, t) = \begin{cases} 1 & \mathbf{x} = (1, 0, 0, 1, 0, 1, 1, 1), t = 600, \\ 1 & \mathbf{x} = (1, 1, 0, 1, 0, 1, 1, 1), t = 449, \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

Fig. 8 shows the plot of $r(t)$ for an input signal pattern $\mathbf{x}_1$ or $\mathbf{x}_2$. When $\mathbf{x}_1$ was given, $r(t)$ became 0 at $t = 600$. On the other hand, when $\mathbf{x}_2$ was given, $r(t)$ became 0 at $t = 449$. Therefore, even when the similarity index is high, the discriminability of activity patterns by output neurons is still preserved.

A simple perceptron is well known to have a computational limitation. For example, the exclusive-OR (XOR) function is a binary function that is not linearly separable and hence not computable by simple perceptrons. Using our network model, it is expected that any Boolean function is represented by a binary signal output from the network at a specific time step. We attempted to embed Boolean functions of two binary input variables, such as AND, OR and XOR, into the network. Specifically, for given $\mathbf{x}_1 = (0, 0, 0, 0, 0, 0, 0, 1)$, $\mathbf{x}_2 = (0, 0, 0, 0, 0, 0, 1, 0)$ and $\mathbf{x}_3 = (0, 0, 0, 0, 0, 0, 1, 1)$, we attempted to operate AND, OR and XOR to the lowest two bits of the input patterns at different time steps as follows:

$$f(\mathbf{x}, t) = \begin{cases} 1 & \mathbf{x} = \mathbf{x}_3, t = 200 \text{ (AND)}, \\ 1 & \mathbf{x} \in \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}, t = 500 \text{ (OR)}, \\ 1 & \mathbf{x} \in \{\mathbf{x}_1, \mathbf{x}_2\}, t = 800 \text{ (XOR)}, \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

The left, middle and right panels in Fig. 9 show output patterns of these three Boolean functions for input patterns $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$, respectively. The outputs of AND, OR and XOR were successfully represented 0 or 1 at $t = 200$, 500 and 800, respectively. This result indicates that our network can represent a nonlinear Boolean function which cannot be computed by simple perceptrons. Moreover, multiple Boolean functions were able to be embedded into the identical network, and the computations were automatically performed in the desired order and timings specified by instruction signals. This automatic computation in the time dimension may be important for coordinating movements of different body parts into a harmoniously integrated body movements. For example, when we produce a phoneme, we need to move a number of muscles around the supralaryngeal vocal tract smoothly and harmoniously at the correct order and timing. Despite the precise control of muscles is required for articulation, we do not feel to attend to individual muscle activities. Any phoneme seems to be produced from its phonetic image automatically.

Taking the results so far together, model Purkinje cells receive the spatiotemporal activity pattern and learn to become inactive at desired timings using instruction signals fed into the climbing fibre. When we reinterpret our cerebellar network model in terms of the liquid state machine, the Purkinje cells correspond to readout neurons, in the sense that readout neurons receive liquid states and have the learning capability of desired input–output relationships.

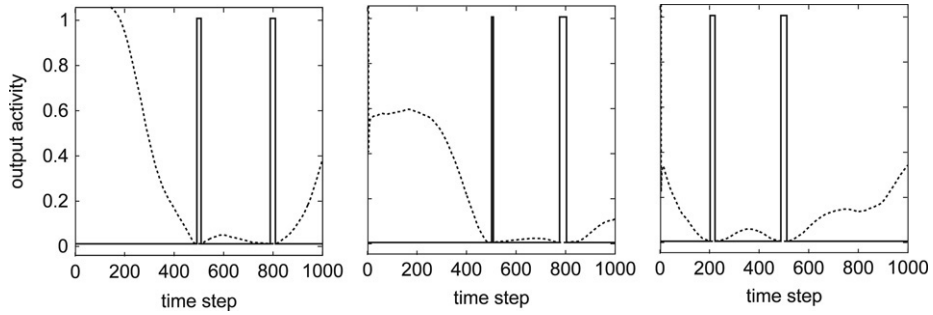

Fig. 9. Plots of $r(t)$ (dotted line) and $\hat{o}(t)$ (solid line) for $\mathbf{x}_1$ (left), $\mathbf{x}_2$ (middle) and $\mathbf{x}_3$ (right) when the function defined by Eq. (16) is learned.

4. Discussion

4.1. The cerebellar cortex as a liquid state machine

In this paper, we studied the computational power of our model cerebellar cortex, which is composed of the model granular layer and the model Purkinje cells. The model granular layer is a random recurrent inhibitory network that generate a spatiotemporal activity pattern of model granule cells when a static input pattern is given through mossy fibres by the model precerebellar nucleus. The recurrent circuit is able to generate different spatiotemporal activity patterns for different static input patterns. Specifically, for 255 static patterns composed of 8 bit binary inputs, the network could generate 255 spatiotemporal activity patterns which do not correlate from each other.

On the other hand, the model Purkinje cells receive the spatiotemporal activity pattern of model granule cells and instruction signals through the climbing fibre. They utilize a stateless learning rule that simulates LTD, in which the synaptic weight of active granule cells at the timing of the arrival of instruction signals is set at 0. By this mechanism, they learn to become inactive at desired timings, and the release from Purkinje cells' inhibition makes output neurons elicit correctly timed binary output signals. We demonstrated that the model cerebellar circuit is able to embed all the 2-bit logic operations.

The organization of the model cerebellar cortex looks similar to that of a liquid state machine (Maass et al., 2002). A liquid state machine is composed of two layers. One layer is usually a recurrent network that generates a spatiotemporal activity pattern of neurons called a "liquid state". The network has to generate different liquid states for different input signals, because the performance of a liquid state machine mostly relies on the quality of the activity patterns. The other layer consists of neurons called "readouts". They receive the liquid state and instruction signals, and learn desired input–output relationships using a stateless learning rule. Thus, if we correspond the model granular layer and the model Purkinje cells respectively to the liquid state generator and the readout neurons, we may regard the model cerebellar cortex as a liquid state machine.

Although several studies have proposed models of the cerebellum that have a recurrent network as a granular layer (Buonomano & Mauk, 1994; Hofstötter et al., 2002; Medina et al., 2000; Yamazaki & Tanaka, 2005a), they did not clarify the computational capability of their models nor the similarity between the cerebellar circuitry and the organization of a liquid state machine. Moreover, the view of the granular layer as a recurrent network is relatively new, whereas the classical perceptron hypothesis seems still popular. This study is the first work which makes a clear connection between the cerebellum and the liquid state machine paradigm, and our liquid state machine hypothesis is expected to supersede the classical perceptron hypothesis.

4.2. Discrimination versus generalization

Information processing is principally to perform discrimination and generalization of input signals, but they are mutually conflicting computational concepts. Discrimination means that different inputs should be transformed to more different outputs, whereas generalization means that similar inputs should be transformed to more similar outputs. A liquid state machine is suitable for discrimination because liquid state generation is sensitive to input bit patterns. This implies in turn that the liquid state machine is not adequate for generalization. Then, how does the cerebellum keep the balance between discrimination and generalization in its action for information processing?

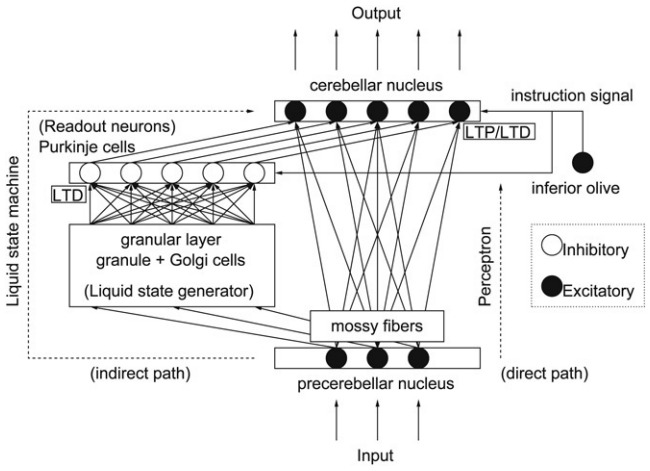

Fig. 10. Hypothetical computational view of cerebellar circuit. The cerebellar circuit is composed of two independent circuits. One is the cerebellar cortex modelled as a liquid state machine, and the other is the precerebellar–cerebellar nuclear circuit that can be regarded as a simple perceptron. Thus, the cerebellum is a hybrid of a liquid state machine and a perceptron.

The cerebellar circuit may be interpreted as the hybrid of two circuits as shown in Fig. 10. One is an indirect pathway circuit represented by a liquid state machine that we have argued so far, and the other is a direct pathway circuit from the precerebellar nucleus to the cerebellar nucleus, which is a two-layer neural network with learnable mossy fibre connections (Hansel, Linden, & D'Angelo, 2001) and thus regarded as a simple perceptron. Both circuits receive instruction signals from the inferior olive which supervise the learning of desired input–output relationship. This view implies that temporally nonspecific information is stored in the perceptron, which is suitable for generalization. On the other hand, the liquid state machine opens or closes gates for the flow of the information in a temporally specific manner. Boyden, Katoh, and Raymond (2004) have suggested that the cerebellar cortex stores timing-related information, whereas the cerebellar nucleus is more important for storing the amplitude of the response. Shutoh, Ohki, Kitazawa, Itohara, and Nagao (2006) have confirmed that memory of learned responses is first formed in the cerebellar cortex and memory only about the response amplitude is transferred to the nuclei and gradually consolidated there. Thus, in the cerebellum, the hybrid construction of a liquid state machine and a perceptron may compromise mutually antagonistic computations of discrimination and generalization.

4.3. Extension to a realistic model

In the present model, we intended to include only the minimal set of components of the cerebellar cortex, which are

necessary to discuss the similarity in structure and computation between the cerebellar cortex and a liquid state machine, because an objective of this paper is to advocate a new picture on the cerebellum as a biological counterpart to a liquid state machine. The simple abstract model discussed here was sufficient for the purpose. In order to keep the simplicity and theoretical clarity, we did not use a realistic spike-based model. This simplistic approach suggests that the liquid state generator is a good picture of the recurrent inhibitory network model of the granular layer in the cerebellum, which was built based on the rate-coding scheme with linearly rectified neuronal responses. It also suggests that Purkinje cells correspond to the readout neurons in terms of the liquid state machine.

Finally, even though such a simple model predicts interesting computational properties of the granular layer and Purkinje cells, it may be that the predicted properties disappear when the model is extended to a realistic network model composed of neurons eliciting spikes with stochastic fluctuation in the timing of spike generation. Using our realistic model of the cerebellum (Yamazaki & Tanaka, 2005b), in which the network structure was carefully constructed on the basis of anatomical and physiological data, we have demonstrated that the nonrecurrent dynamics and capability of time representation were reproduced. Furthermore, we have found that our model cerebellum was robust against Poisson spike input signals, and that the robustness was improved by increasing the number of granule cells. In spite of the simplicity of the present model, the basic dynamic features found in the realistic cerebellar circuit model were retained.

In conclusion, the picture of the cerebellum as a liquid state machine is a promising viewpoint for obtaining a better understanding of cerebellar computational principles.

References

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

Boyden, E. S., Katoh, A., & Raymond, J. L. (2004). Cerebellum-dependent learning: The role of multiple plasticity mechanisms. *Annual Review of Neuroscience*, *27*, 581–609.

Buonomano, D. V., & Mauk, M. D. (1994). Neural network model of the cerebellum: Temporal discrimination and the timing of motor responses. *Neural Computation*, *6*, 38–55.

Eccles, J. C., Ito, M., & Szentágothai, J. (1967). *The cerebellum as a neuronal machine*. New York: Springer-Verlag.

Hansel, C., Linden, D. J., & D'Angelo, E. (2001). Beyond parallel fiber LTD: The diversity of synaptic and non-synaptic plasticity in the cerebellum. *Nature Neuroscience*, *4*(5), 467–475.

Haykin, S. (1999). *Neural networks: A comprehensive foundation* (2nd ed.). Prentice-Hall.

Hofstötter, C., Mitz, M., & Verschure, P. F. M. J. (2002). The cerebellum in action: A simulation and robotics study. *European Journal of Neuroscience*, *16*(2), 1361–1376.

Ito, M. (1984). *The cerebellum and neuronal control*. New York: Raven Press.

Ito, M. (1989). Long-term depression. *Annual Review of Neuroscience*, *12*, 85–102.

Ito, M. (2002a). Historical review of the significance of the cerebellum and the role of Purkinje cells in motor learning. *Annals of the New York Academy of Sciences*, *978*, 273–288.

Ito, M. (2002b). The molecular organization of cerebellar long-term depression. *Nature Reviews: Neuroscience*, *3*(11), 896–902.

Maass, W., Natschläger, T., & Markram, H. (2002). Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Computation*, *14*(11), 2531–2560.

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

Medina, J. F., Garcia, K. S., Nores, W. L., Taylor, N. M., & Mauk, M. D. (2000). Timing mechanisms in the cerebellum: Testing predictions of a large-scale computer simulation. *The Journal of Neuroscience*, *20*(14), 5516–5525.

Mugnaini, E., Atluri, R. L., & Houk, J. C. (1974). Fine structure of granular layer in turtle cerebellum with emphasis on large glomeruli. *Journal of Neurophysiology*, *37*, 1–29.

Rosenblatt, M. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*, 386–408.

Shutoh, F., Ohki, M., Kitazawa, H., Itohara, S., & Nagao, S. (2006). Memory trace of motor learning shifts transsynaptically from cerebellar cortex to nuclei for consolidation. *Neuroscience*, *139*(2), 767–777.

Yamazaki, T., & Tanaka, S. (2005a). Neural modeling of an internal clock. *Neural Computation*, *17*(5), 1032–1058.

Yamazaki, T., & Tanaka, S. (2005b). Realistic modeling of the cerebellum, program No. 413.6. 2005. Abstract Viewer/Itinerary Planner. Washington, DC: Society for Neuroscience.