Computational Models of Timing Mechanisms in the Cerebellar Granular Layer

Tadashi Yamazaki · Shigeru Tanaka

Abstract A long-standing question in neuroscience is how the brain controls movement that requires precisely timed muscle activations. Studies using Pavlovian delay eyeblink conditioning provide good insight into this question. In delay eyeblink conditioning, which is believed to involve the cerebellum, a subject learns an interstimulus interval (ISI) between the onsets of a conditioned stimulus (CS) such as a tone and an unconditioned stimulus such as an airpuff to the eye. After a conditioning phase, the subject's eyes automatically close or blink when the ISI time has passed after CS onset. This timing information is thought to be represented in some way in the cerebellum. Several computational models of the cerebellum have been proposed to explain the mechanisms of time representation, and they commonly point to the granular layer network. This article will review these computational models and discuss the possible computational power of the cerebellum.

T. Yamazaki
Laboratory for Motor Learning Control,
RIKEN Brain Science Institute,
2-1 Hirosawa,
Wako, Saitama 351-0198, Japan

S. Tanaka (✉)
Laboratory for Visual Neurocomputing,
RIKEN Brain Science Institute,
2-1 Hirosawa,
Wako, Saitama 351-0198, Japan
e-mail: shigeru@riken.jp

Present Address:
T. Yamazaki
Strategic Planning Unit, RIKEN BSI-TOYOTA Collaboration
Center, RIKEN Brain Science Institute,
2-1 Hirosawa,
Wako, Saitama 351-0198, Japan

Introduction

Motor control consists of the control of two distinct quantities: gain (the amplitude of movement) and timing (the onset of movement). The cerebellum, playing a central role in motor control [1], is involved in control of both the gain and timing [2]. Cerebellar mechanisms of timing control have been studied in depth using an experimental paradigm of Pavlovian delay eyeblink conditioning ([3–6] for review), in which a subject is exposed to paired presentation of a sustained tone (conditioned stimulus or CS) and an airpuff (unconditioned stimulus or US) that induces an eyeblink reflex. After the conditioning by repeated CS–US presentation with a fixed interstimulus interval (ISI) between the tone (CS) and airpuff (US) onsets, the subject learns to close the eyes with an appropriate delay after the CS onset but prior to the US onset (conditioned response or CR), indicating that the ISI between CS and US onsets is memorized. It is reported [3] that the most reliable CR is elicited when ISI is set at 250 ms. Longer ISIs result in less reliable CR generation [3]. Figure 1a illustrates a schematic of the cerebellar circuit with a flow diagram of neural signals of the CS, US, and CR. The neural signal of CS is conveyed through mossy fibers (MFs) from the precerebellar nucleus (PN) to the cerebellar nucleus (CN, the site that generates the motor command for an eyeblink) and the granule cells. Simultaneously, the CN is inhibited by Purkinje cells in the cerebellar cortex. Before conditioning, the neural activity in CN is constant in response to the CS, because both the
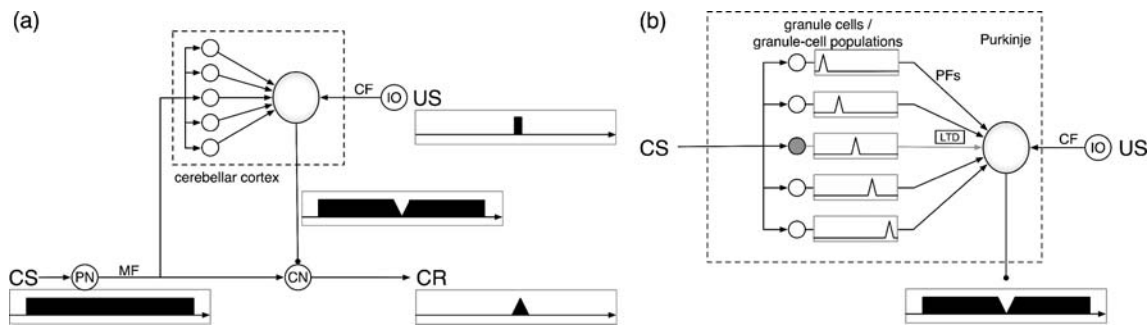
Fig. 1. Schematic diagram of the cerebellar circuit involved in Pavlovian delay eyeblink conditioning. **a** The flow diagram of neural signals representing CS, US and CR. Neural signals of CS and US are fed into the PN and IO, respectively. Purkinje cells (*large circle*) in the cerebellar cortex receive CS signals through granule cells (*small circles*) in the granular layer and inhibit the CN tonically during the CS presentation before conditioning. After the conditioning, they stop firing in advance of the US onset, which disinhibits the CN and enables the CN to elicit spikes representing the CR. *Dotted boxes* illustrate the temporal patterns of CS and US signals, Purkinje cells' activity, and CN neurons' activity. **b** A hypothetical mechanism for Purkinje cells to learn the time to stop firing starting slightly earlier than the US onset. When the CS is given, granule cells or granule-cell populations (*small circles*) become active one by one sequentially. *Dotted boxes* illustrate the activities of these cells/cell populations. At the US onset, the granule cell/granule-cell population shown by the gray circle becomes active, and by conjunctive activation with the CF, the PF synapses of this cell/cell population are depressed by LTD. After the conditioning, the active granule cell/cell population at the US onset cannot transmit the activity to the Purkinje cell due to LTD, thereby the net excitatory drives to the Purkinje cell are decreased around the US onset during the CS presentation. Thus, the Purkinje cell stops firing around the US onset. Abbreviations: *CF* climbing fiber, *CN* cerebellar nucleus, *CR* conditioned response, *CS* conditioned stimulus, *IO* inferior olive, *LTD* long-term depression, *PF* parallel fiber, *PN* precerebellar nucleus, *US* unconditioned stimulus

CS signal and Purkinje cells' firing are constant during CS presentation. After conditioning, Purkinje cells learn to pause starting slightly earlier than the US onset. This pause disinhibits the CN so that the neural activity in CN increases transiently around the US onset. This transient increase of the neural activity in CN induces the CR.

How do Purkinje cells learn the timing of the cessation of firing? Here, we assume that the CS in delay eyeblink conditioning evokes temporally constant neural activity in PN and does not contain any temporal information. Therefore, our question is how the ISI from the CS onset to the US onset is represented in the cerebellar cortex. A working hypothesis is that any time measured from the CS onset is represented by the sequential activation of granule cells or granule-cell populations: there should be one-to-one correspondence between the passage-of-time (POT) from the CS onset and a granule cells' temporal activation pattern. A specific ISI is determined by the cessation of firing activities of some Purkinje cells that do not receive inputs from granule cells that are active at that timing. On the basis of this hypothesis, the mechanism by which Purkinje cells learn the time to stop firing is explained as illustrated in Fig. 1b. At the onset of CS presentation, a sequence of active granule cells or granule-cell populations starts. Let us assume that the active granule cell or the population of active granule cells at the US onset is uniquely determined. At the US onset, the activity of the inferior olive (IO) conveyed by the climbing fiber (CF) induces strong depolarization in a Purkinje cell which receives at the same time signals from the active granule cell/cell population through parallel fibers (PFs). The

conjunctive excitation of PF and CF induces long-term depression (LTD) in those PF–Purkinje cell synapses. Only the PF–Purkinje cell synapses activated by the granule cell/cell population at the US onset are depressed and the other synapses are unaffected. Because the active granule cell/cell population changes gradually with time, the net excitatory drives to the Purkinje cell starts to decrease in advance of the US onset and becomes the minimum at the US onset, which results in the cessation of firing of the cell starting slightly earlier than the US onset. Therefore, the most important aspect of computational modeling is how the granular layer generates sequential activities of granule cells/cell populations without recurrence.

We have classified the current computational models of POT representation in the cerebellar granular layer into four types: (1) delay line [7–10], (2) spectral timing [11], (3) oscillator [12, 13], and (4) random projection [14–17]. In the following section, we will review and evaluate each type of model separately.

Models of the Cerebellar Granular Layer for POT Representation

Delay Line Model

The delay line model implements the sequential activation of neurons in quite a literal manner. Desmond and Moore [7] and Moore et al. [8] posited that a "delay line" is constructed by the sequential linkage of neurons in the precerebellar nuclei such as pontine nuclei, an input stage

to feed signals to the granular layer (Fig. 2a). Owing to this anatomical organization, granule cells receiving afferent inputs from these nuclear cells also become active sequentially during the CS presentation. No such anatomical organization in the precerebellar nuclei has been found thus far, however. Another delay line model assumes large conduction delays at parallel fibers, rather than at mossy fibers or pre-mossy fibers [9, 10]. The conduction velocity and the length of parallel fibers were found to be 0.24 m/s [18] and 4–5 mm [19], respectively. Using these values, we can estimate the maximum conduction delay to be about 21–25 ms. Because delay eyeblink conditioning occurs for

ISIs set maximally at 2–3 s [3], the estimated maximum conduction delay is evidently too short to represent POTs for delay eyeblink conditioning. Thus, it is unlikely that the delay line model sufficiently accounts for POT representation in the cerebellum.

Spectral Timing Model

The spectral timing model proposed by Bullock et al. [11] is based on the assumption of a wide distribution of membrane time constants for different granule cells. This assumption leads to various delays for granule cells to be
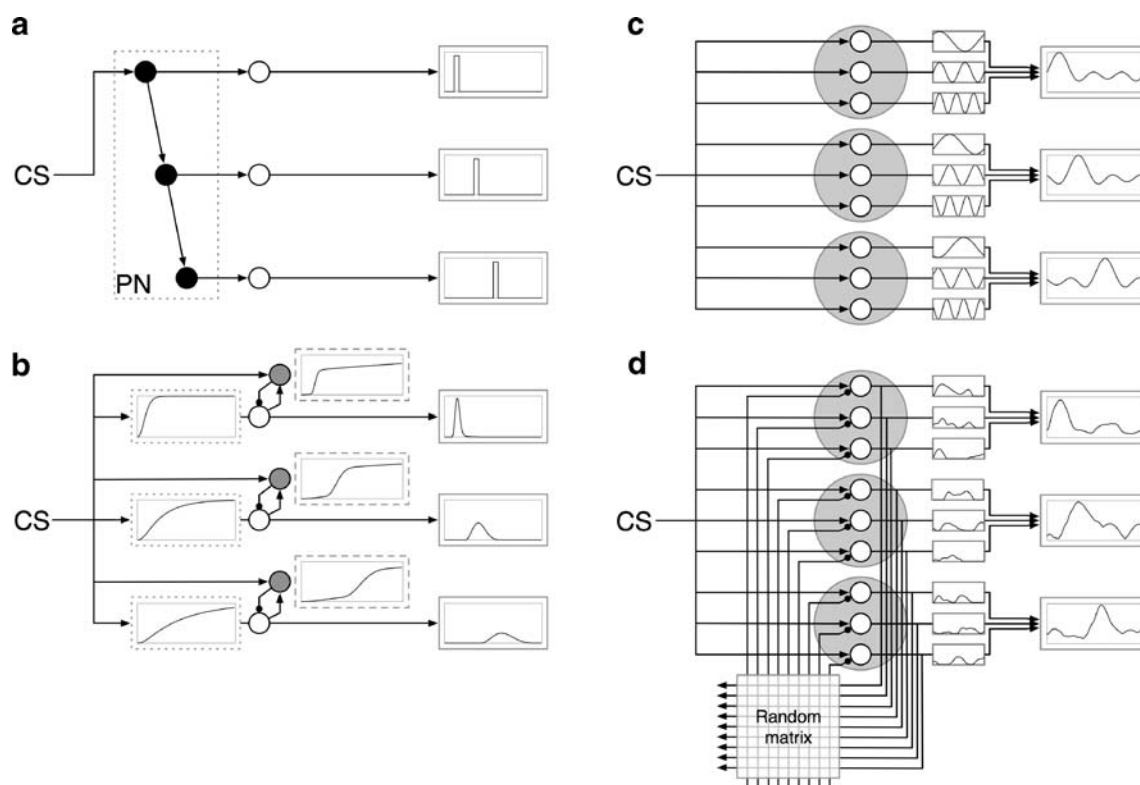

Fig. 2. Schematics of the four main models. **a** Delay line model. Neurons in the precerebellar nuclei (*PN*, *black circles in the dotted box*) are assumed to be connected sequentially so as to construct a "delay line". They become active one by one sequentially in response to a CS. Because connections from PN neurons to granule cells (*white circles*) are one-to-one, granule cells also become active one by one sequentially in response to a CS (*right solid boxes*). **b** Spectral timing model. Granule cells (*white circles*) are assumed to have a variety of membrane time constants so that they become active with various delays (*dotted boxes*). Each granule cell is also assumed to have a companion Golgi cell (*gray circles*), which inhibits the corresponding granule cell after its activation (*dashed boxes*). Because of the delayed inhibition, the activity pattern of granule cells becomes transient. Hence, granule cells become active transiently with various delays from the CS onset (*solid boxes*). Notice that the peak activity of granule cells decreases as the delay increases. **c** Oscillator model. Granule cells are assumed to become active in an oscillating manner, with different frequencies and time lags (*small solid boxes*). *Large*

shaded circles represent populations of granule cells that have common activity peaks at certain times. For each population, the total activity of the cells exhibits a peak at a time that is characteristic of the population (*large solid boxes*). As a result, these populations of granule cells, not individual granule cells, become active sequentially in response to CS. **d** Random projection model. Granule cells' activities are fed back via a random matrix which represents the recurrent connections from granule to Golgi cells and Golgi to granule cells. The elements of the matrix represent effective inhibitory synaptic weights via Golgi cells. This feedback inhibition produces randomly repetitive transitions of active and inactive states of the cells (*small solid boxes*). *Large shaded circles* represent populations of granule cells that are commonly active at a certain time. For each population, the total activity of the cells exhibits a peak at the time characteristic of the population (*large solid boxes*). Therefore, these populations of granule cells become active sequentially in response to a CS

activated. Once a granule cell becomes active, the cell is inhibited by a companion Golgi cell as long as the CS signal is sustained. Consequently, the granule cell becomes active only once during the CS presentation, which results in a spectrum of activity peaks of granule cells throughout the CS–US interval (Fig. 2b). In the spectral timing model, to generate a variety of spectral activity patterns, the membrane time constant of granule cells has to vary widely up to a few seconds to cover the range of ISIs (2–3 s) representable in delay eyeblink conditioning [3], which seems biologically implausible. In addition, there is no experimental evidence that a precise, one-to-one connectivity pattern exists between granule cells and Golgi cells.

Oscillator Model

The above-mentioned models are based on the sequential activation of individual granule cells. In contrast, the following models rely on the sequential activation of granule-cell populations. In the oscillator model (Fig. 2c), granule cells are assumed to behave as oscillators with different frequencies and phases in response to a CS. Although these individual cells become active repeatedly during the CS presentation, the summed activity over granule cells within some cluster can be localized at a certain time; this is guaranteed by mathematics, stating that any localized function can be given uniquely by the sum of Fourier components with various frequencies. Thus, the population of these cells can appear only once during the CS presentation. Such an active granule-cell population may be found at each time step, which constructs a sequence of populations of active granule cells instead of the sequential activation of individual granule cells.

An oscillator model called the "phase encoder model" developed by Fujita [20] was originally proposed as a model for the adaptation of the eye-movement reflex in the vestibulo-ocular reflex. In this model, individual granule cells become oscillators with a fundamental frequency specified by MF signals representing the head rotation velocity and with different time lags (phases). This may be possible, as shown in Fig. 2c, when granule cells oscillate with a slow frequency. Thus, granule cells become active one by one, resulting in the sequential activation. The time lags are developed by the randomness of MF-granule cell synaptic weights and the temporal integration of the feedback inhibition of granule cells via a Golgi cell.

Gluck et al. [12] extended Fujita's idea by describing granule cells as oscillators with different frequencies and time lags in response to temporally constant CS signals, and applied their model to delay eyeblink conditioning, without explaining the mechanism of how granule cell activities behave as oscillators with different frequencies over the broad range. They also assumed that PF–Purkinje cell

synaptic weights are complex values to represent the information on time lag. More recently, Garenne and Chauvet [13] suggested that MF-granule cell synapses show a wide range of efficacy, enabling granule cells to fire from 1 to 100 Hz. In oscillator models, the lowest firing rate among granule cells determines the longest representable ISI: their model represents the POT up to 1 s. To generate a variety of time lags, their model assumed a wide distribution of MF-granule synaptic efficacies. However, if the synaptic efficacy changes, different sequences of active granule-cell populations can be generated for the same mossy fiber signal, which in turn leads to unstable time representation. Indeed, a recent experimental study showed that MF-granule cell synapses are highly plastic [21]. Thus, their model is unlikely to represent the POT robustly.

Random Projection Model

Buonomano and Mauk [14] have proposed a POT representation model using a sequence of populations of active granule cells in a different way from that described in the oscillator model, and the same group later elaborated on their model [15]. In this model (Fig. 2d), granule cells repeatedly undergo random transitions between active and inactive states during the CS presentation. Because of the dynamics of a granular layer network with random recurrent connections, the same population of active granule cells does not appear more than once. This enables a population of active granule cells to encode a unique time after the CS onset. Therefore, POT is represented by a sequence of populations of active granule cells. Buonomano and Mauk argued for the first time that this dynamic, aperiodic activity pattern of granule cells emerges from the granule–Golgi–granule-negative feedback with settings of realistic network structure and parameters [14].

Yamazaki and Tanaka [16] paid a special attention to the importance of randomness in the recurrent network for the generation of a sequence of active granule-cell populations, and proposed the following toy model, which describes the essential dynamics of Buonomano and Mauk's model:

$$z_i(t+1) = \left[I_i(t) - \sum_{s=0}^{t} e^{-\frac{t-s}{\tau}} \sum_{j=1}^{N} w_{ij} z_j(s) \right]^+,$$

where $z_i(t)$ denotes the activity of granule cell i at time t, $I_i(t)$ the MF input signal at time t, τ the time constant of Golgi cells. The granule cell activity is rectified by the function $[.]^+$, where $[x]^+ = x$ if $x > 0$ and 0 otherwise. The second term in the brackets of the right-hand side represents the temporally integrated effective inhibition by other granule cells via Golgi cells, where τ represents the range of the temporal integration and set to be large. w_{ij} represents the effective recurrent connection from granule cell j to i,
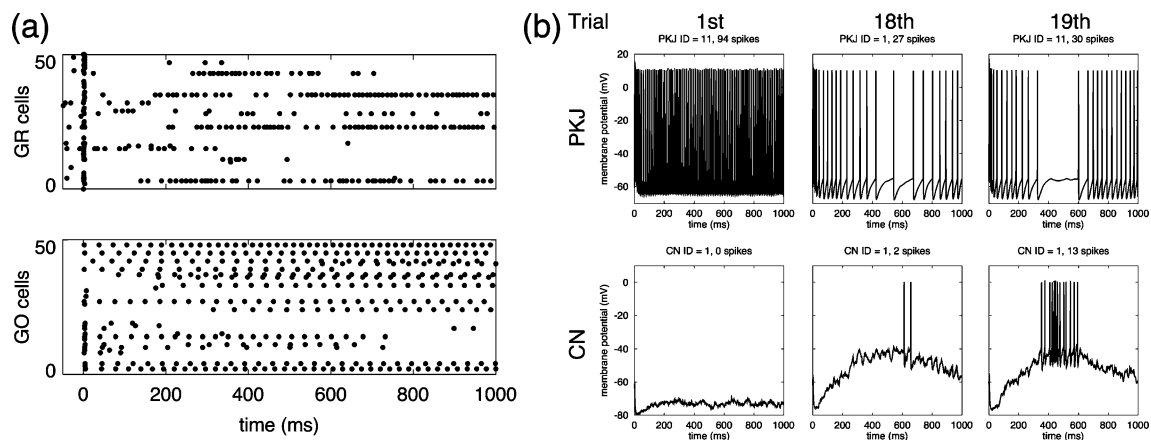
Fig. 3. Dynamics of the elaborated random projection model in which spiking model neurons are implemented [18]. **a** Spike patterns of 50 model granule cells (*top*) and 50 model Golgi cells (*bottom*) during CS presentation in the elaborated model. *Abscissa* and *ordinate* represent time from the CS onset and neuron index, respectively. A *dot* represents a single spike event. Granule cells undergo random repetition of transitions between sustained burst spiking states and silent states. Different granule cells exhibit different temporal spiking patterns. Thus, the population of spiking granule cells changes gradually with time, so that a sequence of populations of spiking granule cells is generated.

Golgi cells elicit spikes rather regularly. **b** Membrane potentials of a model Purkinje cell (*top*) and a model CN neuron (*bottom*) at the first, 18th, and 19th trials of simulated delay eyeblink conditioning (*left to right*). For each panel, the *abscissa* and *ordinate* respectively represent the time from CS onset and the membrane potential. The US onset was set at 500 ms after the CS onset. With training, the Purkinje cell learns to stop firing around the US onset. Consequently, the CN neuron is released from the Purkinje cell's tonic inhibition to elicit spikes representing the CR. Reprinted from [18]

which gives an element of the random matrix of connectivity. The biological counterpart of τ has been later interpreted as a long decay constant of NMDAR-mediated EPSPs at Golgi cells [17]. The polysynaptic circuits of granule–Golgi–granule cells implicitly correspond to the random inhibitory recurrent connections among granule cells in our model.

The essence of our model is to project a population of active granule cells to another population of active granule cells by the matrix transformation, where the matrix elements represent effective inhibitory connection weights via Golgi cells (Fig. 2d). This "random projection[1]" is repeated during CS presentation, and thereby a random sequence of populations of active granule cells is generated. Hereafter, we call this type of model the random projection model. The assumed random recurrent connections are biologically supported by the presence of a recurrent inhibitory network of granule cells via Golgi cells. Furthermore, the generation of the random sequence is reproducible across trials: when the same CS is presented, the same sequence is generated [16].

We later extended our toy model to build a large-scale spiking network model of the cerebellum [17]. Figure 3a represents spike patterns of 50 out of 102,400 model

granule cells and 50 out of 1,024 model Golgi cells during CS presentation. Granule cells undergo repetitive transitions between burst spiking states and silent states, and different cells exhibit different temporal spike patterns. These state transitions occur deterministically, but the patterns of spiking granule cells are random because of the random recurrent connections. The population of spiking granule cells gradually changes with time, and the same spiking granule-cell populations do not appear more than once during the CS presentation. These observations demonstrate that a sequence of populations of active granule cells is generated even in the elaborated spiking network model. Figure 3b shows the membrane potentials of a model Purkinje cell and a model CN neuron at the first, 18th, and 19th trials of the simulated delay eyeblink conditioning, where the US onset was set at 500 ms after the CS onset. The Purkinje cell, therefore, learns to stop firing 100 ms in advance to the US with training, whereas the CN neuron is released from the Purkinje cells' inhibition to elicit spikes around 500 ms after the CS onset as the CR. This result seems inconsistent with experimental findings: Purkinje cell pause starts as early as 250 ms before the US onset [23]. However, another choice of parameter values (e.g., longer time window of LTD) and/or addition of accessory circuits (e.g., circuits composed of molecular layer interneurons that inhibit Purkinje cells) to our model will fit the experimentally observed timing more closely. We demonstrated that ISIs were represented reliably up to 750 ms. The maximal ISI for reliable CRs is determined

---

[1] Random projection is the name of an algorithm that maps a vector to another vector via a random matrix [22]. Random projection is usually used to reduce dimensionality in data, whereas this term is used here to map a population of active granule cells to another population without changing the dimension.

by the reproducibility of the temporal sequence of active granule-cell populations across trials. This issue will be explored using the more elaborated spiking network model in our future studies. The successful application of the random projection model to delay eyeblink conditioning mentioned above suggests that "random projection" provides one of the most important neural bases for cerebellar function.

The Marr–Albus–Ito theory advocates that granule cells act as a spatial pattern discriminator that represents MF with a sparse code [24–26]. This idea has been denied by a theoretical study in which a large-scale granular layer model exhibited memory capacity much smaller than expected [27]. Nevertheless, the random projection model not only supports the hypothesis of spatial discrimination by granule cells in the Marr–Albus–Ito theory but also creates a novel concept of spatiotemporal discrimination as an information processing capability of the cerebellum [28].

Granular Layer Models for Other Functions

So far, we have introduced granular layer models that transform temporally constant MF input signals into a sequence of active granule cells or granule-cell populations that represents POT. Here, we briefly introduce granular layer models for other functions.

Three models have been proposed, in which the granular layer works as decorrelation filters that perform principal component analysis [29], independent component analysis [30], and maximization of information transfer between MFs and granule cells [31]. The decorrelation filter extracts a few meaningful components from a number of noisy and redundant signals. Such decorrelation filter models acquire sparse representation of MF signals, as envisioned by Marr [24] and Albus [26]. These models generate time-varying granule cell activity in response to time-varying MF signals, but they do not generate time-varying granule cell activity in response to constant MF signals. Therefore, as long as we assume that the CS does not contain any temporal information, these models may not be able to account for delay eyeblink conditioning.

Related to this issue, Freeman and Muckler [32] have reported that neurons in the pontine nuclei sending MF inputs to the cerebellum exhibit three different temporal discharge patterns. Phasic neurons elicit spikes transiently in response to the CS onset. Such phasic responses may have a function of resetting the temporal sequences of granule cells' activities [18]. Sustained and late neurons elicit spikes tonically during the CS presentation. Sustained neurons keep their firing rates relatively constant, whereas late neurons increase their firing rates gradually. Thus, it is possible that MF inputs as a whole convey temporally modulated signals. However, the representation of POT

requires the precise and robust relationship between the timing and temporal discharge patterns. Therefore, the sustained and late discharge patterns of the pontine neurons should have fine temporal structure and be reproducible across different trials of CS presentation. Although Figs. 4, 5, and 6 in their paper [32] illustrate temporal modulation in the firing rate of sustained and late neurons during the CS presentation, it is unclear whether the temporal modulation is reproducible across trials or simply irreproducible noise. On the other hand, Aitkin and Boyd [33] have also reported that some neurons in cats' lateral pontine nuclei elicit spikes transiently in response to the onset of the acoustic stimulation and some other neurons elicit spikes tonically during the tone presentation. The tonic discharge gradually decreases with time, and exhibits little temporal modulation. This study suggests that the fine temporal modulation in discharge patterns, if any, is just noise. Thus, the presence of three types of neurons in the pontine nuclei may not contribute to the POT representation, but further experiments are needed for clarifying the possibility that pontine neurons represent POT during CS presentation.

Another issue is that the decorrelation filter models require learning of connection weights between granule and Golgi cells for realizing sparse coding. The learning changes the activity pattern of these cells gradually across repeated trials. On the other hand, a recent experimental study demonstrated that the activity pattern of Golgi cells does not change across 300 trials of saccade adaptation [34], suggesting that learning for sparse coding may not take place in the granular layer.

Extra Granular Layer Models for POT Representation

We may also need brief introduction of some other models that proposed possible mechanisms of POT representation in the outside of the granular layer. Fiala et al. have proposed a Purkinje cell model based on the slow process of intracellular signal transduction mediated by the metabotropic glutamate receptor (mGluR) [35]. The variation in the number of mGluRs expressed on the dendrites for different Purkinje cells results in different latencies in the elevation of intracellular Ca^{2+} concentration for those cells, producing a spectrum of intracellular Ca^{2+} transients across different Purkinje cells. The transient increase of Ca^{2+} causes inward flow of Ca^{2+}-dependent K^+ currents, and the channel conductance was increased by PF–CF pairing. Thereby, after repeated pairings of stimulation to PFs and CFs, a particular Purkinje cell learns to pause at a particular ISI. Fiala et al. assumed that the amount of expressed mGluRs was constant. On the other hand, Steuber and Willshaw assumed that the amount of mGluRs adaptively changes by learning [36]. PF–CF pairing stimulation adjusted

Ca^{2+} response latency to match a particular ISI, which causes the timed pause of the Purkinje cell. On the contrary, Schreurs and colleagues have reported in a series of experiments [37–39] that K^+ currents including Ca^{2+}-dependent K^+ current rather decreased during delay eyeblink conditioning, which resulted in the increase of excitability of the Purkinje cell. To address this contradiction, Hong and Optican have recently proposed another Purkinje cell model coupled with stellate cells [40]. This model is composed of many pairs of a Purkinje cell and a stellate cell. Both Purkinje and stellate cells increase their excitability by PF–CF pairing stimulation. However, the increase of stellate cells' excitability is faster and stronger than that of Purkinje cells, and thereby inhibition from stellate cells to Purkinje cells becomes to dominate excitation from PFs, resulting in the timed pause. This model is another version of spectral timing models: in this model, the temporal spectrum is generated by many pairs of Purkinje and stellate cells, whereas in Bullock and Grossberg's model, the spectrum is generated by many pairs of granule and Golgi cells.

Comparison of Random Projection Model with Experiments

We have computationally demonstrated that the involvement of NMDAR-mediated EPSPs with a long decay time is an important requisite for granule cells to exhibit randomly repetitive transition between burst and silent states during persistent sensory stimulation [18]. To our knowledge, thus far, there has not been enough number of studies of *in vivo* granule cell recording to find experimental evidence supporting the random projection model. Unfortunately, some studies [41–44] used ketamine for anesthesia, which is a blocker of NMDARs ([45] and references therein). Two studies [41, 42] used very brief sensory stimulation that sustains for 50 ms. In delay eyeblink conditioning, the ISI must be longer than 100 ms to generate robust CRs (cf. Fig. 1D of [3]), suggesting that the sensory stimulation should be at least twice longer to observe spatiotemporal dynamics of granule cell activity. We also confirmed that using a brief CS shorter than 100 ms, our cerebellar model failed to learn to elicit anticipatory CRs (unpublished observation).

Svensson and Ivarsson [46] have reported that short-lasting CSs elicited CRs after acquisition training with sustained CSs, suggesting that animals somehow learned to bridge the off-stimulus trace interval as well as the conditioning. Their experimental paradigm seems to obey trace eyeblink conditioning rather than delay eyeblink conditioning, in which the hippocampus and/or the prefrontal cortex are regarded to play more important role than the cerebellum. However, the hippocampus and the pre-frontal cortex did not work because their animals were decerebrated. These authors interpreted their findings by hypothesizing that the cerebellar cortex contains the neural substrate for keeping some information on CS signals even after the cessation of the CS presentation. They justified the hypothesis by the observation of Larson-Prior et al. [47] that granule cells in slice preparations sustained firing activities in response to MF stimulation hundreds of milliseconds after the offset of the stimulation. Kotani et al. [48] have also investigated the trace conditioning in decerebrate guinea pigs. They reported that even decerebrate animals can acquire and express eyeblink conditioning in a trace paradigm. They discussed that the feedback loop from the interpositus nuclei to the pontine nuclei retains the activity during the stimulus-free trace interval. These two studies suggest that decerebrate animals can elicit CRs in response to transient CSs, at least after the training in a delay paradigm. Therefore, they challenge the classical hypothesis on the neural substrates of trace eyeblink conditioning, in which the prefrontal cortex and/or hippocampus sustains the CS information during the off-stimulus period and transmit the activity to the pontine nuclei [49]. Based on these studies, it is expected that if animals are trained in both delay and trace paradigms alternately session by session, they should exhibit intact expression of CRs in both paradigms after sufficient training, even if their hippocampus or prefrontal cortex are inactivated. Nevertheless, Kalmbach et al. [50] have performed such experiments, and obtained results against the expectation: when the prefrontal cortex was inactivated after the conditioning, CRs were disrupted in the trace paradigm whereas CRs was intact in the delay paradigm. This suggests that the prefrontal cortex plays a role in bridging the stimulus-free trace interval, thereby supporting the classical hypothesis. Therefore, we may still be far from complete identification of neural substrates for delay and trace conditionings.

For the functional role of MF signals, there is another experiment performed by Jörntell and Ekerot [51] using awake cats. The top-left and bottom-right panels of Fig. 3c in their paper show an irregular temporal modulation of granule cell activities during stimulation, which seems to support our random projection model. However, the authors suggested that granule cells simply work to enhance the signal-to-noise ratio, because the temporal patterns of granule cell spike activity appeared to follow the activity in the presynaptic mossy fibers. This interpretation is regarded to derive from their PSTHs with a large bin size. Therefore, to clarify whether the random projection model is biologically plausible, it is desired that single-unit recordings of cerebellar granule cells in awake animals will be performed with a higher temporal resolution during persistent sensory stimulation.

It is noteworthy that there are experimental studies to explore the activity of Golgi cells rather than granule cells ([52] and references therein) to elucidate the granular layer dynamics. These studies are performed on the basis of the idea that the spatiotemporal activity pattern of granule cells is shaped by time-varying inhibition from Golgi cells [53].

A New Trend in Neural Computation

The random projection model is a member of liquid-state machines [54] or echo-state networks [55], which has been recently proposed to provide a new paradigm of modern neural computation (see [56, 57] for recent work). A liquid-state machine consists of a large random recurrent network called a reservoir that maps input signals into a higher dimensional space, and a set of neurons called readouts that receive inputs from reservoir neurons to extract time-varying information. A learning rule adopted by the simple perceptron model is applied to a supervised learning of readouts. The network architecture and the learning rule of the liquid-state machine are much simpler than those of conventional recurrent networks, yet the computational power is shown to be versatile. A major advantage of the liquid-state machine over the conventional recurrent networks is their fast learning [58]. Furthermore, there are a variety of real-world applications using liquid-state machines including speech perception [59], robot control [60], and financial prediction [61], which indicates the strong computational power of liquid state machines. Our random projection model is mathematically equivalent to the liquid-state machine, when the granular layer, Purkinje cells, and LTD of PF–Purkinje cell synapses, respectively, are replaced with the reservoir, readouts, and the learning rule. This suggests that the cerebellar cortex, in so far as it can be represented by the random projection model, possesses a versatile computational power [29].

Owing to its huge computational power, the random projection model is capable of constructing internal models that are believed to exist in the cerebellum [62]. Therefore, we believe that the random projection model is a good candidate computational model of the cerebellum.

Conclusion

In this article, we reviewed four types of computational models of the cerebellar granular layer for POT representation: delay line, spectral timing, oscillator, and random projection. We also briefly introduced granular layer models for other functions and extra granular layer models

for POT representation. We pointed out the similarity between the random projection model and the liquid state machine, and thereby suggesting that the random projection model possesses a versatile computational power. Furthermore, we argued what experiments are needed to clarify whether the random projection model is feasible for POT representation in the cerebellum.

Models presented in this article have been implemented in C and C++ independently of the original research. The source codes are available at the Cerebellar Platform [63]. Simulation of these models can be carried out online at the Cerebellar Simulator [64].

References

1. Ito M (1984) The cerebellum and neuronal control. Raven, New York
2. Ivry RB, Spencer RMC (2004) The neural representation of time. Curr Opin Neurobiol 14:225–232
3. Mauk MD, Donegan NH (1997) A model of Pavlovian eyelid conditioning based on the synaptic organization of the cerebellum. Learn Mem 3:130–158
4. Christian KM, Thompson RF (2003) Neural substrates of eyeblink conditioning: acquisition and retention. Learn Mem 11:427–455
5. Hesslow G, Yeo CH (2002) The functional anatomy of skeletal conditioning. In: Moore JW (ed) A neuroscientist's guide to classical conditioning. Springer, New York, pp 86–146
6. De Zeeuw CI, Yeo CH (2005) Time and tide in cerebellar memory formation. Curr Opin Neurobiol 15:667–674
7. Desmond JE, Moore JW (1988) Adaptive timing in neural networks: the conditioned response. Biol Cybern 58:405–415
8. Moore JW, Desmond JE, Berthier NE (1989) Adaptively timed conditioned responses and the cerebellum: a neural network approach. Biol Cybern 62:17–28
9. Chapeau-Blondeau F, Chauvet G (1991) A neural network model of the cerebellar cortex performing dynamic associations. Biol Cybern 65:267–279
10. Braitenberg V, Heck D, Sultan F (1997) The detection and generation of sequences as a key to cerebellar function: experiments and theory. Behav Brain Sci 20:229–277
11. Bullock D, Fiala JC, Grossberg S (1994) A neural model of timed response learning in the cerebellum. Neural Netw 7:1101–1114

12. Gluck MA, Reifsnider ES, Thompson RF (1990) Adaptive signal processing and the cerebellum: models of classical conditioning and VOR adaptation. In: Gluck MA, Rumelhart DE (eds) Neuroscience and connectionist theory. Erlbaum, Hillsdale, New Jersey, pp 131–186

13. Garenne A, Chauvet GA (2004) A discrete approach for a model of temporal learning by the cerebellum: *in silico* classical conditioning of the eyeblink reflex. J Int Neurosci 3: 301–318

14. Buonomano DV, Mauk MD (1994) Neural network model of the cerebellum: temporal discrimination and the timing of motor responses. Neural Comput 6:38–55

15. Medina JF, Garcia KS, Nores WL, Taylor NM, Mauk MD (2000) Timing mechanisms in the cerebellum: testing predictions of a large-scale computer simulation. J Neurosci 20:5516–5525

16. Yamazaki T, Tanaka S (2005) Neural modeling of an internal clock. Neural Comput 17:1032–1058

17. Yamazaki T, Tanaka S (2007) A spiking network model for passage-of-time representation in the cerebellum. Eur J Neurosci 26:2279–2292

18. Vranesic I, Iijima T, Ichikawa M, Matsumoto G, Knöpfel T (1994) Signal transmission in the parallel fiber–Purkinje cell system visualized by high-resolution imaging. Proc Natl Acad Sci, USA 91:13014–13017

19. Pichitpornchai C, Rawson JA, Rees S (1994) Morphology of parallel fibres in the cerebellar cortex of the rat: an experimental light and electron microscopic study with biocytin. J Comp Neurol 342:206–220

20. Fujita M (1982) Adaptive filter model of the cerebellum. Biol Cybern 45:195–206

21. Hansel C, Linden DJ, D'Angelo E (2001) Beyond parallel fiber LTD: the diversity of synaptic and non-synaptic plasticity in the cerebellum. Nat Neurosci 4:467–475

22. Vempala SS. The random projection method, volume 65. Am Math Soc, 2004.

23. Jirenhed D, Bengtsson F, Hesslow G (2007) Acquisition, extinction, and reacquisition of a cerebellar cortical memory trace. J Neurosci 27:2493–2502

24. Marr D (1969) A theory of cerebellar cortex. J Physiol (Lond) 202:437–470

25. Ito M (1970) Neurophysiological aspects of the cerebellar motor control system. Int J Neurol 7:162–176

26. Albus JS (1971) A theory of cerebellar function. Math Biosci 10:25–61

27. Tyrrell T, Willshaw D (1992) Cerebellar cortex: its simulation and the relevance of Marr's theory. Phil Trans R Soc Lond B 336:239–257

28. Yamazaki T, Tanaka S (2007) The cerebellum as a liquid state machine. Neural Netw 20:290–297

29. Jonker HJJ, Coolent ACC, van der Gon JJD (1998) Autonomous development of decorrelation filters in neural networks with recurrent inhibition. Network 9:345–362

30. Philipona D, Coenen OJMD (2004) Model of granular layer encoding in the cerebellum. Neurocomp 58-60:575–580

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

32. Freeman JH, Muckler AS (2003) Developmental changes in eyeblink conditioning and neuronal activity in the pontine nuclei. Learn Mem 10:337–345

33. Aitkin LM, Boyd J (1978) Acoustic input to the lateral pontine nuclei. Hear Res 1:67–77

34. Prsa M, Dash S, Cats N, Dicke PW, Their P (2008) Characteristics of responses of Golgi cells and mossy fibers to eye saccades and saccadic adaptation recorded from the posterior vermis of the cerebellum. J Neurosci 29:250–262

35. Fiala JC, Bullock D, Grossbers S (1996) Metabotropic glutamate receptor activation in cerebellar Purkinje cells as substrate for adaptive timing of the classically conditioned eye-blink response. J Neurosci 16:3760–3774

36. Steuber V, Willshaw D (2004) A biophysical model of synaptic delay learning and temporal pattern recognition in a cerebellar Purkinje cell. J Comput Neurosci 17:149–64

37. Schreurs BG, Gusev PA, Tomsic D, Alkon DL, Shi T (1998) Intracellular correlates of acquisition and long-term memory of classical conditioning in Purkinje cell dendrites in slices of rabbit cerebellar HVI. J Neurosci 18:5498–5507

38. Freeman JH, Shi T, Schreurs BG (1998) Pairing-specific long-term depression prevented by blockade of PKC or intracellular Ca2+. Neuro Report 9:2237–2241

39. Freeman JH, Scharenberg AM, Olds JL, Schreurs BG (1998) Classical conditioning increases membrane-bound protein kinase C in rabbit cerebellum. NeuroReport 9:2669–2673

40. Hong S, Optican LM (2008) Interaction between Purkinje cells and inhibitory interneurons may create adjustable output waveforms to generate timed cerebellar output. PLoS ONE 3:e2770

41. Chadderton P, Margrie TW, Häusser M (2004) Integration of quanta in cerebellar granule cells during sensory processing. Nature 428:856–860

42. Rancz EA, Ishikawa T, Duguid I, Chadderton P, Mahon S, Häusser M (2007) High-fidelity transmission of sensory information by single cerebellar mossy fibre boutons. Nature 450:1245–1249

43. Barmack NH, Yakhnitsa V (2008) Functions of interneurons in mouse cerebellum. J Neurosci 28:1140–52

44. Arenz A, Silver RA, Schaefer AT, Margrie TW (2007) The contribution of single synapses to sensory representation *in vivo*. Science 321:977–980

45. Bengtsson F, Jörntell H (2007) Ketamine and xylazine depress sensory-evoked parallel fiber and climbing fiber responses. J Neurophysiol 98:1697–1705

46. Svensson P, Ivarsson M (1999) Short-lasting conditioned stimulus applied to the middle cerebellar peduncle elicits delayed conditioned eye blink responses in the decerebrate ferret. Eur J Neurosci 11:4333–4340

47. Larson-Prior LJ, Morrison PD, Bushey RM, Slater NT (1995) Frequency dependent activation of a slow N-methyl-D-aspartate-dependent excitatory postsynaptic potential in turtle cerebellum by mossy fibre afferents. Neurosci 67:867–879

48. Kotani S, Kawahara S, Kirino Y (2003) Trace eyeblink conditioning in decerebrate guinea pigs. Eur J Neurosci 17: 1445–1454

49. Woodruff-Pak DS, Disterhoft JF (2008) Where is the trace in trace conditioning? Trends Neurosci 31:105–12

50. Kalmbach BE, Ohyama T, Kreider JC, Riusech F, Mauk MD (2009) Interactions between prefrontal cortex and cerebellum revealed by trace eyelid conditioning. Learn Mem 16:86–95

51. Jörntell H, Ekerot CF (2006) Properties of somatosensory synaptic integration in cerebellar granule cells *in vivo*. J Neurosci 26: 11786–97

52. D'Angelo E, De Zeeuw CI (2009) Timing and plasticity in the cerebellum: focus on the granular layer. Trends Neurosci 32:30–40

53. Maex R, De Schutter E (1998) Synchronization of Golgi and granule cell firing in a detailed network model of the cerebellar granule cell layer. J Neurophysiol 80:2521–2537

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

55. Jaeger H, Haas H (2004) Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. Science 304:78–80

56. Jaeger H, Maass W, Principe J, editors. Echo state networks and liquid state machines, volume 20 of Neural Netw, 2007.

57. Buonomano DV, Maass W (2009) State-dependent computations: spatiotemporal processing in cortical networks. Nature Rev Neurosci 10:113–125

58. Jaeger H. Tutorial on training recurrent neural networks, covering BPPT, RTRL, EKF and the echo state network approach. GMD report 159, 2002.

59. Skowronski MD, Harris JG (2007) Automatic speech recognition using a predictive echo state network classifier. Neural Netw 20:414–423

60. Antonelo EA, Schrauwen B, Van Campenhout J (2007) Generative modeling of autonomous robots and their environments using reservoir computing. Neural Process Lett 26:233–249

61. Neural Forecasting Competition 3. http://www.neural-forecasting-competition.com/NN3/

62. Wolpert DM, Miall RC, Kawato M (1998) Internal models in the cerebellum. Trends Cogn Sci 2:338–347

63. Cerebellar platform. http://platform.cerebellum.neuroinf.jp/.

64. Cerebellar simulator. http://capsule.brain.riken.jp/~cerebellum/model.cgi.