

Chapter 19

The Input-Output Organization of the Cerebrocerebellum as Kalman Filter



Shinji Kakei, Hirokazu Tanaka, Takahiro Ishikawa, Saeka Tomatsu, and Jongho Lee

19.1 Recognition of the Cerebrocerebellum as Loci of Internal Models

The seminal publication of the book on the cerebellum by Eccles et al. (1967) inspired the publication of major theories of the cerebellum and motor control by David Marr (1969), James S. Albus (1971), and Masao Ito (1970). Their legendary papers marked the beginning of the ongoing effort to understand the relationship between the cerebellar neuron circuitry and motor/cognitive control. In particular, Ito (1970) first proposed how the cerebrocerebellum participates in acquiring skilled movements in terms of a control system model (Fig. 19.1a). In voluntary *unskilled* movements (Fig. 19.1a(1)), the initial instruction arising from the association cortex (AC) is transferred to the motor cortex (MX) and transformed into the motor command and relayed to the spinal motor system (SM) through the pyramidal tract (PT). The outcome of the motor command is evaluated by AC using information relayed by the sensory

S. Kakei (✉)

Department of Anatomy and Physiology, Faculty of Life Sciences, Jissen Women's University, Tokyo, Japan

e-mail: kakei-shinji@jissen.ac.jp

H. Tanaka

Faculty of Information Technology, Tokyo City University, Tokyo, Japan

T. Ishikawa

Laboratory for Movement Disorders, Tokyo Metropolitan Institute of Medical Science, Tokyo, Japan

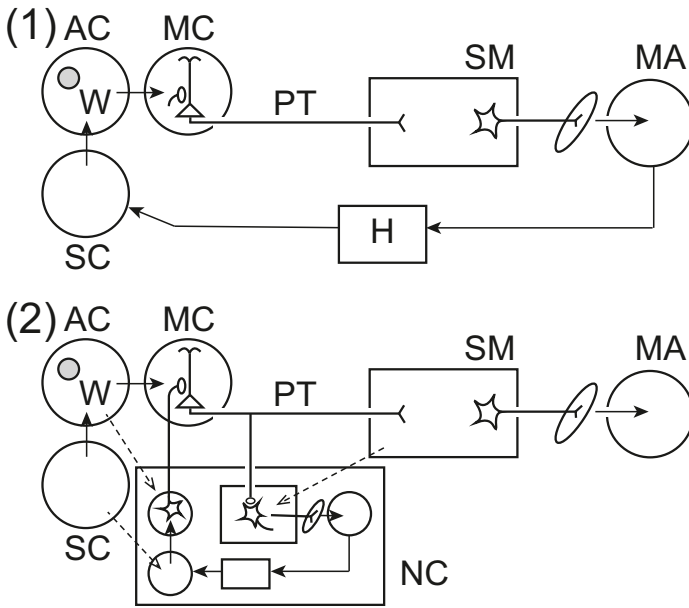
S. Tomatsu

Division of Behavioral Development, National Institute of Physiological Sciences, Okazaki, Aichi, Japan

J. Lee

Faculty of Health Sciences, Komatsu University, Komatsu, Ishikawa, Japan

A. Ito (1970)



B. Allen & Tsukahara (1974)

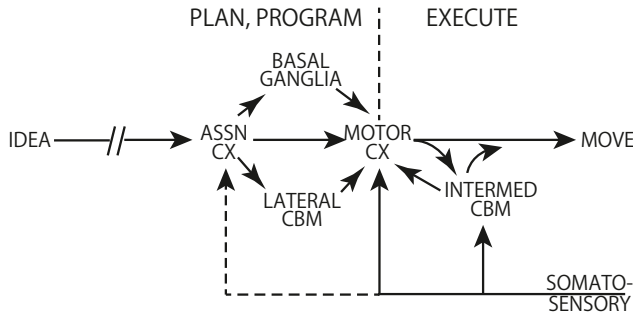


Fig. 19.1 Two types of cerebellar internal models. Model by Ito (1970) (a) is consistent with a forward model, while the model by Allen and Tsukahara (1974) (b) is consistent with an inverse model. (a) Diagram of the possible control system for voluntary movements (Reproduced from Ito (1970), Fig. 19.7). Note the caption is also original. (1) Feedback system used in *unskilled* movement. AC cerebral association area, Small gray circle (W) indicates the origin of the *will*. SC cerebral sensory area, MC cerebral motor area, PT pyramidal tract, SM spinal motor system, MA motor activity. H, feedback pathway through the external world. (2) Feedforward system formed after learning. NC neocerebellum in which SM, MA, H, SC, and AC in A are indicated in a minimized form (Note Ito (1970) assumed that AC, SC, SM, MA, and H are *all* modeled in *neocerebellum* (NC) (= *cerebrocerebellum*)). (b) Scheme showing proposed roles of several brain structures in movement (Reproduced from Allen and Tsukahara (1974), Fig. 9). Note the caption is also original.

feedback loops (H + SC in Fig. 19.1a). However, with practice, the movement becomes more skilled and predictive while less dependent on the sensory feedback information. In other words, as the learning progresses, the long loop through the *external* world may be effectively replaced by an *internal* loop passing through the cerebrocerebellum (= *neocerebellum* (NC) in Fig. 19.1a), which would serve as a model of the combination of SM, the external world, and the sensory pathways (Fig. 19.1a(2)). According to Ito (1970), it is possible to understand this arrangement as a type of *model inference adaptive control system*. Cerebellar ataxia, such as dysmetria or intention tremor, could be explained as impairment or loss of *the internal model* in the cerebrocerebellum, just as in the stage of unskilled movements before motor learning.

Note that in Ito's model (Fig. 19.1a(2)), the cerebrocerebellum (NC) receives the efference copy from MX and returns its output to the same MX. Therefore, it may play a role that is equivalent to a *forward model*. A forward model provides the controller (i.e., the motor cortex) with a state prediction (Todorov, 2004) to compensate for sensory feedback delays and stabilize movements. To the best of our knowledge, it was the first proposal of a forward model in neuroscience.

Few years after Ito's pioneering paper, Allen and Tsukahara (1974) proposed a different type of cerebrocerebellar organization (Fig. 19.1b) to explain skilled voluntary movements. They envisaged a two-stage planning-execution system between the cerebral cortex and the cerebellum to control voluntary movements (Fig. 19.1b). The schema features the idea that the association cortices (ASSN CX) translate the intention to move into a proper spatiotemporal activation of the motor cortex (MOTOR CX), resulting in the intended movement. ASSN CX that project to the cerebrocerebellum (LATERAL CBM) are among those in the *premotor* circuit. Because LATERAL CBM appears to lack direct sensory inputs, it is more suited for planning the movement than in actual execution and correction of the movement, which was more suitable for the intermediate zone (INTERMED CBM) function. Once the movement has been prepared in ASSN CX, with the help of LATERAL CBM (i.e., the cerebrocerebellum) (Fig. 19.1b), MOTOR CX generates the motor command as the common node. At this point, INTERMED CBM updates the movement based on the difference (i.e., error) between the actual movement and intended movement.

The two schemes (Fig. 19.1a(2), Ito and b, Allen, and Tsukahara) may look similar. But they are exclusive to each other from a functional point of view. In the former scheme (Fig. 19.1a(2)), the cerebrocerebellum provides a long feedback loop model and virtually replaces it, and the cerebrocerebellum resides *outside of the controller*. In contrast, in the latter scheme (Fig. 19.1b), the cerebrocerebellum is *a part of the controller* that translates the intention to move into the motor command. Thus, the cerebrocerebellum in Fig. 19.1b is suitable to play a role that is equivalent to an *inverse model* or a part of it.

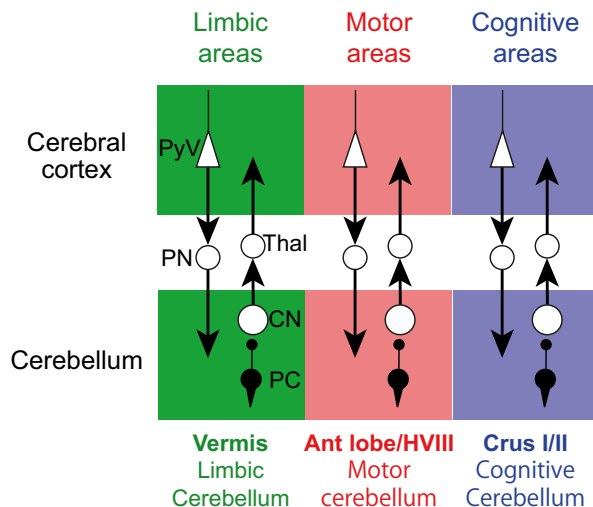
←
Fig. 19.1 (continued) Dashed line represents a pathway of unknown importance. It is proposed that basal ganglia and cerebellar hemisphere are involved with association cortex in programming of volitional movements. At the time that the motor command descends to motoneurons, engaging the movement, the pars intermedia updates the intended movement, based on the motor command and somatosensory description of limb position and velocity on which the movement is to be superimposed. Follow-up correction can be performed by motor cortex when cerebellar hemisphere and pars intermedia do not effectively perform their functions

The two schemes are also exclusive in terms of the neuroanatomical organization between the cerebellum and the cerebral cortex. In the former scheme (Fig. 19.1a(2)), the cerebrocerebellum is connected *reciprocally* with the motor cortex. In contrast, the connectivity in the latter scheme is *non-reciprocal*, collecting its input from the association cortices and returning its output to the motor cortex rather than to the association cortices, the source of the cortical input. Therefore, it is possible to select one from the other by identifying the neuroanatomical organization in theory. Unfortunately, the neuroanatomical techniques available in the 1970s and 1980s, such as the Nauta method or simple neuronal tracers, were not effective enough for this purpose.

19.2 Parallel Organization of the Cerebrocerebellar Communication System Revealed with Transneuronal Tracing Technique

In the 1990s, Peter L. Strick and his colleagues established a revolutionary technique to trace neuron circuitry: a transneuronal tracing technique with neurotropic viruses. It was revolutionary because they enabled the use of transneuronal transport of rabies viruses to reveal the connections of three or more synaptically linked neurons (Kelly & Strick, 2003), which was impossible with conventional neuron tracers. With the new method, analysis of the cerebrocerebellar communication system was within their reach. Their experiments demonstrated that the regions of the cerebellar cortex that receive input from the motor cortex are the same as those that project to the motor cortex. Similarly, the regions of the cerebellar cortex that receive input from area 46 (a part of the prefrontal cortex) are the same as those that project to area 46. Thus, their observations demonstrated that parallel closed-loop circuits represent a fundamental feature of cerebrocerebellar interactions (Fig. 19.2).

Fig. 19.2 The parallel organization of the cerebrocerebellar loops. The cerebellum contains anatomically separate and functionally distinct motor and non-motor domains. This figure was prepared based on Kelly and Strick (2003). CN cerebellar nuclei cells, PC Purkinje cells, PN pontine nuclei cells, PyV layer V pyramidal cells, Thal thalamus



The closed-loop architecture of the cerebrocerebellar communication system is compatible with Ito's closed-loop scheme (Fig. 19.1a(2)), a forward model. But it is not consistent with Allen and Tsukahara's scheme that assumes an open-loop architecture. The cerebellum integrates its various cortical inputs and returns the output to the heteronymous motor cortex. Looking back, Ito's pioneering forward model hypothesis was almost 30 years ahead of his time, but unfortunately, it somehow remained almost unnoticed until very recently.

19.3 Contests Among Control Laws to Explain Movement Trajectories

In the 1980s and 1990s, there was a vigorous debate about putative control laws that govern movement trajectories in reaching movements. The discussion focused on how the central nervous system selects one specific movement trajectory among an infinite number of possible trajectories that lead to the goal. In other words, the competition was over control laws to reduce excess degrees of freedom. Several candidate theories included minimum jerk theory (Flash & Hogan, 1985), minimum energy theory, minimum mean squared velocity theory, minimum mean squared force theory (Stein et al., 1994), or minimum torque change theory (Uno et al., 1989). Each theory predicts an *ideal* movement trajectory that maximizes or minimizes (optimizes) some *criterion* from the start point to the endpoint. For each ideal trajectory, the causal motor command was determined for the entire path in a feedforward manner. It was also assumed that no noise disturbs the movement's execution because these optimization methods did not make noises into consideration.

19.4 Awareness of Noise in Motor Control

Unfortunately, there certainly *is* noise in the real world. Indeed, Harris and Wolpert (1998) pointed out the critical role of inherent noises to determine the final control signal, i.e., muscle activities. We cannot achieve the *ideal* trajectory however hard we may practice because the control signal is always corrupted with the intrinsic noise. Moreover, there are also various noises or disturbances from the environment. Awareness of these intrinsic and extrinsic noises in motor control dramatically changed our approach to feedforward control. For instance, feedforward control for the entire path makes sense only when everything goes as planned during the movement. In reality, the unpredictable noises force trajectories to deviate from the desired path, increasing uncertainty toward the goal. Therefore, there is no guarantee to optimize the criterion as planned. The best we can hope to do is to maximize or minimize the expected value of the criterion.

19.5 Introduction of Stochastic Optimal Control

Intrinsic and extrinsic noises are a typical condition where “stochastic optimal control” or “optimal control” comes into play. Optimal control was developed initially in engineering to control complex multiple-input, multiple-output systems, which were not amenable to classical control theories (Kirk, 1970). It evaluates the system’s random behavior and attempts to optimize responses or stability on the average rather than with assured precision (Stengel, 1994). A stochastic control system performs two functions: first, it controls the system (controller, in Fig. 19.3) and, second, it predicts the current state of the system (estimator, in Fig. 19.3) to provide the best feedback information for the controller (Stengel, 1994). Such an estimator takes efference copy and sensory inputs into account, and it weighs these pieces of information depending on their reliability (i.e., optimally). In modeling practice, one may use a Kalman filter (Kalman & Bucy, 1961), an optimal estimator when the dynamics and sensory measurements are linear and the noise is Gaussian (Todorov, 2004). In Fig. 19.3, the estimator and the controller are in a loop; thus, they can continue to generate time-varying commands recursively without preparing a whole set of motor commands in a feed-forward manner. Then where is the estimator in the central nervous system?

19.6 Difficulty in the Identification of the Cerebellar Forward Model

Previous reviews repeatedly suggested the cerebellum as a potential site of the estimator or forward model mainly based on neuroanatomical data and clinical observations (for instance, Miall et al., 1993; Haggard & Wing, 1995; Wolpert & Miall,

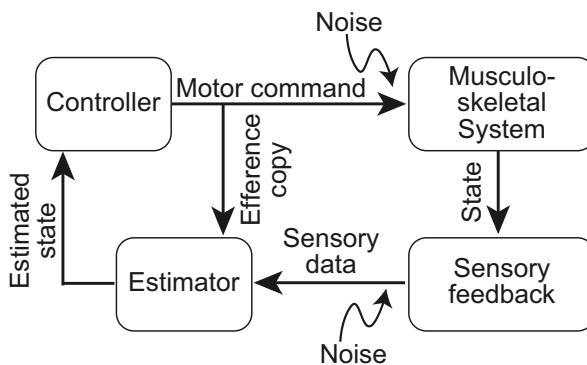


Fig. 19.3 Schematic of closed-loop optimization. (Modified from Todorov (2004)) The estimator needs an efference copy of concurrent motor commands and delayed sensory feedback data in order to compensate for sensory delays. Note that the estimator and controller are in a loop; thus, they can continue to generate time-varying motor commands even when sensory feedback becomes unavailable or unreliable

1996; Bastian, 2006; Ebner & Pasalar, 2008). As mentioned above, a forward model requires two major inputs: (1) a set of sensory feedback signals, which are necessary to update the forward model, and (2) the copy of descending motor commands. These two inputs are integrated in the forward model to generate the state estimate. Indeed, the cerebellum receives both of these inputs. It receives substantial inputs from cortical motor areas via the pontine nuclei (PN) (Brodal and Bijaal, 2003; Schmahmann et al., 2004), and these inputs represent the efference copy of the descending motor commands (Ishikawa et al., 2014, 2016; Tomatsu et al., 2016). The cerebellum also receives substantial somatosensory inputs directly from the ascending spinocerebellar tracts and indirectly via brain stem nuclei, such as the cuneate nucleus or lateral reticular nucleus. These sensory inputs could provide an update on the state of the motor apparatus. The above argument may appear to support the cerebellar forward model hypothesis. But in reality, it is on insufficient grounds because the two lines of inputs are primarily separated in the cerebellar cortex. The mossy fiber (MF) inputs from the cortical motor areas (via PN) distribute mainly in the hemispheric (i.e., lateral) part (Na et al., 2019), while the sensory MF inputs from the spinal cord or the brain stem nuclei distribute in more rostral and medial part (the anterior lobe and the intermediate zone) (e.g., Wu et al., 1999) of the cerebellar cortex. Therefore, we may expect a convergence of the two MF inputs only in a minor part of the intermediate zone. More importantly, even if the nominal convergence has some role to play, the simple summation of the two MF inputs is *not* consistent with their *asymmetric* roles in the forward model. The efference copy plays an essential role in a state prediction, while the sensory input plays a critical role in an *update* of the prediction, as will be discussed later.

As for the output from a forward model, we expect it to correlate with the future state of the motor apparatus (Wolpert and Miall, 1996). In principle, we should examine the output from the cerebrocerebellum in the dentate nucleus (DN) because it is the sole output node from the cerebrocerebellum. Nevertheless, previous studies tried to address this issue by analyzing the Purkinje cell (PC) activities. Note that PCs' activity represents an *intermediate representation* of the cerebellar circuitry and is not ideal for characterizing the *output* of a forward model. In this regard, few studies are eligible to discuss the output of the cerebellar forward models (Thach, 1975, 1978; Thier & Markanday, 2019).

19.7 Movement Representation in the Cerebrocerebellum

To identify movement representations of a forward model, we need to satisfy two requirements: (1) identification of cerebellar neural elements and (2) identification of movement or sensory coordinate frames for activities of each component. Fortunately, the cerebellum provides an ideal place to achieve the first goal (Ishikawa et al., 2014; Tomatsu et al., 2016). Indeed, it is possible in the cerebellum to isolate single-unit activities of MFs (primary cerebellar inputs), PCs (the sole output from the cerebellar cortex), and DN cells (DNCs) (the sole output from the

cerebrocerebellum) (Ishikawa et al., 2014; Tomatsu et al., 2016). Furthermore, it is also possible to achieve the second goal by employing our previous experimental design (Kakei et al., 1999, 2001). With this setup, we recorded activities of MFs, PCs, and DNCs, while monkeys perform wrist movements for eight different directions in two different forearm postures (Ishikawa et al., 2014; Tomatsu et al., 2016). This task design enabled us to dissociate intrinsic coordinate frames from an extrinsic coordinate frame for the wrist movement, depending on the posture-dependent changes in neuron activities. The results revealed distinct steps of movement representation from the input to the output of the cerebrocerebellum.

First, MFs demonstrated temporal and directional properties that were surprisingly similar to those of neurons in the primary motor cortex (M1)/the premotor cortex (PM) (Kakei et al., 1999, 2001). Namely, these MFs relay copies of the M1/PM motor commands to the cerebellum. Besides, their posture-dependent change of directional tuning demonstrated a *bimodal* distribution of shifts in the preferred direction (PD) for the 180° rotation in the forearm posture (Fig. 10a in Tomatsu et al., 2016), much like M1/PM neurons (Tomatsu et al., 2016) – one group with smaller shifts in PD (i.e., *extrinsic-like* neurons) and the other group with larger shifts in PD (*muscle-like* or *joint-like* neurons).

Second, PCs demonstrated much more complex spatiotemporal patterns of activity than MFs. The complexity of PC activities appeared to reflect rapidly changing properties of the peripheral motor apparatus during movement. Also, intricate spatiotemporal patterns of PC activities changed significantly for a change in forearm posture regarding the directional tuning and the gain modulation (Tomatsu et al., 2016). In particular, PCs showed a unimodal distribution of shift in PD that differed from the bimodal distribution of that of MFs (Fig. 10b in Tomatsu et al., 2016). The posture-dependent changes of PC activities indicate that the activities of these PCs encode intrinsic parameters and provide another support that the cerebrocerebellum works as a forward model to predict the state of the motor apparatus (Tomatsu et al., 2016).

Lastly, activities of DNCs, to our great surprise, appeared to *recover* those properties that were typical for MFs (Ishikawa et al., 2014). Namely, DNCs recovered simpler spatiotemporal activity patterns, much like MFs, despite substantial direct inputs from PCs. Also, the posture-dependent shift in PD for DNCs *recovered* a bimodal distribution for the change in the forearm posture (Fig. 19.4a), much like MFs – one group with smaller (i.e., *extrinsic-like*) shifts in PD and the other group with more extensive (i.e., *muscle-like* or *joint-like*) shifts in PD.

Fig. 19.4 (continued) range per 1 ms of the cursor on the monitor controlled by wrist joint movement. See Ishikawa et al. (2014) for the details of the experimental procedures. **(2)**: Optimal delay between the movement speed and $|\Sigma SS_{decl}$ and $|\Sigma SS_{incl}$ for the data shown in A. We calculated the R^2 value for the correlation between them for each 1 ms shift of movement speed from -150 to 50 ms relative to movement onset. Upper panel: R^2 values between the movement speed and $|\Sigma SS_{decl}$ for each delay. The value was the highest (= 0.847) when the movement speed profile was shifted by -61 ms (i.e., optimal delay). Lower panel: R^2 values between the movement speed and $|\Sigma SS_{incl}$ for each delay. The value was the highest (= 0.732) when the movement speed profile was shifted by -7 ms. (Modified from Ishikawa et al. (2016))

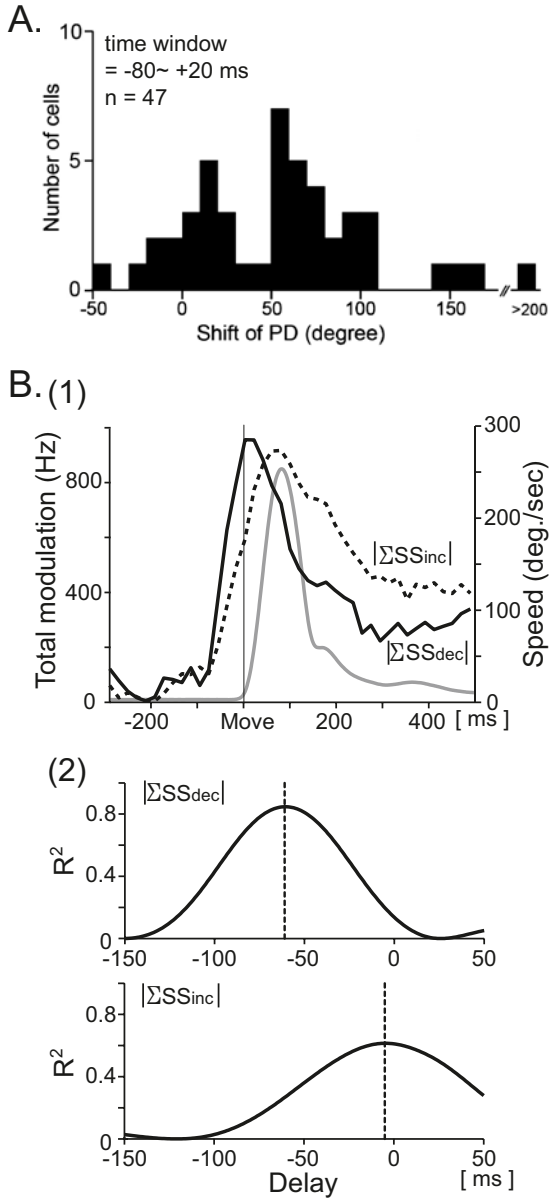


Fig. 19.4 Some spatiotemporal features of dentate nucleus cells (DNCs) and PCs. (a) Distribution of shifts in PD from PRO to SUP for DNCs in a time window of -25 to 0 ms relative to movement onset. Bin width = 10° . Note the bimodal distribution (*Ishikawa and Kakei, unpublished data*). (b) Correlation between the population modulation of Purkinje cells (PCs) and movement kinematics. (1) Temporal patterns of the sum of the decrease ($|\Sigma SS_{dec}|$, solid line) and increase ($|\Sigma SS_{inc}|$, dashed line) of the simple spike (SS) activity of all movement-related PCs and the average speed of the wrist movement (gray line) in a monkey. To obtain $|\Sigma SS_{dec}|$ and $|\Sigma SS_{inc}|$, we summed all decreases and increases of SS activity relative to a reference period (200 – 260 ms before movement onset) separately in each 20 ms bin. The speed profile was calculated from a displacement

In summary, the cerebrocerebellum appears to transform copies of cortical motor commands (i.e., MF inputs) into similar movement representations (i.e., DNC output) through fundamentally distinct representations of PCs in a posture-dependent manner.

19.8 Timing of Movement-Related PC Activities in the Cerebrocerebellum

The timing of the task-related activities of PCs was also compatible with the cerebellar forward model hypothesis. Fig. 19.4b depicts a comparison between the speed profile and PCs' population activity recorded in the cerebrocerebellum of three monkeys during a rapid wrist movement in our recent study (Ishikawa et al., 2014). In this analysis, we summated the increase (ΣSS_{incl}) and decrease (ΣSS_{decl}) of simple spike activity of all movement-related PCs separately. As shown in Fig. 19.4b(2), ΣSS_{decl} demonstrated the highest correlation with the speed profile of the movement when the speed profile was shifted by -60 ms. Namely, the population activity of PCs *precedes* the actual movement by about 60 ms. Indeed, the lead times of PC activities were comparable to the average onset of muscle activities in the same animals (Tomatsu et al., 2016).

On the other hand, the onset latencies of the PCs lagged behind those of M1 and PMv neurons reported in our previous studies (-97.0 ± 15.3 ms for 44 extrinsic-like M1 neurons, -93.6 ± 20.8 ms for 28 muscle-like M1 neurons, and -124.3 ± 30.6 ms for 55 extrinsic-like PMv neurons, Kakei et al., 1999, 2001). Therefore, the PCs' population activity *follows* that of the cortical motor command ($p < 0.001$, Mann-Whitney U-test). Thus, PC activities appear to represent the future states of the motor apparatus rather than motor commands or external sensory feedback. Overall, our observations suggest that the cerebrocerebellum could work as a forward model in terms of timing, representation, and transformation of activities.

19.9 System Identification of the Transformation in the Cerebrocerebellum: Its Similarity to Kalman Filter

If the cerebrocerebellum functions as a forward model, it is expected that the current output from DN should contain predictive information about the future MF input. Therefore, in our previous study (Tanaka et al., 2019), we examined the relationship between activities of MFs (cerebellar inputs), PCs (intermediate representation), and DNCs (cerebellar outputs). Briefly, we found that the activities of individual PCs were reconstructed precisely as a weighted sum of those of MFs. Similarly, the activities of individual DNCs were reconstructed strictly as a weighted sum of those

of PCs and MFs. We further proved that the activities of DNCs contained predictive information about future MF inputs (Tanaka et al., 2019). Namely, the output from the cerebrocerebellum is capable of predicting 200 ms into the future to compensate for the delay of sensory feedback. We finally note that the linear relationship between MF, PC, and DNC activities resembles an optimal linear estimator known as the Kalman filter (Kalman & Bucy, 1961; Tanaka et al., 2019).

The functional similarity of the cerebellum to the Kalman filter has already been suggested in some previous reviews. Most notably, Paulin (1989, 1997) indicated that the cerebellum could be a neural analog of a Kalman filter. Droulez and Cornilleau-Pérez (1993) drew attention to the relevance of multisensory integration in the moving organism to the Kalman filter. Nevertheless, the suggested analogy was only at the functional level and lacked correspondence to the cerebellar network. In our study, we demonstrated the three computational steps in the cerebellar circuit that are compatible with the Kalman filter (Tanaka et al., 2019): (1) the PCs compute a predictive state from a current estimate conveyed by the MFs (prediction step); (2) the DNCs combine the predicted state from the PCs and sensory feedback from the MFs (filtering step); and (3) the DNCs represent future activities of MFs (cerebellar prediction).

Note that even a pair of an excitatory granule cell and an inhibitory Golgi cell that receive the same MF input can function as a neural oscillator (Hoppensteadt & Izhikevich, 1997; Wilson & Cowan, 1973). It can show nonlinear input-output organizations and various types of bifurcations of activities depending on system parameters (Izhikevich, 2007). Therefore, these *linear* steps of the cerebellar information processing were unexpected and surprising, considering the complexity of the whole neuron network of the cerebellum.

Overall, the cerebellum appears to perform not only an internal forward model prediction but also an optimal integration of a predicted state and sensory feedback signals, in a way that is equivalent to Kalman filter as summarized below (a) (Tanaka et al., 2019):

$$\hat{X}_{r,t} = \hat{X}_{r,t-1} + K \left(z_t - C\hat{X}_{r,t-1} \right) = (I - KC) \hat{X}_{r,t-1} + Kz_t \quad (19.a)$$

where a filtered state $\hat{X}_{r,t}$ (DNC output) is generated by combining a predicted state $\hat{X}_{r,t-1}$ (=PC input) and an observed state z_t (=MF collateral input). We speculate that the weights from PC to DNC and the weights from MF to DNC correspond to the matrices $I - KC$ and K in Eq. (19.a), respectively (Tanaka et al., 2019).

19.10 Morphologic Substrata of the Cerebrocerebellum for Kalman Filter

Nevertheless, we realized that the conventional circuit diagram of the cerebellum (Fig. 19.5a) is *not* compatible with the Kalman filter (a). In this diagram (Fig. 19.5a), a MF projects *both* to PC (via granule cell (GC)) and DNC as collaterals, implying

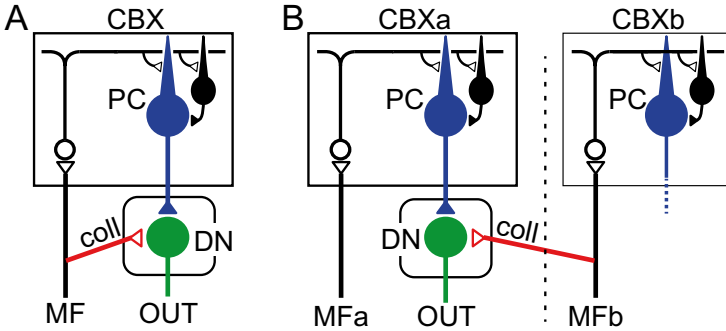


Fig. 19.5 Two schematics of corticonuclear organization. (a) Conventional scheme in which the same MF projects to the cerebellar cortex (CBX) and DN, both of which belong to the same corticonuclear complex (Ito, 1984). (b) Proposed scheme that one MF (MFa) from pontine nuclei (PN) projects to the cerebellar cortex (i.e., cerebrocerebellum) (CBXa) *without* collateral projection to DN, whereas another separate MF (MFb) projects to DN with a collateral. Note that MFa and MFb have distinct projection areas in the cerebellar cortex, CBXa and CBXb, respectively. Only Scheme B is consistent with the requirements of the Kalman filter model and the latest neuroanatomical data for the cerebrocerebellum. (Adapted from Tanaka et al. (2020) under CC BY license)

that PC and DNC of the same corticonuclear microcomplex (Ito, 1984) share the *same* MF input. In contrast, a Kalman filter (a) requires two distinct MF inputs. One MF input originates from cortical motor areas. It contributes to the prediction step in the cerebellar cortex to generate the *current estimate* ($\hat{X}_{t|t-1}$) (i.e., PC activity). The other MF input conveys sensory feedback input to DNC through the collateral and contributes to the filtering step in DN. Most importantly, the contributions of the two MF inputs in (a) are *asymmetrical* and *uninterchangeable*. Therefore, the neuron circuit (Fig. 19.5a), in which the current estimate and current measurement are indistinguishable (i.e., *interchangeable*), cannot function as a Kalman filter.

Incongruent with the conventional diagram, extant anatomical studies suggest that the cerebrocerebellum receives respective MF inputs to PC and DNC (Fig. 19.5b). The first requirement of the Kalman filter is the cortical MFs project to the cerebrocerebellum *without* collaterals to DNC. Na et al. (2019) recently demonstrated that MFs from PN virtually *lack* collaterals to DNC on their way to the cerebrocerebellum. Namely, the first requirement is satisfied with the input from cortical motor areas to the cerebrocerebellum. The second requirement for the Kalman filter is that MFs conveying sensory input give off collaterals to DN. Indeed, Wu et al. (1999) demonstrated that MFs originated from the lateral reticular nucleus (LRN), which receives strong somatosensory inputs from the spinal cord, have an abundant collateral projection to DN and other cerebellar nuclei on their way to the vermis and the intermediate zone (see Figs. 8, 9, and 10 in Wu et al., 1999). Note that the two MF inputs from PN and LRN have only minor overlap in the cerebellar cortex (Na et al., 2019; Wu et al., 1999). Figure 19.5b summarizes these observations and demonstrates the *asymmetrical* relationship of the two lines of MF inputs,

which is consistent with the Kalman filter. We have already pointed out the defect of the symmetrical MF inputs in the previous cerebellar forward model hypothesis (see Section “Difficulty in Identifying the Cerebellar Forward Model” in this article). In this way, the defect has been removed.

Under these anatomical data, we found two functionally distinct populations of MFs in our data (Tanaka et al., 2019). One population of MFs contributed selectively to the reconstruction of PC activities and dominated the prediction step, while the other population of MFs contributed selectively to the reconstruction of DNC activities and dominated the filtering step (Tanaka et al., 2019). The average correlation coefficient between weights of MF–PC and MF–DNC projections was no more than 0.060. A statistical test based on resampling verified that the correlation between the two MF populations was statistically significant ($p < 10^{-5}$). Therefore, we concluded that PCs and DNCs received inputs from distinct populations of MFs, thereby satisfying the Kalman filter model’s requirements.

19.11 Inference About the Primordial Operation of the Cerebellum

The “corticonuclear microcomplex” depicted in Fig. 19.5b is most likely specific for the dentate nucleus and corresponding cerebrocerebellum (i.e., the *newer* part of the cerebellum). In contrast, the other older parts of the cerebellar nuclei receive MF inputs in different ways (Ito, 1984). For example, in the vestibular nucleus, neurons are driven primarily by direct (i.e., primary afferent) MF inputs, whereas PCs activated by the same MF inputs exert modulatory action on the nuclear neurons (Fig. 19.5a; see also Fig. 92a in Ito, 1984). In the fastigial nucleus, however, the PC input plays the primary role. At the same time, collaterals of MFs provide a *background excitation* on which PCs can impose efficient bidirectional modulation (Fig. 19.5a; see also Fig. 92b in Ito, 1984). In these phylogenetically *older* cerebellar regions, the corticonuclear microcomplex (Ito, 1984) is *not* consistent with the Kalman filter, where PC and cerebellar nuclear cells share the same MF input. Overall, even if the local neuron circuitry is common for the entire cerebellar cortex, different regions may perform computationally different operations depending on the organization of microcomplex (Ito, 1984: pp. 195–199 and Fig. 92).

Nevertheless, because of the superb “crystal-like” homogeneity of the neuron circuitry, all these regions of the cerebellar cortex most likely hold the *prediction step* in common. Even if the presumed prediction step alone remains suboptimal due to lack of the filtering step, it could still play an invaluable role in improving its owner’s survival. Indeed, the cerebellum-like structure of fishes with electroreception systems has been suggested as a neural analog of a dynamical state estimator (Bastian & Zakon, 2005; Paulin, 1989; 1997). According to Paulin, the cerebellum is a *sensory* processing structure with a specific role in the state estimation of dynamical systems. He further suggested that the cerebellum has a common

underlying role in sensorimotor, perceptual, and cognitive processes consistent with the state estimator hypothesis (1997). The cerebellar contribution to sensory processing is not surprising if we remember the fact that the cerebellum emerges in the *alar plate* (i.e., sensory domain of the embryonic neural tube) of the rhombencephalon of old jawless fishes (Sugahara et al., 2016). It collects multimodal inputs, including exteroceptive (lateral line, vestibular, acoustic, visual) and somatosensory inputs (Larsell, 1967). The cerebellum has further gained access, in mammals, to cortical information from association areas and motor and sensory areas. Overall, throughout its long history of evolution, the cerebellum has been a unique *hub* to collect afferent, efferent, and finally *internal* (i.e., association) information from the entire brain.

19.12 Extension of Cerebellar Kalman Filter Hypothesis to the Non-motor Cerebrocerebellum

The critical question arises whether the Kalman filter mechanism for the *motor* part of the cerebrocerebellum (Tanaka et al., 2019) generalizes to its *cognitive/affective* part. Our dataset recorded during the motor task may not generalize directly to the cerebellum's contribution to prediction in cognitive/affective domains. Nevertheless, it is possible to search for the *Kalman filter-specific* corticonuclear microcomplex (Fig. 19.5b) in the *non-motor* part of the cerebrocerebellum. There are two requirements: (1) the primary *MFa* input to the cerebrocerebellum is originated from a non-motor cortical area and relayed by PN cells (PNCs) and (2) the filtering *MFb* input is derived from a distinct cortical or subcortical source and relayed by a *non-PN* nucleus with a significant collateral projection to DN (Fig. 19.5b). Requirement (2) is the key because requirement (1) is common for most, if not all, non-motor cortical areas, including prefrontal areas (Schmahmann & Pandya, 1997), parietal association areas (Schmahmann & Pandya, 1989), superior temporal areas (Schmahmann & Pandya, 1991), and occipitotemporal and parahippocampal areas (Schmahmann & Pandya, 1993). There are a few known sources of collateral MF inputs to DN, most notably the lateral reticular nucleus (LRN) (Wu et al., 1999) and the nucleus reticularis tegmenti pontis (NRTP) (Gerrits and Voogd, 1987) in the reticular formation. The LRN receives the main inputs from the spinal cord (Alstermark and Ekerot, 2013) and additional inputs from the sensorimotor areas and the red nucleus (Bruckmoser et al., 1969; Matsuyama and Drew, 1997). The NRTP receives inputs mainly from the sensorimotor areas, the prefrontal areas, and the parietal association areas (Schmahmann et al., 2004). In summary, the Kalman filter model (Fig. 19.5b) is also applicable to the non-motor part of the cerebrocerebellum if the MF collateral to non-motor parts of DN (*MFb*) and MF inputs to the PCs (*MFa*) have distinct sources and *causal* relationship from *MFa* to *MFb*. In that sense, NRTP is a major candidate for the filtering inputs to the non-motor parts of DN (Fig. 19.6, left).

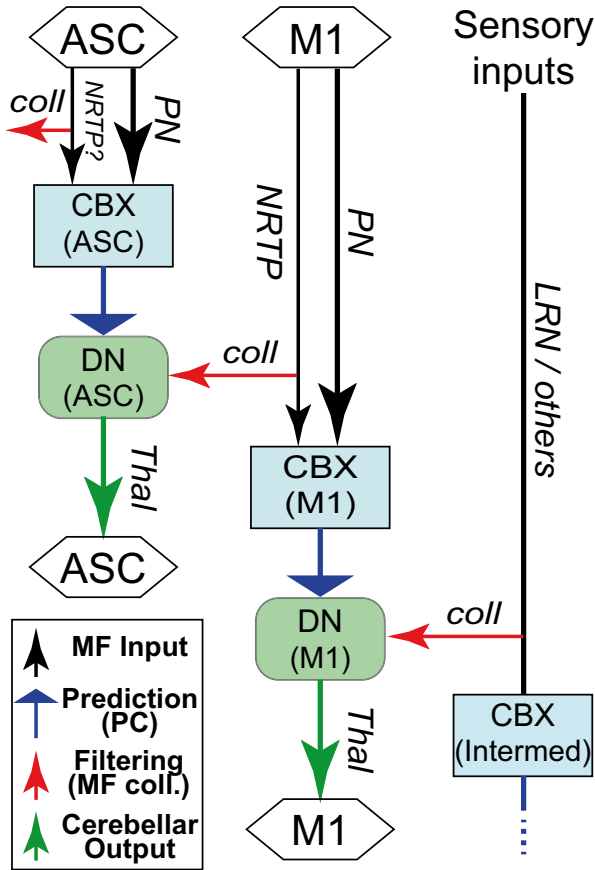


Fig. 19.6 A hypothetical cascade of Kalman filters in the cerebrocerebellar communication loop. The Kalman filter model that predicts M1 activity (*center*) is capable to form a cascade with another Kalman filter that predicts ASC activity (*left*), if M1 sends a filtering input (*coll*) to distinct region of DN (ASC) via NRTTP (*center*). Note the collateral input to DN (ASC) from M1 does not project to M1 region of DN (DN (M1)) (see Tanaka et al., 2019). In this way, the filtering input may play a critical role to make two Kalman filters to work together. This model may also explain how parallel forward models in the cerebrocerebellar communication loops function together in a coordinated manner and may provide a partial explanation for unity of mind. ASC association cortex, CBX cerebellar cortex, coll collateral of MFs, LRN lateral reticular nucleus, M1 the primary motor cortex, NRTTP nucleus reticularis tegmenti pontis, PN pontine nuclei, Thal thalamus

It should be pointed out that the Kalman filter model that predicts M1 activity (Fig. 19.6, *center*) is capable to form a cascade with another Kalman filter that predicts activity of association cortex (ASC) (Fig. 19.6, *left*), if M1 sends a filtering input (*coll*) to distinct region of DN (ASC) via NRTTP (Fig. 19.6, *left*). In this way, the filtering input may play a critical role to make two Kalman filters work together. This model also explains how parallel domains in the cerebrocerebellar

communication loops (Kelly & Strick, 2003) are coordinated in a cascadic manner, providing a partial explanation for unity of mind.

19.13 Compressed Prediction of the Cerebellar Internal Model

Few paid attention to the asymmetry of the cerebrocerebellar loop in terms of the number of output neurons. The number of axons in the cerebral peduncle (CP) conveying cortical outputs to PN and other precerebellar nuclei is estimated as 21 million in humans (Tomasch, 1969). In contrast, the number of axons in the return path (i.e., the superior cerebellar peduncle) relaying the cerebellar output to the thalamus is no more than 0.8 million in humans (Heidary & Tomasch, 1969). Namely, the cerebrocerebellum returns its output to the cerebral cortex after significant compression (1:20) (Tanaka et al., 2020). Therefore, the cerebellar output appears to represent a predicted state of cortical activities in a compressed format. The mapping between the cortical output and the cerebellar output may be compatible with a *homomorphism* (<https://en.wikipedia.org/wiki/Homomorphism>). A homomorphism has a distinguished advantage for an internal model because it enables the model to perform an operation equivalent to the original while using a more simplified representation.

Although there is no consensus on the compressed representation so far (Sanger et al., 2019), the compact (i.e., low-resolution) prediction of the cortical state may help assign more attention to the task currently in focus by minimizing the computational load for the other peripheral tasks. It also reminds us that the cerebellum contributes most to trained and automated repertoires of both motor and cognitive functions with reduced attention.

We cannot spare another important consequence of the relative paucity of DNCs. The MF collateral input to DN (Fig. 19.5b, *coll*) appears far from massive compared (e.g., Wu et al., 1999) to the massive MF input to the cerebrocerebellum (Fig. 19.5b, MFa). Therefore, one may argue that the modest projection of the MF collaterals to DN cannot be effective enough to play such an important function as *filtering* of the Kalman filter. Nevertheless, the limited number of the *target* DNCs appears to help in *amplifying* efficacy of the collateral input.

19.14 Clinical Evidence for the Internal Model Hypothesis of the Cerebellum

Finally, we searched for clinical evidence that supported the cerebellar forward model hypothesis (e.g., Bastian, 2006; Miall et al., 2007). A series of studies from our group confirmed the impaired predictive control in movements of patients with

cerebellar ataxia (CA). We first decomposed the muscle activities for the wrist movement into a low-frequency (≤ 0.5 Hz) component (F1) and a high-frequency (>0.5 Hz) component (F2), each of which represented the predictive control and the feedback correction, respectively (Kakei et al., 2019). Then for each component, we identified a recipe of muscle activities by analyzing a relationship between the muscle tension and movement kinematics (the wrist angle $\theta(t)$ and the wrist angular velocity $\dot{\theta}(t)$) weighted by the coefficients of K_r (the elastic term) and B_r (the viscous term) (Kakei et al., 2019; Lee et al., 2012; Mitoma et al., 2016). Importantly, we found that the ratio of B_r/K_r characterized the recipe of muscle activities for each component. In control subjects, the B_r/K_r ratio for the predictive (F1) component demonstrated a higher value (Fig. 19.5 in Kakei et al., 2019) (Fig. 7a), suggesting the velocity control dominance. On the other hand, the B_r/K_r ratio for the corrective (F2) component demonstrated a much smaller value (Fig. 19.5 in Kakei et al., 2019) (Fig. 19.7a), suggesting the role of F2 component in correction of positional errors (Kakei et al., 2019). In contrast, CAs showed a selective decrease of the B_r/K_r ratio for the predictive (F1) component (Fig. 5 in Kakei et al., 2019) (Fig. 19.7a), suggesting poor recruitment of the predictive velocity control and compensatory dependence on the position-dependent pursuit (Kakei et al., 2019). The loss of component-specific differences in the B_r/K_r ratio suggests impairment of predictive control in CA. Indeed, the B_r/K_r ratio decrease correlated with the increase of error in the predictive (F1) movement (Fig. 19.7b) (Kakei et al., 2019). Another critical

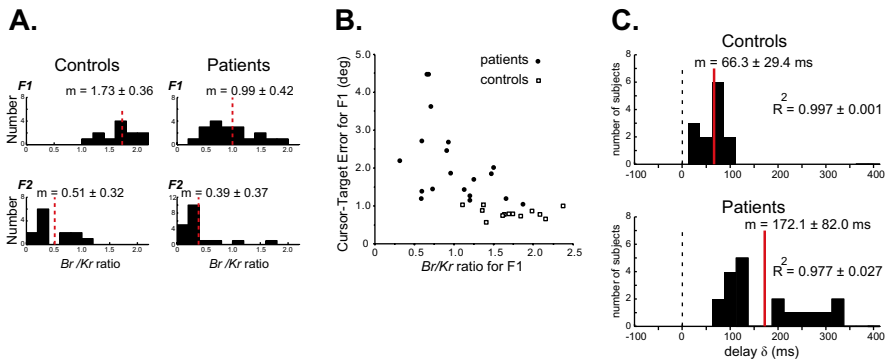


Fig. 19.7 Difference of muscle activity – movement kinematics relationship between controls and cerebellar patients. **(a)** Comparison of the B_r/K_r ratios that represent recipe of the motor commands for the F1 and F2 components between the controls and the cerebellar patients. **Controls**: B_r/K_r ratios of the control subjects for the F1 component (top) and the F2 component (bottom) ($n = 13$). Note the highly significant difference between the two components. **Patients**: B_r/K_r ratios of the patients for the F1 (top) and the F2 (bottom) components ($n = 19$). Note the selective decrease of B_r/K_r ratios for the F1 component in the patients. **(b)** Correlation between the B_r/K_r ratios for F1 component and cursor-target error for F1 (F1 error, in short). The F1 error is defined as an average error between the target motion and the F1 component of the movement. Note the negative correlation. **(c)** Delay of the predictive (F1) component of the movement relative to the target motion calculated with a cross-correlation analysis for controls ($n = 13$) and patients ($n = 19$). (Adapted from Kakei et al. (2019) under CC BY license)

difference between the control and CA was the increased delay of the predictive (F1) component in CA (Fig. 19.7c). In the control subjects, the predictive (F1) movement lagged the target motion only by 66 ms, which was too small to be a visual feedback delay (i.e., a proof of prediction) (Kakei et al., 2019). In contrast, in patients with CA, the delay increased by more than 100 ms, as much as 172 ms. The increased delay (i.e., 172 ms) is comparable to a visual feedback delay, demonstrating lack of compensation of feedback delay in CA patients. In summary, ataxic movements are consistent with an impairment of a forward model in terms of accuracy and delay of state prediction.

19.14.1 *Postscript*

The most primitive cerebellum emerged in the alar plate of the rhombencephalon of old jawless fishes as a sensory *hub* to which multimodal sensory inputs converge. The cerebellum later acquired efference copy inputs, which is essential for active sensing. Indeed, we can see its example in the cerebellum-like structure of some fishes that process information from electroreception systems. We speculate that the active sensing evolved to detect causality and finally led to a more sophisticated state prediction in a primitive forward model. Next, in mammals, the cerebellum acquired a strong loop with the cerebral cortex: the cerebrocerebellar communication loop. In this way, the cerebellum developed into the primary hub in the entire CNS. In particular, the acquisition of the DN filtering step evolved the existing dynamic prediction in the cerebellar cortex to a Kalman filter. This revolutionary event gave each region of the cerebrocerebellum a privilege to predict the state of its counterpart in the cerebral cortex, which includes motor areas, parietal association areas, prefrontal association areas, and limbic areas (Ito, 2008). This Kalman filter model also explains how parallel domains in the cerebrocerebellum operate in a cascadic manner and may provide a partial explanation for unity of mind. Finally, we should not forget the morphological asymmetry of the cerebrocerebellar communication loop. Namely, the cerebrocerebellum returns its output from DN back to the cerebral cortex after significant compression (1:20) (Tanaka et al., 2020). The low-resolution prediction of the cortical state may help assign more attention to the task currently in focus by reducing computational load for peripheral tasks. Given the fact that the cerebellum contributes most to trained and automated repertoires with less effort and attention, this asymmetry appears to make perfect sense.

Acknowledgments We thank Profs. Hiroshi Mitoma and Koji Ito for their valuable comments and discussions.

Support This work was supported by Grants-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology in Japan (MEXT) (<http://www.mext.go.jp/>) (no. 26120003, no. 14580784, no. 15016008, no. 16015212, no. 20033029, and no. 21500319 to SK; no. 25430007, no. 26120005, and no. 16 K12476 to HT; no. 21700229 and no. 24650304 to JL; and no. 24650224

to TI), the Japan Science and Technology Agency (A-STEP) to SK (<http://www.jst.go.jp/>), the Japan Science and Technology Agency (PRESTO: Intelligent Cooperation and Control) (SK), NBRP “Japanese Monkeys” through the National BioResource Project of MEXT, the JSPS Programs (Program for Advancing Strategic International Networks to Accelerate the Circulation of Talented Researchers, and Embodied-Brain Systems Science) (HT), and the Hitachi-Kurata and the Tateishi Science Foundations (HT). This research was also supported by AMED under grant number 16k0109048h0003 to SK. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

- Albus, J. S. (1971). A theory of cerebellar function. *Mathematical Biosciences*, *10*, 25–61.
- Allen, G. I., & Tukahara, N. (1974). Cerebrocerebellar communication systems. *Physiological Reviews*, *54*, 957–1006.
- Alstermark, B., & Ekerot, C. F. (2013). The lateral reticular nucleus: a precerebellar centre providing the cerebellum with overview and integration of motor functions at systems level. A new hypothesis. *Journal of Physiology*, *591*, 5453–5458.
- Bastian, A. J. (2006). Learning to predict the future: The cerebellum adapts feedforward movement control. *Current Opinion in Neurobiology*, *16*, 645–649.
- Bastian, J., & Zakon, H. H. (2005). Plasticity of sense organs and brain. In T. H. Bullock, C. D. Hopkins, A. N. Popper, & R. R. Fay (Eds.), *Electroception* (pp. 195–228). Springer.
- Brodal, P., & Bijaaliev, J. G. (2003). Organization of the pontine nuclei. *Neuroscience Research*, *13*, 83–118.
- Bruckmoser, P., Hepp, M. C., & Wiesendanger, M. (1969). Cortical influence on the lateral reticular nucleus of the cat. *Brain Research*, *15*, 556–558.
- Droulez, J., & Cornilleau-Pérez, V. (1993). Application of the coherence scheme to the multisensory fusion problem. In A. Berthoz (Ed.), *Multisensory control of movement* (pp. 485–501). Oxford University Press.
- Ebner, T. J., & Pasalar, S. (2008). Cerebellum predicts the future motor state. *Cerebellum*, *7*, 583–588.
- Eccles, J. C., Ito, M., & Szentágothai, J. (1967). *The cerebellum as a neuronal machine*. Springer-Verlag.
- Flash, T., & Hogan, N. (1985). The coordination of arm movements: an experimentally confirmed mathematical model. *Journal of Neuroscience*, *5*, 1688–1703.
- Gerrits, N. M., & Voogd, J. (1987). The projection of the nucleus reticularis tegmenti pontis and adjacent regions of the pontine nuclei to the central cerebellar nuclei in the cat. *Journal of Comparative Neurology*, *258*, 52–69.
- Haggard, P., & Wing, A. (1995). Coordinate responses following mechanical perturbations of the arm during prehension. *Experimental Brain Research*, *102*, 483–494.
- Harris, C. M., & Wolpert, D. M. (1998). Signal dependent noise determines motor planning. *Nature*, *394*, 780–784.
- Heidary, H., & Tomasch, J. (1969). Neuron numbers and perikaryon areas in the human cerebellar nuclei. *Acta Anatomica (Basel)*, *74*, 290–296.
- Hoppensteadt, F. C., & Izhikevich, E. M. (1997). *Weakly connected neural networks*. Springer-Verlag. (Page 23).
- Ishikawa, T., Tomatsu, S., Tsunoda, Y., Lee, J., Hoffman, D. S., & Kakei, S. (2014). Releasing dentate nucleus cells from Purkinje cell inhibition generates output from the cerebrocerebellum. *PLoS One*, *9*, e108774. <https://doi.org/10.1371/journal.pone.0108774>
- Ishikawa, T., Tomatsu, S., Izawa, J., & Kakei, S. (2016). The cerebro-cerebellum: Could it be loci of forward models? *Neuroscience Research*, *104*, 72–79.

- Ito, M. (1970). Neurophysiological aspects of the cerebellar motor control system. *International Journal of Neurology*, 7, 162–176.
- Ito, M. (1984). *The cerebellum and neural control*. New York: Raven Press.
- Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Reviews. Neuroscience*, 9, 304–313.
- Izhikevich, E. M. (2007). *Dynamical Systems in Neuroscience – The geometry of excitability and bursting*. The MIT Press.
- Kakei, S., Hoffman, D. S., & Strick, P. L. (1999). Muscle and movement representations in the primary motor cortex. *Science*, 285, 2136–2139.
- Kakei, S., Hoffman, D. S., & Strick, P. L. (2001). Direction of action is represented in the ventral premotor cortex. *Nature Neuroscience*, 4, 1020–1025.
- Kakei, S., Lee, J., Mitoma, H., Tanaka, H., Manto, M., & Hampe, C. S. (2019). Contribution of the cerebellum to predictive motor control and its evaluation in ataxic patients. *Frontiers in Human Neuroscience*, 13, 216. <https://doi.org/10.3389/fnhum.2019.00216>
- Kalman, R. E., & Bucy, R. S. (1961). New results in linear filtering and prediction. *ASME Journal of Basic Engineering*, 83, 95–108.
- Kelly, R. M., & Strick, P. L. (2003). Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. *The Journal of Neuroscience*, 23, 8432–8444.
- Kirk, D. E. (1970). *Optimal control theory: An introduction*. Prentice-Hall.
- Larsell, O. (1967). *The comparative anatomy and histology of the cerebellum*. University of Minnesota Press.
- Lee, J., Kagamihara, Y., Tomatsu, S., & Kakei, S. (2012). The functional role of the cerebellum in visually guided tracking movement. *Cerebellum*, 11, 426–433.
- Marr, D. (1969). A theory of cerebellar cortex. *The Journal of Physiology*, 202, 437–470.
- Matsuyama, K., & Drew, T. (1997). Organization of the projections from the pericruciate cortex to the pontomedullary brainstem of the cat: a study using the anterograde tracer Phaseolus vulgaris-leucoagglutinin. *Journal of Comparative Neurology*, 389, 617–641.
- Miall, R. C., Weir, D. J., Wolpert, D. M., & Stein, J. (1993). Is the cerebellum a smith predictor? *Journal of Motor Behavior*, 25, 203–216.
- Miall, R. C., Christensen, L. O. D., Cain, O., & Stanley, J. (2007). Disruption of state estimation in the human lateral cerebellum. *PLoS Biology*, 5, e316.
- Mitoma, H., Adhikari, K., Aeschlimann, D., Chattopadhyay, P., Hadjivassiliou, M., Hampe, C. S., Honnorat, J., Joubert, B., Kakei, S., Lee, J., Manto, M., Matsunaga, A., Mizusawa, H., Nanri, K., Shanmugarajah, P., Yoneda, M., & Yuki, N. (2016). Consensus paper: Neuroimmune mechanisms of cerebellar ataxias. *Cerebellum*, 15, 213–232.
- Na, J., Sugihara, I., & Shinoda, Y. (2019). The entire trajectories of single pontocerebellar axons and their lobular and longitudinal terminal distribution patterns in multiple aldolase C-positive compartments of the rat cerebellar cortex. *The Journal of Comparative Neurology*, 527, 2488–2511.
- Paulin, M. (1989). A Kalman filter theory of the cerebellum. In M. A. Arbib & S. Amari (Eds.), *Dynamic interactions in neural networks: Models and data* (pp. 239–259). Springer.
- Paulin, M. (1997). Neural representations of moving systems. In J. D. Schmahmann (Ed.), *The cerebellum and cognition* (pp. 515–533). Academic Press.
- Sanger, T. D., Yamashita, O., & Kawato, M. (2019). Expansion coding and computation in the cerebellum: 50 years after the Marr-Albus codon theory. *Journal of Physiology*, 598, 913–928.
- Schmahmann, J. D., & Pandya, D. N. (1989). Anatomical investigation of projections to the basis pontis from posterior parietal association cortices in rhesus monkey. *The Journal of Comparative Neurology*, 289, 53–73.
- Schmahmann, J. D., & Pandya, D. N. (1991). Projections to the basis pontis from the superior temporal sulcus and superior temporal region in the rhesus monkey. *The Journal of Comparative Neurology*, 308, 224–248.
- Schmahmann, J. D., & Pandya, D. N. (1993). Prelunate, occipitotemporal, and parahippocampal projections to the basis pontis in rhesus monkey. *The Journal of Comparative Neurology*, 337, 94–112.
- Schmahmann, J. D., & Pandya, D. N. (1997). Anatomic organization of the basilar pontine projections from prefrontal cortices in rhesus monkey. *The Journal of Neuroscience*, 17, 438–458.

- Schmahmann, J. D., Rosene, D. L., & Pandya, D. N. (2004). Motor projections to the basis pontis in rhesus monkey. *The Journal of Comparative Neurology*, *478*, 248–268.
- Stein, R. B., Oguztoreli, M. N., & Capaday, C. (1994). What is optimized in muscular movements? In Stengel R. F. *Optimal Control and Estimation*. New York: Dover
- Stengel, R. F. (1994). *Optimal control and estimation*. Dover.
- Sugahara, F., Pascual-Anaya, J., Oisi, Y., Kuraku, S., Aota, S., Adachi, N., Takagi, W., Hirai, T., Sato, N., Murakami, Y., & Kuratani, S. (2016). Evidence from cyclostomes for complex regionalization of the ancestral vertebrate brain. *Nature*, *531*, 97–100.
- Tanaka, H., Ishikawa, T., & Kakei, S. (2019). Neural evidence of the cerebellum as a state predictor. *Cerebellum*, *18*, 349–371. <https://doi.org/10.1007/s12311-018-0996-4>
- Tanaka, H., Ishikawa, T., Lee, J., & Kakei, S. (2020). The cerebro-cerebellum as a locus of forward model; a review. *Frontiers in Systems Neuroscience*, *14*, 19.
- Thach, W. T. (1975). Timing of activity in cerebellar dentate nucleus and cerebral motor cortex during prompt volitional movement. *Brain Research*, *88*, 233–241.
- Thach, W. T. (1978). Correlation of neural discharge with pattern and force of muscular activity, joint position, and direction of intended next movement in motor cortex and cerebellum. *Journal of Neurophysiology*, *41*, 654–676.
- Thier, P., & Markanday, A. (2019). Role of the vermal cerebellum in visually guided eye movements and visual motion perception. *Annual Review of Vision Science*, *5*, 247–268.
- Todorov, E. (2004). Optimality principles in sensorimotor control. *Nature Neuroscience*, *7*, 907–915.
- Tomasch, J. (1969). The numerical capacity of the human cortico-pontocerebellar system. *Brain Research*, *13*, 476–484.
- Tomatsu, S., Ishikawa, T., Tsunoda, Y., Lee, J., Hoffman, D. S., & Kakei, S. (2016). Information processing in the hemisphere of the cerebellar cortex for control of wrist movement. *Journal of Neurophysiology*, *115*, 255–270.
- Uno, Y., Kawato, M., & Suzuki, R. (1989). Formation and control of optimal trajectory in human multijoint arm movement: Minimum torque-change model. *Biological Cybernetics*, *61*, 89–101.
- Wilson, H. R., & Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik*, *13*, 55–80.
- Wolpert, D. M., & Miall, R. C. (1996). Forward models for physiological motor control. *Neural Networks*, *9*, 1265–1279.
- Wu, H., Sugihara, I., & Shinoda, Y. (1999). Projection patterns of single mossy fibers originating from the lateral reticular nucleus in the rat cerebellar cortex and nuclei. *The Journal of Comparative Neurology*, *411*, 97–118.