

|              |   |
|--------------|---|
| Title        | Neural Evidence of the Cerebellum as a State Predictor  |
| Author(s)    | Tanaka, Hirokazu; Ishikawa, Takahiro; Kakei, Shinji   |
| Citation     | The Cerebellum, 18(3): 349-371  |
| Issue Date   | 2019-01-09  |
| Type         | Journal Article   |
| Text version | publisher   |
| URL          | <a href="http://hdl.handle.net/10119/16072">http://hdl.handle.net/10119/16072</a>   |
| Rights       | <p>© The Author(s) 2019. Hirokazu Tanaka, Takahiro Ishikawa, Shinji Kakei, The Cerebellum, 18(3), 2019, 349-371. DOI:10.1007/s12311-018-0996-4.</p> <p>This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<a href="http://creativecommons.org/licenses/by/4.0/">http://creativecommons.org/licenses/by/4.0/</a>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.</p> |
| Description  |   |

# *Neural Evidence of the Cerebellum as a State Predictor*

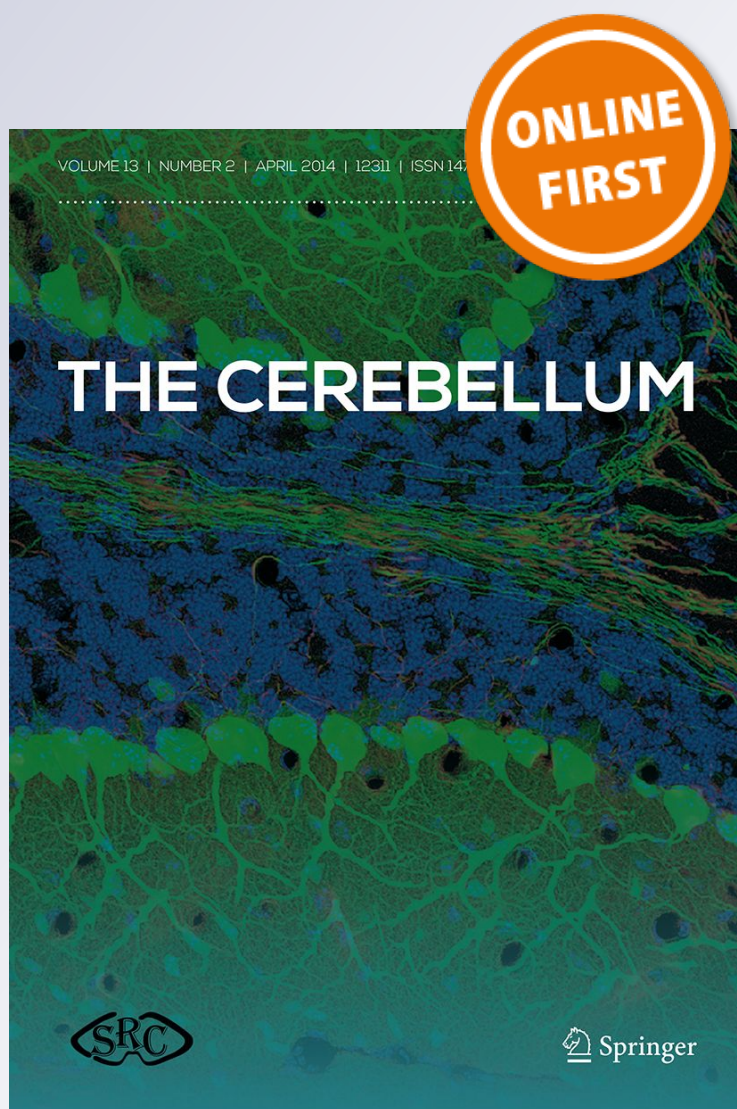
**Hirokazu Tanaka, Takahiro Ishikawa & Shinji Kakei**

**The Cerebellum**

ISSN 1473-4222

Cerebellum

DOI 10.1007/s12311-018-0996-4



**Your article is published under the Creative Commons Attribution license which allows users to read, copy, distribute and make derivative works, as long as the author of the original work is cited. You may self-archive this article on your own website, an institutional repository or funder's repository and make it publicly available immediately.**



# Neural Evidence of the Cerebellum as a State Predictor

Hirokazu Tanaka<sup>1</sup> · Takahiro Ishikawa<sup>2</sup> · Shinji Kakei<sup>2</sup>

© The Author(s) 2019

## Abstract

We here provide neural evidence that the cerebellar circuit can predict future inputs from present outputs, a hallmark of an internal forward model. Recent computational studies hypothesize that the cerebellum performs state prediction known as a forward model. To test the forward-model hypothesis, we analyzed activities of 94 mossy fibers (inputs to the cerebellar cortex), 83 Purkinje cells (output from the cerebellar cortex to dentate nucleus), and 73 dentate nucleus cells (cerebellar output) in the cerebro-cerebellum, all recorded from a monkey performing step-tracking movements of the right wrist. We found that the firing rates of one population could be reconstructed as a weighted linear sum of those of preceding populations. We then went on to investigate if the current outputs of the cerebellum (dentate cells) could predict the future inputs of the cerebellum (mossy fibers). The firing rates of mossy fibers at time  $t + t_1$  could be well reconstructed from as a weighted sum of firing rates of dentate cells at time  $t$ , thereby proving that the dentate activities contained predictive information about the future inputs. The average goodness-of-fit ( $R^2$ ) decreased moderately from 0.89 to 0.86 when  $t_1$  was increased from 20 to 100 ms, hence indicating that the prediction is able to compensate the latency of sensory feedback. The linear equations derived from the firing rates resembled those of a predictor known as Kalman filter composed of prediction and filtering steps. In summary, our analysis of cerebellar activities supports the forward-model hypothesis of the cerebellum.

**Keywords** Motor control · Internal forward model · Mossy fiber · Purkinje cell · Dentate cell · Kalman filter

## Introduction

The cerebellum plays a critical role in the control and coordination of body movements, adaptation to novel environments, and acquisition of new motor skills. Evidence from clinical observations and psychophysical experiments indicates that impairments of the cerebellum lead to motor ataxia characterized by incoordination and dysmetria in multijoint movements. Clinical evidence pioneered by the seminal work of Holmes suggests that impairments in the cerebellum could lead to symptoms characterized by lack of coordination across multiple degrees of freedom in motor control, collectively known as cerebellar ataxia [1]. The cerebellum also plays an

essential role both in adapting to external perturbations such as visual rotation or external force fields and in acquiring a new skill [2–6]. The cerebellum is one of the central nervous system whose anatomical structure, cytoarchitecture, and electrophysiological properties have been thoroughly studied. Despite the plethora of clinical and psychophysical evidence, the precise mechanisms by which the cerebellum coordinates body movements are not yet understood.

Recent computational studies suggest that the cerebellum predicts current and future states of the body by solving the dynamics with given efference copy of motor commands, known as the computation of an internal forward model [7–10]. Sensory feedback signals from the periphery have certain delays in reaching the central nervous system, on the orders of a few tens to 100 ms. Therefore, the brain always receives the “past” state of the body. It is known in engineering and mechanics that feedback control based on time-delayed state can behave in an oscillatory and often unstable way if the delay is on the order of time constants of system dynamics. Fortunately, physical laws that govern body movements allow the brain to predict a current state from a previous state and an efference copy of motor command, essentially by solving the Newtonian mechanics. This predictive mechanism

---

✉ Hirokazu Tanaka  
hirokazu@jaist.ac.jp

<sup>1</sup> School of Information Science, Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi, Ishikawa 923-1211, Japan

<sup>2</sup> Movement Disorders Project, Tokyo Metropolitan Institute of Medical Science, 2-1-6 Kamikitazawa, Setagaya, Tokyo 156-8506, Japan



allows stable and dexterous control of body movements. Although the cerebellum has been suggested as a locus of forward-model computation from psychophysical, neuroimaging, and stimulation studies [11–13], the neural mechanism of how the cerebellum performs predictive computation has not yet been understood [10].

We therefore set out to determine the computation in the cerebellar circuit in a monkey during wrist step-tracking movements and provide neural evidence that current outputs of the cerebellar circuit contain predictive information about future inputs, a hallmark of an internal forward model. We analyzed firing rates of mossy fibers (MFs) (inputs to the cerebellar cortex), Purkinje cells (PCs) (outputs from the cerebellar cortex to the dentate nucleus (DN)), and dentate cells (DCs) (output of the cerebellum) of a monkey performing step-tracking wrist movements. Unlike the cerebral cortex in which multiple inputs and multiple outputs are related in a highly complex way, the cerebellar MF input and DN output are well defined and basically organized in a characteristic feedforward manner. By exploiting the feedforward structure, firing rates of one population were reconstructed from those of other populations that innervate to the target population. Also, we investigated if firing rates of cerebellar output (i.e., DCs) at time  $t$  contained information that was predictive for future inputs to the cerebellar cortex (i.e., MFs) at time  $t + t_1$ . Our analyses targeted to test the forward-model hypothesis of the cerebellum.

## Methods

### Behavioral Task and Electrophysiological Recording

We used firing rates of the cerebellar cells reported in our previous publications. Here, a brief description about the behavioral task and the electrophysiological recording is provided, and for a full account, please refer to the previous publications [14, 15]. All surgical and experimental protocols were approved by the Animal Care and Use Committee of Tokyo Metropolitan Institute of Medical Science. A male macaque monkey participated in a step-tracking movement task of the right wrist. The monkey gripped a manipulandum that measured the wrist movements and controlled a cursor on a computer screen with the manipulandum toward one of eight targets that were uniformly located on a circle of radius of  $8^\circ$ , corresponding to a wrist movement of  $20^\circ$ . An initial posture of the wrist was either pronated or supinated, so there were 16 experimental conditions (eight movement directions  $\times$  two initial postures).

During the step-tracking movement task, neural activities of MFs, PCs, and DCs were recorded. Their anatomical locations and physiological characteristics allowed us to identify these cells with confidence. Spike data were sorted on the

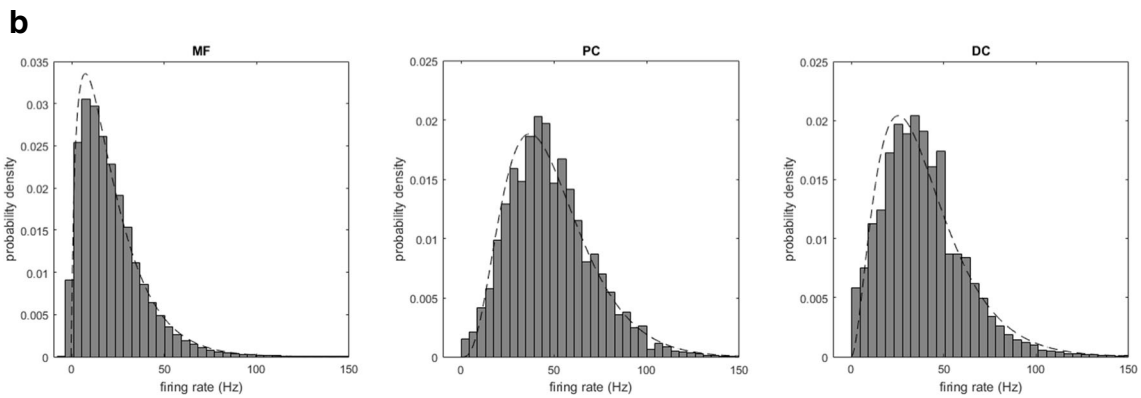
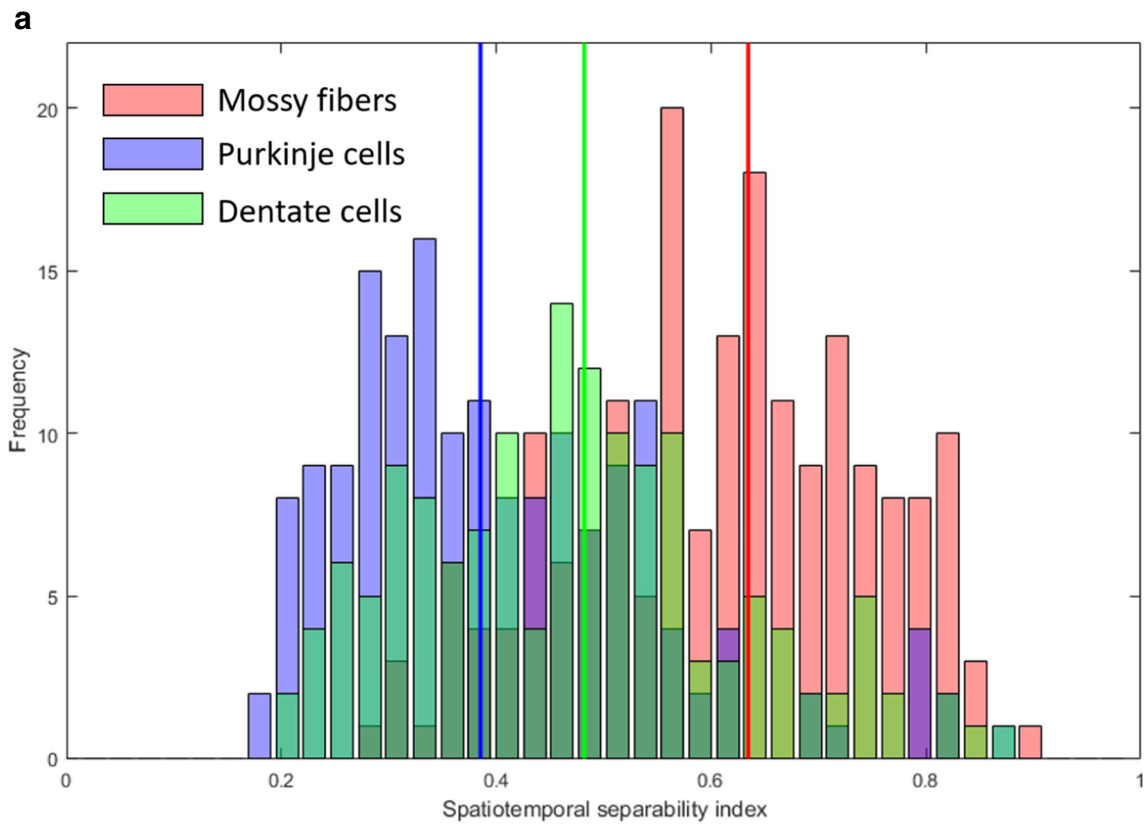
timing of movement onsets and binned into firing rates of a time window of 20 ms averaged over 10 to 20 trials for one condition. The dataset included 94 MFs, 83 PCs, and 73 DCs from monkey 1. PCs were identified by the coexistence of simple and complex spikes, and MFs were identified by the occurrence of a short positive–negative potential followed by a longer negative afterwave [14, 15]. DCs were identified, in addition to anatomical separation from the cerebellar cortex, by the characteristic of large negative–positive spike waveforms [14]. Note that only simple spikes were analyzed for PCs because the monkey had been trained for the task for years and no apparent improvement in task performance occurred during the recordings. These cells were recorded on different experimental sessions or days, so the firing rates but not spikes were analyzed in this study.

### Characterization of Firing Rates: Spatiotemporal Separability Index and Distributions

We here examine the spatiotemporal pattern of activities of single neurons. Specifically, we introduce first an index of spatiotemporal separability and then examine its distributions for the different neuron groups. Let  $\text{MF}_i^{p,d}(t)$  denote the firing rates at  $t$ -th time bin in  $[-1\text{ s}, +1\text{ s}]$  recorded from  $i$ -th MF for direction  $d$  and posture  $p$ . Similarly,  $\text{PC}_i^{p,d}(t)$  and  $\text{DC}_i^{p,d}(t)$  are defined for PCs and DCs, respectively. We first characterize the spatiotemporal patterns of activities of MFs, PCs, and DCs and introduce an index that quantifies how spatiotemporally separable the activities are. If a cell has stationary directional tuning, the firing rates are spatiotemporally separable as a product of a function of movement direction and a function of time (Fig. 1a). If directional tuning of a cell is not stationary but rather exhibits time-varying preferred direction, the firing rates are spatiotemporally nonseparable. For a given posture, the firing rates of  $i$ -th neuron are summarized into a matrix form as  $\mathbf{R} \in \mathbb{R}^{D \times T}$ , where  $D$  is the number of movement directions,  $T$  is the number of time windows of recording, and  $D \ll T$ . To level off the difference in firing rates, each row of the matrix  $\mathbf{R}$  is normalized to zero mean and unit variance.  $\mathbf{R}$  may be decomposed into a factorial form as

$$\mathbf{R} = \mathbf{U}\Sigma\mathbf{V}^\top = \sum_{d=1}^D \sigma_d \mathbf{u}_d \mathbf{v}_d^\top \quad (1)$$

Here  $\{\sigma_d\}_{d=1}^D$  are singular values of a descending order,  $\sigma_1 \geq \dots \geq \sigma_D$ .  $\{\mathbf{u}_d\}$  and  $\{\mathbf{v}_d\}$  are  $D$ -dimensional and  $T$ -dimensional orthonormal vectors characterizing directional tuning and temporal profile of firing rates, respectively. If directional tuning is invariant during experimental duration, then  $\mathbf{R}$  is expressed as a rank-one matrix (i.e.,  $\mathbf{R} = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^\top$ ). On the other hand, if directional tuning varies during movement



**Fig. 1 a** Distribution of the spatiotemporal separability index (STSI) computed for mossy fibers (MFs) (red), Purkinje cells (PCs) (blue), and dentate cells (DCs) (green). Three colored vertical lines depict the median values of the three cell populations. **b** Probability densities of firing rates

of MFs (left), PCs (middle), and DCs (right). The densities are binned into 50 bins. Dashed lines overlapped with the histograms represent best-fit Gamma distributions

duration, then  $\mathbf{R}$  cannot be a rank-one matrix and contains multiple rank-one matrices. Therefore, how well  $\mathbf{R}$  is reconstructed by a rank-one matrix is a candidate measure for characterizing properties of firing rates. Once the firing-rate matrix is decomposed into a sum of rank-one matrices, the degree of spatiotemporal separability is quantified by an index (spatiotemporal separability index, STSI),

$$STSI = \frac{\sigma_1^2}{\sum_{d=1}^D \sigma_d^2} \tag{2}$$

STSI takes a value ranging from 0 to 1; STSI is 1 if  $\mathbf{R}$  is spatiotemporally separable, whereas STSI takes a smaller value if  $\mathbf{R}$  contains multiple rank-one matrices. This index was computed for each cell.

To further characterize the firing rates, we also examined distributions of firing rates of the three populations. For each population, a histogram was constructed by counting the frequency of firing rates in all time bins of cells in that population both for pronated and supinated postures. These histograms were fitted with Gaussian, Gamma, Rayleigh, and inverse

Gaussian distributions using a maximum likelihood method, and the values of Akaike information criterion (AIC) were computed for these distributions.

### Linear Reconstruction of Firing Rates of PCs

We then attempted to reconstruct the firing rate as a weighted sum of those of preceding layers by exploiting the connectivity of the cerebellum (Fig. 1 inset). Namely, PCs receive inputs from MFs through granule cells and inhibitory interneurons (stellate cells and basket cells), and that DCs receive inhibitory and excitatory inputs from PCs and collaterals of MFs, respectively.

An activity of  $i$ -th PC of time  $t$  was reconstructed as a weighted sum of those of MFs at the same time as

$$\widehat{PC}_i^{p,d}(t) = \sum_{j=1}^{N_{MF}} w_{ij}^{MF \rightarrow PC} MF_j^{p,d}(t) \quad (3)$$

Here we used a notation that hatted firing rates were estimates whereas nonhatted variables were experimentally observed ones.  $(\mathbf{w}^{MF \rightarrow PC})_{ij} = w_{ij}^{MF \rightarrow PC}$  is the  $N_{PC} \times N_{MF}$  weight matrix, that is, for each PC, there are  $N_{MF}$  adjustable weights. Note that these weights cannot be interpreted directly as synaptic strengths between MFs and PCs because they are indirectly connected via granule cells as well as inhibitory interneurons. The weights should rather be interpreted not as anatomical but functional connections between MFs and PCs. The weights are assumed to be either positive or negative, reflecting the anatomical fact that, between MFs and PCs, there are excitatory (granule cells) and inhibitory interactions (Fig. 1 inset). The weight matrix was optimized so as to minimize a squared error between actual firing rate  $PC_i(t)$  and reconstructed firing rates  $\widehat{PC}_i(t)$  averaged over experimental duration and movement directions as

$$\frac{1}{2} \sum_d \sum_{-1 \leq t \leq 1} \left\{ PC_i^{p,d}(t) - \widehat{PC}_i^{p,d}(t) \right\}^2 \quad (4)$$

The optimization was performed for the two postures separately so that we investigated how the weights trained for one posture generalized to the other posture. Note that the reconstructed firing rates of a PC in one condition (i.e., movement direction and posture) used those of MFs in the same condition. For optimizing the weights, the entire experimental duration  $[-1 \text{ s}, +1 \text{ s}]$  was used. The fitting to the data was evaluated by computing the coefficient of determination ( $R^2$ ) in a time window of  $[-0.2 \text{ s}, +0.5 \text{ s}]$  because task-related modulation of firing rates occurred mainly in this time window.

### Linear Reconstruction of Firing Rates of DCs

It is known from the cerebellar anatomy that DCs receive excitatory inputs from MFs and inhibitory inputs from PCs, respectively. As in the case for PCs, firing rates of a DC were reconstructed as a weighted sum of those of MFs and PCs,

$$\widehat{DC}_i^{p,d}(t) = \sum_{j=1}^{N_{MF}} w_{ij}^{MF \rightarrow DC} MF_j^{p,d}(t) + \sum_{k=1}^{N_{PC}} w_{ik}^{PC \rightarrow DC} PC_k^{p,d}(t) \quad (5)$$

Collaterals of MFs send excitatory inputs and PCs send inhibitory inputs to DCs, so  $\{w_{ij}^{MF \rightarrow DC}\}$  and  $\{w_{ik}^{PC \rightarrow DC}\}$  were assumed to be nonnegative and nonpositive, respectively. For each DC, the linear model contained  $N_{MF} + N_{PC}$  weights, and the squared error between  $DC_i(t)$  and  $\widehat{DC}_i(t)$  was minimized under the nonnegative and nonpositive constraints. These weights were optimized using a standard quadratic programming algorithm.

Additionally, to investigate how the MF and PC firings contributed to the reconstruction of DC firings, we attempted to reconstruct the DC firings using only the MF firings as in

$$\widehat{DC}_i^{p,d}(t) = w_i^{MF \rightarrow DC} + \sum_{j=1}^{N_{MF}} w_{ij}^{MF \rightarrow DC} MF_j^{p,d}(t) \quad (6)$$

or using only the PC firings as in.

$$\widehat{DC}_i^{p,d}(t) = w_i^{PC \rightarrow DC} + \sum_{k=1}^{N_{PC}} w_{ik}^{PC \rightarrow DC} PC_k^{p,d}(t) \quad (7)$$

To explain spontaneous activities of DCs, the constant terms ( $w_i^{MF \rightarrow DC}$  in Eq. (6) and  $w_i^{PC \rightarrow DC}$  in Eq. (7)) were included. Values of the goodness-of-fit using both MF and PC firings (Eq. (5)), MF firings only (Eq. (6)), or PC firings only (Eq. (7)) were compared.

### Statistical Tests of Linear Reconstruction

To test statistical significance of linear reconstructions of PCs, two tests were performed. First, we assessed the goodness-of-fit of the linear model to the data by comparing other models. This was a model comparison *within* a posture. Two standard nonlinear models (a threshold model and a quadratic model) were fit to the data. One was a linear-threshold model where the MF firing rates are thresholded by zeroing the activity equal to or smaller than a threshold as

$$\widehat{PC}_i(t) = \sum_{j=1}^{N_{MF}} w_{ij}^{MF \rightarrow PC} [MF_j(t) - \theta_j]_+ \quad (8)$$

Here  $\theta_j$  denotes an activity threshold of  $j$ -th MF. The linear-threshold model is often used for modeling nonlinear amplification, multiplicative gain modulation, and winner-takes-all

selection [16]. In our context, the thresholding operation may be regarded as a nonlinear processing of granule cells between MFs and PCs. The other was a quadratic model which is analogous to an energy model of complex cells in the visual cortex [17]:

$$\widehat{PC}_i(t) = \sum_{j=1}^{N_{MF}} w_{ij}^{MF \rightarrow PC} MF_j(t) + \sum_{j=1}^{N_{MF}} v_{ij}^{MF \rightarrow PC} [MF_j(t) - \theta_j]^2 \tag{9}$$

This model has been used for modeling the properties of complex cells whose responses are orientation selective but not phase selective. The above models are instantaneous, i.e., the firing rate of a PC at time  $t$  is reconstructed with that of a MF at the same time  $t$ . In addition, we also considered a finite-impulse-response (FIR) model of order 1 defined as

$$\widehat{PC}_i^{p,d}(t) = \sum_{j=1}^{N_{MF}} w_{ij}^{(0) MF \rightarrow PC} MF_j^{p,d}(t) + \sum_{j=1}^{N_{MF}} w_{ij}^{(1) MF \rightarrow PC} MF_j^{p,d}(t-1) \tag{10}$$

The second term may be interpreted as a reverberation in the recurrent circuit composed of MFs, granule cells, and Golgi cells, so this FIR model is the simplest example of the adaptive filter model.

For a single PC cell, the linear, the threshold, and the quadratic models contain  $N_{MF}$ ,  $2N_{MF}$ , and  $3N_{MF}$  adjustable parameters, respectively. For each cell, the three models were compared according to the AIC. In the same way, the linear reconstruction model of DCs was tested with a linear-threshold model,

$$\widehat{DC}_i^{p,d}(t) = \sum_{j=1}^{N_{MF}} w_{ij}^{MF \rightarrow DC} [MF_j^{p,d}(t) - \theta_j] + \sum_{k=1}^{N_{PC}} w_{ik}^{PC \rightarrow DC} [PC_k^{p,d}(t) - \varphi_k] \tag{11}$$

and a quadratic model,

$$\widehat{DC}_i^{p,d}(t) = \sum_{j=1}^{N_{MF}} w_{ij}^{MF \rightarrow DC} MF_j^{p,d}(t) + \sum_{j=1}^{N_{MF}} w_{ij}^{-MF \rightarrow DC} [MF_j^{p,d}(t) - \theta_j]^2 + \sum_{k=1}^{N_{PC}} w_{ik}^{PC \rightarrow DC} PC_k^{p,d}(t) + \sum_{k=1}^{N_{PC}} w_{ik}^{-PC \rightarrow DC} [PC_k^{p,d}(t) - \varphi_k]^2 \tag{12}$$

The linear, linear-threshold, and quadratic models contain  $N_{MF} + N_{PC}$ ,  $2N_{MF} + 2N_{PC}$ , and  $3N_{MF} + 3N_{PC}$  parameters, respectively. As in Eq. (9),  $\{w_{ij}^{MF \rightarrow DC}\}$  and  $\{w_{ik}^{PC \rightarrow DC}\}$  were constrained to be nonnegative and nonpositive, respectively, and other parameters were unconstrained. These models were also compared according to the AICs.

Second, how the reconstruction of linear and other models learned at one posture generalized to those at the other posture. Here, we expected that a reconstruction model that appropriately describes the firing rates should reconstruct not only at a trained posture but also at an untrained posture. Hence, this was a model comparison *across* the two postures. The degree of generalization was evaluated for the four models of PC firing-rate fitting and for the three models of DC firing-rate fitting. Specifically, for one target cell, the weights optimized at one posture were used for a linear reconstruction at the other posture, and the goodness-of-fit between the linear reconstruction and the data was computed at the untrained posture. The goodness-of-fit of multiple models was statistically compared by one-way ANOVA.

### Statistical Comparison of MF–PC and MF–DC Projections

MFs project both to PCs via granule cells and parallel fibers and to DCs. Correspondingly, the linear equations we consider include the two terms  $\{w_{ij}^{MF \rightarrow PC}\}$  in Eq. (3) and  $\{w_{ij}^{MF \rightarrow DC}\}$  in Eq. (5), which represent functional projections from MFs and PCs and from MFs to DCs, respectively. Each projection was characterized by a column of  $\{w_{ij}^{MF \rightarrow PC}\}$  or  $\{w_{ij}^{MF \rightarrow DC}\}$ , which we refer to a projection vector. Similarity of projections was quantified by computing a correlation coefficient between two projection vectors. Specifically, we asked whether the two functional projections statistically differed. Correlation coefficients of all possible pairs of MF–PC projection vectors and MF–DC projection vectors were computed, and their average served as an index of similarity between the two projections. Note that the MF–PC projection vectors took either positive or negative values while the MF–DC projection vectors were nonnegative. Absolute values of the MF–PC projection vectors were used for the computation of correlation coefficients because only the magnitudes of projections were of our interest. To assess the statistical significance of the average of correlation coefficients, a resampling test was performed based on a null hypothesis that there was no statistical difference between MF–PC and MF–DC projections. Specifically, a bootstrap distribution of correlation coefficients was computed by randomly permuting the labels of projection vectors for 100,000 times. The probability of observing the average of correlation coefficients was assessed in terms of the bootstrap distribution.

### Linear Predictions of Future MF Activities from Current DC Activities

To test the forward-model hypothesis, we then explored to investigate whether the current outputs from the cerebellum



(i.e., DCs) could predict the future inputs to the cerebellum (i.e., MFs). Specifically, the firing rates of each MF at time  $t + t_1$  were reconstructed as a weighted sum of the firing rates of DCs at time  $t$ , as

$$\widehat{\text{MF}}_i(t + t_1) = \sum_{j=1}^{N_{\text{DC}}} w_{ij}^{\text{DC} \rightarrow \text{MF}} \text{DC}_j(t) \quad (13)$$

This linear prediction model contained  $N_{\text{DC}}$  weights for one MF. As in the linear reconstruction cases, the squared error between the predicted and the actual firing rates was minimized to compute the optimal values of the weights. The weights  $\{w_{ij}^{\text{DC} \rightarrow \text{MF}}\}$  cannot be interpreted as functional connectivity from DCs to MFs, as there are no anatomically direct connections between DCs and MFs. Rather, the linear prediction model was introduced to test if the current DC activity contained predictive information about the future MF activity.  $t_1$  is a parameter of time advance and was varied from 0 to 200 ms in steps of 20 ms (the window of time bins). The goodness-of-fit of the linear model was evaluated in a time window of  $[-200 \text{ ms}, 500 \text{ ms}]$ .

One may suspect that the linear prediction in Eq. (13) is possible if the terms of the right-hand side span a wide variety of waveforms, as in the case of Fourier expansion whereby any time series can be fit with a set of sinusoidal waveforms. To assess the fit statistically and to exclude that possibility, a bootstrap test was performed by shuffling the movement directions of DCs on the right-hand side of Eq. (13), while the movement directions of MFs on the left-hand side were fixed to those used experimentally. Randomizing the movement directions on the right-hand side maintained the diversity of waveforms but eliminated directional relationship between both sides of Eq. (13). The bootstrap test was designed based on a null hypothesis that any time series of the same degree of similarity could predict the future inputs to the cerebellum. With one sequence of shuffled targets, the goodness-of-fit was computed for all MFs and then averaged. This was repeated for 10,000 times to produce a bootstrap distribution of goodness-of-fit, and then the probability of the experimental goodness-of-fit was computed. This bootstrap test was performed with  $t_1$  ranging from 20 to 100 ms in step of 20 ms for the two postures, separately.

## Sparse Linear Analyses

Finally, to further investigate what connectivity structure underlies the linear reconstruction and prediction described above, a sparse linear analysis was performed. Specifically, we aimed to discover how many nonzero weights sufficed to reconstruct or predict neural firing patterns of a target cell. In addition to the squared errors between actual and

reconstructed firing rates, sparsity imposing terms were incorporated. Specifically, an L1 norm of a weight vector (i.e., a sum of absolute values of weight coefficients) was used by minimizing the following cost functions,

$$\frac{1}{2T} \sum_{-1 \leq t \leq 1} \left\{ \text{PC}_i(t) - \sum_{j=1}^{N_{\text{MF}}} w_{ij}^{\text{MF} \rightarrow \text{PC}} \text{MF}_j(t) \right\}^2 + \lambda \sum_{j=1}^{N_{\text{MF}}} |w_{ij}^{\text{MF} \rightarrow \text{PC}}| \quad (14)$$

$$\frac{1}{2T} \sum_{-1 \leq t \leq 1} \left\{ \text{DC}_i(t) - \sum_{j=1}^{N_{\text{MF}}} w_{ij}^{\text{MF} \rightarrow \text{DC}} \text{MF}_j(t) - \sum_{k=1}^{N_{\text{DC}}} w_{ik}^{\text{PC} \rightarrow \text{DC}} \text{PC}_k(t) \right\}^2 + \lambda \left( \sum_{j=1}^{N_{\text{MF}}} w_{ij}^{\text{MF} \rightarrow \text{DC}} - \sum_{j=1}^{N_{\text{MF}}} w_{ij}^{\text{MF} \rightarrow \text{PC}} \right) \quad (15)$$

$$\frac{1}{2T} \sum_{-1 \leq t \leq 1} \left\{ \text{MF}_i(t + t_1) - \sum_{j=1}^{N_{\text{DC}}} w_{ij}^{\text{DC} \rightarrow \text{MF}} \text{DC}_j(t) \right\}^2 + \lambda \sum_{j=1}^{N_{\text{DC}}} |w_{ij}^{\text{DC} \rightarrow \text{MF}}| \quad (16)$$

Here  $T$  denotes the length of data. Note that Eqs. (14) and (16) are the cost functions used in a standard sparse linear analysis known as LASSO [18]. Equation (15) slightly differs from LASSO because of the nonnegative constraint of MF  $\rightarrow$  DC weights and the nonpositive constraint of PC  $\rightarrow$  DC weights. The time advance parameter  $t_1$  in Eq. (16) was fixed to 40 ms in this sparse analysis. The parameter  $\lambda$  determines the tradeoff between the squared error and the sparseness and must be optimized for individual target cells. For each target cell, we varied  $\lambda$  in a range from 0.05 to 5 and chose the value that exhibited the smallest generalization error for test data in tenfold cross validation. The weight coefficients computed with the optimized  $\lambda$  were assessed in two ways: the proportion of nonzero weights and the proportion of significantly contributing weights. Here, the significantly contributing weights were defined by counting the number of weights whose cumulative sum exceeded 90% of the sum of total weights.

## Results

First, statistical characteristics of firing rates of the three populations were computed and compared (see “[Characterization of Firing Rates: Spatiotemporal Separability Index and Distributions](#)” in the “[Methods](#)” section). Then, the firing rates of PCs and DCs were reconstructed with linear weighted models of MF and PC firing rates (see “[Linear Reconstruction of Firing Rates of PCs](#)” and “[Linear Reconstruction of Firing Rates of DCs](#)” in the “[Methods](#)” section). The reconstructions of the linear models were statistically compared to those of nonlinear and FIR models by computing AICs within a single posture and the degree of generalization across two postures (see “[Statistical Test of Linear Reconstruction](#)”). Finally, the

forward-model hypothesis was tested by applying a linear-weighted model that predicted the MF activity at time  $t + t_1$  from the DC activity at time  $t$  (see “[Linear Prediction of Future MF Activities from Current DC Activities](#)” in the “[Methods](#)” section).

### Characterization of Firing Rates

The STSI, defined in Eq. (2), was introduced to assess the complexity of spatiotemporal patterns of firing rates of a given neuron (see “[Characterization of Firing Rates: Spatiotemporal Separability Index and Distributions](#)” in the “[Methods](#)” section). STSI exhibited clearly separable values for MFs and PCs (Fig. 1a). Our previous papers reported that task-related activities recorded from MFs showed a unimodal and directionally tuned phasic modulation around movement onset that was analogous to those observed in activities of the primary motor cortex (M1) [19–21], and that simple-spike activities of PCs showed dynamic and time-varying directional tuning before and after movement onset [14, 15]. Consistent with these observations, STSIs of MFs were significantly larger than STSIs of PCs (unpaired  $t$  test,  $p < 0.01$ ). The population of MFs had the largest value (median  $0.63 \pm 0.14$  SD), followed by the population of DCs (median  $0.48 \pm 0.15$  SD), and the population of PCs had the smallest values (median  $0.38 \pm 0.15$  SD). There was a statistically significant difference between STSIs of the three populations as determined by one-way ANOVA ( $F(2, 497) = 95.3$ ,  $p = 9.18 \times 10^{-36}$ ). A Tukey post hoc test revealed that the STSIs of MFs were statistically significantly larger than those of PCs ( $p = 9.6 \times 10^{-10}$ ) and those of DCs ( $p = 9.6 \times 10^{-10}$ ) and that the STSIs of DCs were statistically larger than those of PCs ( $p = 2.7 \times 10^{-5}$ ). Therefore, the firing rates of MFs exhibited spatiotemporally separable and simple characteristics, the firing rates of PCs exhibited spatiotemporally nonseparable and complex characteristics, and the firing rates of DCs exhibited characteristics intermediate between MFs and PCs.

The distributions of firing rates were computed for the three populations, respectively (Fig. 1b). They were unimodal (MF: mean  $21.1 \pm 17.6$  SD; PC: mean  $48.2 \pm 22.7$ ; DC: mean  $38.6 \pm 22.7$ ). These distributions were fitted with Gaussian, Gamma, Rayleigh, and inverse Gaussian distributions, which were compared according to the values of AICs. We found that the Gamma distribution provided the smallest values of AICs for the MF ( $1.26 \times 10^6$  (Gaussian),  $1.18 \times 10^6$  (Gamma),  $1.25 \times 10^6$  (Rayleigh),  $5.5 \times 10^6$  (inverse Gaussian)), PC ( $1.201 \times 10^6$  (Gaussian),  $1.191 \times 10^6$  (Gamma),  $1.194 \times 10^6$  (Rayleigh),  $1.216 \times 10^6$  (inverse Gaussian)), and DC firing rates ( $1.045 \times 10^6$  (Gaussian),  $1.023 \times 10^6$  (Gamma),  $1.025 \times 10^6$  (Rayleigh),  $1.052 \times 10^6$  (inverse Gaussian)). A Gamma distribution has a property that a sum of two independent random variables from a Gamma distribution

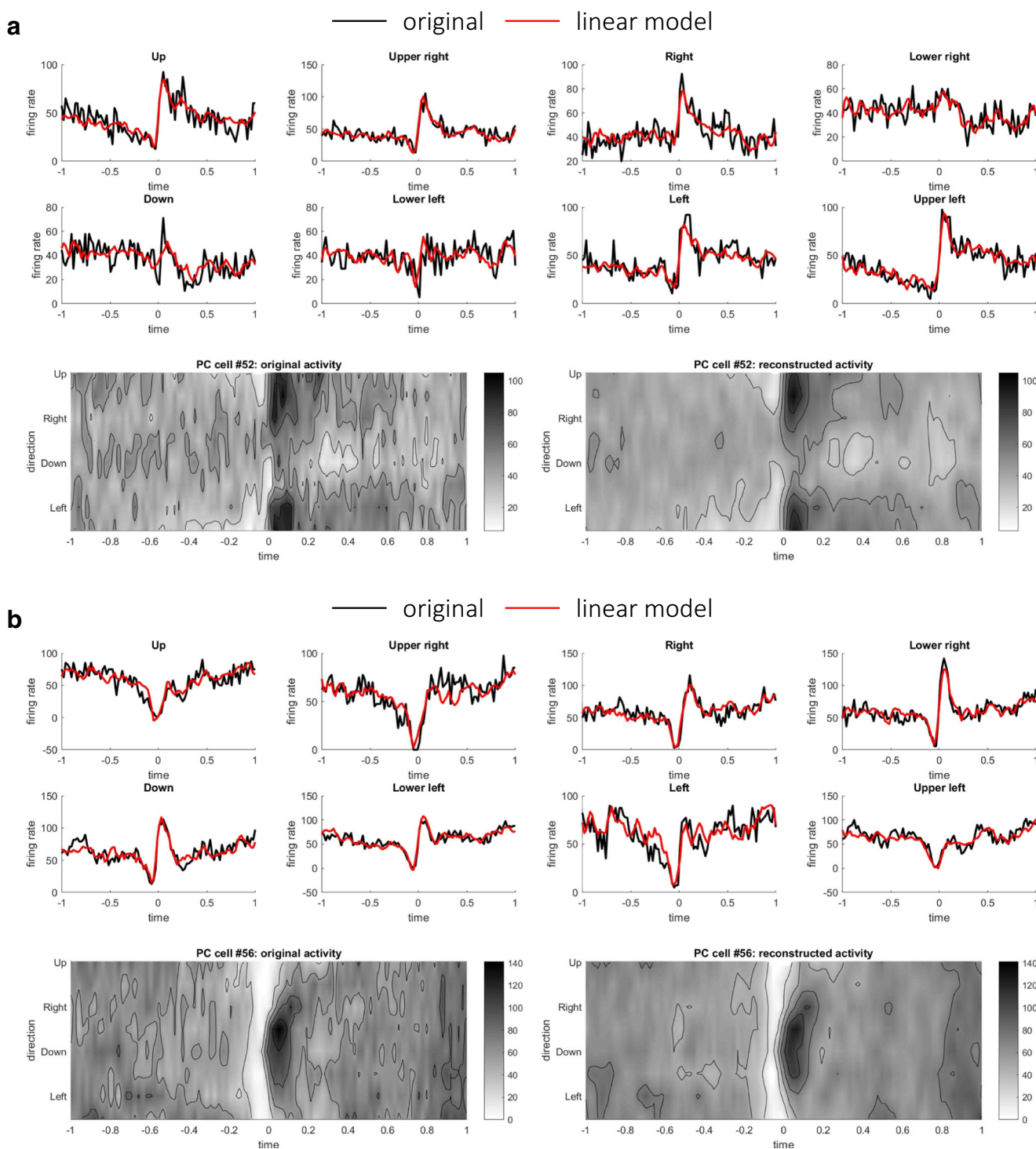
obeys also a Gamma distribution. Actual neural activities are correlated to each other, so the fact that the firing rates were all distributed as Gamma distributions does not necessarily support simple summation from one population to another, but rather imply some simple transformations among the populations.

The two analyses above found that the three populations had distinct degrees of spatiotemporal complexity as quantified in terms of STSIs, while the firing rates were all Gamma distributed. These results, taken together, led us to a hypothesis that the transformation from one population to another might be linear, so we went on to test the hypothesis by linearly fitting the firing rates of one population with those of preceding populations.

### Linear Reconstruction of Firing Rates of PCs and DCs

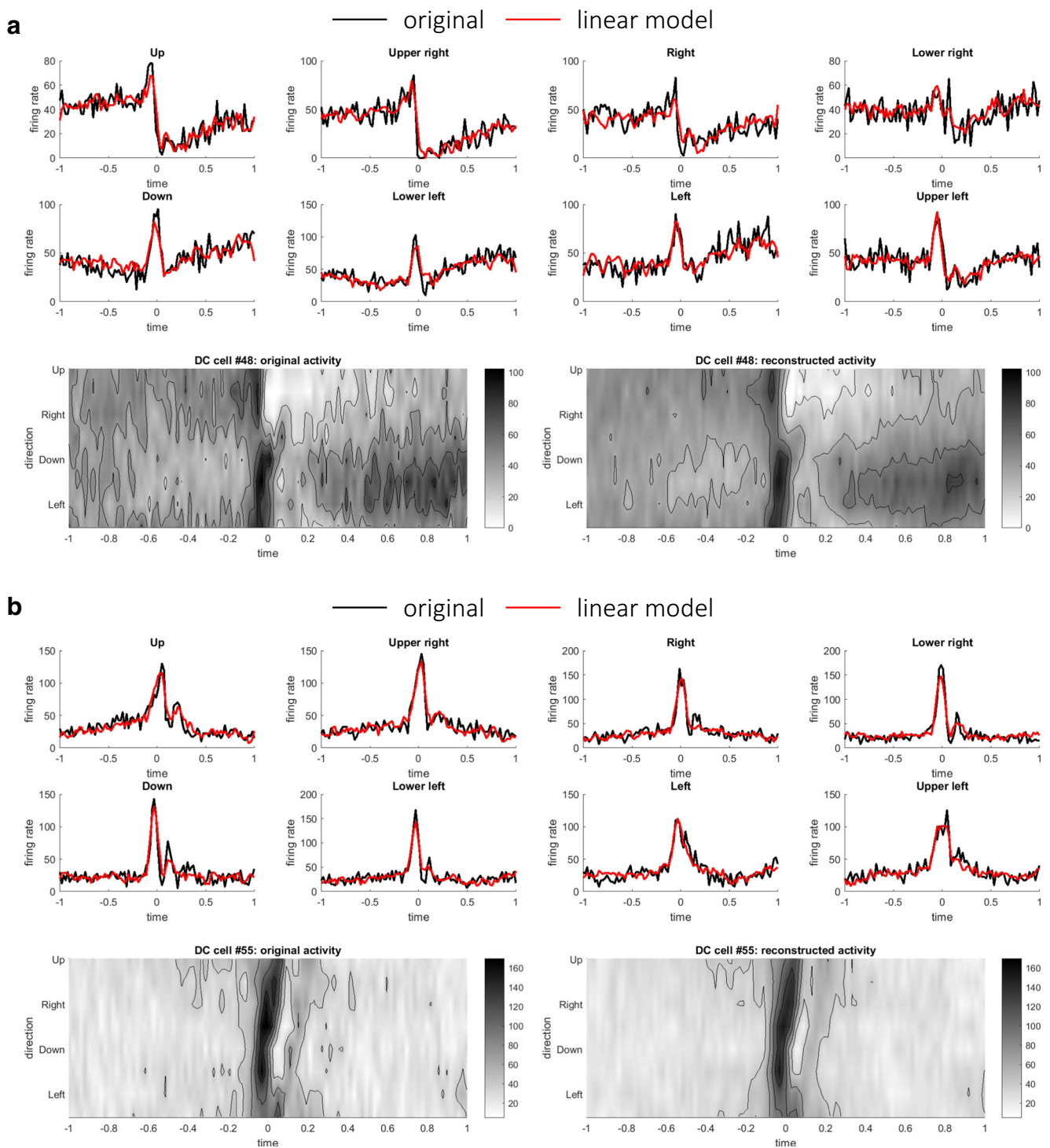
Based on the results of STSI and distributions of firing rates, we inferred that the firing rates of PCs and DCs might consist of weighted sum of the firing rates of input populations. We here attempted to reconstruct the firing rates of PCs as weighted sums of those of MFs using a linear model (3) (see “[Linear Reconstruction of Firing Rates of PCs](#)” in the “[Methods](#)” section). The weight parameters were optimized so that the squared error between actual firing rate  $PC_i(t)$  and reconstructed firing rates  $\hat{PC}_i(t)$  (Eq. (4)) was minimized separately for the pronated and supinated postures. The results of the two postures were similar; we present the results of the pronated posture below. Figure 2 illustrates time series and contour plots of the firing rates of two representative PCs that exhibited the highest  $R^2$  values between the original and reconstructed firing rates. In one PC (Fig. 2a), firing rates underwent a suppression just before the movement onset and an increase after the movement onset around movement directions 1 and 8, therefore exhibiting the reversal of its preferred direction. The reconstructed firing rates captured the reversal of preferred direction. In another PC (Fig. 2b), there was a uniform suppression of firing rates over the movement direction before the movement onset, followed by the emergence of directional tuning around directions 4 and 5 after the movement onset. As a population, the linear reconstruction model explained the original firing rates of PCs, as evidenced by  $R^2$  values for the pronated posture (mean  $0.95 \pm 0.023$  SD) and the supinated posture (mean  $0.96 \pm 0.023$  SD).

To further examine the hypothesis, we then proceeded to reconstruct the firing rates of DCs as linear weighted sums of MFs and PCs using Eq. (5) (see “[Linear Reconstruction of Firing Rates of DCs](#)” in the “[Methods](#)” section). The firing rates of DCs were reconstructed as weighted sum of excitatory (nonnegative) inputs from MFs and inhibitory



(nonpositive) inputs from PCs (Eq. (5)). Figure 3 depicts two representative DCs that exhibited the highest  $R^2$  values between the original and reconstructed firing rates. As in the PC

case, the firing rates of DCs were well reconstructed as linear sums of MFs and DCs. As a population, the median values of  $R^2$  of linear model were  $0.94 \pm 0.048$  SD for the pronated



**Fig. 3** Two representative examples of linear reconstructions of DCs: **a** DC cell #48 ( $R^2$  value 0.95) and **b** DC cell #55 ( $R^2$  value 0.96), in the same format of Fig. 2

posture and  $0.93 \pm 0.0420$  SD for the supinated posture. For a comparison, the DC firing rates were reconstructed using only the MF or PC firings minimizing Eq. (6) or (7), respectively. The median values of  $R^2$  using only MFs were  $0.92 \pm 0.069$  (pronated) and  $0.91 \pm 0.067$  (supinated), and the

median values of  $R^2$  using only MFs were  $0.91 \pm 0.060$  (pronated) and  $0.91 \pm 0.050$  (supinated). There was a statistically significant difference between the  $R^2$  values of the linear reconstructions using MF + PC, MF only, or PC only as determined by one-way ANOVA ( $F(2, 216) = 5.02, p = 7.4 \times$



$10^{-3}$  (pronated);  $F(2, 216) = 6.82$ ,  $p = 1.3 \times 10^{-3}$  (pronated)). A Tukey post hoc test revealed that the  $R^2$  values of MF + PC reconstructions were significantly better than those of MF reconstructions or those of PC reconstructions ( $p = 0.047$  for MF only and  $p = 0.008$  for PC only (pronated),  $p = 0.0098$  for MF only and  $p = 0.0018$  for PC only (supinated)).

### Comparison of AIC for Multiple Models of PCs and DCs

The quality of linear reconstruction indicated that the firing rates of MFs and PCs contributed linearly and instantaneously to those of PCs and DCs. Previous models of the cerebellum, on the contrary, suggest nonlinear or noninstantaneous models; the perceptron model assumes thresholding nonlinearity, and the adaptive filter model assumes finite impulse responses that depend on not only current inputs but also previous inputs. To verify whether the linear reconstruction sufficed or not, two simple nonlinear reconstruction models (a threshold model (8) and a quadratic model (9)) and an FIR model of order 1 (10) were fitted to the firing rates of PCs (see “Statistical Test of Linear Reconstruction”). The results of the two postures were similar; the results of the pronated posture are shown here. A typical PC reconstruction was presented in the format of time series (Fig. 4a) and contour plots (Fig. 4b). A visual inspection suggested that the reconstruction of the threshold model was indistinguishable from that of the linear model and that the reconstructions of quadratic model and FIR model were marginally better. The values of AIC are summarized in Fig. 4c and Table 1. As a population, AIC took the smallest value for the linear model and the second smallest value for the FIR model. For individual PCs, the values of AIC were minimum for the linear model for a majority of PCs (77 for pro and 71 for sup) and for the FIR model for a minority of PCs (6 for pro and 12 for sup). The thresholding model or the quadratic model was not selected for any PCs. Therefore, the linear model best captured the relationship between firing rates of MFs and PCs.

Similarly, for DC firing-rate fitting, the linear model fitting was statistically compared with those of two nonlinear models (the thresholding model (11) and the quadratic model (12)). The linear and nonlinear models were fitted to the firing rates of DCs, and AIC was computed for all DCs. A typical DC reconstruction was presented in the format of time series (Fig. 5a) and contour plots (Fig. 5b). The three models reconstructed the original activities at almost the same quality. The values of AIC are summarized in Fig. 5c and Table 2. The comparison of AIC indicated that the linear model could explain the DC firings most parsimoniously for all the DCs. Again, the linear model best described the firing rates of DCs with firing rates of MFs and PCs. In summary, the linear reconstruction model was selected for both reconstructions of PC and DC firing rates.

### Generalization Across Postures of Fitted Models of PCs and DCs

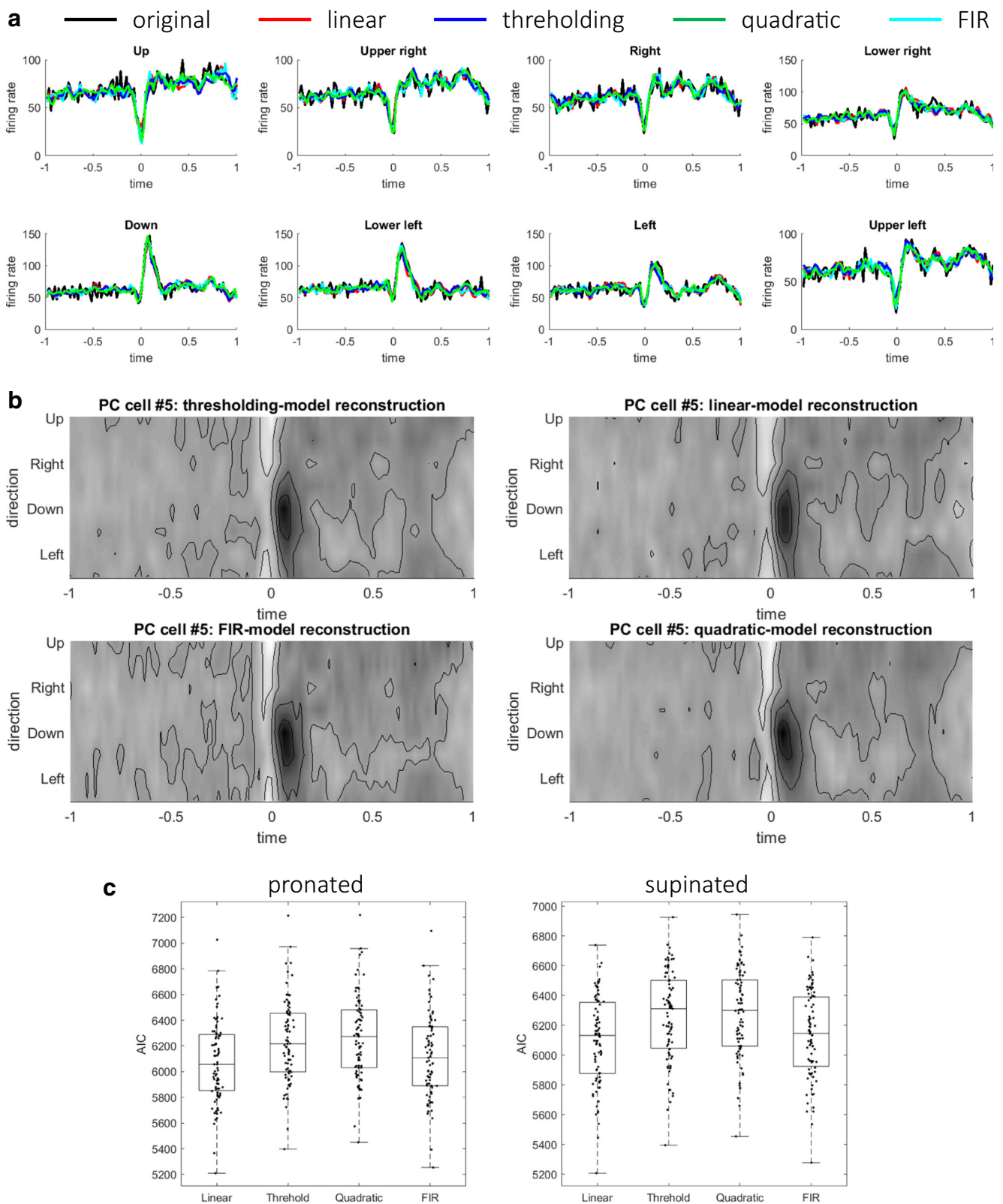
The AIC test was performed to select the most parsimonious model *within* a posture. We then proceeded to ask which model best generalized firing-rate fitting *across* two postures (see “Statistical Test of Linear Reconstruction”). Specifically, the weights trained in one posture were used to reconstruct the firing rates in the other posture that were not used for training. The degrees of generalization were summarized in box plots of Fig. 6a: (left panel) from supinated to pronated posture, and (right panel) from pronated to supinated posture, and Table 3. The goodness-of-fit was almost equal for the linear, threshold, and FIR models. In contrast, the goodness-of-fit was significantly worse for the quadratic model than those for the other models, indicating that the quadratic model overfit to the trained posture and did not generalize properly to the untrained posture. There was a statistically significant difference between the four models as determined by one-way ANOVA ( $F(3, 328) = 19.8$ ,  $p = 7.57 \times 10^{-12}$  (trained in pro and tested in sup);  $F(3, 328) = 20.9$ ,  $p = 2.1 \times 10^{-12}$  (trained in sup and tested in pro)). A Tukey post hoc test revealed that the goodness-of-fit of the quadratic model was significantly worse than the three models ( $p = 3.6 \times 10^{-9}$  for the linear model,  $p = 5.1 \times 10^{-9}$  for the thresholding model, and  $p = 1.3 \times 10^{-7}$  for the FIR model (trained in pro and tested in sup);  $p = 3.8 \times 10^{-9}$  for the linear model,  $p = 3.8 \times 10^{-9}$  for the thresholding model, and  $p = 1.4 \times 10^{-7}$  for the FIR model (trained in sup and tested in pro)). There was no significant difference between the linear, thresholding, and FIR models. In summary, the linear, thresholding, and FIR models generalized from one posture to another almost equally, whereas the generalization of the quadratic model was significantly worse than those of the other models.

Next, we investigated how the weights trained in one posture reconstructed the DC activities in another posture, as summarized in Fig. 6b (left panel) from supinated to pronated posture, and (right panel) from pronated to supinated posture and Table 4. The goodness-of-fit was almost equal for the three models. There was no statistically significant difference between the three models as determined by one-way ANOVA ( $F(2, 246) = 0.10$ ,  $p = 0.90$  (trained in pro and tested in sup);  $F(2, 246) = 2.5 \times 10^{-3}$ ,  $p = 0.9975$  (trained in sup and tested in pro)). Therefore, the degree of generalization from one posture to another was not different between the three models.

### Statistical Comparison of MF–PC and MF–DC Projections

In the linear equations derived from the experimental firing rates, there are two distinct projections from the MFs: from MFs to PCs ( $w^{\text{MF} \rightarrow \text{PC}}$  in (3)) and from MFs to DCs ( $w^{\text{MF} \rightarrow \text{DC}}$  in (5)) (see “Statistical Comparison of MF–PC and MF–





**Fig. 4** Comparison of Akaike information criterion (AIC) between the linear, threshold, quadratic, and FIR models for PC cell fitting. A typical example of model fitting to the firing rates of a PC (#52) shown in **a** time series and **b** contour plots. In **a**, the firing rates of the original, linear, threshold, quadratic, and FIR models are shown with black, red, blue,

green, and cyan lines, respectively. **c** Box plots of the goodness-of-fit for the four models at the pronated (left) and supinated (right) postures. On each box, the central mark is the median, and the edges of the box are the 25th and 75th percentiles, respectively

**Table 1** Summary of AIC statistics for PC firing-rate fitting

|                                 | Linear model                                 | Thresholding model                           | Quadratic model                              | FIR model                                    |
|---------------------------------|--|--|--|--|
| Median (STD)<br>for pro posture | $6.05 \times 10^3$<br>( $3.19 \times 10^2$ ) | $6.22 \times 10^3$<br>( $3.23 \times 10^2$ ) | $6.27 \times 10^3$<br>( $3.23 \times 10^2$ ) | $6.11 \times 10^3$<br>( $3.23 \times 10^2$ ) |
| Median (STD)<br>for sup posture | $6.10 \times 10^3$<br>( $2.99 \times 10^2$ ) | $6.26 \times 10^3$<br>( $3.02 \times 10^2$ ) | $6.27 \times 10^3$<br>( $2.97 \times 10^2$ ) | $6.14 \times 10^3$<br>( $2.97 \times 10^2$ ) |

**DC Projections**” in the “**Methods**” section). We asked whether a common population of MFs projected both to PCs and DCs or separate populations of MFs projected to PCs and DCs. The average of correlation coefficients between MF–PC and MF–DC projection vectors was 0.060. We assessed a statistical significance of this value of correlation by a re-sampling test with a null hypothesis that there was no statistical difference between the two types of projections. A 99% confidence interval was [0.1027, 0.1089], and the average value of experimental correlation coefficients was significantly small ( $p < 10^{-5}$ ). Therefore, the null hypothesis was rejected, and PCs and DCs did not receive projections from the same population of MFs.

### Linear Predictions of Future Inputs from Current Outputs

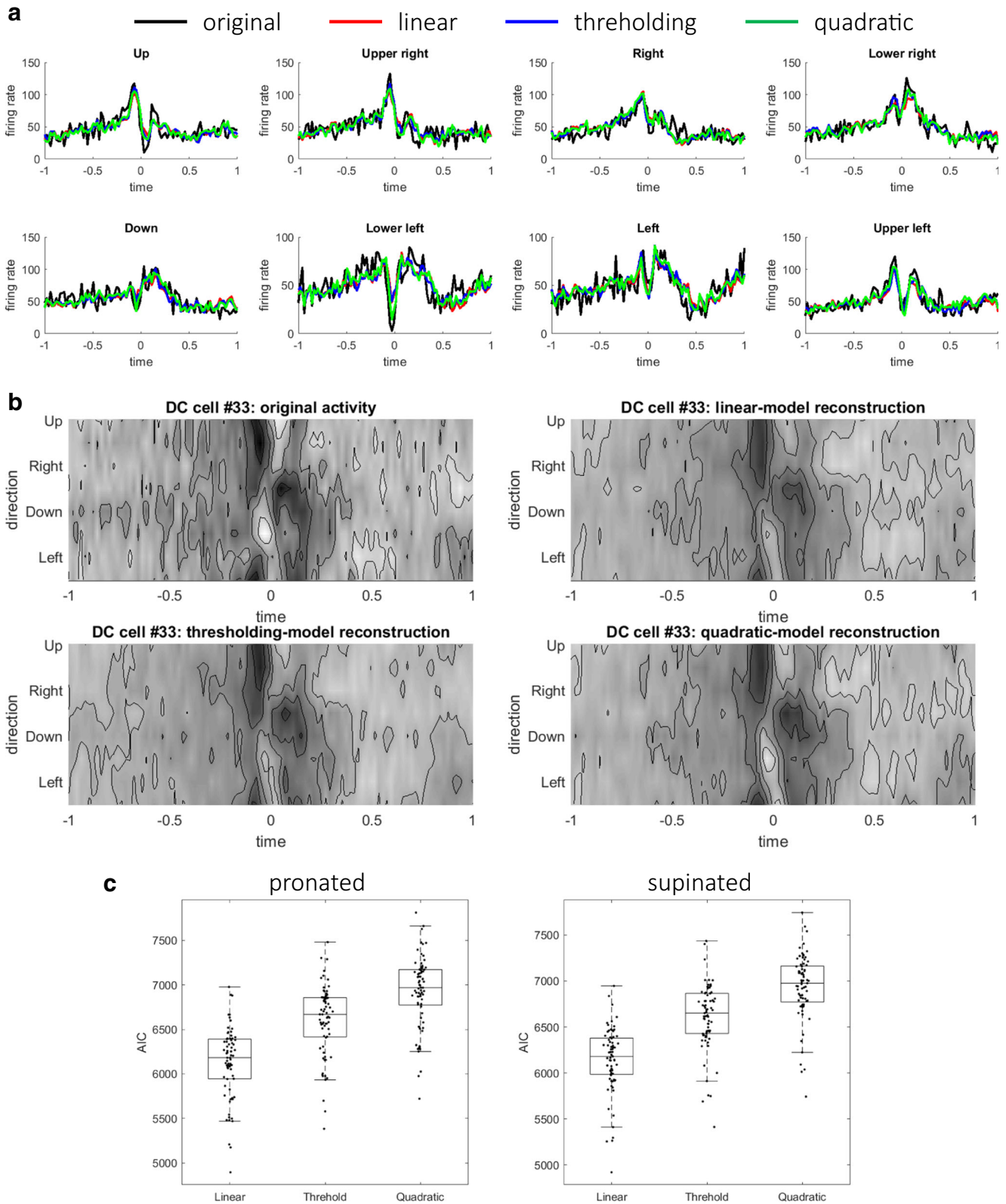
Finally, we tested the forward-mode hypothesis of the cerebellum by predicting future inputs to the cerebellum (MFs) at time  $t + t_1$  from current outputs from the cerebellum (DCs) at time  $t$  (see “**Linear Prediction of Future MF Activities from Current DC Activities**”). The time advance  $t_1$  was varied from 20 to 200 ms in steps of 20 ms, and the reconstructions with  $t_1 = 40$  ms (Fig. 7) and  $t_1 = 80$  ms (Fig. 8) were presented. When  $t_1 = 40$  ms, spatiotemporal patterns of firing rates of MFs were captured by linear predictions without any noticeable delay. When  $t_1 = 80$  ms, although fitting to peaks and troughs becomes less accurate, the overall patterns of firing rates were preserved. In fact, the goodness-of-fit decreased moderately when  $t_1$  was increased from 20 to 200 ms (Fig. 9). One may suspect that any time series of similar complexity could predict the future input reasonably, so we proceeded on to test whether the performance of the linear prediction was a statistical change or not. Statistical significance of the goodness-of-fit was assessed by a bootstrap test based on a null hypothesis that any time series of similar complexity could predict the future MF activities. A bootstrap distribution of goodness-of-fit was constructed by shuffling the movement directions of DCs on the right-hand side of Eq. (13) for 10,000 times. Intuitively, DC activities with shuffled movement directions were of the same complexity of those with the experimental directions but did not retain movement-

specific information. We found that the goodness-of-fit of the original data was significantly better than those of the shuffled data for  $t_1$  ranging from 20 to 100 ms for both the pronated ( $p < 5 \times 10^{-3}$  for all  $t_1$  ranging from 20 to 100 ms, Bonferroni corrected) and supinated ( $p < 5 \times 10^{-3}$  for all  $t_1$  ranging from 20 to 100 ms, Bonferroni corrected) postures. Therefore, the current output from the cerebellum contained predictive information about the future input to the cerebellum.

### Weight Distributions of Linear Reconstruction and Prediction Models

Finally, we computed the distributions of weights of the linear reconstruction models of PCs (3) and DCs (5) and the linear prediction model of MFs (13). For the MF  $\rightarrow$  PC connectivity, the weights exhibited an exponential distribution rather than a Gaussian (Fig. 10a), indicating that a relatively few MFs contributed dominantly to the reconstruction of each PC. For the MF  $\rightarrow$  DC connectivity, most of the weights from MFs to DCs were zero reflecting the nonnegative constraint, and nonzero weights were distributed exponentially (Fig. 10b, left panel). Similarly, most of the weights from PCs to DCs were zero reflecting the nonpositive constraint, and nonzero weights were distributed exponentially (Fig. 10b, right panel). Therefore, relatively small number of MFs and PCs contributed to the reconstruction of firing rates of each DC. Finally, for the linear prediction model from DCs to MFs, the weights were again distributed exponentially (Fig. 10c), although there appeared outliers at large values of weights on both positive and negative ends. In summary, the weight distributions were all exponential but not Gaussian, indicating some structured projections between the separate populations in the cerebellar circuits.

Motivated by these exponential distributions of the weights, we further investigated the characteristics of the weights that enabled the linear reconstruction and prediction (see “**Sparse Linear Analyses**”). Specifically, we imposed explicitly sparseness on weights and performed sparse linear analyses that minimized the cost functions composed of a sum of squared error and a sparsity term as in Eqs. (14), (15), and (16). There were two findings. Approximately 60–80% of the input cells had nonzero weights, and about one third of input cells exhibited significantly



**Fig. 5** Comparison of AIC between the linear, threshold, and quadratic models for DC cell fitting. A typical example of model fitting to the firing rates of a DC (#33) shown in time series (**a**) and contour plots (**b**). In **a**, the firing rates of the original, linear,

threshold, and quadratic models are shown with black, red, blue, and green lines, respectively. **c** Box plots of the goodness-of-fit for the three models at the pronated (left) and supinated (right) postures

**Table 2** Summary of AIC statistics for DC firing-rate fitting

|                              | Linear model                              | Thresholding model                        | Quadratic model                           |
|------------------------------|---|---|---|
| Median (STD) for pro posture | $6.18 \times 10^3$ ( $4.01 \times 10^2$ ) | $6.67 \times 10^3$ ( $4.01 \times 10^2$ ) | $6.97 \times 10^3$ ( $3.98 \times 10^2$ ) |
| Median (STD) for sup posture | $6.18 \times 10^3$ ( $3.62 \times 10^2$ ) | $6.65 \times 10^3$ ( $3.63 \times 10^2$ ) | $6.97 \times 10^3$ ( $3.59 \times 10^2$ ) |

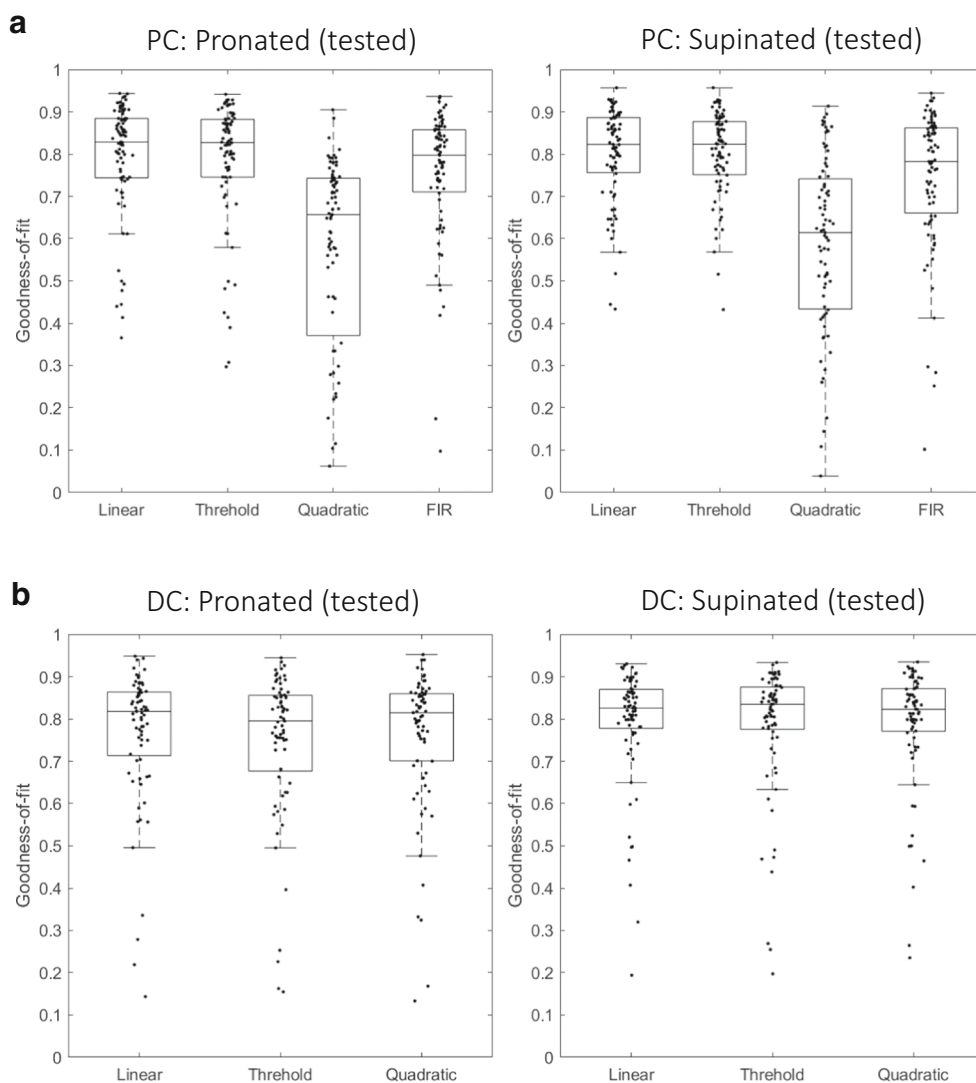
large weights (Table 5). This analysis revealed that, within our dataset, only a few dozens of input cells contributed mainly to the linear reconstruction and prediction of a target cell.

## Discussion

There are three main findings in this study. First, the distributions of firing rates of the three populations were all Gamma distributed, and they exhibited various degrees of spatiotemporal complexity. This indicated that the activities of the three populations represented functionally distinct roles in computation within the cerebellar circuit. Second, the firing rates of PCs are

reconstructed linearly as a weighted sum of the firing rates of MFs, and the firing rates of DCs are reconstructed linearly as a weighted sum of the firing rates of MFs and PCs. Finally, the firing rates of DCs at time  $t$  linearly predict the firing rates of MFs at time  $t + t_1$ , so the current output from the cerebellum contains predictive information about the future input to the cerebellum. These findings reveal the linear computation from one population to another and support the forward-model hypothesis of the cerebellum. It is worth mentioning that no nonlinearity that is expected from the perceptron model or dependence on previous inputs that is expected from the adaptive filter model was necessary to explain our data. Our results provide a strikingly simple picture of linear transformations for the cerebellar computation.

**Fig. 6** Generalization from fitting at one posture to another posture of multiple models. **a** Box plots of goodness-of-fit (left) trained in supinated posture and tested in pronated posture and (right) trained in pronated posture and tested in supinated posture for PCs. **b** Box plots of goodness-of-fit (left) trained in supinated posture and tested in pronated posture and (right) trained in pronated posture and tested in supinated posture for DCs



**Table 3** Summary of degrees of generalization across two postures for PC firing-rate fitting

|   | Linear model | Thresholding model | Quadratic model | FIR model   |
|---|--------------|--------------------|-----------------|-------------|
| Fig. 6a, PC: pronated (trained in supinated)<br>Median (SD)   | 0.82 (0.18)  | 0.82 (0.17)        | 0.61 (0.34)     | 0.78 (0.17) |
| Figure 6a, PC: supinated (trained in pronated)<br>Median (SD) | 0.83 (0.19)  | 0.83 (0.12)        | 0.66 (0.37)     | 0.78 (0.20) |

In the following, we branch out to discuss our findings in consideration of previous studies and speculate implications for the computation of internal models in the cerebellum.

We should remark that all the results in this study were obtained from one animal and could have reflected an idiosyncrasy of that animal, so the conclusion in this study must be confirmed with another animal in a future study.

### Previous Electrophysiological Studies

There are two kinds of internal models hypothesized for motor control: a forward model that performs a state prediction from a current estimate and an efference copy, and an inverse model that transforms a desired goal of movement into the necessary motor commands [22]. There has long been a controversy over whether the cerebellum functions as a forward model or an inverse model. Previous single-unit recording studies of PC activities during hand movements provided controversial results for the internal-model hypothesis of the cerebellum [23, 24]. These studies examined the correlation between activities of PCs and movement kinematics and/or dynamics. The underlying assumption was that kinematic and dynamic representations of PCs relate to forward and inverse models, respectively. As kinematic variables (e.g., hand trajectory) and dynamic variables (e.g., muscle activities) are highly correlated under unperturbed conditions, one approach is to make a monkey perform the same movement trajectory with different loads on the hand to dissociate dynamics from kinematics ([25, 26]. The monkey compensated the varying load in order to keep the same hand path, thereby dissociating the dynamics from the kinematics. For instance, Pasalar et al. [23] recorded simple-spike activities of task-related PCs while monkeys performed a circular manual tracking task under varying viscous and elastic loads. The simple-spike firing rates and spatial tuning did not change significantly under various load conditions, which supported a kinematic representation of arm movements in the cerebellar cortex. Their results appeared to

be compatible with the forward-model hypothesis of the cerebellum, which predicts movement kinematics.

Similarly, Yamamoto et al. [24] recorded simple-spike activities of PCs while monkeys performed elbow extension or flexion movements under assistive or resistive forces. In contrast with the findings of Pasalar et al., the simple-spike activities did change according to the load condition and correlated with the change in muscle activities, thereby seemingly consistent with the inverse-model hypothesis. Although the two studies examined PC firing rates in similar experiments, their conclusions were opposite to each other. Other studies described simple-spike activities correlated with eye-movement dynamics [27] or cursor-movement kinematics [28]. Therefore, to date, these single-unit-recording studies seem inconclusive about whether the cerebellum plays a role of an internal forward model or inverse model.

These studies rely on an assumption that kinematic and dynamic representations of PCs relate to forward and inverse models, respectively. This assumption, however, does not hold because an internal forward model should include dynamical variables such as efference copies of motor control signals. Also, disentangling predicted state signals, sensory feedback signals from the periphery, and motor commands is rather difficult because these signals resemble each other [8]. Therefore, a mere correlative comparison of PC activities with one or other behavioral parameters would not lead to conclusive evidence for either of forward or inverse models.

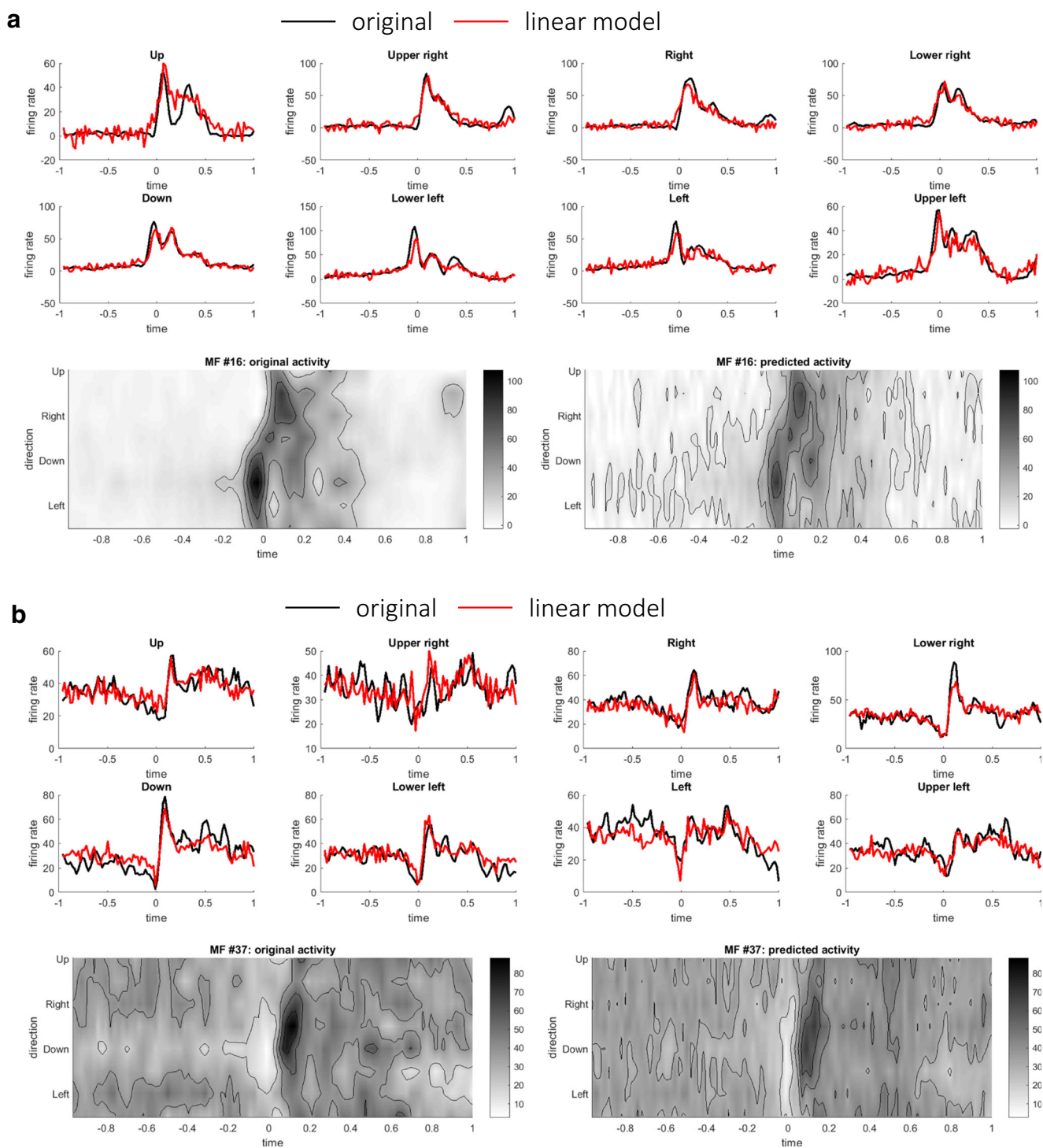
### State Prediction as a Prerequisite for a Forward Model

To resolve the limitation of the single-unit studies that correlated firing rates of one cell population and behavioral measures, we believe it essential to analyze network-level computation across multiple cell populations, as suggested by Wolpert and Miall [8]. The current study was designed to circumvent the abovementioned difficulty of disentangling multiple representations and targeted the transformation through the cerebellum

**Table 4** Summary of degrees of generalization across two postures for PC firing-rate fitting

|   | Linear model | Thresholding model | Quadratic model |
|---|--------------|--------------------|-----------------|
| Fig. 6b, DC: pronated (trained in supinated)<br>Median (SD) | 0.81 (0.29)  | 0.82 (0.29)        | 0.81 (0.29)     |
| Fig. 6b, DC: supinated (trained in pronated)<br>Median (SD) | 0.80 (0.29)  | 0.78 (0.29)        | 0.80 (0.29)     |



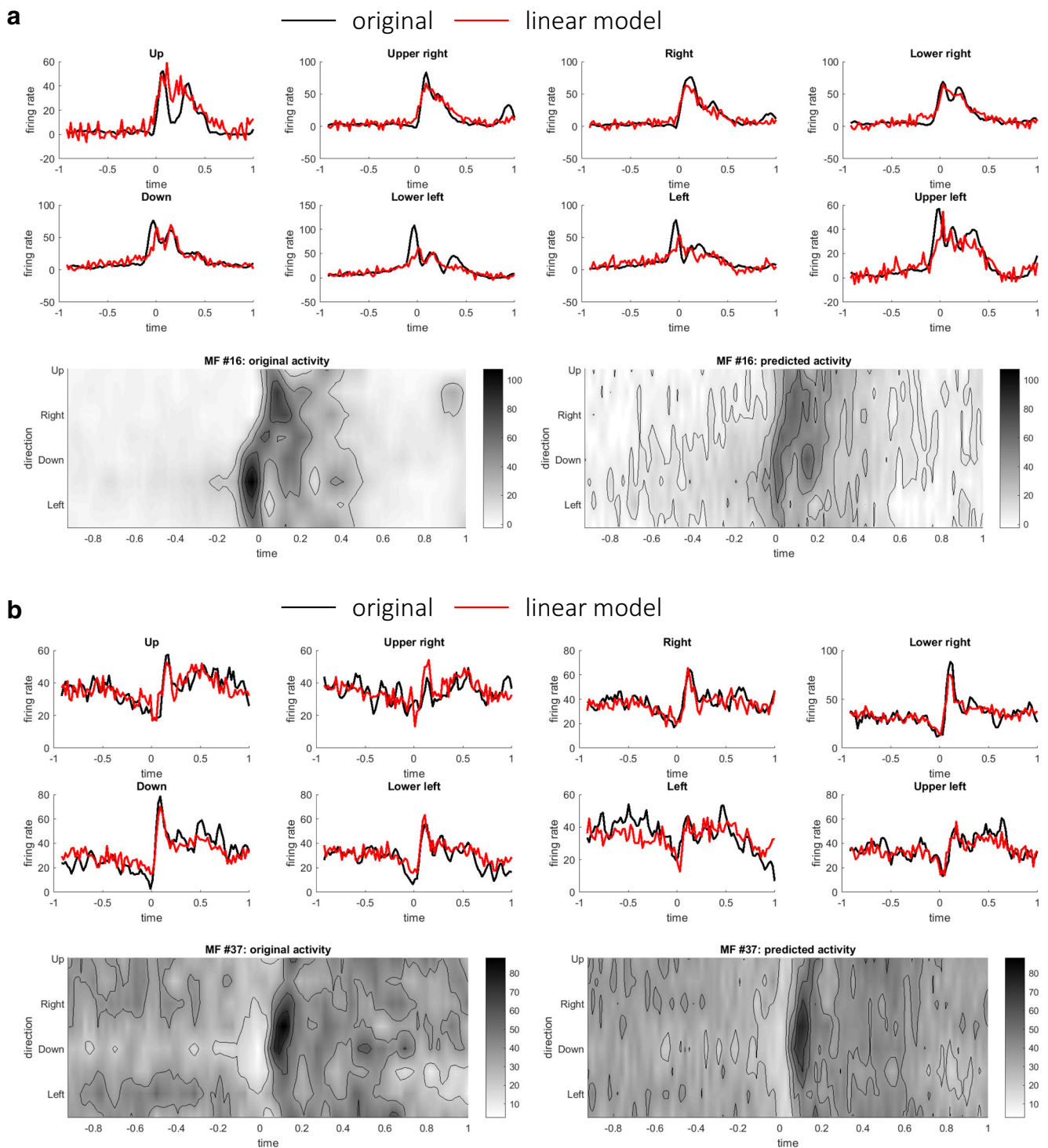


**Fig. 7** Two representative examples of linear predictions of MFs (#16 in **a** ( $R^2$  value 0.91) and #37 in **b** ( $R^2$  value 0.93)) at time  $t + t_1$  from DCs at time  $t$ . Here  $t_1$  was set to 40 ms. In each panel, the original activities and the predicted activities were compared in terms of time series and contour plots

from MF (input to the cerebellum) to DC (output from the cerebellum) via PC (output from the cerebellar cortex to the cerebellar nuclei), revealing the linear computation from one cell population to another. Furthermore, we found that the current output of the cerebellum predicts the future inputs to the cerebellum, a distinctive feature of an internal forward model. These findings

could not have been achieved with the analysis of correlation between activities of one cell population and behavioral measures.

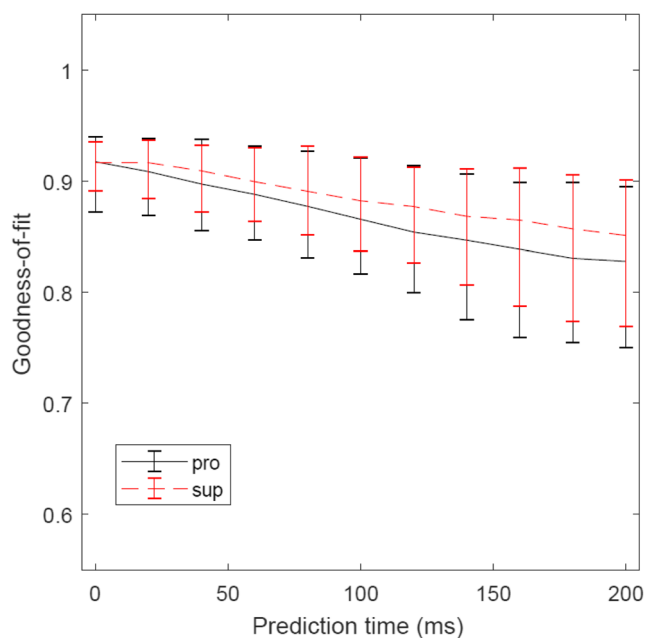
A critical test of the forward-model hypothesis of the cerebellum is whether the prediction performed in the cerebellum can offset delayed sensory feedback. Delays in sensory



**Fig. 8** Two representative examples of linear predictions of MFs at time  $t + t_1$  from DCs at time  $t$ . These are the same MFs (#16 in **a** ( $R^2$  value 0.85) and #37 in **b** ( $R^2$  value 0.89)) presented in Fig. 7. Here  $t_1$  was set to 80 ms

feedback can differ from one sensory modality to another; proprioceptive feedback takes of the order of 50 ms from muscle spindles to the somatosensory and motor cortices, and visual feedback takes about 50 ms from the retina to the primary visual cortex and 100 ms to the higher visual cortices [29]. These delays can deteriorate the performance of rapid

movements of the order of a few hundred milliseconds employed in this study. Our analysis revealed that the current DC activity contained predictive information about the future MF activity for a range of time advance. Therefore, our result supports that the cerebellum is capable of compensating the sensory delays of the order of 100 ms, supporting the forward-

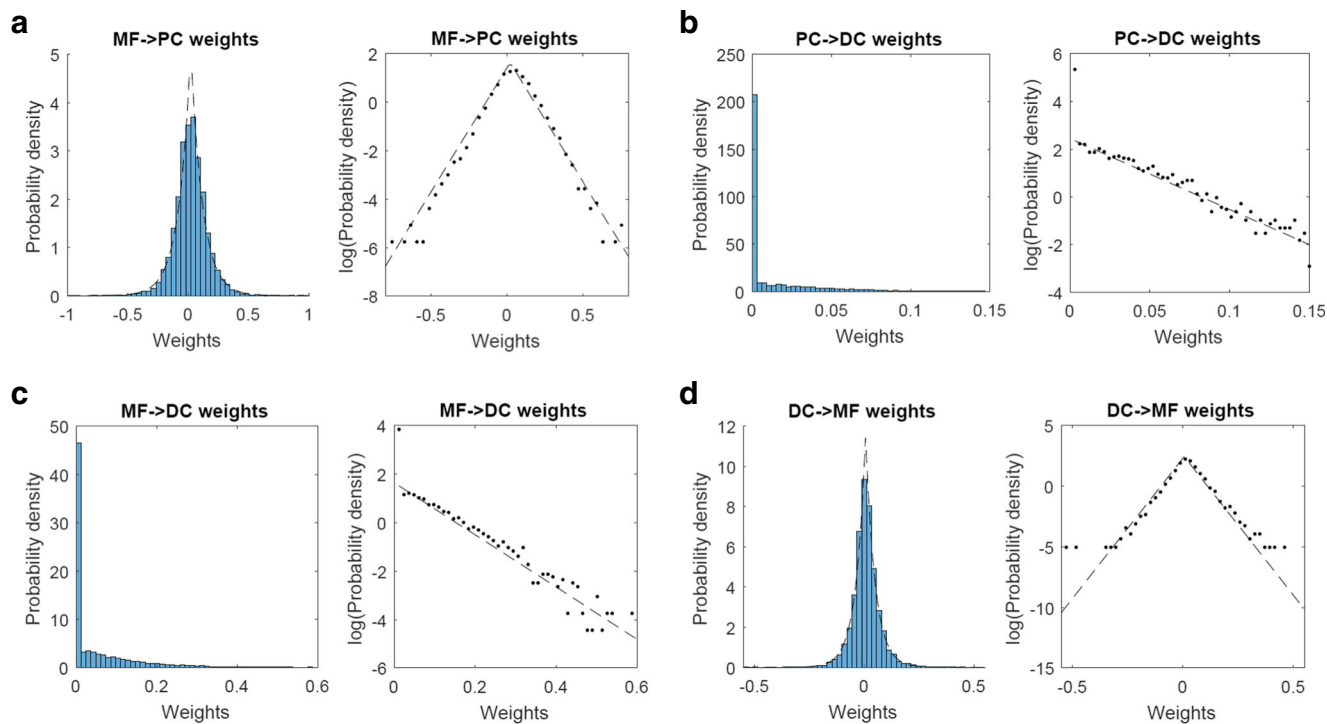


**Fig. 9** Goodness-of-fit of linear predictions with an increasing time advance  $t_1$  ranging from 0 to 200 ms with an interval of 20 ms. Error bars indicate standard deviations at each time advance. The goodness-of-fit was computed separately for the two postures: pronated (black solid line) and supinated (red dashed line)

model hypothesis. A previous electrophysiological study reported that activity of postcentral neurons changed on the

average of about 60 ms before the onset of agonist elbow muscles in voluntary elbow movements [30]. The early onset of activity of postcentral neurons is within the timescale of prediction in the cerebellum found in this study.

Morphological and physiological evidence accumulated over decades suggests that a region of the cerebro-cerebellum that forms a closed-loop circuit with M1 appears to satisfy the basic requirements for a forward model that generates a prediction of the outcome of a motor command [31]. First, this region of the cerebro-cerebellum receives a putative efference copy as well as a strong somatosensory input [14, 15], and these inputs are presumed to be integrated in the cerebellar cortex. Second, the activities of PCs in this region lag behind those of M1 neurons, while they precede the movement onset [15]. The timing of activity is compatible with the idea that it works as a forward model that predicts an outcome of the motor command. As a result, the output of this region of the cerebro-cerebellum may help M1 to generate a suitable motor command for the next moment depending on the predicted consequence before a feedback signal is available for the current motor command. We note that there are in general two input pathways to the MFs: one from the cerebral cortex through the pons and another from the peripheral sensory organs. Our single-unit recording did not allow to identify the origin of MFs and thus to discuss what information the MF activities encoded.



**Fig. 10** Distributions of weights of linear reconstruction models: **a** MF  $\rightarrow$  PC connectivity, **b** PC  $\rightarrow$  DC, and **c** MF  $\rightarrow$  DC connectivity. Note that the PC  $\rightarrow$  DC weights are nonpositive, and the MF  $\rightarrow$  DC weights are nonnegative. The signs of PC  $\rightarrow$  DC weights were flipped for a visual presentation. **d** DC  $\rightarrow$  MF weights of linear prediction model

trained with the time-advance parameter  $t_1 = 40$  ms. These distributions were normalized as probability density functions and were plotted in linear (left) and logarithmic (right) scales. Dashed lines indicate exponential distributions best fitted to the experimental distributions

**Table 5** Summary of weight characteristics obtained in sparse linear analyses

|                | Proportion of nonzero weights | Proportion of significantly contributing weights |
|----------------|-------------------------------|--|
| PC (pronated)  | 0.60 (0.14)                   | 0.32 (0.082)                                     |
| PC (supinated) | 0.68 (0.13)                   | 0.36 (0.096)                                     |
| DC (pronated)  | 0.82 (0.34)                   | 0.38 (0.26)                                      |
| DC (supinated) | 0.73 (0.354)                  | 0.37 (0.14)                                      |
| MF (pronated)  | 0.77 (0.10)                   | 0.35 (0.063)                                     |
| MF (supinated) | 0.81 (0.12)                   | 0.38 (0.067)                                     |

Our analyses assume the feedforward anatomical structure of the cerebellar circuit, but it is known that the cerebellar circuit contains recurrent anatomical connections that form a closed-loop circuit within the cerebellum, such as those composed of Golgi and granule cells and those composed of Purkinje and basket cells. Among these recurrent connections, the most relevant to this study is the nucleocortical projections from the cerebellar nuclei to the granular layer as MFs [32]. A recent study reported that excitatory output cells in the interposed nucleus provide efference copy signals via MFs to the cerebellar cortical zones and that an eye-blink conditioning training increased the local density of nucleocortical MF terminals [33]. One may suspect, therefore, that the linear prediction from DCs to MFs reported in this study could be attributed to the nucleocortical recurrent connections. We note, however, that this nucleocortical projection per se does not explain the longer time scale of linear prediction up to 100 ms reported in this study. Also, the nucleocortical pathway comprises only approximately 5% of the total of cerebellar MF inputs [34]. We hence expect that recurrent connections in the nucleocortical projections have minor contributions to our results of linear prediction from DCs to MFs.

### Linear Computation in the Cerebellar Circuit

We have shown the linear transformations from MFs to PCs and from MFs and PCs to DCs explained the observed firing rates recorded during the wrist movement task. In addition, the future MF activities were linearly predicted from the current DC activities. The success of linear modeling of PC and DC activities was unexpected and intriguing for the following three reasons. First, a PC receives parallel fiber inputs of the order of 100,000, while our dataset from monkey 1 contained only 94 MFs. Second, these firing rates were recorded at different sessions or even across different recording days separated by years, so the firing rates of the three populations were of no direct causal relation. Nonetheless, the computation in the cerebellar circuit turned out to be linear.

A missing piece in our study is the granule cell activity. Our results demonstrated that the transformation from MFs to PCs is linear, implying another linear computation in the granular layer. Two possibilities of computation in the granular layer are suggested. The first possibility is that each granule cell performs linear computation by linearly summing up the inputs from MFs. A previous study reported linear computation from MFs to medium ganglion cells in the cerebellum-like structure of electric fish [35]. They reported that linear weighted sums of sparsely and randomly mixed MF inputs reconstructed the membrane potentials of granule cells. The reconstructed granule cell activities exhibited a rich repertoire of temporal bases, which in turn constitute a negative image of sensory inputs. Their study suggests the linear computation from MFs to granule cells, in line with our results. It is interesting to note that, in their study, sparsity of linear weights was explicitly incorporated into an error function. In contrast, our results demonstrated that the sparse distributions of weights emerged spontaneously without a sparseness term in the cost function. Another possibility is that each granule cell performs nonlinear computation of MF inputs and the population of granule cells as a whole encodes inputs linearly. Recent studies demonstrated that individual granule cells were more narrowly tuned to the whisker angle of a rat than EPSC, thereby exhibiting nonlinear computation of granule cells sharpening their inputs [36], while individual PCs encoded whisker position linearly [37]. Interestingly, the population of narrowly tuned granule cells provides a linear excitatory drive across a range of whisker positions to PCs [36]. Because our dataset does not contain granule cell activities, we are not certain which may be the case.

Despite the fact that there are interneurons with recurrent connections in the cerebellar circuit, our finding indicates that the computation in the cerebellum is unexpectedly linear. We here speculate two possible explanations for the success of our linear modeling. One reason is that the performance of the monkeys was stable because they have been trained over years for this wrist movement task. Therefore, we expect that the response properties of



cerebellar cells remained stable across experimental sessions once the monkeys had achieved stable task performance. Another reason is that the cells in the dataset were selected if they showed task-related modulations of firing rates. Among numerous parallel-fiber inputs to a PC, it is conceivable that only a fraction of task-related inputs determines the response properties of that PC, as revealed by the sparse linear analyses. The stability of task performance and the selected sampling of task-related cells could explain the success of our linear modeling.

In line with our findings of linear transformations and predictions, an increasing number of recent literatures have reported that firing rates of cerebellar cells encode movement-related parameters linearly [38]. In the case of MFs [39, 40], Laurens et al., for example, reported a linear monotonic relationship between firing rate of MFs and eye position [39]. Similarly, for PCs [27, 37, 41–47], Hong et al. found that regularly firing spikes perform linear encoding of eye movement velocity by firing rate [43]. Finally, neurons in the cerebellar nuclei [48–51], the timing, and kinematics of motor output were modulated by linearly graded disinhibition of neurons in the deep cerebellar nuclei [49]. Along with these reports of linear encoding of behavioral parameters in the cerebellar cells, our findings reinforce the perspective of linear computation within the cerebellar circuits. Whereas there is evidence for linear encoding in the cerebellum as referenced above, we note that there is also evidence for nonlinear coding of saccade onset timings by spikes of PCs that are related to the period of occasional pauses [43], suggesting a possibility of multiplexed encoding by cerebellar cells.

## Linear Equations and Interpretation as Kalman Filter

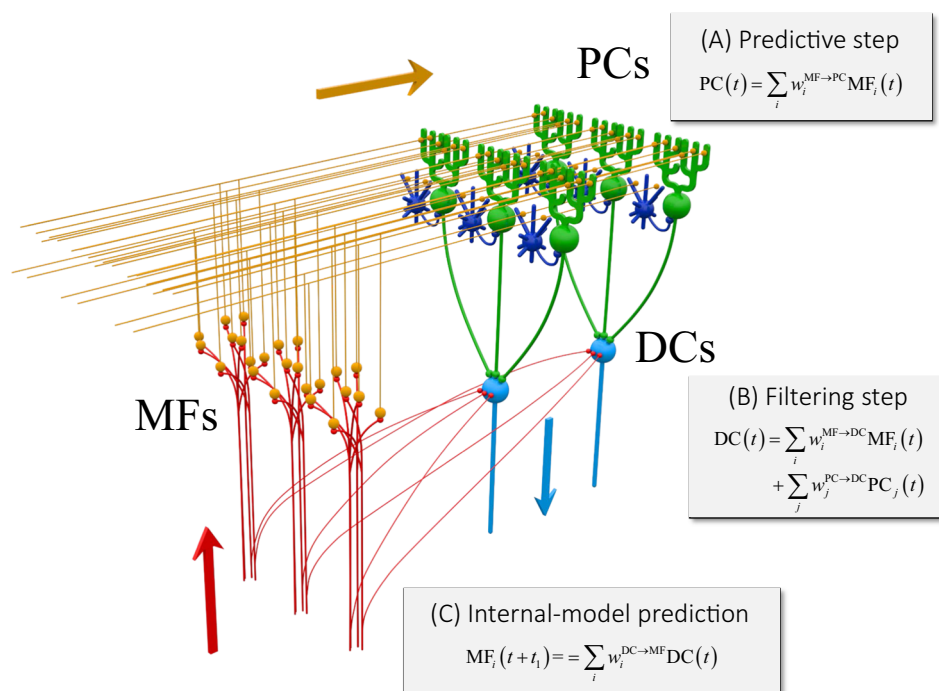
The three linear equations derived from the firing rates are summarized as follows,

$$\begin{aligned}\widehat{PC}_i(t) &= \sum_{j=1}^{N_{MF}} w_{ij}^{MF \rightarrow PC} MF_j(t) \\ \widehat{DC}_i(t) &= \sum_{j=1}^{N_{MF}} w_{ij}^{MF \rightarrow DC} MF_j(t) + \sum_{k=1}^{N_{PC}} w_{ik}^{PC \rightarrow DC} PC_k(t) \\ \widehat{MF}_i(t+t_1) &= \sum_{j=1}^{N_{DC}} w_{ij}^{DC \rightarrow MF} DC_j(t)\end{aligned}\quad (17)$$

and as in the schematics (Fig. 11). The activities of PCs are linear summations of activities of MFs, and the activities of DCs are linear summations of activities of MFs and PCs. There are two functional projections from MFs in the linear equations: one from MFs to PCs and the other from MFs and DCs. The two projections demonstrated little overlap, indicating that these two projections might convey functionally distinct information.

We realized that the chain of linear equations of neuron activities resembles those of an estimator known as Kalman filter. If we assume that MFs represent a current estimate of state and sensory feedback, PCs represent a prediction of state, and DCs represent a filtered state, then the linear equations can be interpreted as Kalman filter as follows. The first equation is a prediction step ((A) in Fig. 11); a current estimate (MFs) is projected to a predicted state (PCs). Then the second equation is a filtering

**Fig. 11** Summary schematic of our findings overlaid on the cerebellar circuit. MF, mossy fiber (red); PC, Purkinje cell (green); DC, dentate cell (light blue). Granule cells (orange) and inhibitory interneurons (blue) that are not analyzed in this work are included to show the basic structure of the cerebellar neuron circuitry. Three stages of linear computation obtained in our analysis are accompanied with the three types of computation of Kalman filter explained in the main text





step; a predicted state (PCs) and a sensory feedback (MFs) are integrated into a filtered state (DCs) ((B) in Fig. 11). Then, a filtered state (DC) contained predictive information about the future inputs (MFs) ((C) in Fig. 11).

In the above analogy with the Kalman filter, we have made an important assumption: PCs and DCs receive information about current estimate of state and sensory afferent signals through MFs, respectively. Our statistical test suggests that MF projections to PCs differ from MF collateral projections to DC in terms of their sources. It is known for the cerebro-cerebellum that MFs originate from the pons (which receives direct cortical projections) and from the brainstem/spinal cord, and that each of the MF inputs reflects either cortical activities or sensory feedback signals. Anatomical studies using an anterograde tracer (HRP-WGA) in cats reported that MFs of cortical origins via pons consist of the main input to the cerebellar cortex, whereas collaterals of those MFs poorly project directly to the dentate nuclei [52, 53]. Therefore, “the bulk of information of cortical origin reaches the cerebellar nuclei only after processing in the cerebellar cortex” (p. 22 of Brodal et al. [52]). This is consistent with our assumption that DCs receive a predicted state not from MFs but from PCs.

One critical aspect of Kalman filter is the Kalman gain which balances a predicted state and an observed state in an optimal way. In Kalman filter, a filtered state  $\hat{\mathbf{x}}_{t|t}$  combines a predicted state  $\hat{\mathbf{x}}_{t|t-1}$  and an observed state  $\mathbf{z}_t$  as in

$$\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + \mathbf{K}(\mathbf{z}_t - \mathbf{C}\hat{\mathbf{x}}_{t|t-1}) = (\mathbf{I} - \mathbf{K}\mathbf{C})\hat{\mathbf{x}}_{t|t-1} + \mathbf{K}\mathbf{z}_t \quad (18)$$

Here  $\mathbf{K}$  is the Kalman gain and  $\mathbf{C}$  is the observation matrix. By comparing the second equation of (17) and Eq. (18), we speculate that the weights from PC to DC ( $w^{\text{PC} \rightarrow \text{DC}}$ ) and the weights from MF to DC ( $w^{\text{MF} \rightarrow \text{DC}}$ ) correspond to the matrices  $\mathbf{I} - \mathbf{K}\mathbf{C}$  and  $\mathbf{K}$  in Eq. (18), respectively. While these weights were assumed to be stable and constant in our analysis because the task performance of the monkey was unchanged for years, the analogy predicts an opposing plasticity of  $w^{\text{PC} \rightarrow \text{DC}}$  and  $w^{\text{MF} \rightarrow \text{DC}}$ , namely,  $w^{\text{PC} \rightarrow \text{DC}}$  and  $w^{\text{MF} \rightarrow \text{DC}}$  should change their strengths in opposite directions when learning occurs. Although the analogy between the cerebellum and Kalman filter presented here is a speculation, we believe that this analogy could serve as a computational proposal that drives future studies of the cerebellum.

## Previous Computational Models of the Cerebellum

There are lines of computational models of the cerebellum in the literatures. The pioneering and most dominant model of the cerebellum is the perceptron model of the cerebellar cortex by Marr [54] and independently by Albus [55]. The perceptron model was first inspired by the analogy of feedforward network structures between the

cerebellar cortex and perceptron. The core hypothesis was that two independent inputs to a PC (MFs and a climbing fiber) represent input pattern signals and supervised error signals, respectively. Later, the climbing fiber inputs were found to induce long-term depression in synapses between parallel fibers and a PC in rabbits' cerebellar slices [56, 57]. The perceptron model contains a non-linear term to threshold a weighted sum of inputs. On the contrary, our results have shown that the linear model sufficed to explain the firing rates of PCs in terms of MFs.

The perceptron model is essentially a static pattern classifier and not designed to handle time-varying, dynamic inputs. Subsequently, the perceptron model was extended to the adaptive filter model which generates a dynamic response by summing various temporal basis patterns [58]. The adaptive filter model considers a recurrent circuit among MFs, Golgi cells, and granule cells which generates resonant temporal patterns with various phase leads and lags. Therefore, the adaptive filter model assumes that the activities of PCs result from the interaction among MFs, Golgi cells, and granule cells. There is supportive evidence of the adaptive filter model [59, 60]. On the contrary, our finding of linear transformation from MFs to PCs suggests that the recurrent circuit plays a negligible role in generating temporal bases and rather that MFs already contain rich temporal repertoire that in turn drives the activities of PCs. The present study cannot exclude a possibility that different parts of the cerebellum may adopt different neural mechanisms for generating temporal patterns; the adaptive filter model has been tested in the flocculus, whereas our data was recorded from the cerebellar hemisphere.

**Acknowledgments** The authors thank Drs. Yoshikazu Shinoda, Hiroyuki Kambara, and Satoshi Hirose for helpful discussions and encouragements and the members of Swartz Center for Computational Neuroscience at the University of California San Diego for their valuable comments, where a part of this work was conducted. The authors also thank two anonymous reviewers for helpful comments in improving a previous version of the manuscript.

**Grants** This work is supported by JSPS KAKENHI Grant-in-Aid for Scientific Research (No. 25430007, No. 26120005, and No. 16K12476 to H.T.; No. 24650224 to T.I.; No. 14580784, No. 15016008, No. 16015212, No. 20033029, No. 21500319, No. 26120003 to S.K.), the Japan Science and Technology Agency (PRESTO: Intelligent Cooperation and Control) and the Ministry of Education, Culture, Sports, Science and Technology (MEXT), NBRP “Japanese Monkeys” through the National BioResource Project of the MEXT Japan, the JSPS Programs (Program for Advancing Strategic International Networks to Accelerate the Circulation of Talented Researchers, and Embodied-Brain Systems Science), and the Hitachi-Kurata and the Tateishi Science Foundations.

**Author Contribution** HT and SK designed the study. TI and SK provided the data. HT analyzed the data. HT, TI, and SK discussed the results. HT drafted the manuscript. HT, TI, and SK revised the manuscript.

## Compliance with Ethical Standards

**Conflict of Interest** The authors declare that they have no conflict of interest.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References

- Holmes G. The symptoms of acute cerebellar injuries due to gunshot injuries. *Brain*. 1917;40(4):461–535.
- Lang CE, Bastian AJ. Cerebellar subjects show impaired adaptation of anticipatory EMG during catching. *J Neurophysiol*. 1999;82(5):2108–19.
- Martin TA, Keating JG, Goodkin HP, Bastian AJ, Thach WT. Throwing while looking through prisms. I. Focal olivocerebellar lesions impair adaptation. *Brain*. 1996;119(Pt 4):1183–98.
- Maschke M, Gomez CM, Ebner TJ, Konczak J. Hereditary cerebellar ataxia progressively impairs force adaptation during goal-directed arm movements. *J Neurophysiol*. 2004;91(1):230–8.
- Morton SM, Bastian AJ. Cerebellar contributions to locomotor adaptations during splitbelt treadmill walking. *J Neurosci*. 2006;26(36):9107–16.
- Tseng YW, Diedrichsen J, Krakauer JW, Shadmehr R, Bastian AJ. Sensory prediction errors drive cerebellum-dependent adaptation of reaching. *J Neurophysiol*. 2007;98(1):54–62.
- Wolpert DM, Miall RC, Kawato M. Internal models in the cerebellum. *Trends Cogn Sci*. 1998;2(9):338–47.
- Wolpert DM, Miall RC. Forward models for physiological motor control. *Neural Netw*. 1996;9(8):1265–79.
- Bastian AJ. Learning to predict the future: the cerebellum adapts feedforward movement control. *Curr Opin Neurobiol*. 2006;16(6):645–9.
- Ishikawa T, Tomatsu S, Izawa J, Kakei S. The cerebro-cerebellum: could it be loci of forward models? *Neurosci Res*. 2016;104:72–9.
- Synofzik M, Lindner A, Thier P. The cerebellum updates predictions about the visual consequences of one's behavior. *Curr Biol*. 2008;18(11):814–8.
- Kawato M, Kuroda T, Imamizu H, Nakano E, Miyauchi S, Yoshioka T. Internal forward models in the cerebellum: fMRI study on grip force and load force coupling. *Prog Brain Res*. 2003;142:171–88.
- Miall RC, Christensen LO, Cain O, Stanley J. Disruption of state estimation in the human lateral cerebellum. *PLoS Biol*. 2007;5(11):e316.
- Ishikawa T, Tomatsu S, Tsunoda Y, Lee J, Hoffman DS, Kakei S. Releasing dentate nucleus cells from Purkinje cell inhibition generates output from the cerebrocerebellum. *PLoS One*. 2014;9(10):e108774.
- Tomatsu S, Ishikawa T, Tsunoda Y, Lee J, Hoffman DS, Kakei S. Information processing in the hemisphere of the cerebellar cortex for control of wrist movement. *J Neurophysiol*. 2016;115(1):255–70.
- Dayan P, Abbott LF. *Theoretical neuroscience*. Cambridge: MIT Press; 2001.
- Ohzawa I, DeAngelis GC, Freeman RD. Stereoscopic depth discrimination in the visual cortex: neurons ideally suited as disparity detectors. *Science*. 1990;249(4972):1037–41.
- Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Ser B Methodol*. 1996;267–88.
- Kakei S, Hoffman DS, Strick PL. Muscle and movement representations in the primary motor cortex. *Science*. 1999;285(5436):2136–9.
- Kakei S, Hoffman DS, Strick PL. Direction of action is represented in the ventral premotor cortex. *Nat Neurosci*. 2001;4(10):1020–5.
- Kakei S, Hoffman DS, Strick PL. Sensorimotor transformations in cortical motor areas. *Neurosci Res*. 2003;46(1):1–10.
- Jordan MI, Rumelhart DE. Forward models: supervised learning with a distal teacher. *Cogn Sci*. 1992;16(3):307–54.
- Pasalar S, Roitman AV, Durfee WK, Ebner TJ. Force field effects on cerebellar Purkinje cell discharge with implications for internal models. *Nat Neurosci*. 2006;9(11):1404–11.
- Yamamoto K, Kawato M, Kotosaka S, Kitazawa S. Encoding of movement dynamics by Purkinje cell simple spike activity during fast arm movements under resistive and assistive force fields. *J Neurophysiol*. 2007;97(2):1588–99.
- Evarts EV. Relation of pyramidal tract activity to force exerted during voluntary movement. *J Neurophysiol*. 1968;31(1):14–27.
- Kalaska JF, Cohen DA, Hyde ML, Prud'homme M. A comparison of movement direction-related versus load direction-related activity in primate motor cortex, using a two-dimensional reaching task. *J Neurosci*. 1989;9(6):2080–102.
- Shidara M, Kawano K, Gomi H, Kawato M. Inverse-dynamics model eye movement control by Purkinje cells in the cerebellum. *Nature*. 1993;365(6441):50–2.
- Liu X, Robertson E, Miall RC. Neuronal activity related to the visual representation of arm movements in the lateral cerebellar cortex. *J Neurophysiol*. 2003;89(3):1223–37.
- Schmolesky MT, Wang Y, Hanes DP, Thompson KG, Leutgeb S, Schall JD, et al. Signal timing across the macaque visual system. *J Neurophysiol*. 1998;79(6):3272–8.
- Soso MJ, Fetz EE. Responses of identified cells in postcentral cortex of awake monkeys during comparable active and passive joint movements. *J Neurophysiol*. 1980;43(4):1090–110.
- Ramnani N. The primate cortico-cerebellar system: anatomy and function. *Nat Rev Neurosci*. 2006;7(7):511–22.
- Houck BD, Person AL. Cerebellar loops: a review of the nucleocortical pathway. *Cerebellum*. 2014;13(3):378–85.
- Gao Z, Proietti-Onori M, Lin Z, Ten Brinke MM, Boele HJ, Potters JW, et al. Excitatory cerebellar Nucleocortical circuit provides internal amplification during associative conditioning. *Neuron*. 2016;89(3):645–57.
- Hamori J, Takacs J. Two types of GABA-containing axon terminals in cerebellar glomeruli of cat: an immunogold-EM study. *Exp Brain Res*. 1989;74(3):471–9.
- Kennedy A, Wayne G, Kaifosh P, Alvina K, Abbott LF, Sawtell NB. A temporal basis for predicting the sensory consequences of motor commands in an electric fish. *Nat Neurosci*. 2014;17(3):416–22.
- Chen S, Augustine GJ, Chadderton P. Serial processing of kinematic signals by cerebellar circuitry during voluntary whisking. *Nat Commun*. 2017;8(1):232.
- Chen S, Augustine GJ, Chadderton P. The cerebellum linearly encodes whisker position during voluntary movement. *elife*. 2016;5:e10509.
- Raymond JL, Medina JF. Computational principles of supervised learning in the cerebellum. *Annu Rev Neurosci*. 2018;41:233–53.
- Laurens J, Heiney SA, Kim G, Blazquez PM. Cerebellar cortex granular layer interneurons in the macaque monkey are functionally

- driven by mossy fiber pathways through net excitation or inhibition. *PLoS One*. 2013;8(12):e82239.
40. Lisberger SG, Fuchs AF. Role of primate flocculus during rapid behavioral modification of vestibuloocular reflex. II. Mossy fiber firing patterns during horizontal head rotation and eye movement. *J Neurophysiol*. 1978;41(3):764–77.
  41. Dugue GP, Tihy M, Gourevitch B, Lena C. Cerebellar re-encoding of self-generated head movements. *elife*. 2017;6.
  42. Herzfeld DJ, Kojima Y, Soetedjo R, Shadmehr R. Encoding of action by the Purkinje cells of the cerebellum. *Nature*. 2015;526(7573):439–42.
  43. Hong S, Negrello M, Junker M, Smilgin A, Thier P, De Schutter E. Multiplexed coding by cerebellar Purkinje neurons. *elife*. 2016;5.
  44. Medina JF. The multiple roles of Purkinje cells in sensori-motor calibration: to predict, teach and command. *Curr Opin Neurobiol*. 2011;21(4):616–22.
  45. Medina JF, Lisberger SG. Variation, signal, and noise in cerebellar sensory-motor processing for smooth-pursuit eye movements. *J Neurosci*. 2007;27(25):6832–42.
  46. Medina JF, Lisberger SG. Encoding and decoding of learned smooth-pursuit eye movements in the floccular complex of the monkey cerebellum. *J Neurophysiol*. 2009;102(4):2039–54.
  47. Sun Z, Smilgin A, Junker M, Dicke PW, Thier P. The same oculomotor vermal Purkinje cells encode the different kinematics of saccades and of smooth pursuit eye movements. *Sci Rep*. 2017;7:40613.
  48. Ebner TJ, Hewitt AL, Popa LS. What features of limb movements are encoded in the discharge of cerebellar neurons? *Cerebellum*. 2011;10(4):683–93.
  49. Heiney SA, Kim J, Augustine GJ, Medina JF. Precise control of movement kinematics by optogenetic inhibition of Purkinje cell activity. *J Neurosci*. 2014;34(6):2321–30.
  50. Kleine JF, Guan Y, Buttner U. Saccade-related neurons in the primate fastigial nucleus: what do they encode? *J Neurophysiol*. 2003;90(5):3137–54.
  51. Ten Brinke MM, Heiney SA, Wang X, Proietti-Onori M, Boele HJ, Bakermans J, et al. Dynamic modulation of activity in cerebellar nuclei neurons during Pavlovian eyeblink conditioning in mice. *elife*. 2017;6.
  52. Brodal P, Dietrichs E, Walberg F. Do pontocerebellar mossy fibres give off collaterals to the cerebellar nuclei? An experimental study in the cat with implantation of crystalline HRP-WGA. *Neurosci Res*. 1986;4(1):12–24.
  53. Shinoda Y, Sugiuchi Y, Futami T, Izawa R. Axon collaterals of mossy fibers from the pontine nucleus in the cerebellar dentate nucleus. *J Neurophysiol*. 1992;67(3):547–60.
  54. Marr D. A theory of cerebellar cortex. *J Physiol*. 1969;202(2):437–70.
  55. Albus JS. A theory of cerebellar function. *Math Biosci*. 1971;10(1–2):25–61.
  56. Ito M, Kano M. Long-lasting depression of parallel fiber-Purkinje cell transmission induced by conjunctive stimulation of parallel fibers and climbing fibers in the cerebellar cortex. *Neurosci Lett*. 1982;33(3):253–8.
  57. Ito M, Sakurai M, Tongroach P. Climbing fibre induced depression of both mossy fibre responsiveness and glutamate sensitivity of cerebellar Purkinje cells. *J Physiol*. 1982;324:113–34.
  58. Fujita M. Adaptive filter model of the cerebellum. *Biol Cybern*. 1982;45(3):195–206.
  59. Dean P, Porrill J. Decorrelation learning in the cerebellum: computational analysis and experimental questions. *Prog Brain Res*. 2014;210:157–92.
  60. Dean P, Porrill J, Ekerot CF, Jorntell H. The cerebellar microcircuit as an adaptive filter: experimental and computational evidence. *Nat Rev Neurosci*. 2010;11(1):30–43.